

Fujitsu Software

Technical Computing Suite V4.0L20

Job Operation Software

Command Reference

J2UL-2549-01ENZ0(08)
March 2023

Preface

Purpose of This Manual

This manual is a command and configuration file reference, describing the commands and files provided by the Job Operation Software for Technical Computing Suite. The manual also describes the command messages of these commands.

Intended Readers

This manual is intended for the administrator who operates and manages interconnects with the Job Operation Software.

The manual assumes that readers have the following knowledge:

- Basic Linux knowledge
- General knowledge of the Job Operation Software from the "Job Operation Software Overview"

Organization of This Manual

This manual is organized as follows.

[Chapter 1 How to Read Command Reference](#)

This chapter describes notes on specifying command options, how to read the command messages.

[Chapter 2 Lists of Commands and Configuration Files](#)

This chapter lists the commands and configuration files provided by the Job Operation Software.

[Chapter 3 Command Reference for End-user](#)

This chapter is a command reference for end-users, describing the commands and command messages. The chapter describes the messages displayed at the job execution time.

[Chapter 4 Command and Configuration File Reference for Administrator](#)

This chapter is a command, command message, and configuration file reference for administrators, describing the commands, messages, and files. The chapter is also a script reference, describing the exit script.

Notation Used in This Manual

Notation of users

The users of the Job Operation Software include the administrators responsible for system management and job operations and the end users who use the system to run programs. Unless otherwise noted, "user" in this manual means an end user.

Administrators

The Job Operation Software has different types of administrators: system administrator, cluster administrator, and job operation administrator. However, they may all be represented as just "administrator" in this document. In such cases, an administrator who manages the system usually means the system administrator or cluster administrator. An administrator who manages job operations means the cluster administrator or job operation administrator.

Representation of units

The following table lists the prefixes used to represent units in this manual. Basically, disk size is represented as a power of 10, and memory size is represented as a power of 2. Be careful about specifying them when displaying or entering commands.

Prefix	Value	Prefix	Value
K (kilo)	10^3	Ki (kibi)	2^{10}
M (mega)	10^6	Mi (mebi)	2^{20}
G (giga)	10^9	Gi (gibi)	2^{30}
T (tera)	10^{12}	Ti (tebi)	2^{40}
P (peta)	10^{15}	Pi (pebi)	2^{50}

Notation of model names

In this manual, the computer that based on Fujitsu A64FX CPU is abbreviated as "FX server", and FUJITSU server PRIMERGY as "PRIMERGY server" (or simply "PRIMERGY").

Also, specifications of some of the functions described in the manual are different depending on the target model. In the description of such a function, the target model is represented by its abbreviation as follows:

[FX]: The description applies to FX servers.

[PG]: The description applies to PRIMERGY servers.

Path names of the commands

In the examples of the operations, the path names of the commands in the directory /bin, /usr/bin, /sbin or /usr/sbin might not be represented by absolute path.

Symbols in This Manual

This manual uses the following symbols.



Note

The Note symbol indicates an item requiring special care. Be sure to read these items.



See

The See symbol indicates the written reference source of detailed information.



Information

The Information symbol indicates a reference note related to Job Operation Software.

Export Controls

Exportation/release of this document may require necessary procedures in accordance with the regulations of your resident country and/or US export control laws.

Trademarks

- Linux(R) is the registered trademark of Linus Torvalds in the U.S. and other countries.
- Red Hat and Red Hat Enterprise Linux are registered trademarks of Red Hat, Inc. in the U.S. and other countries.
- Intel is trademark of Intel Corporation or its subsidiaries in the U.S. and/or other countries.
- All other trademarks are the property of their respective owners.

Date of Publication and Version

Version	Manual code
March 2023, Version 1.8	J2UL-2549-01ENZ0(08)
September 2022, Version 1.7	J2UL-2549-01ENZ0(07)
March 2022, Version 1.6	J2UL-2549-01ENZ0(06)
November 2021, Version 1.5	J2UL-2549-01ENZ0(05)
August 2021, Version 1.4	J2UL-2549-01ENZ0(04)
March 2021, Version 1.3	J2UL-2549-01ENZ0(03)

Version	Manual code
September 2020, Version 1.2	J2UL-2549-01ENZ0(02)
June 2020, Version 1.1	J2UL-2549-01ENZ0(01)
February 2020, First version	J2UL-2549-01ENZ0(00)

Copyright

Copyright FUJITSU LIMITED 2020-2023

Update history

Changes	Location	Version
Improved description of Device under the Mount entry in the pxsystem.conf file.	4.3.16	1.8
Also fixed errata.	-	
Changed the description when specifying "del-node" in the Type item of the Operation section in the pxsystem.conf file and added an example of deleting nodes.	4.3.16	1.7
Changed plexec command messages as follows: Message ID 0050: Added pjaexe command for Meaning and Action. Message ID 0052: Added a message.	3.2.1	1.6
Added description of job statistical information about node power consumption on FX server.	3.3.2	
Added the descriptions for the -P option of the papwrctl command.	4.1.15	
Added notes to [NOTE] for the pasnap command.	4.1.21	
Added notes to [NOTE] for the pasnap command.	4.1.28	
The following fixes have been made for end-user and administrator commands. - Unification of descriptions in [SYNOPSIS]]. - Unify the --help option description.	3.1.2 3.1.3 3.1.5 3.1.7 3.1.9 3.1.11 3.1.13 3.1.14 4.1.5 4.1.12 4.1.13 4.1.25 4.1.26 4.1.27 4.1.30 4.1.31 4.1.32 4.1.33 4.1.38	1.5
Added the description of the mswk option to the pjsb command to indicate that only torus mode can be specified as the node allocation method for the master worker type job.	3.1.14	
The message ID 0045 of the pjsb command has been changed in order to provide a mechanism for continuing job operations by using other file systems, even if some file systems are unavailable.	3.1.14	
Also fixed errata.	-	

Changes	Location	Version
Fixed the "Action" for the following command messages: - pjalter command Message ID 0557 - pjsub command Message ID 0057 - pmalter command Message ID 1057	3.1.3 3.1.14 4.1.25	1.4
Added an example of using the pjstat command to specify values for the --filter option. Also, fixed the table description for item which can be specified with the --choose, --filter, and --sort options.	3.1.13	
Changed the description of the --mail-list option for the pjsub command.	3.1.14	
Changed the default values for the setting items JobSchedulingTargetLimit and JobSchedulingTargetMode in the papjm.conf and pmpjm.conf files.	4.3.5 4.3.13	
Also fixed errata.	-	
Improved the action for pjshowrsc command message ID 3299.	3.1.11	1.3
Fixed the description for the -S option of the pjstat command.	3.1.13	
Added the actions for plexec command message IDs 0022 and 0611.	3.2.1	
Added the description for the --force-all option of the papwrctl command.	4.1.15	
Added or changed the following to support "I/O-exclusive mode" in how nodes are allocated to jobs: - pjalter command Added the message ID: 0575 - pjsub command - node = quantity _ of _ resources for option {-L --rsc-list} "resource = value [,.]" Added the options for specifying I/O-exclusive mode for FX servers ": io-exclusive" and ": no-io-exclusive". - Changed message 0002. - pjstatsinfo - Limit on node allocation (item name: nalimit) Added the I/O-exclusive mode flag. - Number and shape of requested nodes (NODE NUM (REQUIRE)) Added the I/O-exclusive mode specified (:io-exclusive) to the output of the number and shape of requested nodes to the node-exclusive job. - Number and shape of allocated nodes (NODE NUM (ALLOC)) Added the I/O-exclusive mode specified (:io-exclusive) to the output of the number and shape of allocated nodes to the node-exclusive job. - pmalter command Added the message ID: 1075 - pmjacladm command Added the definition items of the following job ACL functions: - define allocation-io-mode - execute pjsub-io-exclusive - execute pjsub-io-exclusive - papjm.conf file Added the item ShelfPackingNodeAllocation. - pmpjm.conf file - Added the item "HelpPackingNodeAllocation" to the ResourceUnit and ResourceGroup sections. - Changed the description of the item ResourceGroupTsha.	3.1.3 3.1.14 3.3.2 4.1.25 4.1.30 4.3.5 4.3.13	1.2
Changed pjsub command messages as follows: - Message ID 0070: Changed the descriptions of meaning and action.	3.1.14	

Changes	Location	Version
Added a new item esnum to the Job statistical information for pjstatsinfo.	3.3.2	
Added the option --check to pachecker command. Changed the action for pachecker command message ID 4102.	4.1.1	
Added the argument all-status to papwrctl command.	4.1.15	
Added the description of the item JobEvaluation in the papjm.conf and pmpjm.conf files.	4.3.5 4.3.13	
Also fixed errata.	-	
Added the action for pjstat command message ID 0291.	3.1.13	1.1
Changed plexec command messages as follows: Message ID 0021: Changed the message ID to 0094 and the message. Added a note. Message ID 0080: Added a message. Message ID 0081: Improved the description of action.	3.2.1	
Added a method for calculating job statistics to the description of the pjstatsinfo items. Also fixed the description for item "Maximum memory usage by assistant cores (item name: ammszu)".	3.3.2	
Fixed the descriptions for the -a option of the paclstmgr command. Added a note on specifying options. Added the message ID 4018.	4.1.3	
Added a note about specifying options for the papwrctl command. Added the following papwrctl command message: Message IDs 2017 and 2018 Fixed the action for message ID 2007.	4.1.15	
Added a note for the pasnap command	4.1.21	
Added the description of the pmsetstats command.	4.1.37	
Changed the descriptions for the pxinst command as follows: <ul style="list-style-type: none"> - Changed the subcommand 'node add' of the pxinst command to cause the pxinst command to terminate abnormally without overwriting the setting if a host name with the same name was registered. - Fixed the description of --rpm option in subcommand 'kmod add'. - The --bootopt option of the pxinst command also supports FX servers. - Added message IDs 1036, 1037, and 1038 of the pxinst command. - Deleted message ID 1021. 	4.1.39	
Fixed the description for item NodeID in CustomResource section in pmpjm.conf file	4.3.13	
Also fixed errata.	-	

All rights reserved.
The information in this manual is subject to change without notice.

Contents

Chapter 1 How to Read Command Reference.....	1
1.1 Notes on Specifying Command Options.....	1
1.2 How to Read Command Messages.....	1
Chapter 2 Lists of Commands and Configuration Files.....	2
Chapter 3 Command Reference for End-user.....	7
3.1 Commands.....	7
3.1.1 pjacl command.....	7
3.1.2 pjaexe command.....	10
3.1.3 pjalter command.....	13
3.1.4 pjdel command.....	22
3.1.5 pjexe command [PG].....	26
3.1.6 pjhold command.....	28
3.1.7 pijbind command [PG].....	33
3.1.8 pjrls command.....	37
3.1.9 pjrrsh command.....	40
3.1.10 pjshowip command.....	43
3.1.11 pjshowrsc command.....	45
3.1.12 pjsig command.....	54
3.1.13 pjstat command.....	58
3.1.14 pjsub command.....	73
3.1.15 pjwait command.....	102
3.2 Messages in job outputs.....	106
3.2.1 plexec command.....	106
3.2.2 Parallel execution environment.....	119
3.2.3 Job management function.....	120
3.2.4 Tofu Interconnect [FX].....	124
3.3 Others.....	125
3.3.1 mpiexec.tcs_intel [PG].....	125
3.3.2 pjstatsinfo.....	126
Chapter 4 Command and Configuration File Reference for Administrator.....	199
4.1 Commands.....	199
4.1.1 pachecker command.....	199
4.1.2 paclone command.....	207
4.1.3 paclstmgr command.....	215
4.1.4 paconsole command.....	227
4.1.5 padeadline command.....	228
4.1.6 padumpmgr command.....	236
4.1.7 paeventadm command.....	240
4.1.8 paledctl command [FX].....	242
4.1.9 pamonitor_notice command.....	244
4.1.10 pamonitoradm command.....	246
4.1.11 pamopluginadm command.....	247
4.1.12 papjmadm command.....	248
4.1.13 papjmstatsadm command.....	252
4.1.14 papwradm command.....	256
4.1.15 papwrctl command.....	257
4.1.16 papwrmgradm command.....	267
4.1.17 parender command.....	271
4.1.18 parscadm command.....	275
4.1.19 paserviceadm command.....	282
4.1.20 pashowclst command.....	283
4.1.21 pasnap command.....	290
4.1.22 pastart command.....	295

4.1.23	pastop command.....	300
4.1.24	pasyspwr command.....	304
4.1.25	pmalter command.....	311
4.1.26	pmdumpjobinfo command.....	320
4.1.27	pmerls command.....	326
4.1.28	pmexe command.....	329
4.1.29	pmgather command.....	337
4.1.30	pmjacladm command.....	342
4.1.31	pmpjmadm command.....	375
4.1.32	pmpjmopt command.....	379
4.1.33	pmresume command.....	384
4.1.34	pmrscadm command.....	387
4.1.35	pmscatter command.....	396
4.1.36	pmsetjobinfo command.....	400
4.1.37	pmsetstats command.....	403
4.1.38	pmsuspend command.....	405
4.1.39	pxinst command.....	409
4.1.40	pxsystemadm command.....	425
4.2	Common command messages.....	429
4.2.1	Common command messages of the system management function.....	429
4.2.2	Common command messages of the install function.....	435
4.3	Configuration Files.....	438
4.3.1	How to Write in the Configuration Files.....	439
4.3.2	paevent.conf file.....	439
4.3.3	pamonitor.conf file.....	443
4.3.4	pamoplugin.conf file.....	446
4.3.5	papjm.conf file.....	449
4.3.6	papjmstats.conf file.....	460
4.3.7	papwr.conf file.....	467
4.3.8	papwrm.conf file.....	470
4.3.9	parsc.conf file.....	479
4.3.10	paservice.conf file.....	481
4.3.11	patrouble.conf file.....	483
4.3.12	pmpjcmd.conf file.....	485
4.3.13	pmpjm.conf file.....	486
4.3.14	pmrsc.conf file.....	501
4.3.15	pxinst.conf file.....	503
4.3.16	pxsystem.conf file.....	505
4.4	Exit Scripts.....	516
4.4.1	postfree.....	516
4.4.2	prealloc.....	520
4.4.3	predel.....	524

Chapter 1 How to Read Command Reference

This manual is a command and configuration file reference, describing the commands and files provided by the Job Operation Software. The manual also describes the command messages displayed at the command execution time. This chapter describes notes on specifying command options, how to read the command messages.



Information

The descriptions in the command and configuration file reference in this manual have the same contents as the man pages. You can view a man page in the following way.

```
$ man command or configuration-file
```

1.1 Notes on Specifying Command Options

Regarding how to specify the options of each command, be careful with the format when specifying multiple values in an argument. This is because how they are specified in an argument of the command itself is different from how they are specified in an argument of an option.

[COMMAND ARGUMENTS]

When specifying multiple arguments, delimit them with a space.

```
$ command value1 value2 value3 ...
```

[OPTION ARGUMENTS]

When specifying multiple arguments, concatenate them with a comma (.).

```
$ command -option value1,value2,value3,...
```

1.2 How to Read Command Messages

Command messages are output in the following format.

```
[priority] component-name message-ID command-name message-text
```

- *priority*: Indication of the level of importance of the message
- *component-name*: Name indicating the respective function of the Job Operation Software
- *message-ID*: Message identifier
- *command-name*: Name of the command for which the message is displayed
- *message-text*: Displayed information or error details

The following table lists details of the priority of command messages.

Table 1.1 Priority of command messages

Priority	Meaning
INFO	Information message: The message displays process-related information.
NOTE	Notice message: The message provides notification of the start or stop of a process.
WARN	Warning message: A process encountered a problem but has continued.
ERR.	Error message: A process encountered a problem and could not continue. (Note) The notation of priority has "ERR" followed by a dot (.).

Chapter 2 Lists of Commands and Configuration Files

This chapter lists the commands and configuration files described in this manual.

[For end-user]

Users and all administrators can use the following commands.

Table 2.1 Commands for referring to job information and commands for job operation

Command name	Expected use scenario	Node where operation is permitted
pjacl	To confirm the contents of the job ACL database	Login node Compute cluster management node
pjaexe	To generate a worker process of a master-worker job, and to execute the generated worker process	Login node Compute node
pjalter	To change the parameter of the job	Login node Compute cluster management node
pjdel	To delete jobs	Login node Compute cluster management node
pjexe [PG]	To execute the sequential program all together on the plural virtual nodes	Compute node
pjhold	To hold submitted jobs	Login node Compute cluster management node
pjpbinding [PG]	To bind CPU cores to processes	Compute node
pjrls	To cancel the job hold	Login node Compute cluster management node
pjrsh	[FX] To execute a remote program at any compute node in the same job from an MPI process [PG] To execute MPI programs using the pjrsh command instead of rsh/ssh command when the programs are created by not using Development Studio	Compute node
pjshowip	To search for the IP address of a compute node allocated to a job	Compute node
pjshowrsc	To confirm the entire amount or the usage status of activated computer resources (*1)	Login node Compute cluster management node System management node
pjsig	To send signals to running jobs	Login node Compute cluster management node
pjstat	To confirm the state of submitted jobs	Login node Compute cluster management node System management node
pjsub	To submit jobs	Login node Compute cluster management node
pjwait	To wait for a job to complete	Login node Compute cluster management node

(*1)

In the pjshowrsc command, there is an option that the job operation administrator and the end-user cannot specify.

(*2)

Aside from messages for the commands described above, messages of the parallel execution environment and job manager function are output at the job execution time. For details on these messages, see "[3.2 Messages in job outputs.](#)"

Table 2.2 Others

Command name	Expected use scenario	Node where operation is permitted
mpiexec.tcs_intel [PG]	To execute Intel MPI by the view similar to MPI of Development Studio	Compute node
pjstatsinfo	To confirm the content of the output of the job statistical information by the pjstat command, the pjsub command, and the pmdumpjobinfo command	-

[For administrator]

Table 2.3 Commands and configuration files for system management

Command/file name	Expected use scenario	Node where operation is permitted
pxinst.conf	To configure the operating environment of the install function and to define the basic settings for installation target nodes	System management node (*)
pxinst	To set up OS installation	System management node (*)
pxsystem.conf	To define cluster configuration settings, such as to register a new cluster or add/delete a node in a registered cluster	System management node (*)
pxsystemadm	To apply the contents of the cluster configuration definition file (pxsystem.conf) to the active system management node and the nodes in a cluster, and also to configure FX server control	System management node (*)
pashowclst	To check whether the cluster node OS and services (functions) are working correctly, and also to check the cluster configuration	System management node
paclstmgr	To prevent the use of nodes during hardware and software maintenance, and also to prevent the assignment of new jobs to a compute node to be maintained	System management node (*)
papwr.conf	To define settings for power control	System management node (*)
papwradm	To reflect the contents of the system control configuration file (papwr.conf) to the power control function, and also to check the set contents	System management node (*)
papwrctl	Use it for power control of the specified node	System management node (*)
pastart	Use it for power-on of the specified cluster and node group	System management node (*)
pastop	Use it for power-off of the specified cluster and node group	System management node (*)
paledctl [FX]	To control and check the state of the location LED for the FX server	System management node
pamonitor.conf	To define settings for the system monitoring function, such as the OS or service check	System management node (*)
pamonitoradm	To reflect the contents of the system monitoring configuration file (pamonitor.conf) used for the system monitoring function, and also to check the set contents	System management node (*)
pamoplugin.conf	To define the settings of a plugin that runs when an error is detected in the state of the OS or a service	System management node (*)
pamopluginadm	To apply the contents of the system monitoring plugin configuration file (pamoplugin.conf) used by the monitoring plugin function, and also to check the set contents	System management node (*)

Command/file name	Expected use scenario	Node where operation is permitted
pamonitor_notice	To isolate the specified node from job operation and to automatically collect a dump, with the node regarded as being down	Arbitrary node
paservice.conf	To define settings for the user definition service	System management node (*)
paserviceadm	To reflect the contents of the user definition service configuration file (paservice.conf) used for the user definition service, and also to check the set contents	System management node (*)
pmexe	To batch execute a command for the specified node and node group	System management node
pmscatter	To batch transfer a file to the specified node and node group	System management node
pmgather	To batch collect a file from the specified node and node group	System management node
paconsole	To connect a console to the specified node	System management node
padumpmgr	To batch process dumps	System management node (*)
pasnap	To batch collect investigation data, and also to collect investigation data according to the trouble events defined in the trouble event definition file (patrouble.conf)	System management node Multiuse node
patrouble.conf	To define trouble events requiring the pasnap command to collect investigation data	System management node
paevent.conf	To define settings for monitoring the special messages output to the system log, Job Operation Software log, etc.	System management node (*)
paeventadm	To reflect the contents of the log monitoring condition definition file (paevent.conf), and also to check the set contents	System management node (*)
parender	To deploy the keywords in a file and save them in information on the node executing the command	Arbitrary node
pachecker	To check the contents of the Job Operation Software package and settings to apply them to the target nodes, and also to define the contents you want to check regarding nodes, packages, and settings in any definition file (check definition file) and to check them	System management node (*)
paclone	To create a disk image of a backup source (node, device, partition, etc.), and also to restore the created disk image	System management node (*)

(*) In a redundant configuration, you can execute commands and place files only on the active system management node.

Table 2.4 Commands and configuration files for power management

Command/file name	Expected use scenario	Node where operation is permitted
papwrn.conf	To configure the power management function	System management node (*)
papwrngradm	To reflect the contents of the power management function configuration file (pwrn.conf), and also to check its settings	System management node (*)
pasyspwr	To check the power consumption of the system, compute nodes, or external devices	System management node (*)

(*) In a redundant configuration, you can execute commands and place files only on the active system management node.

Table 2.5 Commands and configuration files for job management

Command/file name	Expected use scenario	Node where operation is permitted
padeadline	To set deadline schedules	System management node Compute cluster management node
papjm.conf	To define the operation of the job operations of an entire cluster	System management node
papjmadm	To set and change the job operations of an entire cluster, and also to check the contents of set definitions	System management node
parsc.conf	To define the operating environment, including the computer resource allocation amounts for the jobs of a running cluster	System management node (*)
parscadm	To set and change computer resources for the jobs of a running cluster. To make settings and changes, and also to check the contents of set definitions	System management node (*)
papjmstats.conf	To define the save settings, etc. of statistical information for an entire cluster.	System management node
papjmstatsadm	To set or update statistical information for an entire cluster. Also use it to check the contents of the set definitions	System management node
pmpjm.conf	To define the operation of job operations by resource unit	System management node
pmpjmadm	To set and change job operations by resource unit, and also to check the contents of set definitions	System management node
pmrsc.conf	To define the operation of computer resources for the jobs in a resource unit	System management node (*)
pmrscadm	To define the operation of the operating environment, including the computer resource allocation amounts for the jobs in a resource unit, and also to check the contents of set definitions	System management node (*)
pmjacladm	To set and change items in the job ACL database in a resource unit, and also to check the contents of set definitions in the job ACL database	System management node Compute cluster management node
pmpjmopt	To set whether jobs can be submitted or can be executed To set the fair share value	System management node
pmpjcmd.conf	To define the operating environment of the command API	System management node Compute cluster management node Login node
pmdumpjobinfo	To output the contents of the statistical information file	System management node Compute cluster management node
pmalter	To change job parameters, such as the elapsed time limit value and the priority of a submitted job	System management node Compute cluster management node
pmerls	To cancel the error state of a job	System management node Compute cluster management node
pmsuspend	To suspend a running job or sub job	System management node Compute cluster management node
pmresume	To resume a suspended job or sub job	System management node Compute cluster management node
pmsetjobinfo	To set job statistical information (REASON item) from the exit script of the job resource manager exit function	Compute node

Command/file name	Expected use scenario	Node where operation is permitted
pmsetstats	To set job statistical information (administrator definition item) from the exit script of the job resource manager exit function	Compute node

(*) In a redundant configuration, you can execute commands and place files only on the active system management node.

Table 2.6 Exit scripts

Script name	Expected use scenario	Node where operation is permitted
postfree	Script for writing the process performed after a resource for a job is released	System management node (*)
prealloc	Script for writing the process performed before a resource for a job is allocated	System management node (*)
predel	Script for writing the process performed when deleting or holding a job	System management node (*)

(*) In a redundant configuration, you can execute commands and place files only on the active system management node.

Chapter 3 Command Reference for End-user

This chapter provides command references and command messages for end-users. Aside from the command messages, the chapter also describes the messages displayed at the job execution time.

3.1 Commands

This section provides a command reference and command messages for end-users.

Commands are listed in alphabetical order.

3.1.1 pjacl command

[NAME]

pjacl - Displays the content of the job ACL function.

[SYNOPSIS]

```
/usr/bin/pjacl [{--rscunit|--ru} rscuname]  
               [{--rscgrp|--rg} rscgname]  
               [--data [--delimiter delimiter]]  
               [-g gname | --gname gname]  
               [-u uname | --uname uname]  
  
/usr/bin/pjacl --help
```

[DESCRIPTION]

The pjacl command displays the content of defined items configured in the job ACL function. The displayed targets include the user who executes this command and users which are in the same group. Besides, other users or groups can be included according to the permission of Administrator.

All kinds of effective defined items of the target user will be displayed. If users have not defined their own items, the default value will become effective and be displayed. If the definition item execute and permit are not defined for the root user, the default value are 'enable' and 'allow all' respectively, regardless of what the pmjacladm or pjacl command displays. This command can be executed on login node and the compute cluster management node.

[OPTION]

{--rscunit|--ru} *rscuname*

Use this option to specify the resource unit which will be the output object. Only one resource unit can be specified.

If this option is omitted, the default resource unit of the user will become effective.

{--rscgrp|--rg} *rscgname*

Use this option to specify the resource group which will be the output object. Only one resource group can be specified.

If this option is omitted, the default resource group of the user will become effective.

--data

Use this option to format the output.

The default delimiter is comma (","), yet it can be changed with the option --delimiter. There are characters such as comma and colon which are included in the displayed items, so be careful with the delimiter.

--delimiter *delimiter*

Use this option to specify the delimiter when the option --data is set.

If no delimiter is specified, comma (",") will be the default one.

A character string beginning with a hyphen ("-") can not be specified.

-u *uname* | --uname *uname*

Use the option *uname* to specify the target user. Only one can be specified. User names which just include digits will be treated as an user ID. If one user name does just include digits, use the option --uname instead of -u.

If this option is omitted, the target will be the user who runs this command. Generally, Common users can only show their own items, however, other users or groups can be included if the authority is granted by the administrator.

-g *gname* | --gname *gname*

Use the option *gname* to specify the target group of users. Only one can be specified. Group names which just include digits will be treated as a group ID. If one user name does just include digits, use the option --gname. The relationship between the specified user and group will not be checked.

If this option is omitted, the target group will be the current group of the user.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[END STATUS]

0: Normal end

1: Abnormal end

[REMARKS]

The output target (user/group) will be different according to the permission of Administrator.

[SEE ALSO]

pmjacladm(8)

[MESSAGES]

If an error occurs during the execution of the pjacl command, the following messages is output to the standard error output:

Error Message

**[ERR.] PJM 3201 pjacl Unknown option or no option argument: *opt*.
Try `pjacl --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 3202 pjacl Invalid combination of options: *opt* and *opt*.
Try `pjacl --help` for more information.**

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

**[ERR.] PJM 3203 pjacl Unknown option argument: *opt arg*.
Try `pjacl --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 3204 pjacl Invalid option argument: *opt arg*.**Meaning**

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

[ERR.] PJM 3240 pjacl User(*uname*) does not exist.**Meaning**

The specified user name *uname* does not exist.

Action

Confirm that the user name *uname* is registered in the system. Alternatively, check whether it can be retrieved from a name service.

[ERR.] PJM 3241 pjacl Group(*gname*) does not exist.**Meaning**

The specified group name *gname* does not exist.

Action

Confirm that the group name *gname* is registered in the system. Alternatively, check whether it can be retrieved from a name service.

[ERR.] PJM 3270 pjacl No execute permission.**Meaning**

The user is not permitted to execute pjacl command.

Action

Confirm the permission to execute the pjacl command by contacting the administrator.

[ERR.] PJM 3271 pjacl No permission: *user group*.**Meaning**

The user is not permitted to display the specified *user* or *group*.

Action

Display the job ACL function settings with the pjacl command. Then, confirm the permission (permit pjacl) for the displayed objects of the pjacl command for the specified user or group in the range of the target resource unit or resource group. If necessary, ask the administrator to add permission.

[ERR.] PJM 3291 pjacl Internal error: *details*.**Meaning**

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 3293 pjacl The operating command cannot be executed except login node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a login node or compute cluster management node.

Action

Execute the command on a login node or compute cluster management node.

[ERR.] PJM 3297 pjacl Resource unit does not exist: *rscuname*.

Meaning

The specified resource unit *rscuname* or the default resource unit *rscuname* defined by the job ACL function does not exist.

Action

Specify the resource unit again correctly. Alternatively, check whether the default resource unit exists.

[ERR.] PJM 3298 pjacl Resource group does not exist: *rscgname*.

Meaning

The specified resource group *rscgname* or the default resource group *rscgname* defined by the job ACL function does not exist.

Action

Specify the resource group again correctly. Alternatively, check whether the default resource group exists.

3.1.2 pjaexe command

[NAME]

pjaexe - Generates a worker process for a master-worker job.

[SYNOPSIS]

```
[Coordinate specification]
/usr/bin/pjaexe --vcoord coord command

[File specification]
/usr/bin/pjaexe --vcoordfile vcoordfile command

[Usage]
/usr/bin/pjaexe --help
```

[DESCRIPTION]

The pjaexe command can be executed only in the job script of a master-worker job or its child process. *command* is executed as a worker process on the node specified in the --vcoord or --vcoordfile option. The pjaexe command returns when the command process is generated. The --vcoord option specifies the node coordinates where *command* starts. Use the --vcoordfile option to collectively specify multiple sets of coordinates. The processes generated directly by the pjaexe command are numbered sequentially from 0. Their numbers are called rank numbers. When using the --vcoord option, the pjaexe command directly generates only one process, so its rank number is always 0. When using the --vcoordfile option, the *command* generates processes in the specified node order, so their set rank numbers are 0 to <number of specified nodes-1>. Worker processes can reference the rank numbers with the environment variable PLE_VPID.

[OPTIONS]

--vcoord *coord*

This option specifies the coordinates of the node executing *command*, in the format of x (one dimension), x,y (two dimensions), or x,y,z (three dimensions). These coordinates are coordinates in the node shape specified by the -L node or --rsc-list node option of the pjsub command. The number of dimensions must be the same as that of the shape. Specify coordinates in *coord* as follows.

- When *coord* is one dimensional (x), specify "(x)" or x.
- When *coord* is two dimensional (x,y), specify "(x,y)" or x,y.
- When *coord* is three dimensional (x,y,z), specify "(x,y,z)" or x,y,z.

The --vcoord option cannot be specified more than once.

--vcoordfile *vcoordfile*

This option executes *command* on the node located at coordinates written in the file *vcoordfile*. The file *vcoordfile* has one set of coordinates written per line in the following format.

Number of dimensions	Format
1 dimension	(x)
2 dimension	(x,y)
3 dimension	(x,y,z)

(Note) The written coordinates for the --vcoordfile option must be enclosed in parentheses.

[Coding example]

```
(0,0,0)
(1,0,0)
(2,0,0)
```

Note the following about writing the coordinates.

- If the number of written sets of coordinates exceeds the number of allocatable nodes, an error occurs.
- Multiple sets of the same coordinates can be written. In such cases, *command* is executed multiple times on the same node.
- Only coordinates (parentheses, commas, and numerical values) can be written. Space and tab characters cannot be written.

The --vcoordfile option cannot be specified more than once.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[NOTE]

/dev/zero is set as the standard input of command executed by the pjaexe command, and /dev/null is set as the standard output and standard error output.

The following coding example directs the standard output and standard error of a.out to a file when the pjaexe command is executed from a job script.

```
mkdir ./stdout ./stderr
pjaexe --vcoord "(0)" './a.out >> ./stdout/0.txt 2>> ./stderr/0.txt'
```

[END STATUS]

0: *command* could start normally. Alternatively, 0 is also returned if *command* does not exist.

1: The specified coordinate values in the --vcoord or --vcoordfile option or their specification method is incorrect. Alternatively, an incorrect option is specified.

255: An error in the parallel execution environment occurred.

If a signal terminates the pjaexe command, the end status is the logical addition (OR) of 0x80 and the signal number. Furthermore, the pjaexe command cannot be used to learn the end status of *command*.

[MESSAGES]

If an error occurs during the execution of the pjaexe command, the following messages is output to the standard error output. Also, information messages that report process completion, etc. are output to the standard output.

Error Message

[ERR.] PLE 9200 pjaexe The --vcoord or --vcoordfile must be specified.

Meaning

Neither the --vcoord option nor the --vcoordfile option is specified in the pjaexe command.

Action

Specify the --vcoord option or --vcoordfile option in the pjaexe command, and execute the job again.

[ERR.] PLE 9201 pjaexe command must be specified.

Meaning

No program is specified in the pjaexe command.

Action

Specify a program in the pjaexe command, and execute the job again.

[ERR.] PLE 9202 pjaexe *opt* option is not supported.

Meaning

The option *opt* specified in the pjaexe command is not supported.

Action

Specify a correct option in the pjaexe command, and execute the job again.

[ERR.] PLE 9203 pjaexe Duplicated an option.

Meaning

An option is specified more than once in the pjaexe command.

Action

Check the options specified in the pjaexe command.

[ERR.] PLE 9230 pjaexe A system error occurred.(code)

Meaning

A system error occurred.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance" and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PLE 9231 pjaexe The plexec cannot be executed.(CODE=X,Y,Z)

Meaning

The plexec command cannot be executed.

X,Y,Z: Internal codes

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance" and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PLE 9232 pjaexe cannot get memory.(CODE=X,Y,Z)

Meaning

There is insufficient memory.

X, Y, Z: Internal codes

Action

Increase the memory amount allocated to the job, and execute the job again. If the same error recurs, contact the administrator. The administrator is requested to review system memory estimates.

Information Message

[INFO] PLE 9249 pjaexe Try 'pjaexe --help' for more information.

Meaning

There is an error in how the pjaexe command is used.

Action

Referring to the message displayed immediately before this message, confirm the correct use of the pjaexe command.

3.1.3 pjalter command

[NAME]

pjalter - Changes the parameters of a submitted job.

[SYNOPSIS]

```
/usr/bin/pjalter [{-L|--rsc-list} {elapse|rscunit|ru|rscgrp|rg}=value[,...]]  
                [-p priority]  
                jobid ...  
/usr/bin/pjalter --help
```

[DESCRIPTION]

The pjalter command changes the following parameters of a submitted job:

- Job elapsed time limit ({-L|--rsc-list} elapse=*elapselimit*)
- Name of the resource unit executing the job ({-L|--rsc-list} {rscunit|ru}=*rscuname*)
- Name of the resource group executing the job ({-L|--rsc-list} {rscgrp|rg}=*rscgname*)
- Priority among users of the job (-p *priority*)

The pjalter command can be executed on the login node and compute cluster management node.

The job ID, a sub job ID of a step job or a sub job ID of a bulk job is specified for *jobid*.

If the job ID of a bulk job or step job is specified, all the sub jobs in the job are targets (except the ended sub jobs in the step job).

You can change the parameters, except resource unit names, of a sub job of a step job by specifying its sub job ID.

When specifying the sub job ID of a bulk job, only the elapsed time limit value can be changed.

The pjalter command cannot be used for interactive jobs.

The pjalter command cannot be used for jobs that specify the minimum value of elapsed time (example: specifying only minimum value 10-, minimum value and maximum value 10-100).

For each parameter, the states of jobs and sub jobs that pjalter command can be accepted are listed as follows. If the targets are not in any of these states, the pjalter command ends with an error.

- Job elapsed time limit
You can change parameters with the pjalter command only if the target jobs and sub jobs are in the QUEUED, RUNNING, HOLD, or ERROR state.
Changing the elapsed time limit value in the RUNNING state is possible only when the function to change the elapsed time limit value for jobs in the RUNNING state is enabled and the jobs are on the FX server compute node and the elapsed time limit value is shortened.

- Name of the resource unit executing the job
- Name of the resource group executing the job
- Priority among users of the job

You can change parameters with the `pjalter` command only if the target jobs and sub jobs are in the QUEUED, HOLD, or ERROR state.

If you have changed the parameters of a job in the QUEUED state, the job is rescheduled.

The specified parameter values are checked according to the job ACL function settings, and an error occurs in the following cases.

- The job elapsed time limit or job priority exceeds the specifiable upper limit.
- The owner of the job is not permitted to use the specified resource unit or resource group.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

1-100

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

100_1-10

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 100)

100[1-10]

When multiple job IDs (sub job IDs) or a range of job IDs (sub job IDs) are specified, the processing is not aborted, and all target jobs (sub jobs) are processed, even if the error occurs in a certain job (sub job).

- Multiple specification
Because the job IDs (sub job IDs) are specified explicitly, the error message is displayed when the error occurs.
- Range specification
Because the error message is not displayed, it is necessary to confirm the job attributes by the `pjstat` command.

[OPTIONS]

The option specification method conforms to that of the `pjsub` command.

{-L|--rsc-list} "*resource=value[,...]*"

This option specifies the upper limit values of the resources required by jobs.

The numerical values or character strings that can be specified in *value* for each *resource* are different.

The resources shown below can be specified for *resource*. When specifying multiple values or strings, delimit them with a comma (",").

elapse=elapsedtimelimit

This changes the job elapsed time limit to *elapsedtimelimit* seconds. The minimum value is 1 second, and the maximum value is unlimited.

It is not possible to specify the minimum value specification of elapsed timelimit (example: specify minimum value only 10-, minimum value and maximum value specified 10-100).

If an acceptance message is output after executing `pjalter` command for the job in the RUNNING state, it does not necessarily mean that the job elapsed time limit value change succeeded. Please confirm whether the job elapsed time limit was correctly changed in the RUNNING state by the `pjstat` command.

Changing the elapsed time limit value in the RUNNING state is possible only when the function to change the elapsed time limit value for jobs in the RUNNING state is enabled and the jobs are on the FX server compute node and the elapsed time limit value is shortened.

{rscunit|ru}=rscuname

This changes the resource unit where the job is submitted to *rscuname*. Permission to submit the job to the changed resource unit is necessary.

The resource unit name can be set with 63 bytes at most.

When a resource group with `rscgrp=rscgname` is not specified, this also changes a resource group to the default resource group of the resource unit *rscuname* according to the job ACL function settings.

It is not allowed to change from a resource unit with FX server to a resource unit with PRIMERGY server. Also, it is not allowed to change from a resource unit with PRIMERGY server to a resource unit with FX server.

`{rscgrp|rg}=rscgname`

This changes the resource group where the job is submitted to *rscgname*. Permission to submit the job to the changed resource group is necessary.

The resource group name can be set with 63 bytes at most.

`-p priority`

This option sets the job priority to *priority*. You can specify a value from 0 to 255 for *priority*. The lowest priority is 0, and the highest priority is 255.

The job priority within jobs from the user is determined according to the specified priority.

`--help`

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[END STATUS]

0: Normal end

1: Abnormal end (When multiple jobs are specified, 1 is returned even if only one of them ends with an error.)

[SEE ALSO]

`pjstat(1)`, `pjsub(1)`, `pmalter(8)`

[MESSAGES]

If an error occurs during the execution of the `pjalter` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 0501 pjalter Unknown option or no option argument: *opt*.
Try `pjalter --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0503 pjalter Unknown option argument: *opt arg*.
Try `pjalter --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0504 pjalter Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0506 pjalter No parameter specified.
Try `pjalter --help` for more information.**

Meaning

No parameter to be changed is specified.

Action

Specify the parameter to be changed.

**[ERR.] PJM 0510 pjalter No jobid(s) specified.
Try `pjalter --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0511 pjalter Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0512 pjalter Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 0513 pjalter Job *jobid* is not in the state that pjalter can be accepted.

Meaning

The command cannot be executed in the state of the job specified by the job (job ID *jobid*).

Action

To change any of the following parameters, execute the command when the state of the job is QUEUED, HOLD, or ERROR:

- Resource unit for executing a job
- Resource group for executing a job
- Priority among jobs of the same user

To change the parameter listed below, execute the command when the state of the job is QUEUED, RUNNING, HOLD, or ERROR. In the RUNNING state, you can change the elapsed time limit value only under all three of the following conditions. First, the function to change the elapsed time limit values of jobs in the RUNNING state is enabled. Second, the job is on an FX server compute node. Third, the elapsed time limit value will be shorter.

- Elapsed time limit value for a job (Elapsed time limit value)

[ERR.] PJM 0514 pjalter Job *jobid* is interactive job. pjalter cannot be executed for interactive jobs.

Meaning

The job specified by job ID *jobid* is interactive job. The pjalter command cannot be executed for interactive jobs.

Action

Specify any job other than an interactive job.

[ERR.] PJM 0515 pjalter It is unable to change resource by specifying sub job ID.

Meaning

You cannot specify a sub job ID for a step job when changing the resource unit name.

You cannot specify a sub job ID for a bulk job when changing the resource unit name or the resource group name.

Action

Review the specified job ID.

[ERR.] PJM 0551 pjalter The specified time is too short: *limittime*, *jobid*.

Meaning

Resource information specified for a job (job ID *jobid*) with the -L|--rsc-list option is incorrect.

- If *limittime* is elapse

The specified elapsed time limit value is shorter than the execution time that has already elapsed.

Action

Specify a limit value higher than the execution time that has already elapsed.

[ERR.] PJM 0552 pjalter Job *jobid* parameter is being changed.

Meaning

The command cannot be executed because a parameter change is being processed for the job specified with the job ID *jobid*.

Action

Wait a moment, and execute the command again.

[ERR.] PJM 0554 pjalter *rscname=value* that have been specified in the job *jobid* is less than the lower limit (*limit-value*).

Meaning

When a job with the job ID *jobid* was submitted, the value *value* for the resource name *rscname* specified in the -L|--rsc-list option was lower than limit-value, the lower limit value defined by the job ACL function of a changed resource unit or resource group.

Action

Specify a resource limit value that is not below the range defined by the job ACL function of the changed resource unit or resource group. Alternatively, specify a resource unit or resource group whose resource limit value is not below the range defined by the job ACL function.

[ERR.] PJM 0557 pjalter *rscname=value* that have been specified in the job *jobid* is greater than the upper limit (*limit-value*).

Meaning

When a job with the job ID *jobid* was submitted, the value *value* for the resource name *rscname* specified in the -L|--rsc-list option exceeded limit-value. *limit-value* is either the upper limit value defined by the job ACL function of a changed resource unit or resource group, or the amount of resources that can be allocated to a job.

Action

Specify a resource amount that does not exceed the range defined in the job ACL as the limit value for the resource. Or, specify a resource unit or resource group whose resource limits do not exceed the range defined in the job ACL feature and whose resource amount (Number of nodes, node shape, etc.) can be assigned to the job.

For details on the amount of resources that can be allocated to jobs on the FX server, see the table "Size of nodes that can be allocated

[FX]" in "Checking resource units and resource groups" in "Chapter 2 Job Operation Procedures" of the manual "Job Operation Software End-User's Guide."

[ERR.] PJM 0558 pjalter *rscname=value* does not exist.

Meaning

- If the resource name *rscname* is *rscunit*
The specified resource unit *value* does not exist.
- If the resource name *rscname* is *rscgrp*
The specified resource group *value* does not exist.

Action

Specify an existing resource unit name or resource group name.

[ERR.] PJM 0559 pjalter *rscname=value* is disabled.

Meaning

- If the resource name *rscname* is *rscunit*
A job currently cannot be submitted to the specified resource unit *value*.
- If the resource name *rscname* is *rscgrp*
A job currently cannot be submitted to the specified resource group *value*.

Action

Check the state of the resource unit or resource group with the `--rsc` option of the `pjstat` command. Then, specify a resource unit or resource group to which the job can be submitted. If necessary, ask the administrator to change the state of the resource unit or resource group.

[ERR.] PJM 0561 pjalter The architecture of compute nodes in the resource unit is different before and after the change.

Meaning

The resource unit for FX servers cannot be changed to the resource unit for PRIMERGY servers. Likewise, the resource unit for PRIMERGY servers cannot be changed to the resource unit for FX servers.

Action

Specify a resource unit that has the same architecture.

**[ERR.] PJM 0570 pjalter No execute permission.
[ERR.] PJM 0570 pjalter No execute permission: *jobid*.**

Meaning

There are two types of messages.
If the message does not display a job ID, `pjalter` command execution is not permitted. If it displays the job ID *jobid*, the user does not have the privileges to submit job *jobid* to the specified resource unit or resource group.

Action

Display the job ACL settings with the `pjacl` command. Check the privileges to execute the `pjalter` command (execute `pjalter`) and the privileges to submit jobs to the specified resource unit or resource group (execute `pjsub` and execute `pjsub(--xxxx)`, where `xxxx` is the option name).
If necessary, ask the administrator to add the `pjalter` command execution privileges and job submission privileges.

[ERR.] PJM 0571 pjalter No permission: *jobid*.

Meaning

The user does not have sufficient permission to operate the job *jobid*.
The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the `pjacl` command. Then, confirm the permission (permit `pjalter`) for the operated objects of the `pjalter` command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0572 `pjalter` Job exceeded the accept limit: *target, jobid*.

Meaning

The following problem was encountered in an attempt made to change the resource unit or group for executing the job specified with the job ID *jobid*: The maximum number of jobs that can be concurrently accepted for the change-to resource unit or group is reached.

The value of *target* indicates the following:

- `ru-accept`: Number of batch jobs that can be concurrently accepted for a resource unit
- `ru-interact-accept`: Number of interactive jobs that can be concurrently accepted for a resource unit
- `rg-accept`: Number of batch jobs that can be concurrently accepted for a resource group
- `rg-interact-accept`: Number of interactive jobs that can be concurrently accepted for a resource group
- `ru-accept-allsubjob`: Number of sub jobs and normal jobs (batch jobs) for bulk and step jobs that can be concurrently accepted for a resource unit
- `ru-accept-bulksubjob`: Number of bulk sub jobs that can be concurrently accepted for a resource unit
- `ru-accept-stepsubjob`: Number of step sub jobs that can be concurrently accepted for a resource unit
- `rg-accept-allsubjob`: Number of sub jobs and normal jobs (batch jobs) for bulk and step jobs that can be concurrently accepted for a resource group
- `rg-accept-bulksubjob`: Number of bulk sub jobs that can be concurrently accepted for a resource group
- `rg-accept-stepsubjob`: Number of step sub jobs that can be concurrently accepted for a resource group

Action

Using the `pjacl` command with the `--rscunit` (`--ru`) or `--rscgrp` (`--rg`) option, confirm the maximum number of jobs that can be concurrently accepted for the change-to resource unit or group.

The maximum number of jobs that can be concurrently accepted for a resource unit is the "acceptablejob", "acceptable job(interact)", "acceptable all-subjob", "acceptable bulk-subjob", or "acceptable step-subjob" value of the item "limit in rscunit (each users)", "limit in rscunit (total users in same group)", or "limit in rscunit (total all users)."

The maximum number of jobs that can be concurrently accepted for a resource group is the "acceptablejob", "acceptable job(interact)", "acceptable all-subjob", "acceptable bulk-subjob", or "acceptable step-subjob" value of the item "limit in rscgroup (each users)", "limit in rscgroup (total users in same group)", or "limit in rscgroup (total all users)."

If necessary, ask the administrator to increase the limit of the number of concurrently acceptable jobs.

[ERR.] PJM 0573 `pjalter` Option value is out of range: *target, jobid*.

Meaning

The value of the option *target* specified for job *jobid* exceeds the valid range defined in the job ACL function.

Action

Display the job ACL function settings with the `pjacl` command, and confirm the definition value of the specified option. If necessary, ask the administrator to change the limit value.

[ERR.] PJM 0575 `pjalter` Invalid attribute of job: *jobid (attr)*.

Meaning

Jobs with the attribute *attr* are not allowed to execute on the resource unit or resource group to which they are modified.

Action

Cannot be changed to resource units or resource groups that have the following settings:.

- For jobs with the attribute io-exclusive
The job ACL function definition item define allocation-io-mode is set to no-io-exclusive and execute io-exclusive is set to disable.
- For jobs with the attribute no-io-exclusive
The job ACL function definition items define allocation-io-mode is set to io-exclusive and execute no-io-exclusive is set to disable.

Use the `pjacl` command to check the attributes of jobs that can execute on the resource units or resource groups to which they are modified.

[ERR.] PJM 0576 pjalter *rscname=value* that have been specified in the job *jobid* cannot be specified. The value must be specified as follows: *strings,...*

Meaning

The value *value* for the resource name *rscname* specified in the `-L|--rsc-list` option for a job with the job ID *jobid* is not among the types that can be specified for the custom resources defined by the job ACL function. *strings,...* is a list of types of specifiable custom resources.

Action

Specify a resource unit or resource group that has a defined custom resource.

[ERR.] PJM 0578 pjalter *rscname=value* that have been specified in the job *jobid* does not exist.

Meaning

The value *value* for the resource name *rscname* specified in the `-L|--rsc-list` option is for a job with the job ID *jobid* does not have a defined custom resource.

Action

Specify a resource unit name or resource group name that has a defined custom resource.

[ERR.] PJM 0579 pjalter *Arbitrary character string*.

Meaning

An error occurred at the job attribute change exit that was set by the administrator. The contents of this message were set by the administrator.

Action

Referring to the message, find a solution to why the parameter change caused the error. If you cannot find out the cause from the message, contact the administrator. The administrator is requested to check why the parameter change caused the error at the job attribute change exit.

[ERR.] PJM 0589 pjalter Not supported: *message*.

Meaning

The current version of Technical Computing Suite does not support the parameter change of the specified job.

Action

If you want to change the specified parameter, delete the target job and submit it again specifying with suitable parameter.

[ERR.] PJM 0590 pjalter PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function and the system state.

[ERR.] PJM 0591 pjalter Internal error: *details*.**Meaning**

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0593 pjalter The operating command cannot be executed except login node and compute cluster management node.**Meaning**

The command cannot be executed on a node other than a login node or compute cluster management node.

Action

Execute the command on a login node or compute cluster management node.

[ERR.] PJM 0596 pjalter Job *jobid* was rejected *opt. message***Meaning**

The parameter change for the job (job ID: *jobid*) specified in the option *opt* was rejected.

opt: -L elapse: Specified elapsed time limit value

message shows the reason why the parameter change was rejected.

Message	Meaning
pjalter cannot extend the elapse time limit of a running job.	The parameter change for the job (job ID: <i>jobid</i>) specified in the option <i>opt</i> was not accepted.
Changing the elapse time limit of a running job is only supported for the job on FX server compute node.	The elapsed time limit values of running jobs executed on servers other than FX servers cannot be changed.
Internal error (<i>details</i>).	An internal error occurred.

Action

The appropriate action varies depending on the message displayed in *message*.

- When "pjalter cannot extend the elapse time limit of a running job." is displayed
The pjalter command cannot extend the elapsed time limit value of a running job. Review the specified elapsed time limit value.
- When "Changing the elapse time limit of a running job is only supported for the job on FX server compute node." is displayed
If you want to change the elapsed time limit value of a running job, execute the job on an FX server.
- When "Internal error (*details*)." is displayed
Re-execute the command because the system may have been temporarily unable to accept the request. If the situation remains unchanged after about 10 minutes, contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk with the collected data together with the output message.

Information Message

[INFO] PJM 0500 pjalter Job *jobid* is accepted *opt*.

Meaning

The change to the job (jobID: *jobid*) parameter specified by the option *opt* was accepted.

Action

No action is necessary.

[INFO] PJM 0581done.

Meaning

The request to change the job parameter waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

3.1.4 pjdel command

[NAME]

pjdel - Deletes jobs.

[SYNOPSIS]

```
/usr/bin/pjdel [-R|--reason reasonmessage]  
               [--enforce] [--no-stats] [--no-history] [--llio-flush]  
               jobid [jobid ...]  
/usr/bin/pjdel --help
```

[DESCRIPTION]

The pjdel command deletes jobs. This command can be executed on the login node or compute cluster management node.

This command deletes jobs by specifying the job ID or the sub jobid *jobid*.

When the job is in ACCEPT status, the command will change its status to REJECT. Otherwise, when the job is in other status, the command will change the status to CANCEL.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

1-100

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

100_1-10

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

101[1-10]

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges
- Jobs in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly (Processing is not aborted, and all specified jobs are processed).

[OPTIONS]

-R|--reason *reasonmessage*

It outputs the delete reason of the job to the REASON column of the pjstat command.

When the job is deleted normally, the specified string by this option is output to the REASON column of the pjstat command. The string must be shorter than 64 characters. The string can only contain single-byte alphabet and numeric character. If this option is specified for a job that is in the HOLD state, an invalid value ('-' or empty) is set to the REASON.

--enforce

The job executing the prologue script or the epilogue script can be canceled by specifying optional **--enforce**.

Moreover, when the pjdel command is executed specifying optional **--enforce** for the job executing the job script, the job is deleted without executing the epilogue script.

--no-stats

If the job being deleted is in the QUEUED state, suppress the output of the job statistical information file (.stats file) for that job.

When this option is specified, suppresses the output of the job statistical information file, even for jobs that use the **--stats** or **--STATS** options with the pjsub command.

--no-history

If the job being deleted is in the QUEUED state, suppress the output of that job to the job history information that is output by the **-H** option of the pjstat command.

--llio-flush

This option defines the wait for flushing unwritten files within the limit for elapsed job time in cases where the job being executed is canceled.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end (When jobs are specified repeatedly, the command return abnormally even if only one error occurs.)

[SEE ALSO]

pjsub(1), pjstat(1)

[MESSAGES]

If an error occurs during the execution of the pjdel command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 0101 pjdel Unknown option or no option argument: *opt*.
Try `pjdel --help` for more information.

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

[ERR.] PJM 0102 pjdel Invalid combination of options: *opt* and *opt*.
Try `pjdel --help` for more information.

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

**[ERR.] PJM 0103 pjdel Unknown option argument: *opt arg*.
Try `pjdel --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0104 pjdel Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0110 pjdel No jobid(s) specified.
Try `pjdel --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0111 pjdel Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0112 pjdel Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 0113 pjdel Job *jobid* is not in the state that pjdel can be accepted.

Meaning

The command cannot be executed in the state of the specified job (job ID: *jobid*).

Action

Execute it when the state of the job is QUEUED, RUNNING-A, RUNNING-P, RUNNING, RUNNING-E, RUNOUT, HOLD, ERROR, SUSPEND, SUSPENDED, or RESUME.

[ERR.] PJM 0170 pjdcl No execute permission.
[ERR.] PJM 0170 pjdcl No execute permission: *opt*, *jobid*.

Meaning

There are two types of messages.

If the message does not display a job ID, the pjdcl command execution is not permitted.

If it displays *opt*, the user does not have the privileges to specify the option *opt* for job *jobid*.

Action

Display the job ACL settings with the pjacl command, and check the permission to execute the pjdcl command (execute pjdcl) or .to specify the option (execute pjdcl(--xxxx)). If necessary, ask the administrator to add permission.

[ERR.] PJM 0171 pjdcl No permission: *jobid*.

Meaning

The user is not permitted to operate the target job (job ID: *jobid*).

The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the pjacl command. Then, confirm the permission (permit pjdcl) for the operated objects of the pjdcl command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0190 pjdcl PJM daemon is not present.

Meaning

The job manager function is not operating, or the communication to the job manager function is unavailable.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator must confirm the job manager function is executing.

[ERR.] PJM 0191 pjdcl Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information on maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0193 pjdcl The operating command cannot be executed except login node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a login node or compute cluster management node.

Action

Execute the command on a login node or compute cluster management node.

Information Message

[INFO] PJM 0100 pjdcl Accepted job *jobid*.

Meaning

The request to delete a job (job ID: *jobid*) was accepted.

Action

No action is necessary.

[INFO] PJM 0119 pjdel Already accepted job *jobid*.

Meaning

The request to delete a job (job ID: *jobid*) has already been accepted.

Action

No action is necessary.

[INFO] PJM 0181done.

Meaning

The request to delete the job waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

3.1.5 pjexe command [PG]

[NAME]

pjexe - Executes process remotely.

[SYNOPSIS]

```
/usr/bin/pjexe --vnode vnodenum command  
/usr/bin/pjexe --help
```

[DESCRIPTION]

The program specified for *command* by the virtual node of a few minutes of *vnodenum* specified with --vnode is executed.
The virtual node from 0 of the virtual nodes ID to *vnodenum* -1 is used.
Virtual node ID from which the process was executed is set to environment variable PLE_VPID of each process.

[OPTIONS]

--vnode *vnodenum*

The command is executed for the virtual nodes of the specified number.
The repetition specification of this option cannot be done.

--help

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[END STATUS]

The end status of the program is returned.
However, the following values are returned in abnormal circumstances.

255: Unexpected error

254: Node failure

(0x80 | Signal number): The program ends with the signal.
1: Use mistake

[NOTE]

- The pjexe command can be used only from a process within a job.
- The pjexe command can be used only on PRIMERGY compute nodes.

[MESSAGES]

If an error occurs during the execution of the pjexe command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PLE 9100 pjexe The --vnode must be specified.

Meaning

The --vnode option is not specified in the pjexe command.

Action

Specify the --vnode option in the pjexe command, and execute the command again.

[ERR.] PLE 9101 pjexe command must be specified.

Meaning

No program is specified in the pjexe command.

Action

Specify a program in the pjexe command, and execute the command again.

[ERR.] PLE 9102 pjexe *opt* option is not supported.

Meaning

The pjexe command does not support the specified option *opt*.

Action

Specify the correct option in the pjexe command, and execute the command again.

[ERR.] PLE 9103 pjexe *opt* option is invalid.

Meaning

The arguments of the option *opt* specified in the pjexe command contain an error.

Action

Confirm the arguments of the option specified in the pjexe command.

[ERR.] PLE 9104 pjexe Duplicated an option.

Meaning

An option specified in the pjexe command is repeated.

Action

Confirm the options specified in the pjexe command.

[ERR.] PLE 9105 pjexe The specified vnode(*vnodearg*) is invalid.

Meaning

The argument *vnodearg* of the `--vnode` option of the `pjexe` command contains an error.

Action

Confirm the argument value in the `--vnode` option of the `pjexe` command.

[ERR.] PLE 9106 `pjexe` The plexec cannot be executed.(CODE=X,Y,Z)

Meaning

The `plexec` command cannot be executed.

X,Y,Z: Internal codes

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9107 `pjexe` cannot get memory.(CODE=X,Y,Z)

Meaning

There is insufficient memory.

X,Y,Z: Internal codes

Action

Increase the amount of memory to be allocated to the job, and execute the job again. If the same error recurs, contact the administrator. The administrator is requested to estimate the system memory again.

Information Message

[INFO] PLE 9149 `pjexe` Try '`pjexe --help`' for more information.

Meaning

The usage of the `pjexe` command is incorrect.

Action

Confirm whether the use of the `pjexe` command is correct referring to the message displayed immediately before.

3.1.6 `pjhold` command

[NAME]

`pjhold` - Holds jobs

[SYNOPSIS]

```
usr/bin/pjhold [-R|--reason reasonmessage]  
               [--enforce] [--llio-flush]  
               jobid [jobid ...]  
/usr/bin/pjhold --help
```

[DESCRIPTION]

The `pjhold` command holds the submitted jobs. When the specified jobs are in RUNNING status, they will be interrupted by the `pjhold` command and held. This command can be executed on the login node or the compute cluster management node.

The statuses of the jobs that can be held by this command are listed as follows:

- QUEUED

- RUNNING-A
- RUNNING
- SUSPEND
- SUSPENDED
- RESUME

When the jobs are in other status, the error message will be reported. In addition, the specified jobs that cannot be executed again are not able to be held in the status except QUEUED.

This command holds jobs by specifying the job ID or sub job ID *jobid*. When several jobs are specified and a part of them cannot be held, the command only operates the jobs that can be held, and then the other jobs make the error message reported.

When the job IDs are specified, all the specified sub jobs will be held if they can. If the status of the sub job is EXIT or ERROR which cannot be held, the jobs will be ignored. If the sub job IDs are specified, the corresponding sub job will be held.

When the status of the job is RUNNING-A or RUNNING, this command will interrupt the job and hold it.

If the command holds the jobs normally, the status of the jobs will change to HOLD status.

When the same job ID is specified more than one time, the operation is accepted, if the job is able to be specified. But the hold operation can only be accepted at the first time.

When job is in RUNNING or RUNNING-A status, the command will return after the beginning of interruption process, and do not need to wait that the process is over.

When the results of the jobs which are going to be held are error, the job status changes to ERROR, and the job sends a mail to the owner of the job to inform the error. But when the pjhold command is failed, the command sends a mail to the command-executed user to notice the hold failure information (each failed command send one mail).

In other situation, the job status changes to HOLD, then the command returns.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

```
1-100
```

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

```
100_1-10
```

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

```
101[1-10]
```

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges
- Jobs in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly (Processing is not aborted, and all specified jobs are processed).

[OPTIONS]

-R|--reason *reasonmessage*

It outputs the hold reason of the job to the REASON column of the pjstat command.

When the job is held normally, the specified string by this option is output in the REASON column of the pjstat command in the form "*command execution user name:reasonmessage*". The string must be shorter than 64 characters. The string can only contain single-byte alphabet and numeric character.

--enforce

The job executing the prologue script or the epilogue script can be held by specifying optional --enforce.

Moreover, when the pjhold command is executed specifying optional --enforce for the job executing the job script, the job is held without executing the epilogue script.

--llo-flush

This option defines the wait for flushing unwritten files within the limit for elapsed job time in cases where the job being executed enters a fixed state.

--help

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following is a use example of the pjhold command:

```
$ pjhold 1 2
[INFO] PJM 0300 pjhold Accepted job 1.
[INFO] PJM 0300 pjhold Accepted job 2.
```

```
$ pjstat 1-2
JOB_ID ... JOB_NAME MD ST  USER  ... REASON
1 ... jobname1 NM HLD user1 ... user1:
2 ... jobname2 NM HLD user2 ... user1:
```

When job ID 1 and 3 can be held, and job ID 2 cannot be held,

```
$ pjhold 1 2 3
[INFO] PJM 0300 pjhold Accepted job 1.
[ERR.] PJM 0313 pjhold Job 2 is not in the state that pjhold can be accepted.
[INFO] PJM 0300 pjhold Accepted job 3.
```

[END STATUS]

- 0: Normal end
- 1: Abnormal end (When jobs are specified repeatedly, the command return abnormally even if only one error occurs.)

[SEE ALSO]

pjrls(1)

[MESSAGES]

If an error occurs during the execution of the pjhold command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 0301 pjhold Unknown option or no option argument: *opt*.
Try `pjhold --help` for more information.

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

[ERR.] PJM 0303 pjhold Unknown option argument: *opt arg*.
Try `pjhold --help` for more information.

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0304 pjhold Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0310 pjhold No jobid(s) specified.
Try `pjhold --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0311 pjhold Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0312 pjhold Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 0313 pjhold Job *jobid* is not in the state that pjhold can be accepted.

Meaning

The command cannot be executed because of the state of the specified job (job ID: *jobid*).

Action

Check that the job is in the state QUEUED, RUNNING-A, RUNNING, SUSPEND, SUSPENDED, or RESUME that the job can be held.

If the automatically re-execution is disabled for the job by --norestart option or job operation setting, the job can be held only at the status of QUEUED.

[ERR.] PJM 0314 pjhold Job *jobid* is interactive job. pjhold cannot be executed for interactive jobs.

Meaning

The job specified with the job ID *jobid* is an interactive job. The pjhold command cannot be executed for interactive jobs.

Action

Specify a job that is not an interactive job.

[ERR.] PJM 0370 pjhold No execute permission.
[ERR.] PJM 0370 pjhold No execute permission: *opt*, *jobid*.

Meaning

There are two types of messages.

If the message does not display a job ID, the pjhold command execution is not permitted. If it displays *opt*, the user does not have the privileges to specify the option *opt* for job *jobid*.

Action

Display the job ACL settings with the pjacl command. Check whether pjhold command execution is permitted (execute pjhold), and check the privileges to specify the option (execute pjhold (--xxxx), where xxxx is the option name). If necessary, ask the administrator to add permission.

[ERR.] PJM 0371 pjhold No permission: *jobid*.

Meaning

The user is not permitted to operate the target job (job ID: *jobid*).

The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the pjacl command. Then, confirm the permission (permit pjhold) for the operated objects of the pjhold command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0390 pjhold PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function or the system state.

[ERR.] PJM 0391 pjhold Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0393 pjhold The operating command cannot be executed except login node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a login node or compute cluster management node.

Action

Execute the command on a login node or compute cluster management node.

Information Message

[INFO] PJM 0300 pjhold Accepted job *jobid*.

Meaning

The request to hold a job (job ID: *jobid*) was accepted.

Action

No action is necessary.

[INFO] PJM 0319 pjhold Already accepted job *jobid*.

Meaning

The request to hold a job (job ID: *jobid*) has already been accepted.

Action

No action is necessary.

[INFO] PJM 0381done.

Meaning

The request to hold the job waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

3.1.7 pjpbinding command [PG]

[NAME]

`pjpbinding` - Binds CPU resources to processes and threads.

[SYNOPSIS]

```
mpiexec [options] /usr/bin/pjpbinding [ --disable-thread ] [ -v ] [a.out]
/usr/bin/pjpbinding [ --help ]
```

[DESCRIPTION]

The `pjpbinding` command binds CPU resources (CPU cores or logical CPUs) to the MPI program *a.out* specified in the argument. The CPU resources allocated to the job are bound to the `pjpbinding` command processes according to priority; a higher priority is assigned to those resources not bound to other processes.

After that, the `pjpbinding` command executes *a.out*.

The `pjpbinding` command can be used on the PRIMERGY compute node.

To use the `pjpbinding` command to bind CPU resources to threads without duplication between multiple processes on one node when executing a multithread MPI program, use a runtime library that can interpret the `GOMP_CPU_AFFINITY` environment variable of GNU gcc.

The rules for binding CPU resources to processes and threads are determined by the environment variables explained below. If the environment variables are not set, CPU resources are bound according to the default values of the variables. These environment variables are ignored when the `pjpbinding` command is executed from `mpiexec` in Development Studio.

[OPTIONS]

`--disable-thread`

This option automatically binds CPU resources to processes but not to threads, in the program to be executed.

-v

This option outputs information on the process-bound CPU resources to the standard error output of the job. The output information is as follows:

Node ID, execution file name, process ID, number of bound CPU resources, bound CPU ID

--help

This option displays the usage of this command.

[ENVIRONMENT]

By using the following environment variables, you can change how CPU resources are bound.

For an image of the binding, see the "End-user's Guide," which is a Job Operation Software manual.

PLE_MPI_PIN_DOMAIN

Specify "omp" or integer value *n*.

If this environment variable is omitted, the operation is the same as with "omp" specified.

- "omp" specified

The number of CPU resources that are bound to processes is the same as the value specified in the PARALLEL or OMP_NUM_THREADS environment variable. If both of the environment variables are specified concurrently, the PARALLEL environment variable has priority.

- Integer value *n* specified

The number of CPU resources that are bound to processes is *n*.

Either CPU cores or logical CPUs are bound, depending on the setting of the PLE_MPI_PIN_CELL environment variable.

PLE_MPI_PIN_CELL

Specify "core" or "unit".

If this environment variable is omitted, CPU resources are bound without duplication between multiple processes on one node.

- "core" specified

CPU resources are bound to processes for each CPU core.

- "unit" specified

CPU resources are bound to processes for each logical CPU.

PLE_MPI_PIN_ORDER

Specify any of "range", "compact", or "scatter".

If this environment variable is omitted, the operation is the same as with "scatter" specified.

- "range" specified

CPU resources are bound to processes in ascending order of CPU ID, so the CPU resource bound to the current process has a higher ID than that for the previous process.

- "compact" specified

CPU resources are bound such that the CPU resource bound to the current process is close to that of the previous process.

- "scatter" specified

CPU resources are bound such that the CPU resource bound to the current process is far from that of the previous process.

In any case, in a job, the starting point of CPU resource binding in a cycle with the above order is the "youngest" CPU resource in terms of CPU ID.

Either CPU cores or logical CPUs are bound, depending on the setting of the PLE_MPI_PIN_CELL environment variable.

[EXAMPLE]

Example 1: pjpbinding command used with an MPI program in 4-process parallel and 16-thread parallel execution

```
#PJM -L "node=4"                (*1)
export OMP_NUM_THREADS=16        (*2)
export PLE_MPI_PIN_DOMAIN=omp    (*3)
export PLE_MPI_PIN_CELL=unit     (*4)
```

```
export PLE_MPI_PIN_ORDER=scatter      (*5)
mpiexec -n 4 pjpbind -v a.out         (*6)
```

(*1) Specifies 4 as number of nodes

(*2) Specifies 16 as number of threads

(*3) Binding based on value of OMP_NUM_THREADS environment variable

(*4) Binding of each logical CPU

(*5) Sets "scatter" for binding order

(*6) Executes with 4 specified as number of processes so that information on bound CPU resources will be output

Example 2: pjpbind command used with a non-MPI multithread program in single-process and eight-thread parallel execution

```
#PJM -L "node=1"                      (*1)
export OMP_NUM_THREADS=8               (*2)
pjpbind a.out                          (*3)
```

(*1) Specifies 1 as number of nodes

(*2) Specifies 8 as number of threads

(*3) Executes

[END STATUS]

The returned value is the end status of the MPI program a.out. However, if an error occurs, the returned value is as follows:

254: Invalid usage

255: Unexpected error

[MESSAGES]

If an error occurs during the execution of the pjpbind command, the following messages is output to the standard error output:

Error Message

[ERR.] PLE 9310 pjpbind PLE service error occurred.(CODE=code1,code2)

Meaning

An error occurred in the parallel execution environment daemon. *code1* and *code2* are internal codes.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9311 pjpbind A system error occurred.(CODE=code1,code2)

Meaning

A system error occurred. *code1* and *code2* are internal codes.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9312 pjpbind cannot get memory.(CODE=code)

Meaning

There is insufficient memory. *code* is an internal code.

Action

Execute the job again. If the same error recurs, contact the administrator. The administrator is requested to estimate the system memory again.

[ERR.] PLE 9313 pjpbinding The program is not specified.

Meaning

No program is specified in the pjpbinding command.

Action

In the pjpbinding command, specify the program to be executed.

[ERR.] PLE 9314 pjpbinding option *opt* is not supported.

Meaning

The specified option *opt* in the pjpbinding command is not supported.

Action

Specify the correct option in the pjpbinding command.

[ERR.] PLE 9315 pjpbinding Duplicated an option.

Meaning

The same option is specified multiple times in the pjpbinding command.

Action

Review the option to be specified in the pjpbinding command.

[ERR.] PLE 9316 pjpbinding Cannot get jobid.

Meaning

Job ID acquisition failed.

Action

Execute the pjpbinding command in a job script.

[ERR.] PLE 9317 pjpbinding The program(*program*) cannot be executed.(*errno=errno*)

Meaning

The system call `execvp(2)` to the user program *program* failed.

Action

The error code *errno* indicates the error of the system call `execvp(2)`. Take appropriate action according to the error code.

[ERR.] PLE 9318 pjpbinding CPU core(s) not allocated in the virtual node is included. *cpuid=cpuid*

Meaning

In an attempt by the pjpbinding command to bind CPU cores to processes, these CPU cores included one that is not allocated to the virtual node of the job.

cpuid displays a comma-separated list of all CPU IDs that the command attempted to bind in the job.

Action

CPU cores other than those allocated to the virtual nodes of the job cannot be bound to processes by the pjpbinding command. This error message is output under the following conditions:

- the specified value for the number of CPU cores (vnode-core) allocated to the virtual nodes is less than the number of CPU cores mounted on the compute node when the job is submitted; and

- the CPU IDs of the CPU cores to be bound to processes are explicitly specified.

The message is output because the CPU ID of a CPU core allocated to a virtual node of a job is undefined.

To explicitly specify the CPU ID of a CPU core to be bound to a process, you need to specify a value for the number of CPU cores (vnode-core) allocated to virtual nodes when submitting the job. To do so, either specify the same value as the number of CPU cores mounted on the compute node and submit the job, or submit the job as a node allocated job using all nodes.

Warning Message

[WARN] PLE 9331 pjpbinding Cannot set env=val.

Meaning

The setting value of *val* in the environment variable *env* referenced by the *pjpbinding* command is invalid. The specified value is ignored, and processing continues with the default value.

Action

Review the setting values of the environment variable *env*.

3.1.8 pjrls command

[NAME]

pjrls - Releases the jobs from HOLD status

[SYNOPSIS]

```
/usr/bin/pjrls jobid [jobid ...]  
/usr/bin/pjrls --help
```

[DESCRIPTION]

The *pjrls* command releases the jobs from HOLD status. This command can be executed on login node and compute cluster management node.

This command can only specify the jobs in HOLD status. Otherwise, when the specified jobs are in other status, the error message will be reported.

When the command is executed, the specified jobs will be changed to QUEUED status and be re-scheduled.

This command releases jobs from HOLD status by specifying the job ID or sub job ID *jobid*. The job ID or the sub job ID can be specified by designated numbers or range.

When the same job ID is specified repeatedly, the jobs which can be specified are all accepted. But the release operation can only be accepted at the first time. Then the following commands are all rejected with errors. When the jobs are released from the HOLD status, the command returns.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

```
1-100
```

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

```
100_1-10
```

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

```
101[1-10]
```

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges
- Jobs in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly (Processing is not aborted, and all specified jobs are processed).

[OPTIONS]

--help

This option displays the usage of this command.

If this option is specified, all the arguments are ignored.

[END STATUS]

0: Normal end

1: Abnormal end (When jobs are specified repeatedly, the command return as the abnormal end even if only one error occurs.)

[SEE ALSO]

pjhold(1)

[MESSAGES]

If an error occurs during the execution of the pjrls command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 0401 pjrls Unknown option or no option argument: *opt*.
Try `pjrls --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0410 pjrls No jobid(s) specified.
Try `pjrls --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0411 pjrls Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0412 pjrls Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 0413 pjrls Job *jobid* status is not hold.**Meaning**

The specified job (job ID: *jobid*) is not in the hold state.

Action

Check the job state.

[ERR.] PJM 0470 pjrls No execute permission.**Meaning**

The user is not permitted to execute the pjrls command.

Action

Display the job ACL settings with the pjacl command, and check whether the pjrls command can be executed (execute pjrls). If necessary, ask the administrator to add permission.

[ERR.] PJM 0471 pjrls No permission: *jobid*.**Meaning**

The user is not permitted to operate the target job (job ID: *jobid*).

The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the pjacl command. Then, confirm the permission (permit pjrls) for the operated objects of the pjrls command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0490 pjrls PJM daemon is not present.**Meaning**

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function or the system state.

[ERR.] PJM 0491 pjrls Internal error: *details*.**Meaning**

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0493 pjrls The operating command cannot be executed except login node and compute cluster management node.**Meaning**

The command cannot be executed on a node other than a login node or compute cluster management node.

Action

Execute the command on a login node or compute cluster management node.

Information Message

[INFO] PJM 0400 pjrIs Job *jobid* released.

Meaning

The hold state of the job (job ID: *jobid*) was canceled.

Action

No action is necessary.

[INFO] PJM 0481done.

Meaning

The request to cancel the holding of the job waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

3.1.9 pjrsh command

[NAME]

pjrsh - Remotely executes a command.

[SYNOPSIS]

```
[IP address specification]
/usr/bin/pjrsh [ -n ] IPaddress command
```

```
[Coordinate specification]
/usr/bin/pjrsh [ -n ] coord command
```

```
[Usage]
/usr/bin/pjrsh --help
```

[DESCRIPTION]

This command executes the program specified in *command* on the compute node corresponding to the *IPaddress* or the coordinates *coord*. When specifying coordinates, use the format of x (one dimension), x,y (two dimensions), or x,y,z (three dimensions) for the specification. Specify coordinates in *coord* as follows.

- When *coord* is one dimensional (x), specify "(x)" or x.
- When *coord* is two dimensional (x,y), specify "(x,y)" or x,y.
- When *coord* is three dimensional (x,y,z), specify "(x,y,z)" or x,y,z.

These coordinates are in the node geometry specified by the -L node or --rsc-listnode option of the pjsb command. The number of dimensions must be the same as in the geometry. The node geometry specified by the pjsb command can be obtained from the values of environment variables PJM_NODE_X, PJM_NODE_Y and PJM_NODE_Z which can be used in a job script.

The pjrsh command transfers the standard input, standard output, and standard error output of the program specified in *command* to the standard input, standard output, and standard error output of the pjrsh command. However, if -n is specified, the standard input is set to /dev/null.

[OPTIONS]

-n

/dev/null is assumed specified as the standard input.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

When the node geometry specified by the pjsb command is one dimensional (node=2) and when you want to execute *command* on the compute node at the one dimensional coordinate "(1)".

```
$ /usr/bin/pjrsh "(1)" command
```

[NOTE]

- The pjrsh command can be used only from a process within a job.
- The list of IP addresses of the allocated nodes to the job can be obtained from the pjshowip command.

[END STATUS]

A program end code is returned. However, the following values are returned for error occurrences.

255: Unexpected error

254: Node down

(0x80 | Signal number): The program ends with the signal.

1: Error in use

[SEE ALSO]

pjshowip(1)

[MESSAGES]

If an error occurs during the execution of the pjrsh command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PLE 9150 pjrsh IP address or virtual coordinate must be specified.

Meaning

No IP address or coordinate is specified in the pjrsh command.

Action

Specify an IP address or coordinate in the pjrsh command, and execute the job again.

[ERR.] PLE 9151 pjrsh command must be specified.

Meaning

No program is specified in the pjrsh command.

Action

Specify a program in the pjrsh command, and execute the command again.

[ERR.] PLE 9152 pjrsh *opt* option is not supported.

Meaning

The `plrsh` command does not support the specified option *opt*.

Action

Specify the correct option in the `plrsh` command, and execute the command again.

[ERR.] PLE 9153 plrsh failed to initialize PSM.

Meaning

Initialization of the system management function failed.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9154 plrsh IP address (*ipaddress*) is invalid.

Meaning

The IP address *ipaddress* specified in the `plrsh` command contains an error.

Action

Confirm the IP address specified in the `plrsh` command.

[ERR.] PLE 9155 plrsh failed to finalize PSM.

Meaning

The end processing of the system management function failed.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9156 plrsh The plexec cannot be executed.(CODE=X,Y,Z)

Meaning

The `plexec` command cannot be executed.

X,Y,Z: Internal codes

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9157 plrsh cannot get memory.(CODE=X,Y,Z)

Meaning

There is insufficient memory.

X,Y,Z: Internal codes

Action

Increase the amount of memory to be allocated to the job, and execute the job again. If the same error recurs, contact the administrator. The administrator is requested to estimate the system memory again.

[ERR.] PLE 9158 pjrsh Virtual coordinate (*vcoord*) is invalid.**Meaning**

The specified coordinate *vcoord* is invalid.

The specified coordinate *vcoord* may exceed the range of nodes allocated to the job, or the number of dimensions may be different from the shape of the nodes.

Action

Review the coordinate specified in the pjrsh command.

[ERR.] PLE 9159 pjrsh Virtual node (*vnodeid*) is invalid.**Meaning**

The specified virtual node ID *vnodeid* is invalid.

The specified virtual node ID *vnodeid* may be outside the range of virtual nodes allocated to the job.

Action

Review the virtual node ID specified in the pjrsh command.

[ERR.] PLE 9198 pjrsh A system error occurred.(nid=own node)(CODE=X,Y,Z)**Meaning**

A system error occurred.

X,Y,Z: Internal codes

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

Information Message

[INFO] PLE 9199 pjrsh Try 'pjrsh --help' for more information.**Meaning**

The usage of the pjrsh command is incorrect.

Action

Confirm whether the use of the pjrsh command is correct referring to the message displayed immediately before.

3.1.10 pjshowip command

[NAME]

pjshowip - Outputs the IP addresses of the compute nodes allocated to a job.

[SYNOPSIS]

```
/usr/bin/pjshowip
```

[DESCRIPTION]

The pjshowip command outputs the IP addresses of the compute nodes allocated to a job, to individual lines.
You can use it from a process within the job, such as a job script.

[EXAMPLE]

Output example

```
$ /usr/bin/pjshowip
10.208.0.11
10.208.0.12
```

[NOTE]

The pjshowip command can be used only from a process within a job.

[END STATUS]

0: Normal end
1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the pjshowip command, the following messages is output to the standard error output. The messages of the pjshowip command may seem as the message of the plestat command which is internal command of the pjshowip command.

Error Message

[ERR.] PLE 9250 pjshowip *opt* option is not supported.

Meaning

The pjshowip command does not support the specified option *opt*.

Action

The option cannot be specified for the pjshowip command. Execute the job again without the option of pjshowip command.

[ERR.] PLE 9006 plestat PLE service error occurred.(CODE=*code*)

Meaning

The error occurred in the parallel execution environment daemon.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9007 plestat Cannot get memory.(CODE=*code*)

Meaning

The node lacks memory resources.

Action

Execute the job again. If the problem persists, contact the administrator. The administrator shall check total memory of the system again.

[ERR.] PLE 9010 plestat PSM function error.(CODE=*code*)

Meaning

Calling the system management function failed.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 9011 plestat Cannot get job id.

Meaning

Job ID retrieval failed.

Action

Execute the pjshowip command in the job script.

[ERR.] PLE 9012 plestat Cannot execute as sequential job.

Meaning

The pjshowip command cannot be executed in the sequential job.

Action

Execute the pjshowip command in the process parallel job. That is, the pjshowip command should be executed in the job which is allocated the two or more nodes, or the job which is specified the option "--mpi proc=num" of the pjsb command.

3.1.11 pjshowrsc command

[NAME]

pjshowrsc - Displays both the entire amount and the usage status of activated compute resources.

[SYNOPSIS]

```
/usr/bin/pjshowrsc [-c clstname] [scope options]
                  [-E|--expand] [-l|-v [level]] [--raw] [--exclusive|--ex]
                  [--data [--delimiter "delimiter"]]
/usr/bin/pjshowrsc [-c clstname] [scope options]
                  -v [level] [--raw] [--exclusive|--ex]
                  [--data [--delimiter "delimiter"]]
                  --status
/usr/bin/pjshowrsc [-c clstname] -n nodeid [,nodeid,...]
                  [--raw] [--data [--delimiter "delimiter"]]
                  --status
/usr/bin/pjshowrsc [-c clstname] [--rscunit|--ru] [rscuname [,rscuname,...]]
                  [--raw] [--data [--delimiter "delimiter"]]
                  [--custom-resource|-C]
/usr/bin/pjshowrsc [-c clstname] [--rscunit|--ru] [rscuname [,rscuname,...]] [--rscgrp|--rg]
                  [rscgname [,rscgname,...]]
                  [--raw] [--data [--delimiter "delimiter"]] [--exclusive|--ex]
                  [--custom-resource|-C]
/usr/sbin/pjshowrsc --help

scope options(exclusive):
    --nodegrp nodegid [,nodegid,...]
    --bootgrp bootgid [,bootgid,...]
    [--rscunit|--ru] [rscuname [,rscuname,...]]
    [--rscunit|--ru] [rscuname [,rscuname,...]] [--rscgrp|--rg] [rscgname [,rscgname,...]]
    -n nodeid [,nodeid,...]
```

[DESCRIPTION]

The pjshowrsc command is a command that displays a whole amount, an empty amount, and the quota of the calculation resource being mounted by the system according to the range specified in the option. This command is for the end user.

This command displays the following compute resources:

- **NODE**
Number of compute nodes.
- **CPU**
Number of CPU cores (The resources in a compute node).

- MEM

Amount of memory (The resources in a compute node).

This command can be used on the system management node, the compute cluster management node, and the login node.

The system administrator or the cluster administrator can view all the information of the cluster.

The job operation manager can view calculation resource information in the managed resource unit.

The end user can view calculation resource information in the resource unit with the authority to execute the job.

[OPTIONS]

-c *clstname*

This option displays the compute resource information on the *clstname* cluster.

When this option is specified except in the system management node, the option is ignored, and only the information on the belonging cluster is output.

When this option is not specified in the system management node, environment variable PXYMYCLST is assumed to be a cluster name.

When environment variable PXYMYCLST is not set, the command output information of all clusters have permission to view.

--nodegrp *nodegid* [,*nodegid*,...]

This option displays the compute resource information on the *nodegid* node group. Please describe *nodegid* by the hexadecimal number.

Plurals can be specified by delimiting *nodegid* by comma (","). (example: 0x01,0x02,0x03)

Moreover, the range can be specified by hyphen ("-"). (example: 0x01-0x03)

When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXYMYCLST or to specify the -c option at the same time.

This option can not be used on the login node.

--bootgrp *bootgid* [,*bootgid*,...]

This option displays the compute resource information on the *bootgid* boot group.

Plurals can be specified by delimiting *bootgid* by comma (","). (example: 0x01,0x02,0x03)

Moreover, the range can be specified by hyphen ("-"). (example: 0x01-0x03)

When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXYMYCLST or to specify the -c option at the same time.

This option can not be used on the login node.

{--rscunit|--ru} [*rscuname* [,*rscuname*,...]]

This option displays the compute resource information on the *rscuname* resource unit.

Plurals can be specified by delimiting *rscuname* by comma (","). (example: unitA, unitB)

When an optional argument is not specified, information on all the resource units is displayed.

This option can be specified at the same time as the --rscgrp option.

When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXYMYCLST or to specify the -c option at the same time.

{--rscgrp|--rg} [*rscgname* [,*rscgname*,...]]

This option displays the compute resource information on the *rscgname* resource group.

Plurals can be specified by delimiting *rscgname* by comma (","). (example: unitA, unitB)

When an optional argument is not specified, information on all the resource group is displayed.

This option can be specified at the same time as the --rscunit option.

For compute resources that are shared by more than one resource group, information on ALLOC is displayed by the value including not only the amount of the resource usage of the specified resource group but also the amount of the resource usage of the resource group that shares the compute resources.

If you want to display only the resource usage of the specified resource group, please specify --exclusive option at the same time.

When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXYMYCLST or to specify the -c option at the same time.

-n *nodeid* [,*nodeid*,...]

This option displays the compute resource information on the *nodeid* node.

Plurals can be specified by delimiting *nodeid* by comma (","). (example: 0x01010010,0x01010011,0x01010012)

Moreover, the range can be specified by hyphen ("-"). (example: 0x01010010-0x01010012)

When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXYMYCLST or to specify the -c option at the same time.

--expand|-E

This option displays the compute resource information on the lower hierarchy of the specified range.
Specifying a node group for this option, for example, displays the information on a boot group in the node group.

--raw

This option displays the amount of resources without converting it to Ki (kibi) or Gi (gibi).

--exclusive|--ex

Only when --rscgrp option is specified, this option is valid.
The amount of the resource that using by the specified resource group with --rscgrp option is displayed.
Please specify if you want to verify the resource use rate of the resource group.
If you want to see a free physical resources amount, this option should be omitted.
The value of the amount of free space resources (FREE) will be displayed hyphen ("-").
When --rscgrp option is not specified, this option is disregarded.

-l

This option lists all the compute resources.
When the -v option is specified when this option is specified, this option is disregarded.

-v [level]

This option outputs detailed information according to level.
You can choose either level as follows. If level is omitted, this value is 0.

level	Description
0	Displays all the compute resources for each node.
1	Displays the job ID of the job being executed in addition to level 0 information. The job ID is displayed following RUNNING_JOBS: .
2	Displays the detailed custom resources information in addition to level 1 information. The custom resources set to resource unit or resource group are displayed following CUSTOM_RESOURCE, and the custom resources set to node are displayed following CUSTOM_RESOURCE(PER_NODE) . The meanings of display items, please refer to the description of --custom-resource option.
3	Displays the job ID of the job that uses this compute node as a communication path, in addition to level 2 information. The job ID is displayed following JOBS_USING_ROUTE: .

If this option is specified, the -l option is ignored.
When this option is specified in the system management node, it is necessary to set the cluster name to environment variable PXMYCLST or to specify the -c option at the same time.

--data

When this option is specified, the output is made a form for the processing.

--delimiter

This option specifies the delimiter when outputting it by the form for the processing.
Comma (",") is used when not specifying this option.
A character string beginning with a hyphen ("-") can not be specified.

--status

When the --status option is specified, compute resources in any state are displayed.
Note that this option must be specified together with -n or -v.
In addition, if this option is used on the system management node, the -c option must also be specified at the same time, or the cluster name must be set in the environment variable PXMYCLST.

--custom-resource|-C

When the --custom-resource option is specified, custom resources (set to resource unit and resource group) are displayed.
The custom resources set to resource unit or resource group are displayed following CUSTOM_RESOURCE, and the custom resources

set to node are displayed following CUSTOM_RESOURCE(PER_NODE).
The meanings of display items are as follows.

Display item	Meanings
RSCNAME	It means custom resource name. When the custom resources are defined as the type, the form of "custom resource name/type name" is displayed.
TOTAL	It means total number of custom resources. When the custom resources are defined as the type, "unlimited" is displayed.
FREE	It means total number of custom resources what is not allocated to jobs. When the custom resources are defined as the type, "unlimited" is displayed.
ALLOC	It means total number of custom resources what is allocated to jobs. When the custom resources are defined as the type, the number is calculated as one custom resource specified.
NODE	It means total number of node defined custom resources. It is displayed in the case of custom resources set to node.

When this option is specified in the system management node or the compute cluster management node, the option must be specified together with --rscunit or --rscgrp.

To specify this option, the -E option, -l option, -v (level 0 or 1 specified) option and --status option are ignored.

In addition, if this option is used on the system management node, the -c option must also be specified at the same time, or the cluster name must be set in the environment variable PXMYCLST.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT VARIABLES]

PXMYCLST

The environment variable PXMYCLST can specify a cluster name. PXMYCLST allows you to omit specifying a cluster name with the -c option. PXMYCLST also displays the same contents as when a cluster name set with PXMYCLST is specified in the -c option. If a cluster name is explicitly specified with the -c option, the information on the cluster name specified with the -c option is displayed. PXMYCLST can specify only one cluster.

An example is shown below.

[EXAMPLE]

```
export PXMYCLST=cluster1
```

[EXAMPLE]

1. Display an overview of the compute resources of a specific cluster in cluster units.

```
$ pjshowrsc -c cluster2
[ CLST: cluster2 ]
RSCUNIT  NODE
          TOTAL    FREE    ALLOC
unit1    9216      9216      0
```

2. Display calculation resource information on the specified calculation node group.

```
$ pjshowrsc -c cluster2 -n 0x02030001-0x02030002
[ CLST: cluster2 ]
[ NODE: 0x02030001 ]
RSC  TOTAL  FREE  ALLOC
cpu   8      0      8
mem  1.4Gi   0    1.4Gi
```


[NODE: 0x02030002]			
RSC	TOTAL	FREE	ALLOC
cpu	8	0	8
mem	1.4Gi	0	1.4Gi

[END STATUS]

The end status is returned as follows:

0: Normal end

Other than 0: Abnormal end

[SEE ALSO]

pashowclst(8)

[MESSAGES]

If an error occurs during the execution of the `pjshowrsc` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PRM 3201 pjshowrsc Duplicated option is specified: *opt*

Meaning

The option *opt* is specified more than once.

opt: Option name

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 3202 pjshowrsc Conflicting option is specified

Meaning

An incompatible option is specified.

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 3203 pjshowrsc Cluster name is not specified

Meaning

No cluster name is specified.

Action

Specify a cluster name, and execute the command again.

[ERR.] PRM 3204 pjshowrsc Invalid option: *opt*

Meaning

The specified option is invalid.

opt: Option name

Action

Review the option, and execute the command again.

[ERR.] PRM 3205 pjshowrsc -v [level]: The level must be 0-3

Meaning

The specified level is incorrect.

Action

Specify the detail level correctly, and execute the command again.

[ERR.] PRM 3206 pjshowrsc Specification of cluster name is wrong: *clstname*

Meaning

The specified cluster name *clstname* is incorrect.

Action

Specify the correct cluster name, and execute the command again.

[ERR.] PRM 3207 pjshowrsc Specification of resource unit name is wrong: *rscuname*

Meaning

The specified resource unit name *rscuname* is incorrect.

Action

Specify the correct resource unit name, and execute the command again.

[ERR.] PRM 3208 pjshowrsc The cluster other than compute cluster is specified: *clstname*

Meaning

The specified cluster name *clstname* is not for a compute cluster.

Action

Specify the cluster name of a compute cluster, and execute the command again.

[ERR.] PRM 3209 pjshowrsc Specification ID (node group ID, boot group ID, or node ID) is wrong *id*

Meaning

id, which is the specified node group ID, boot group ID, or node ID, is incorrect.

Action

Specify the correct node group ID, boot group ID, or node ID, and execute the command again.

[ERR.] PRM 3210 pjshowrsc Specification method or specified range is wrong: *range*

Meaning

The method of specifying the range or the specified range *range* is incorrect.

Action

Specify the correct range, and execute the command again.

[ERR.] PRM 3211 pjshowrsc The same ID (node group ID, boot group ID, or node ID) is specified: *id*

Meaning

id, which is the specified node group ID, boot group ID, or node ID, is the same as that already specified.

Action

Specify the correct node group ID, boot group ID, or node ID, and execute the command again.

[ERR.] PRM 3212 pjshowrsc Specified node group ID, boot group ID, or node ID doesn't exist: *id*

Meaning

id, which is the specified node group ID, boot group ID, or node ID, does not exist.

Action

Specify the correct node group ID, boot group ID, or node ID, and execute the command again.

[ERR.] PRM 3213 pjshowrsc There is no resource in the resource unit: *rscuname*

Meaning

The specified resource unit name *rscuname* does not exist.

Action

Specify the correct resource unit name, and execute the command again.

[ERR.] PRM 3214 pjshowrsc In this node, pjshowrsc cannot be used.

Meaning

The pjshowrsc command cannot be used on this node.

Action

Execute it on the active system management node, compute cluster management node, or login node.

[ERR.] PRM 3215 pjshowrsc In the standby node, pjshowrsc cannot be used.

Meaning

The pjshowrsc command cannot be used on a standby node.

Action

Execute it on the system management node, compute cluster management node, or login node.

[ERR.] PRM 3216 pjshowrsc Permission denied ID: *id*

Meaning

The user does not have the privileges to display the state of the specified node (node ID: *id*).

Action

Confirm that the specified node ID is correct. If the specified node ID is correct, the user does not have the display privileges for that node ID. If necessary, contact the administrator to confirm the display privileges for the relevant node.

[ERR.] PRM 3218 pjshowrsc Job resource management function is not starting: *ret*

Meaning

The job resource management function has not been started.

ret: Maintenance information

Action

Wait a moment, and execute the command again.

If the same error recurs, contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PRM 3219 pjshowrsc The same resource unit name is specified: *rscuname*

Meaning

The same resource unit name is already specified.

Action

Specify the correct resource unit name, and execute the command again.

[ERR.] PRM 3220 pjshowrsc No memory (*details*)

Meaning

Memory acquisition failed.

details: Detailed information for maintenance

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PRM 3221 pjshowrsc Internal error (*details*)

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PRM 3222 pjshowrsc Permission denied rscunitname: *rscuname*

Meaning

The user does not have the privileges to display the state of the specified resource unit.

rscuname: Resource unit name

Action

Confirm that the specified resource unit name is correct. If the specified resource unit name is correct, the user does not have the display privileges for that resource unit. If necessary, contact the administrator to confirm the display privileges for the relevant resource unit.

[ERR.] PRM 3226 pjshowrsc There is no information that can be displayed with specified node group ID, boot group ID, or node ID: *id*

Meaning

There is no node information that can be displayed for the specified node group ID, boot group ID, or node ID *id*.

Action

The computer resources for the specified node group ID, boot group ID, or node ID *id* may be inactive. Ask the administrator to check the node status. The administrator should check the node status for the specified node group ID, boot group ID, or node ID *id* by using the pashowclst command.

If the error occurs even though the node status is normal, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PRM 3227 pjshowrsc There is no information that can be displayed with specified resource unit: *rscuname*

Meaning

There is no node information that can be displayed for the specified resource unit name *rscuname*.

Action

The computer resources for the specified resource unit *rscuname* may be inactive. Ask the administrator to check the node status. The administrator should check the node status of the specified resource unit *rscuname* by using the `pashowclst` command.

If the error occurs even though the node status is normal, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or the Support Desk with the collected data together with the output message.

[ERR.] PRM 3228 pjshowrsc A required option is not specified

Meaning

An indispensable option is not specified.

Action

Confirm the correct option, and execute the command again.

[ERR.] PRM 3229 pjshowrsc The specified resource group name is wrong: *rscgname*

Meaning

The specified resource group name *rscgname* does not exist. Alternatively, the user does not have the privileges to display it.

Action

Confirm that the specified resource group name is correct. If the specified resource group name is correct, the user does not have the display privileges for that resource group. If necessary, contact the administrator to confirm the display privileges for the relevant resource group.

[ERR.] PRM 3230 pjshowrsc The same resource group name is specified: *rscgname*

Meaning

The same resource group name *rscgname* is already specified.

Action

Execute the command again without specifying the same resource group name as one already specified.

[ERR.] PRM 3231 pjshowrsc There is no information that can be displayed with specified resource group: *rscgname*

Meaning

There is no node information that can be displayed for the specified resource group name *rscgname*.

Action

The computer resources for the specified resource group *rscgname* may be inactive. Ask the administrator to check the node status. The administrator should check the node status of the specified resource group *rscgname* by using the `pashowclst` command.

If the error occurs even though the node status is normal, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or the Support Desk with the collected data together with the output message.

[ERR.] PRM 3232 pjshowrsc No execute permission : *command*

Meaning

The user is not permitted to execute the command *command*.

Action

Display the job ACL settings with the `pjacl` command, and check whether the command *command* can be executed. If necessary, ask the administrator to add permission.

[ERR.] PRM 3299 pjshowrsc Internal error (details)

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

Warning Message

[WARN] PRM 3101 pjshowrsc Connection error: *clstname*

Meaning

Resource information display failed because communication was not possible with the active compute cluster management node in the cluster *clstname*.

If two or more clusters exist, resource information within the communicated range is displayed.

Action

Ask the cluster administrator to check the system state. The cluster administrator is requested to check the state of the compute cluster management node in the cluster *clstname* by using the `pshowclst` command. If the error occurs even though the compute cluster management node is in the normal state, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

Information Message

[INFO] PRM 3001 pjshowrsc There are no resources can be displayed

Meaning

No resource can be displayed.

Action

No action is necessary.

3.1.12 pjsig command

[NAME]

`pjsig` - Sends signal to running jobs.

[SYNOPSIS]

```
/usr/bin/pjsig -s signal jobid ...  
/usr/bin/pjsig --help
```

[DESCRIPTION]

The `pjsig` command sends signals to running jobs (status RUNNING). This command can be executed on the login node and compute cluster management node.

If non-running jobs are specified, errors will be prompted. Under the situation that multiple jobs are specified and part of these are running, only the running ones can receive signals. The others will be prompted with error messages.

Job ID or sub job ID *jobid* is specified so that the signals can be sent. Multiple or ranged job ID or sub job ID of step job can be specified. If a job ID is specified, the signal should be sent to all the running sub jobs belong to this job (Except the non-running sub jobs). If a sub job ID is specified, the signal should be sent to this sub job (If it is not running, an error occurs).

If the command is executed more than once for the same job, signals should be accepted on condition that this job is running. However, only

the signals which can be handled will be sent to the job.
After that the signals are sent, the command will not wait for the result and returns at once.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

```
1-100
```

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

```
100_1-10
```

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

```
101[1-10]
```

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges
- Jobs in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly (Processing is not aborted, and all specified jobs are processed).

[OPTIONS]

-s *signal*

With this option, signal number or signal name can be specified.
The signal number should be from 1 to 64, and the signal name should be less than 15 characters.

--help

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following gives examples of pjsig command.

```
$ pjsig -s 9 1
[INFO] PJM 0700 pjsig Accepted job 1.
```

```
$ pjsig -s SIGKILL 2
[INFO] PJM 0700 pjsig Accepted job 2.
```

Under the situation that jobs with ID 1 and 3 are running, and job with ID 2 is not running.

```
$ pjsig -s 9 1 2 3
[INFO] PJM 0700 pjsig Accepted job 1.
[ERR.] PJM 0713 pjsig Job 2 status is not running.
[INFO] PJM 0700 pjsig Accepted job 3.
```

[END STATUS]

- 0: Normal end
- 1: Abnormal end (If multiple jobs are specified, even only one becomes an error, the return value should also be this one.)

[MESSAGES]

If an error occurs during the execution of the pjsig command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 0701 pjsig Unknown option or no option argument: *opt*.
Try `pjsig --help` for more information.

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

[ERR.] PJM 0702 pjsig No signal specified.
Try `pjsig --help` for more information.

Meaning

No signal number or signal name is specified.

Action

Specify a signal number or signal name.

[ERR.] PJM 0703 pjsig Unknown option argument: *opt arg*.
Try `pjsig --help` for more information.

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0704 pjsig Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

[ERR.] PJM 0710 pjsig No jobid(s) specified.
Try `pjsig --help` for more information.

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0711 pjsig Invalid jobid :*jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0712 pjsig Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 0713 pjsig Job *jobid* status is not running.

Meaning

The specified job (job ID: *jobid*) is not in the RUNNING state.

Action

Check the job state.

[ERR.] PJM 0770 pjsig No execute permission.

Meaning

The user is not permitted to execute the pjsig command.

Action

Display the job ACL settings with the pjacl command, and check whether the pjsig command can be executed (execute pjsig). If necessary, ask the administrator to add permission.

[ERR.] PJM 0771 pjsig No permission: *jobid*.

Meaning

The user is not permitted to operate the target job (job ID: *jobid*).

The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the pjacl command. Then, confirm the permission (permit pjsig) for the operated objects of the pjsig command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0790 pjsig PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function or the system state.

[ERR.] PJM 0791 pjsig Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0793 pjsig The operating command cannot be executed except login node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a login or compute cluster management node.

Action

Execute the command on a login or compute cluster management node.

Information Message

[INFO] PJM 0700 pjsig Accepted job *jobid*.

Meaning

The request to send a signal to a job (job ID: *jobid*) was accepted.

Action

No action is necessary.

[INFO] PJM 0781done.

Meaning

The request to send a signal to a job waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

3.1.13 pjstat command

NAME

`pjstat` - Displays information about jobs.

SYNOPSIS

```
/usr/bin/pjstat [-c clstname]
                  [{--rscunit|--ru} [rscuname[,...]]]
                  [{--rscgrp|--rg} [rscgname[,...]]]
                  [-A|--all]
                  [--choose item[,...]]
                  [--data [--delimiter delimiter]]
                  [--filter item=value[+value...][,item=...]]
                  [-H|--history [day=value
                                |start=sdate,end=edate
                                |start=sdate,period=value]]
                  [-v [--pattern=value]|-s|-S]
                  [-E|--expand]
                  [--sort item:kind[,...]]
                  [--summary|--with-summary]
                  [jobid ...]
/usr/bin/pjstat [-c clstname] --rsc
                  [{--rscunit|--ru} rscuname[,...]]
                  [{--rscgrp|--rg} rscgname[,...]] [--shape]
/usr/bin/pjstat [-c clstname] --rsc [unit=rscuname[:...]] [--shape]
/usr/bin/pjstat [-c clstname] --limit
                  [{--rscunit|--ru} rscuname] [{--rscgrp|--rg} rscgname]
                  [--user {uid|uname}] [--group {gid|gname}]
```

```
/usr/bin/pjstat [-c clstname] -X [-E|--expand] [jobid ...]
/usr/bin/pjstat --help
```

DESCRIPTION

The pjstat command outputs information about jobs. This pjstat command can be executed on the login node, compute cluster management node and system management node.

If the pjstat command has specified the *jobid*, then it outputs information of the corresponding jobs. If the command has not specified the *jobid*, then it outputs information of jobs submitted by the executor. But if the executor is the manager, then the command outputs information of all the submitted jobs.

If the specified job ID does not exist, the command will ignore the job ID and do not output the error message.

The followings are the items and contents that are output as the standard information about jobs:

JOB_ID	JOB_NAME	MD	ST	USER	START_DATE	ELAPSE_LIM	NODE_REQUIRE	VNODE
CORE	V_MEM							
XXXXXXXXXX	XXXXXXXXXX	XX	XXX	XXXXXXXXXX	MM/DD hh:mm:ss	hhhh:mm:ss-hhhh:mm:ss	nnnnnn:XXxYYxZZ	nnnnnn
nnn	nnnnnnnnnnMiB							

Each line displays detailed information about a job. The output items can be changed depending on the specified options.

Moreover, when the -E option is specified, not only the information corresponding to job ID (It is called summary information on the bulk/step job at the following) but also the information corresponding to the sub job IDs are displayed for the bulk/step job.

The following table shows items (standard items) and their contents displayed by the command executed with no options specified.

Name	Description
JOB_ID	Job id (10-digit decimal number) When the -E option is specified, the information on the sub jobs are displayed. For the sub job, the sub job ID is output. (jobID[BulkNumber] or jobID_StepNumber)
JOB_NAME	Job name (Only first 10 characters)
MD	Job model NM: normal job ST: step job BU: bulk job MW: master-worker job
ST	Current processing state of the job ACC: Accepted job submission RJT: Rejected job submission QUE: Waiting for job execution RNA: Acquiring resources required for job execution RNP: Executing prologue RUN: Executing job RNE: Executing epilogue RNO: Waiting for completion of job termination processing SPP: Suspend in progress SPD: Suspended RSM: Resume in progress EXT: Exited job end execution CCL: Exited job execution by interruption HLD: In fixed state due to users ERR: In fixed state due to an error
USER	Name of user who executes job (Only first eight characters)
START_DATE	If job is yet to be executed, planned execution start time is output. If job is being executed or has been executed, actual start time is output. (A planned execution start time is enclosed in parentheses.) Planned execution start time "(MM/DD hh:mm)" If planned execution start time is after one year or later, it is output in the format of "(YYYY/MM/DD)". After the execution is started "MM/DD hh:mm:ss"

Name	Description
	<p>As for jobs to which backfill is applied, "<" is added after the time. <code>"(MM/DD hh:mm)<"</code> <code>"MM/DD hh:mm:ss<"</code></p> <p>If the job has specified the start time, "@" is added after the time. <code>"(MM/DD hh:mm)@"</code> " <code>"MM/DD hh:mm:ss@"</code></p> <p>If planned execution start time exceeds the schedule period, "#" is added after the time. <code>"(MM/DD hh:mm)@"</code></p> <p>In this case, the planned execution start time is after <code>"MM/DD hh:mm"</code>.</p> <p>For the summary information on a step job, the information on the sub job under execution is output. (When there is no sub job under execution, the information on the sub job executing next is output.)</p> <p>For the summary information on a bulk job, "-" will be output.</p>
ELAPSE_LIM	<p>Elapsed time limit "<code>hhhh:mm:ss</code>", or Minimum value of the elapsed time limit-Maximum value of the elapsed time limit value "<code>hhhh:mm:ss-hhhh:mm:ss</code>".</p> <p>When column overflow occurs, output is executed with ss omitted. For the summary information on a step job, "-" will be output. For the summary information on a bulk job, "-" will be output when the elapsed time limit value differs for each sub job.</p>
NODE_REQUIRE	<p>Shape and number of nodes when job is submitted: "<code>nnnnnn:XXxYYxZZ</code>" [Only FX server of the shape is effective]</p> <p>If the format above is too short to hold information, only number of nodes is output. For the node allocated job executed on PRIMERGY server compute node, only number of nodes is output. For the summary information on a step job, "-" will be output.</p>
VNODE	<p>Number of virtual nodes "<code>nnnnnn</code>"</p> <p>For the job executed on FX server compute node, "-" will be output. For the node allocated job executed on PRIMERGY server compute node, "-" will be output. For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.</p>
CORE	<p>Number of CPU cores for each virtual node "<code>nnn</code>"</p> <p>For the job executed on FX server compute node, "-" will be output. For the node allocated job executed on PRIMERGY server compute node, "-" will be output. For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.</p>
V_MEM	<p>Amount of memory for each virtual node(vnode-mem) "<code>nnnnnnnnnnMiB</code>"</p> <p>When cpu-mem is specified, the amount of the memory is converted into vnode-mem (Multiply it by the number of cores), and the amount of the memory is output by the unit of MiB. For the node allocated job, "-" will be output. For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.</p>

Specify the --with-summary option to also display the number of jobs by state as a summary.

```
$ pjstat --with-summary
ACCEPT QUEUED RUNING RUNOUT HOLD   ERROR  SUSPND REJECT EXIT   CANCEL  TOTAL
nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn
s nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn nnnnnn
JOB_ID   JOB_NAME   MD ST   USER      START_DATE      ELAPSE_LIM      NODE_REQUIRE    VNODE
CORE V_MEM
XXXXXXXXXX XXXXXXXXXXXX XX XXX XXXXXXXX MM/DD hh:mm:ss  hhhh:mm:ss-hhhh:mm:ss nnnnnn:XXxYYxZZ nnnnnn
nnn  nnnnnnnnnnnMiB
```

The summary on the number of jobs displays the number of jobs submitted by the user who executed the pjstat command (or all users when the -A option is specified). The jobs have been added up by state as follows.

Name	Description
ACCEPT	Number of jobs being accepted
QUEUED	Number of jobs waiting for job execution
RUNING	Number of jobs being executed
RUNOUT	Number of jobs waiting job end
HOLD	Number of jobs in fixed state due to a user
ERROR	Number of jobs in fixed state due to an error
SUSPND	Number of jobs in the job suspended period (total number of jobs in one of SPP, SPD, RSM state)
REJECT	Number of jobs rejected and not accepted
EXIT	Number of jobs that have ended
CANCEL	Number of jobs suspended by the user or administrator
TOTAL	Number of total jobs

The summary on the number of jobs appears on two rows.

At the top, bulk jobs and step jobs have been added up in units of jobs.

At the bottom (row starting with "s"), the numbers of sub jobs for the bulk jobs and step jobs have been added up.

When the command assigns the additional information option -v, additional information is output.

For the items that are output, see the description of the -v option.

```
$ pjstat -v
JOB_ID      JOB_NAME  MD ST  USER      GROUP   START_DATE      ELAPSE_TIM ELAPSE_LIM
RANK       LST EC   PC  SN PRI ACCEPT      RSC_GRP  REASON
XXXXXXXXXX XXXXXXXXXX XX  XXX XXXXXXXX XXXXXXXX MM/DD hh:mm:ss< hhhh:mm:ss hhhh:mm:ss-hhhh:mm:ss
XXXXXXXXXX XXX  XXX  XXX  XX  XXX MM/DD hh:mm:ss XXXXXXXX XXXXXXXXXXXXXXXXX
```

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify job IDs 1 to 100 by range

```
1-100
```

Example: Specify step number 1 to 10 of step job (job ID = 100) by range

```
100_1-10
```

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

```
101[1-10]
```

If job IDs are specified by range, even if jobs like the following are included, they are not handled as errors but ignored, and the entire processing continues.

- Jobs that do not exist
- Jobs for which the user does not have privileges

[OPTIONS]

```
{--rscunit|--ru} [rscuname[,...]]
```

This option displays information on the job in each resource unit.

When the resource unit name *rscuname* is specified, jobs in the specified resource unit are output.

When the resource unit name *rscuname* is omitted, jobs of all the resource units are output.

Two or more resource unit names *rscuname* can be specified by delimiting it by comma (",").

If you specify the --sort option at the same time, jobs are sorted in each resource unit.

{--rscgrp|--rg} [*rscgname*[,...]]

This option displays information on the job in each resource group.

When the resource group name *rscgname* is specified, jobs in the specified resource group are output.

When the resource group name *rscgname* is omitted, jobs of all the resource groups are output.

Two or more resource group names *rscgname* can be specified by delimiting it by comma (",").

If you specify the --sort option at the same time, jobs are sorted in each resource group.

-A|--all

The summary object is all clusters. The command displays all the jobs in the cluster.

-c *clstname*

The option displays various information of the jobs in cluster *clstname*. When the command is executed on the system management node and the environment parameter PXMYCLST has been set, then this option can be omitted.

When the command is executed on the node with other types, this option will be ignored.

--choose *item*[,...]

This option outputs only the items specified in *item* in the specified order as job information. If the same item is specified in duplicate, the one specified last is valid.

A table shown below lists which items specified after *item* in the --choose option are valid.

Any specified item that is invalid or not saved as job statistical information will not be output.

--data

The option outputs in the processed form.

--delimiter *delimiter*

When the pjstat outputs in the processed form, this option specified the delimit character.

When this option has not been specified, the output use comma as the delimit character.

A character string beginning with a hyphen("-") can not be specified.

--filter *item=value*[+*value*...][,*item=*...]

This option outputs only the jobs that have the value in *value* for the item in *item*. If you specify multiple items, only the jobs that satisfy all those conditions are output.

If the same item is specified in duplicate, the one specified last is valid. A table shown below lists which items specified after *item* in the --filter option are valid.

For any specified item that is invalid or not saved as job statistical information, it is regarded as a condition that no job satisfies.

[How to specify value]

- Specify by value

If the item matches the value specified, it will be output.

Multiple values can be specified by using + (plus).

Example:

`pjstat --filter "jid=100"`

Specifies information with the job ID set to 100.

`pjstat --filter "jid=1+10"`

Specifies information with the job ID set to 1 or 10.

`pjstat --filter "jnam=1\+10"`

Specifies information with the job name set to "1+10".

To specify + (plus) as a value, place a \ (back slash) before the +.

`pjstat --filter "jnam=1\,10"`

Specifies information with the job name set to "1,10".

To specify , (comma) as a value, place a \ (back slash) before the ,.

- Specify by range

If the item is included in the range specified, it will be output.

The range is delimited by - (hyphen). Multiple ranges can be specified by using + (plus).

Example:

```
pjstat --filter "jid=1-10"
```

Specifies information with the job ID from 1 to 10.

```
pjstat --filter "jid=1-"
```

Specifies information with the job ID of 1 or more.

```
pjstat --filter "jid=-10"
```

Specifies information with the job ID of 10 or less.

```
pjstat --filter "jid=1-10+21-30"
```

Specifies information with the job ID from 1 to 10 and from 21 to 30.

- Specify by wildcard

If the character specified is included in the item, it will be output.

The following regular expressions can be used for specification.

Sign	Description
. (period)	One arbitrary character except line feed
* (asterisk)	Arbitrary character string of zero or more characters except line feed characters

Multiple wild cards can be specified by using + (plus).

Example:

```
pjstat -v --filter "grp=*abc"
```

Specifies jobs whose end of group name is "abc".

```
pjstat -v --filter "grp=abc*"
```

Specifies jobs whose beginning of group name is "abc".

```
pjstat -v --filter "grp=*abc*"
```

Specifies jobs of which some part of group name is "abc".

```
pjstat -v --filter "grp=*abc*+*efg*"
```

Specifies jobs of which some part of group name is "abc" or "efg".

Example:

```
pjstat --filter "jnam=1\.*+10"
```

Specifies information with the job name "1.*+10".

To specify "." (period), "*" (asterisk), and "+" (plus) as values, place "\" (back slash) before each character.

-H | --history [day=value | start=sdate,end=edate | start=sdate,period=value]

This option displays only the information on jobs that finished processing (in the REJECT, EXIT, or CANCEL state).

If neither the -H nor --history argument is specified, information on jobs that finished in the past three days is output.

day=value

Outputs information on jobs that finished within the past number of days represented by *value*.

The *value* is an integer [1-365].

start=sdate,end=edate

Outputs information on jobs that finished from *sdate* to *edate*.

Specify the date in the "YYYYMMDD[hhmm]" format. (YYYY: year, MM: month, DD: day, hh: hour, mm: minute).

If a future date is specified, today is assumed to be specified.

start=*sdate*,period=*value*

Outputs information on jobs that finished within the number of days represented by *value* after *sdate*.

Specify the date in the "YYYYMMDD[hhmm]" format. (YYYY: year, MM: month, DD: day, hh: hour, mm: minute).

If a future date is specified, today is assumed to be specified.

When you also specify the -E option, information on the sub jobs of bulk jobs and step jobs is also displayed. A finished sub job is displayed only when the bulk job or step job has finished.

When you do not specify the -H option in a case where just some sub jobs are finished and the bulk job or step job is not yet finished, the finished sub jobs are displayed.

-v [--pattern=*value*]

Outputs the following items in addition to the standard items.

Name	Description
GROUP	Group name of user executing job (Only first 8 characters) For the summary information on a step job, the information on the sub job under execution is output. (When there is no sub job under execution, the information on the sub job executing next is output.)
ELAPSE_TIM	Elapsed time limit "hhhr:mm:ss" When column overflow occurs, output is executed with ss omitted. The execution elapsed time does not include the time that the job status was SPP, SPD or RSM. For the summary information on a step job or bulk job, output the total of completed sub jobs.
V_POL	Arrangement policy of virtual node A_PCK : ABS_PACK PACK : PACK A_UPK : ABS_UNPACK U_PCK : UNPACK For the job executed on FX server compute node, "-" will be output. For the node allocated job executed on PRIMERGY server compute node, "-" will be output. For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.
E_POL	Execution mode policy SHARE : SHARE SMPLX: SIMPLEX For the job executed on FX server compute node, "-" will be output. For the node allocated job executed on PRIMERGY server compute node, "-" will be output. For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.
RANK	The allocation rule of the rank is output For the job executed on PRIMERGY server compute node, the following is output: rank-map-bynode : "bynode" rank-map-bychip[= <i>n</i>]: "bychip <i>n</i> " When "= <i>n</i> " is specified for "--mpi rank-map-bychip", " <i>n</i> " is output. When option has been omitted, for the node allocated job, for the summary information on a step job, "-" will be output. For the job executed on FX server compute node, the following is output: rank-map-bynode : "bynode" rank-map-bychip : "bychip" For the summary information on a step job, "-" will be output. This item may be configured by the administrator to not be displayed.
LST	Last processing state of the job (prior state from which the "current processing state of the job" is shifted) The meaning of the value is the same as the standard information item "ST" of the job.
EC	Job script exit code

Name	Description
	For the summary information on a step job or a bulk job, "-" will be output.
PC	<p>PJM code</p> <p>Code that indicates the job manager processing result of the job execution</p> <p>0: Normal end of job</p> <p>1: Cancellation due to pjdel command</p> <p>2: Rejection due to reject the job</p> <p>3: Rejection to execute by gate check</p> <p>4: Hold due to pjhold command</p> <p>6: Cancellation due to step job depending relationship</p> <p>7: Cancellation due to specified deadline</p> <p>8: Cancellation due to gate check</p> <p>9: Specification of norestart job, so job is EXIT when rebuild job information.</p> <p>11: Job execution timeout due to elapsed time limit violation</p> <p>12: Forced termination due to use of excessive amount of memory</p> <p>16: Job termination due to an current directory or stdin/stdout/stderr file inaccessible</p> <p>18: A job that was being executed exceeding the minimum value of elapsed time limit was terminated by the start of another job or the start of a deadline.</p> <p>20: Node failure</p> <p>21: Failure in shell execution</p> <p>22: ICC error</p> <p>23: Termination due to OOM Killer</p> <p>25: Failure in HA</p> <p>26: Error in prologue/epilogue</p> <p>27: Error in resource management exit script</p> <p>28: Error in the job execution environment</p> <p>29: Specified job execution environment information that is invalid</p> <p>30: Abort due to failure of suspend or resume operation</p> <p>100: Internal error in job manager</p> <p>120: Internal error in job scheduler</p> <p>140: Internal error in job resource management</p> <p>160: Internal error in tofu library [FX server]</p> <p>180: Internal error in layered storage system</p> <p>For the summary information on a step job or a bulk job, "-" will be output.</p>
SN	<p>Signal number</p> <p>For the summary information on a step job or a bulk job, "-" will be output.</p>
PRI	<p>Job priority (0-255)</p> <p>A larger number has a higher priority.</p> <p>For the summary information on a step job, "-" will be output.</p>
ACCEPT	Job submission date and time "MM/DD hh:mm:ss"
RSC_GRP	Resource group when job is submitted
REASON	<p>Error message</p> <p>Message corresponding to result code for some processing of a job regardless of whether job is executed (Only first 16 characters) The following messages are output:</p> <p>-:</p> <p>No error</p> <p>MM/DD hh:mm DELAY: Estimated time of execution is delayed from the specified execution time(MM/DD hh:mm)</p> <p>ANOTHER JOB STARTED: A job that was being executed exceeding the minimum value of elapsed time due to the start of another job ended.</p>

Name	Description
	<p>DEADLINE SCHEDULE STARTED: A job that was being executed exceeding the minimum value of elapsed time limit due to the start of the deadline schedule ended.</p> <p>ELAPSE LIMIT EXCEEDED: Exceed the elapse time limit</p> <p>FILE IO ERROR: Cannot access the current directory at the time of the job submitted</p> <p>GATE CHECK: Canceled by gate control</p> <p>IMPOSSIBLE SCHED: Disabled scheduling</p> <p>INSUFF CPU: Physically insufficient number of CPUs</p> <p>INSUFF MEMORY: Physically insufficient memory amount</p> <p>INSUFF NODE: Physically insufficient number of nodes</p> <p>INSUFF <i>CustomRscName</i>: Insufficient custom resource defined with the resource name <i>CustomRscName</i></p> <p>INTERNAL ERROR: Internal error</p> <p>INVALID HOSTFILE: The hostfile is invalid [FX server]</p> <p>LIMIT OVER MEMORY: Exceed the memory in job execution</p> <p>LOST COMM: All-to-all communication of parallel process not guaranteed</p> <p>NO CURRENT DIR: The current directory that submitted a job, or stdin/stdout/stderr file does not exist, or there is no authorization to access.</p> <p>NOT EXIST <i>CustomRscName</i>: No custom resource with the resource name <i>CustomRscName</i> is defined.</p> <p>RESUME FAIL: The job failed to resume.</p> <p>RSCGRP NOT EXIST: Resource group doesn't exist, or the number of compute nodes assigned to the resource group is zero.</p> <p>RSCGRP STOP: Stopped resource group</p> <p>RSCUNIT NOT EXIST: Resource unit doesn't exist</p> <p>RSCUNIT STOP: Stopped resource unit</p> <p>RUNLIMIT EXCEED: Number of simultaneous executions exceeds limit.</p> <p>SUSPEND FAIL: The job failed to suspend.</p>

Name	Description
	<p>USELIMIT EXCEED: Waiting for execution by simultaneous number of use node limitations or use core limitations</p> <p>USER NOT EXIST: There is not the job submitted user in a system</p> <p>WAIT SCHED: The job is exempt from scheduling because the limit on the number of jobs subject to scheduling has been reached.</p> <p>Other character strings: Following message is output.</p> <ul style="list-style-type: none"> - Reason specified by pjdel, pjhold, or pmsuspend command. Or a message set by the job operation administrator. - The administrator set the message with the job manager exit function, job scheduler exit function, or resource management exit function. <p>When the --reason option is specified for the pjhold and pmsuspend commands, reason is displayed in the form of "<i>command execution user: specified message</i>".</p> <p>When the option (--reason) is not specified, "<i>command execution user:</i>" is displayed.</p> <p>For the summary information on a step job, "-" will be output.</p>

--pattern=*value*

The output content will be changed according to the value specified by --pattern option.

Only 1 is currently supported for *value*. The following information is additionally output.

Section name	Description
SUB_JOB_NUM	Number of sub jobs
NODE_ALLOC	<p>Allocated node shape and count (<i>N:Xx YxZ</i>)</p> <p>For the summary information on a bulk job, output the total of completed sub jobs.</p> <p>For the summary information on a step job, "-" will be output.</p>
RSC_UNIT	Resource unit

When the command also specifies the -s, -S options, the command ends with outputting the error message.

-s

This option outputs the detail information of resource use condition and the resource limit value. For the output information, see pjstatsinfo(7).

When options -v, -S are specified at the same time, the command ends with outputting the error message.

-S

-S option outputs the result information that -s outputs and it also shows additional information about the job set node unit. For the output information, see pjstatsinfo(7).

When options -v, -s are specified at the same time, the command ends with outputting the error message.

-E|--expand

When there are sub jobs, this option outputs the sub job list.

--sort *item:kind* [*,item:kind...*]

This option outputs jobs after sorting them with the item in *item* as a key. If you specify multiple items, all of them are used as keys for sorting.

If the same item is specified in duplicate, the item specified last is valid.

The table shown below lists which items specified after *item* in the --sort option are valid. If a specified item is invalid or not saved as job statistical information, it is ignored.

The following characters can be specified for *kind*.

A: Ascending D: Descending

When this option has been omitted, the command sorts the output result by the job ID.

--rsc

This option displays the information of resource unit and resource group.

{--rscunit|--ru} *rscuname*[,*rscuname*...]

This option displays information on the specified resource unit *rscuname*.

Two or more resource unit names *rscuname* can be specified by delimiting it by comma (",").

Resource units / resource groups without permission to submit jobs are not displayed.

When this option is omitted, all the resource units with permission to submit jobs are displayed.

unit=*rscuname*[:*rscuname*...]

This option displays information on the specified resource unit *rscuname*.

Two or more resource unit names *rscuname* can be specified by delimiting it by colon (":").

Resource units / resource groups without permission to submit jobs are not displayed.

When this option is omitted, all the resource units with permission to submit jobs are displayed.

{--rscgrp|--rg} *rscgname*[,*rscgname*...]

This option displays information on the specified resource group *rscgname*.

Two or more resource group names *rscgname* can be specified by delimiting it by comma (",").

Resource units / resource groups without permission to submit jobs are not displayed.

When this option is omitted, all the resource groups with permission to submit jobs are displayed.

--shape

This option displays the information of maximum shape in resource group.

--limit

This option outputs the system limit value.

{--rscunit|--ru} *rscuname*

The system limit values of the resource unit specified by resource unit name *rscuname* are output. The resource unit name can be specified only once.

When this option has been omitted, the default resource unit of user who executes the command will be applied.

{--rscgrp|--rg} *rscgname*

The system limit values of the resource group specified by resource group name *rscgname* are output. The resource group name can be specified only once.

When this option has been omitted, the system limit value of the resource unit is outputted.

When --rscunit option and --rscgrp are specified at the same time, the system limit values of the resource group *rscgname* in the resource unit *rscuname* are displayed.

When both of --rscunit and --rscgrp options are omitted, the system limit values of the default resource unit of user who submit the command will be displayed.

When --rscunit option is omitted and --rscgrp option is specified, the system limit values of the resource group *rscgname* in the default resource unit of user who submit the command will be displayed.

--user {*uid* | *uname*}

This option outputs the limit value information of the specified user *uid* | *uname*.

--group {*gid* | *gname*}

This option outputs the limit value information of the specified group *gid* | *gname*.

-X

If job IDs are specified, this option displays the correspondence table of the allocated compute node ID about the jobs in RUNNING and the rank number of the process generated when the program started. (However, information on the process dynamically generated is not displayed.)

This option is used to know which node is being allocated by the jobs which are attached to the debugger.

If job IDs are not specified, this option displays the allocated compute node ID and rank number about all RUNNING jobs.

This option does not support master-worker jobs. If executed for a master-worker job, the results for the job will be incorrect.

--summary

This option displays only the summary of jobs.

--with-summary

This option displays the summary and the detail information of jobs.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[Specification of items]

Item names that can be specified with --choose option, --filter option, and --sort option are as follows.

The items shown with "o" in the table are valid. Note that the valid items depend on an option (-v, -s, -S) being specified at the same time.

item	Description	pjstat, pjstat -v	pjstat s/-S
jid	Job ID or sub job ID	o	o
jnam	Job name	o	o
jtyp	Job type		o
jmdl	Job model	o	o
rnum	Retry count		o
snum	Number of sub jobs	o	o
usr	Name of user executing job	o	o
grp	Name of group executing job	o	o
rscu	Resource unit	o	o
rscg	Resource group	o	o
pri	Job priority	o	o
sh	Shell path name		o
cmt	Comment		o
lst	Previous processing state of job	o	o
st	Current processing state of job	o	o
prmdt	PRM data collection time (<i>YYYY/MM/DD hh:mm:ss</i>)		o
ec	End code of shell script	o	o
sn	Signal number	o	o
pc	PJM code	o	o
errmsg	Error message	o	o
mail	E-mail send flag		o
adr	E-mail send destination address		o
sde	Step job dependency relational expression		o
mask	umask value of user submitting job		o
std	Path name of the standard output file		o
stde	Path name of the standard error output file		o
infop	Statistical information file path		o
adt	Job submission time (<i>MM/DD hh:mm:ss</i>)	o	o

item	Description	pjstat, pjstat -v	pjstat s/-S
qdt	Last queuing time		o
exc	EXIT/CANCEL state transition time		o
lhusr	Last hold user name		o
holnm	Hold count		o
thldtm	Accumulated hold time		o
sdt	Job execution start time	o	o
edt	Job execution end time		o
nnumr	Node shapes and count at job submission (<i>N:Xx YxZ</i> or <i>N:Xx Y</i> or <i>N</i>)	o	o
cnumr	Requested number of CPUs		o
elpl	Elapsed time limit or maximum value of elapsed time limit (<i>hhhh:mm:ss</i>)	o	o
mszl	Physical memory amount limit by node		o
pcl	CPU usage time limit (sec) by process		o
pcfl	Core file limit by process		o
pcpl	Max user process count limit by process		o
pdl	Data segment limit by process		o
prml	Lock memory size limit by process		o
pmql	POSIX message queue size limit by process		o
pofl	File descriptor limit by process		o
ppsl	Signal count limit by process		o
ppl	File size limit by process		o
psl	Stack segment limit by process		o
pvml	Virtual memory size limit by process		o
nnuma	Allocated node shape and count (<i>N:Xx YxZ</i> or <i>N</i>)	o	o
msza	Physical memory amount allocated to a node		o
cnumat	Total number of CPUs allocated		o
elp	Execution elapsed time (<i>hhhh:mm:ss</i>)	o	o
nnumv	Number of unavailable nodes within the allocated node range [FX server]		o
nnumu	Number of nodes used [FX server, or node allocated jobs on PRIMERGY server compute node]		o
nidlu	Node ID list of the nodes used (hexadecimals delimited by single-byte space)		o
tofulu	Tofu coordinate list ((<i>X,Y,Z</i>)) of the nodes used [FX server]		o
mmszu	Total max physical memory usage		o
cnumut	Total number of CPUs used		o
uctmut	Total user CPU time (ms)		o
sctmut	Total system CPU time (ms)		o
usctmut	Total user CPU time and total of system CPU time (ms)		o
vnid	Virtual node ID [For virtual node allocated jobs on PRIMERGY server compute node]		o(*1)
vnuma	Allocated number of virtual nodes [For virtual node allocated jobs on PRIMERGY server compute node]	o	o

item	Description	pjstat, pjstat -v	pjstat s/-S
vcnuma	Number of cores of each virtual nodes [For virtual node allocated jobs on PRIMERGY server compute node]	o	
vmema	Allocated memory amount for each virtual node	o	o(*2)
vmszu	Maximal amount of virtual memory used		o
vpol	Policy that arranges virtual node [For virtual node allocated jobs on PRIMERGY server compute node]	o(*1)	
epol	Execution mode policy [For virtual node allocated jobs on PRIMERGY server compute node]	o(*1)	
rankm	Rank map	o(*1)	

(*1) Only --choose option is valid.

(*2) Only --filter and --sort options are valid.

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

pjsub(1), pjdel(1), papjstats.conf(5), pjstatsinfo(7)

[MESSAGES]

If an error occurs during the execution of the pjstat command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 0201 pjstat Unknown option or no option argument: *opt*.
Try `pjstat --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0202 pjstat Invalid combination of options: *opt* and *opt*.
Try `pjstat --help` for more information.**

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

**[ERR.] PJM 0203 pjstat Unknown option argument: *opt arg*.
Try `pjstat --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0204 pjstat Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

[ERR.] PJM 0211 pjstat Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0270 pjstat No execute permission.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Display the job ACL settings with the `pjacl` command, and check whether the `pjstat` command can be executed (execute `pjstat`). If necessary, ask the administrator to add permission.

[ERR.] PJM 0290 pjstat PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function and the system state.

[ERR.] PJM 0291 pjstat Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information on maintenance

Action

- "details" is "code=99 detail=110"

Job operation management function may not have been set correctly. Contact the administrator. The administrator should check that the "MariaDB Settings" of "Job Operation Software Administrator's Guide for Job Management" is properly implemented, and if not, follow the steps.

- Other

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0295 pjstat No cluster specified on system management node.**Meaning**

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the pjstat command, or set a cluster name in the environment variable PXMYCLST.

[ERR.] PJM 0297 pjstat Resource unit does not exist: *rscuname*.**Meaning**

The specified resource unit *rscuname* does not exist.

Action

Specify the resource unit again correctly.

Information Message

[INFO] PJM 0281 pjstat Data loading...**Meaning**

Display preparation is in progress. Please wait.

Description

This message appears when display does not start within three seconds. One dot (.) appears every three seconds.

Action

No action is necessary.

3.1.14 pjsub command

[NAME]

pjsub - Submits jobs.

[SYNOPSIS]

```
[Normal Job]
/usr/bin/pjsub [options] [scriptfile]

[Bulk Job]
/usr/bin/pjsub --bulk --sparam startbulkno-endbulkno [options] [scriptfile]

[Step Job]
/usr/bin/pjsub --step [--sparam parameter[,...]] [options] [{scriptfile[ ...]|scriptfile[,...]}]

[Master-Worker Job]
/usr/bin/pjsub --mswk [options] [scriptfile]

[Interactive Job]
/usr/bin/pjsub --interact [--sparam wait-time=waittime] [options] [scriptfile]

[Usage Display]
/usr/bin/pjsub --help
```

[DESCRIPTION]

The pjsub command submits jobs for the job manager. This command can be executed on the login node and compute cluster management node.

Specify a script file containing the content of the jobs which is going to be executed. If no script file is specified, this command will read the content of a script from standard input. When the necessary script has been typed, the standard input will be closed by pressing the Ctrl +D keys, the pjsub command is completed.

The content of a script is internally held by the job manager. Therefore, even though the script file is modified after submitting a job, the submitted job will not be affected.

The options that can be specified in the command lines can also be written at the beginning of the contents of a shell script.

When the option is written in the script, the prefix must be written at the beginning of the script file. Otherwise, if other lines (except comment lines) have appeared, the option cannot be executed even if the directive prefix has been written afterwards. (When the prefix is #PJM, the content will be processed as the comment) If the same option as that written in a script file is specified in the command lines, the option (or argument) in the command lines has higher priority.

If an option that can set multiple arguments is specified in a script file and the command lines, the arguments specified in the script file are effective when they are not specified in the command lines.

When an option without parameter has been specified repeatedly, it will be processed as being specified only once with no error. When a parameter of option has been specified repeatedly, only the last specified parameter is valid. When the option can set multiple arguments, then all the parameters are valid by OR condition (such as --sparam sd=xxx for the step job).

When there is incompatible relationship between options, if the relationship has not been indicated particularly, the last specified option will be valid. The incompatible relationship which will be error and invalid should be indicated.

The job manager evaluates the content specified with the option of the pjsub command, according to each of the limit values that the job ACL defines, and then determines whether the job is accepted. If no options are specified in the pjsub command, the job manager uses the default value that the job ACL defines.

The result files shown below are created in the current directory as of the time of job submission.

- The files are created using the format of the data written to the standard output by the job ("keyword" + job ID + .out).
- The files are created using the format of the data written to the standard error output by the job ("keyword" + job ID + .err).

Keyword is specified by the option -o, --out, -e, --err, or -N|--name. If these options have not been specified, the keyword will be set as the script file name. If the job is submitted from the standard input, the keyword will be set as "STDIN".

If the first character of the script file name is a numeric character, the "keyword" is created by the script file (J + script file name). The number of characters of "keyword" is assumed to be 63 characters in maximums. When the number of characters of script file name or "J + script file name" exceeds 63 characters, only first 63 characters are set to the "keyword".

[OPTIONS]

--at *datetime*

This option executes the job after the specified execution start time *datetime*. When this option has not been specified, the job will search the available resource to execute from now on, and begin to execute at the time when the resource is assigned.

The format of execution start time *datetime* is *YYYYMMDD[hhmm]*. *hhmm* can be omitted. When *hhmm* is omitted, 0:00:00 will be specified. (Note: the value of second is always "00", so it do not need to be set). The pjsub command and shell cannot analyze the other *datetime* format correctly. This rule can also be used --at option in script file *scriptfile*.

When the job cannot execute at the specified time, as specified execution time is during maintenance period, then it will reserve the time after the specified execution time and the resource is available.

--bulk --sparam *startbulkno-endbulkno*

This option executes bulk job. The bulk job is submitted once, and executes a number of jobs (sub jobs) which have changed parameter script file.

When the bulk job is executing, the --sparam option must be specified.

startbulkno specifies the assigned ID of the first submitted sub job.

endbulkno specifies the assigned ID of the last submitted sub job.

When the option --interact, --step or --mswk has also been specified, the error occurs (Incorrect specified option combination).

--comment *comment*

Add the comment to job.

This option specifies any string that can be seen when confirming the job content. The comment can be 255 characters at most.

If this option has not been specified, but the environment variable PJM_COMMENT has been set when job is submitted, then the comment will be the string set in PJM_COMMENT.

{-C|--dir-prefix} *directiveprefix*

This option set directive prefix when describing the job script with any characters.

When this option has not been specified, the directive prefix will be "#PJM".

The maximum length of the directive prefix is 63 characters. And the character can only be single-byte alphabet, numeric character or other symbols.

This option cannot be written in the script file. When the option has been written, it will be invalid.

{-e|--err} *pathname*

This option directs the output file to the standard error output *pathname*. When this option has not been used, the standard error output file name will be the combination of job name and the job ID (or sub job ID when the job is a sub job) and string ".err".

pathname can use following metacharacter:

Metacharacter	Description
%j	Expand the job ID
%J	Expand the sub job ID
%b	Expand the bulk number
%s	Expand the step number
%n	Expand the job name

Note

When some sub jobs have specified the standard error output files with the same name, the output content cannot be guaranteed, because there is no file lock among sub jobs.

Do not specify the file name among sub jobs repeatedly.

{-g|--gname} *gname* | {-g|--gid} *gid*

When the job is executing, this option specifies the group which the process of job belongs to. The group can be specified with the group name *gname* or the group ID *gid*. If the specified group does not exist in the compute cluster management node, the job submission has an error.

--interact [*--sparam parameter*]

This option executes interactive job.

The option *--sparam* can specify the parameter *parameter* of the interactive job.

When *--sparam* option has not been specified, the parameters are set as the default values.

The parameter *parameter* can be specified as follows:

wait-time=*waittime*

This parameter specifies the time waiting for computing resource assignment in seconds. For interactive job, after submission, the computing resource is assigned at once and the job tries to execute. If the resource is insufficient, the job cannot execute at once, and then the job will wait for the specified time *waittime* and resource assignment. The range of *waittime* can be set from 0 to 36000. When the parameter has not been set, the *waittime* will be 0 to process. The *waittime* can also be set as unlimited, so the job will wait until the computing resource is assigned.

When the interactive job is canceled due to waiting for the computing resource assignment. Then an error message will be displayed and end the *pjsub* command.

When the option *--bulk*, *--step* or *--mswk* has also been specified, the error occurs (incorrect specified option combination).

--fs *filesystem*

An arbitrary character string of the filesystem name (*filesystem*) etc. is specified.

The specified filesystem name (*filesystem*) is output to statistical information.

The number of characters that can be specified for *filesystem* is up to 63.

If this option has not been specified, but the environment variable *PJM_FSNAME* has been set when job is submitted, then the filesystem name will be the string set in *PJM_FSNAME*.

--appname *applicationname*

An arbitrary character string of the application name (*applicationname*) etc. is specified.

The specified application name (*applicationname*) is output to statistical information.

The number of characters that can be specified for *applicationname* is up to 63.

If this option has not been specified, but the environment variable PJM_APPNAME has been set when job is submitted, then the applicationname will be the string set in PJM_APPNAME.

-j

The standard error output will be output to the standard output.

--llo *"param=value[,...]"*

This option specifies parameters for the Lightweight Layered IO-Accelerator (LLIO) function.

The numerical values or character strings that can be specified in *value* vary depending on *param*.

You can specify the following patterns for *param=value*. When specifying multiple parameters, delimit them with a comma (",").

sharedtmp-size=*stmpsize*

This parameter specifies the size of a shared temporary area. The size of the shared temporary allocated to a job is obtained by multiplying *stmpsize* by the number of allocated nodes. For the format of the value written in *stmpsize*, see the binary prefix in [LIMIT VALUE]. You can specify a value in mebibytes.

Minimum: 0 bytes, Maximum: 2147483647 mebibytes

localtmp-size=*ltmpsize*

This parameter specifies the size of a temporary area on a node. The temporary area on a node is configured by node. For the format of the value written in *ltmpsize*, see the binary prefix in [LIMIT VALUE].

Minimum: 0 bytes, Maximum: 2147483647 mebibytes

auto-readahead={on|off}

This parameter specifies whether to enable or disable the automatic read-ahead function.

on: Enable, off: Disable

async-close={on|off}

This parameter specifies whether to enable or disable the asynchronous close function.

on: Enable, off: Disable

cn-cached-write-size=*cncwsize*

If the output size of data to write to a file is smaller than the specified value, the data is not immediately output to storage but instead stored temporarily in cache on the compute node. For the format of the value written in *cncwsize*, see the binary prefix in [LIMIT VALUE]. You can specify a value in 4 kibibytes.

Minimum: 0 bytes, Maximum: 2147483647 kibibytes

cn-cache-size=*cncsize*

This parameter specifies the size of memory allocated to cache on the compute node.

For the format of the value written in *cncsize*, see the binary prefix in [LIMIT VALUE].

Minimum: 4 mebibytes, Maximum: 2147483647 mebibytes

cn-read-cache={on|off}

This parameter specifies whether data is cached on the compute node.

on: Cached, off: Not cached

sio-read-cache={on|off}

This parameter specifies whether data read from first-layer storage is cached in second-layer storage.

on: Cached, off: Not cached

stripe-count=*scnt*

This parameter specifies the number of stripes per file used for distributed allocation of a file in first-layer storage.

Minimum: 1, Maximum: 2147483647

stripe-size=*ssize*

This parameter specifies the stripe size for distributed allocation of a file in tier-1 storage. For the format of the value written in *ssize*, see the binary prefix in [LIMIT VALUE]. You can specify a value in kibibytes.

Minimum: 64 kibibytes, Maximum: 4194240 kibibytes

uncompleted-fileinfo-path=pathname

This parameter outputs unwritten file information to the file specified in *pathname*. If omitted, the default value defined by the job ACL function is adopted. *pathname* can use the following metacharacters.

Metacharacter	Description
%j	Expand the job ID
%J	Expand the sub job ID
%b	Expand the bulk number
%s	Expand the step number
%n	Expand the job name
%o	Expand the standard output file path
%e	Expand the standard error output file path

perf

This parameter outputs the LLIO performance information to a file.

perf-path=pathname

This parameter outputs the LLIO performance information to the file specified in *pathname*. If omitted, the default value defined by the job ACL function is adopted. Also, the parameter is ignored when perf is not specified. *pathname* can use the following metacharacters.

Metacharacter	Description
%j	Expand the job ID
%J	Expand the sub job ID
%b	Expand the bulk number
%s	Expand the step number
%n	Expand the job name
%o	Expand the standard output file path
%e	Expand the standard error output file path

{-L|--rsc-list} "resource=value[,...]"

This option specifies the upper limit value of the resource required by jobs.

Each *resource* has different numerical values or strings that can be specified in *value*.

In *resource=value*, these values or strings can be specified using the following pattern. When specifying multiple values or strings, delimit them with a comma (",").

The optional argument of {-L|--rsc-list} is recommended to be enclosed with double quotation (") because the shell metacharacters (such as "(", ";", and so on) are used in the optional argument.

node=quantity_of_resources

This option specifies the amount of the resource of the node assigned to the job. This option is used when you submit the "node allocated job" (node-exclusive job).

To specify the memory size limitation for each node for the job, use the node-mem option.

[FX server]

node=*N*[:torus|:mesh|:noncont][[:io-exclusive|:no-io-exclusive],
node=*N*/*xN2*[:torus|:mesh|:noncont][[:io-exclusive|:no-io-exclusive] or
node=*N*/*xN2xN3*[:torus|:mesh|:noncont][[:strict|:strict-io][[:io-exclusive|:no-io-exclusive]

This option specifies the number of nodes and shapes that are assigned to jobs.

For the MPI job, the largest shape of assigned nodes is specified. The specified shape must be sufficiently large so that all the processes that actually exist during execution, including those to be dynamically created, can be stored.

- For one-dimensional shapes, set " $N1$ ".
- For two-dimensional shapes, set " $N1 \times N2$ ".
- For three-dimensional shapes, set " $N1 \times N2 \times N3$ ".

You can specify the node shape followed by node assignment method.

- "torus" means the torus mode that allocates computer resources to a job in units of Tofu (12 nodes).
- "mesh" means the mesh mode that allocates computer resources to a job in units of node.
- "noncont" means the non-contiguous mode that allocates computer resources to jobs in units of node.

If omitted, the default value defined in the job ACL function applies.

When three-dimensional shape has been specified, option "strict" or "strict-io" can be specified. When "strict" has been specified, the node will be assigned as the specified shape (" $N1$ =(node number in X-axis)", " $N2$ =(node number in Y-axis)", " $N3$ =(node number in Z-axis)") without rotating the axes. When "strict-io" has been specified, the node will be assigned as the specified shape (" $N1$ =(node number in X-axis)", " $N2$ =(node number in Y-axis)", " $N3$ =(node number in Z-axis)") without rotating the axes and be assigned at a rack origin. The option "strict" or "strict-io" can be specified only the torus mode.

When two-dimensional or one-dimensional shape is going to be assigned, the three-dimensional shape will be used as following methods:

- For one-dimensional shapes, set " $N2=1$ ", " $N3=1$ ".
- For two-dimensional shapes, set " $N3=1$ ".

You can set from 1 to 2147483647 nodes for $N1$, $N2$ and $N3$.

When the number of nodes exceeds the specified value, the nodes are not assigned to jobs.

This option cannot be specified with vnode-core, vnode-mem and core-mem option at the same time.

:io-exclusive or :no-io-exclusive are specify the io-exclusive mode to the job using FX server. The io-exclusive mode is effective only FX server. When specify to the job using NOT FX server, these options are omitted.

:io-exclusive

Specify io-exclusive mode to job. The I/O to the first-layer storage of the job specified this option does not have impacts from another job's I/O. The nodes for io-exclusive mode job are round up to a unit of shelf (4 Tofu = 48 nodes). This option can be specified only the torus mode.

:no-io-exclusive

Specify io-sharing mode to job. The I/O to the first-layer storage of the job specified this option have impacts from another job's I/O.

[PRIMERGY server]

node= $N1$

This option specifies the number of nodes $N1$ assigned to the job.

Minimum value: 1. Maximum value: 2147483647

vnode=*quantity_of_resources*

This option specifies the amount of the resource of the virtual node assigned to the job. This option is used when you submit the "virtual node allocated job".

This option cannot be used with the --mpi max-proc-per-node option.

Number of virtual nodes, number of CPU cores per virtual node, and amount of memory per CPU core(or amount of memory per virtual node) are specified for an amount of the resource of the virtual node.

There are two specified forms as follows;

The specified form 1 and 2 cannot be specified at the same time.

1. Form that encloses the number of CPU cores per virtual node and the amount of memory with parentheses

vnode=*num*[(core=*value*][:core-mem=*value*|mem=*value*)]

2. Form for which the number of CPU cores per virtual node and the amount of memory are specified by vnode-core, core-mem and vnode-mem options.

vnode=*num*[,vnode-core=*value*][,{core-mem=*value*|vnode-mem=*value*}]

The following section describes two specified forms.

1. `vnode=num[(core=value)[;core-mem=value|mem=value]]`

Form	Description
<i>num</i>	Number of virtual nodes assigned to the job. Minimum value: 1. Maximum value: 2147483647
<i>core=value</i>	Number of CPU cores per virtual node (optional). Minimum value: 1. Maximum value: 2147483647
<i>core-mem=value</i>	Amount of memory per CPU core (optional). Minimum value: 1 mebibyte. Maximum value: unlimited (No limitation)
<i>mem=value</i>	Amount of memory per virtual node (optional). Minimum value: 1 mebibyte. Maximum value: unlimited (No limitation)

Example: "vnode=1(core=3;mem=50Mi)"

The optional argument is recommended to be enclosed with double quotation (") because the shell metacharacters (such as "(", ";", and so on) are used in the optional argument.

The format of the *value* can be referred to "LIMIT VALUE".

When the core, core-mem or mem options are omitted, the default values defined respectively by job ACL function are used. The core option is same as vnode-core shown by a specified form 2, the core-mem option is same as core-mem shown by a specified form 2, and the mem option is same as vnode-mem shown by a specified form 2.

2. `vnode=num[,vnode-core=value][,{core-mem=value|vnode-mem=value}]`

Form	Description
<i>num</i>	Number of virtual nodes assigned to the job. Minimum value: 1. Maximum value: 2147483647
<i>vnode-core=value</i>	Number of CPU cores per virtual node (optional). Minimum value: 1. Maximum value: 2147483647
<i>core-mem=value</i>	Amount of memory per CPU core (optional). When this option and vnode-mem option (Amount of memory per virtual node) are specified at the same time, the option specified behind is effective. Minimum value: 1 mebibyte. Maximum value: unlimited (No limitation)
<i>vnode-mem=value</i>	Amount of memory per virtual node (optional). When this option and core-mem option (Amount of memory per CPU core) are specified at the same time, the option specified behind is effective. Minimum value: 1 mebibyte. Maximum value: unlimited (No limitation)

Example: `vnode=1,vnode-core=3,vnode-mem=50Mi`

The format of the *value* can be referred to "LIMIT VALUE".

When the core, core-mem or mem options are omitted, the default values defined respectively by job ACL function are used.

The vnode-core option is same as core shown by a specified form 1, the core-mem option is same as core-mem shown by a specified form 1, and the vnode-mem option is same as mem shown by a specified form 1.

[FX server: Note]

The vnode option does not support FX servers.

[PRIMERGY server: Note]

When node option and vnode option are specified at the same time, the job will be submitted as "virtual node allocated job". The number of compute nodes that the job allocates may be limited depending on the arrangement policy of a virtual node(vn-policy).

- When vn-policy is "unpack":
The number of compute nodes that the job allocates will be less than or equal to the number specified in the node option.
- When vn-policy is not "unpack":
The number of compute nodes that the job allocates will not be limited (node option will be ignored).

elapse=elapsedtime

elapsedtime is specified in one of *elapsedtimelimit*, *elapsedtimemin*-, *elapsedtimemin-elapsedtimemax* format.

- **elapse=elapsedtimelimit**
For all the processes that configure jobs, this option sets *elapsedtimelimit* (executable time) for each job.
If the job is executing it, though the *elapsedtimelimit* has passed, the job is terminated by force.
The format of the value of *elapsedtimelimit* can be referred to "LIMIT VALUE".
Minimum value: 1 second. Maximum value: unlimited (No limitation)
- **elapse=elapsedtimemin-, elapsedtimemin-elapsedtimemax**
For jobs that are allocated compute node of FX servers, you can also specify an upper limit for the elapsed time as a range from the minimum value *elapsedtimemin* to the maximum value *elapsedtimemax*.
For jobs that specify *elapsedtimemin*, depending on the status of resources allocated to the job, continue to run beyond the time specified by *elapsedtimemin*. At the longest, it is executed up to the system setting or *elapsedtimemax* specified by the user.
If the job is executing it, though the *elapsedtimemax* has passed, the job is terminated by force.
When the executed job exceeds the *elapsedtimemin* and the other job is allocated to the same resources, the job is terminated by sending SIGTERM. After sending SIGTERM, if the job is running after grace is set by administrator, SIGKILL is sent to the job.
The format of the value of *elapsedtimemin* and *elapsedtimemax* can be referred to "LIMIT VALUE".
Minimum value of *elapsedtimemin*: 1 second. Maximum value: unlimited (No limitation)
Minimum value of *elapsedtimemax*: 2 second. Maximum value: unlimited (No limitation)

node-mem=pernodememorylimit

This option sets the memory limit for the memory used for each node for the job. Only node allocated job is effective.
When this option has been set, the memory size per node is limited to *pernodememorylimit*. The format of the value of *pernodememorylimit* can be referred to "LIMIT VALUE".

Minimum value: 1 MiB. Maximum value: unlimited (No limitation)

{rscunit|ru}=rscuname

This option sets the resource unit name *rscuname* where the job is submitted. The resource unit name can be set with 63 bytes at most.

{rscgrp|rg}=rscgname

This option sets the resource group name *rscgname* where the job is submitted. The resource group name can be set with 63 bytes at most.

proc-core=perprocesscorefilelimit

This option sets the maximum core file size limit for all the processes that constitutes the job.
The format of the value of *perprocesscorefilelimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

proc-cpu=perprocesscpulimit

This option sets the maximum CPU time limit for each process of all the processes constitutes the job.
The format of the value of *perprocesscpulimit* can be referred to "LIMIT VALUE".

Minimum value: 1 second. Maximum value: unlimited (No limitation)

proc-crproc=perprocesscreateprocesseslimit

This option sets the user process number limit in the node which created by the executed user ID which can call the process for each process of all the processes constitutes the job.
The format of the value of *perprocesscreateprocesseslimit* can be referred to "LIMIT VALUE".

Minimum value: 0. Maximum value: unlimited (No limitation)

proc-data=*perprocessdatalimit*

This option sets the maximum data segmentation size limit for each process of all the processes constitutes the job.
The format of the value of *perprocessdatalimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

proc-lockm=*perprocesslockedmemorylimit*

This option sets the maximum lock memory size limit on RAM for each process of all the processes constitutes the job.
The format of the value of *perprocesslockedmemorylimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

proc-msgq=*perprocessmessagequeuelimit*

This option sets the maximum POSIX message queue size limit in the node which can be sure to the executed user ID which can call the process for each process of all the processes constitutes the job.
The format of the value of *perprocessmessagequeuelimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

proc-openfd=*perprocessopenfileslimit*

This option sets the maximum file descriptor number limit for each process of all the processes constitutes the job.
The format of the value of *perprocessopenfileslimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: 1048576

proc-psig=*perprocesspendingsignalslimit*

This option sets the pending signal number limit in the node which can enter the queue, and for the executed user ID which can call the process for each process of all the processes constitutes the job.
The format of the value of *perprocesspendingsignalslimit* can be referred to "LIMIT VALUE".

Minimum value: 0. Maximum value: unlimited (No limitation)

proc-filesz=*perprocesspermfilelimit*

This option sets the maximum file size limit for each process of all the processes constitutes the job.
The format of the value of *perprocesspermfilelimit* can be referred to "LIMIT VALUE".

Minimum value: 2048 Byte. Maximum value: unlimited (No limitation)

proc-stack=*perprocessstacklimit*

This option sets the maximum stack segmentation size limit for each process of all the processes constitutes the job.
The format of the value of *perprocessstacklimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

proc-vmem=*perprocessvirtualmemorylimit*

This option sets the maximum virtual memory size limit for each process of all the processes constitutes the job.
The format of the value of *perprocessvirtualmemorylimit* can be referred to "LIMIT VALUE".

Minimum value: 0 Byte. Maximum value: unlimited (No limitation)

vnode-core=*value*

This option sets the number of CPU cores per virtual node. It is valid only when the vnode option is specified. Please refer to the optional vnode explanation.

vnode-mem=*value*

This option sets the amount of memory per virtual node. It is valid only when the vnode option is specified. Please refer to the optional vnode explanation.

core-mem=*value*

This option sets the amount of memory per CPU core. It is valid only when the vnode option is specified. Please refer to the optional vnode explanation.

custom-resource-name[=*value*]

This option specifies the number or type of a custom resource to allocate to the job.

- When specifying the number of a custom resources

For the custom resources associated with nodes, specifies the required amount per required node or per required virtual node. Total required amount is (# of custom resource) x (# of nodes or virtual nodes). For the custom resources not associated with nodes, specifies the total required amount per the job.

If *value* is not specified, a default value will be used. The default value is defined by an administrator with using the job ACL function.

If the default value does not defined in job ACL function, the lower bound value will be used.

- When specifying the type of a custom resource

Specifies the type of custom resource requested by the job.

If *value* is not specified, a default value will be used. The default value is defined by an administrator with using the job ACL function.

If the default value does not defined in job ACL function, the job fails to submit.

[LIMIT VALUE]

Job manager supports the job resource limit function.

Option ({-L|--rsc-list} "*resource= value*":resource can be any string or --llo "*param= value*") sets the resource limit value. There are 3 types of the resource limit value format: limit about time, limit about size and limit about number. If the limit value has no limitation, then set "unlimited".

The followings are the description about 3 types of limit.

[Limit about time]

There are two types of the time format in CPU time limit, elapse time limit.

- [[*HOURL*]:]*MINUTE*[:]*SECOND*

Note

There cannot be space before or after the ":".

Example of setting limit values:

1234:58:21 (1234 hours 58 minutes 21 seconds)

12345 (12345 seconds)

59:01 (59 minutes 1 second)

If "unlimited" has not been specified, the maximum value can be 2147483647 seconds.

- [[*HOURLh*]*MINUTEm*].*SECONDSs*

Note

There cannot be space before or after the "h", "m", "s".

"h", "m", "s" also can be capital letters("H", "M", "S").

Example of setting limit values:

1234h58M (1234 hours 58 minutes)

12345s (12345 seconds)

59H01s (59 hours 1 second)

If "unlimited" has not been specified, the maximum value can be 2147483647 seconds.

[Limit about size]

The format is displayed as follows:

- *Integer*[*unit*]

Note

There cannot be space between value and *unit*.

The units can be set as following strings:

For memory: binary prefix

Unit	Description
Ki	kibibyte (2 ¹⁰)

Unit	Description
Mi	mebibyte (2 ²⁰)
Gi	gibibyte (2 ³⁰)
Ti	tebibyte (2 ⁴⁰)
Pi	pebibyte (2 ⁵⁰)

If "unlimited" has not been specified, the maximum value can be 2147483647 mebibytes. If unit has not been described, the unit will be Mi (mebibyte).

For disk: SI prefix

Unit	Description
K	kilobyte (10 ³)
M	megabyte (10 ⁶)
G	gigabyte (10 ⁹)
T	terabyte (10 ¹²)
P	petabyte (10 ¹⁵)

If "unlimited" has not been specified, the maximum value can be 2147483647 megabytes. If unit has not been described, the unit will be Mi (megabyte).

[Limit about number]

When the set number is integer, the unit is needless.

If "unlimited" has not been specified, the maximum value can be 2147483647.

-P "*parameter=policy*"

The node allocation method (node selection policy) is specified.

This option is effective only for virtual node allocated jobs on PRIMERGY servers.

The following are specified for *parameter*.

vn-policy=*policy*

The arrangement policy of a virtual node is specified for *policy*.

For policy, specify any of the following arrangement policies of virtual nodes.

Arrangement policy of a virtual node	Description
abs-pack	Absolutely PACK (All the virtual nodes are arranged in one physical node.)
pack	PACK (The virtual node is arranged in physical nodes of a little number as much as possible.)
abs-unpack	Absolutely UNPACK (One virtual node is arranged in one physical node.)
unpack	UNPACK (As much as possible one virtual node is arranged in one physical node.)

With Absolutely UNPACK, you can also define a specification where *N* (number) virtual nodes is always placed on different physical nodes.

In this case, specify the -P "vn-policy=abs-unpack=*N*" option (the number *num* of virtual nodes specified by -L vnode=*num* must be a multiple of *N*).

With UNPACK, you can also define a specification where *N* (number) virtual nodes is placed on different physical nodes if at all possible.

In this case, specify the -P "vn-policy=unpack=*N*" option (the number *num* of virtual nodes specified by -L vnode=*num* must be a multiple of *N*).

exec-policy=*policy*

The execution mode policy is specified for *policy*.

For policy, specify any of the following execution mode policies.

Execution mode policy	Description
simplex	SIMPLEX (A physical node is monopolized by one job.)
share	SHARE (A physical node is shared with other "SHARE" jobs.)

-m *mailoption*[,...]

This option informs the information about job status and so on by mail.

When job begins, ends or restarts, the information about changing the job status should be informed by mail. And the statistical information can also be informed by mail. Description of *mailoption* is displayed as follows. When the option is specified more than 1 time, a comma (",") is used to delimit. The statistical information can be specified with either "s" or "S".

REMARKS

Bulk jobs and step jobs have information corresponding to the job ID and information corresponding to the sub job ID. The former in particular is called summary information of the bulk or step job.

- Normal Job

Item	Description
b (begin)	Beginning of job execution
e (end)	End of job execution
r (restart)	Restart of job
s (info)	Inform statistical information (without node information).
S (info node)	Inform statistical information (with node information).

- Step Job

If "b" or "r" is set, mail is notified in each sub job.

If the rest is set, mail is notified in both each sub job and each job.

Item	Description
b (begin)	Beginning of sub job execution
e (end)	End of sub job execution
r (restart)	Restart of sub job
s (info)	Inform statistical information and summary information (without node information).
S (info node)	Inform statistical information and summary information (with node information).

- Bulk Job

In the following, mail is notified in not each sub job but each job.

Item	Description
b (begin)	Beginning of first sub job execution
e (end)	End of last sub job execution
r (restart)	No mail information
s (info)	Inform summary information (without node information).
S (info node)	Inform summary information (without node information).

If this option has not been specified, the information is not informed. But when the job ends abnormally, the information must be informed even the option is set.

--mail-list *mailaddress*[,...]

This option sets the sending address of mail.

Mail address is set as *mailaddress*. When the option is specified more than 1 time, a comma (",") is used to delimit. The size of the

specified string is 255 characters at most. The format of *mailaddress* is *login-name* (not including @ character) or *login-name@machine*. Both are right.

--mpi "*parameter*[,...]"

This option specifies the parameters of an MPI job.

The following values are set for *parameter*.

The specified parameter is different in FX servers and PRIMERGY servers.

When the option specifies multiple values, they will be delimited with a comma (",").

[FX server]

shape={X|Xx Y|Xx YxZ}

This option specifies the shape of process generated when program starts. One-dimensional, two-dimensional, and three-dimensional shapes can be specified. Set the same number of dimensions as that specified with the node of the -L option. If the number of dimensions is not set, the same value as that specified with the node will be used.

proc=*procnum*

This option specifies the number of maximum processes generated when program starts. When specification is omitted, the product of the value specified with *shape* (number of nodes obtained from shape) is used. When a value that is larger than the upper bounds of number of processes generable in one node (= number of CPU cores in one node) times the product of the value of *shape* is specified, the acceptance of the job is refused.

max-proc-per-node=*mppnnum*

This option specifies the number of maximum processes per node. When the specification is omitted, the number of maximum processes per node is the quotient of the number specified with "proc" option (*procnum*) divided by the number of nodes specified with "shape" option (if not specified with shape option, the number of nodes specified with -L "node" option of pjsub command is adopted). When the specified value *mppnnum* exceeds the number of CPU cores per node, or when the quotient of the number specified with "proc" option (*procnum*) divided by the number of nodes exceeds *mppnnum*, the acceptance of the job is refused.

rank-map-bynode[=*rankmap*]

rank-map-bychip[:*rankmap*]

rank-map-hostfile=*filename*

This option specifies the allocation rule of the rank of the generated process. The rank-map-bynode and rank-map-bychip cannot be specified at the same time. If neither is specified, the pjsub command specifies rank-map-bychip.

1. *rank-map-bynode*[=*rankmap*]

XY|YX|XYZ|XZY|YXZ|YZX|ZXY|ZYX is specified for *rankmap*. The *rankmap* specifies the same one as the number of dimension specified with shape.

The *rankmap* cannot be specified for one dimension.

This command will create one process for a node, move to the next node, then automatically assign the rank using round-robin processing. If the job shape is two-dimensional, rankmap should be specified by combining characters X and Y. And in case of three-dimensional, rankmap should be specified by combining characters X, Y and Z. The order of the assigning of the rank moves to the axis of the following character when the coordinate origin is assumed to be rank 0, the rank is arranged axially of the first character, and it reaches the upper bound.

When the specification of the *rankmap* is omitted, rank number set as follows.

- one-dimensional: The rank number is the same as the one-dimensional coordinates.
- two-dimensional: The rank number is the same as [XY].
- three-dimensional: The rank number is the same as [XYZ].

2. *rank-map-bychip*[:*rankmap*]

XY|YX|XYZ|XZY|YXZ|YZX|ZXY|ZYX is specified for *rankmap*. Refer to "1.rank-map-bynode" for *rankmap*. The *rankmap* cannot be specified for one dimension.

First of all, the number of processes allocated to a node, where the number of processes is [proc divided by shape] (rounding up to the nearest whole number), and move to the next node, then assign the rank.

3. *rank-map-hostfile*=*filename*

The rank of a created process is allocated according to *filename*. The *filename* can be set execution nodes with one-dimensional, two-dimensional, or three-dimensional coordinates. Please specify the same number of dimension as the number

of dimension specified with shape.

The description form of the *filename* file is shown as follows:

- one-dimensional shapes, specified (*x*).
- two-dimensional shapes, specified (*x,y*).
- three-dimensional shapes, specified (*x,y,z*).

The *x*, *y*, *z* in the parenthesis "()" shows numerical value.

A blank line in the *filename* file and the line where the first letter begins in "#" are ignored.

If the description format is not according to the format above, an error occurs.

assign-online-node

This option guarantees there are no breakdown nodes in the resource assigned nodes.

[PRIMERGY server]

proc=*procnum*

This option specifies the number of maximum processes generated when program starts.

When this option is not specified, the number of virtual nodes for virtual node allocated job, or the number of nodes for node allocated job is used.

max-proc-per-node=*mppnnum*

This option specifies the number of maximum processes per node. When this option is not specified, the number of maximum processes per node is the quotient of the number specified with "proc" option (*procnum*) divided by the number of nodes specified with "shape" option (if not specified with shape option, the number of nodes specified with -L "node" option of pjsub command is adopted). When the quotient of the number specified with "proc" option (*procnum*) divided by the number of nodes exceeds *mppnnum*, the acceptance of the job is refused.

When vnode option is specified, the acceptance of the job is refused.

rank-map-bynode

rank-map-bychip=*n*

This option is effective only for virtual node allocated jobs. This option specifies the allocation rule of the rank of the generated process.

The rank-map-bynode and rank-map-bychip=*n* cannot be specified at the same time.

When neither is specified, virtual nodeid is set in order of the virtual node arrangement for virtual node allocated job. And, the same value as the rank number is set as for virtual nodeid.

1. rank-map-bynode

Virtual nodeid (rank number) of the allocated virtual node is set by one node by UNPACK (round robin) form.

2. rank-map-bychip=*n*

Virtual node ID (rank number) of the allocated virtual node is set to one node by "*n*" piece.

It is necessary to allocate the number of virtual nodes of multiples of the specified value in each node. Specify the rank-map-bychip=*n* parameter together with unpack=*m* or abs-unpack=*m*, where *m* must be a multiple of *n*.

When the description form is not described as it is the above-mentioned, pjsub returns the error.

--mswk

This option executes master-worker job.

The master-worker job is continue a computing task by creating a mechanism to re-execute a worker process on a different node when a compute node is down or the worker process has ended abnormally. The parallel execution environment of the job operation software determines the node that will generate a worker process in a master-worker job.

Alternatively, the user specifies the node when generating the process. Therefore, specification of the --mpi rank-map-hostfile option of the pjsub command has no meaning. This option is ignored even if specified.

The option --mswk can be specified only the torus mode.

When the option --bulk, --step or --interact has also been specified, the error occurs (incorrect specified option combination).

{-N|--name} *name*

This option sets the job name.

Job name is set in 63 bytes at most.

When this option has not been specified, the job name will be set as the script file name specified by the command line. If the script file

name has not been specified, the job name will be set as "STDIN"

The first character of the job name must be single-byte alphabet, and *name* can not be the string including "/".

--net-route *value*

This option specifies whether to change the communication path when a Tofu interconnect link goes down.

value is set as following.

Value	Description
dynamic	Changes the communication path when a Tofu interconnect link goes down, and job execution continues.
static	Does not change the communication path when a Tofu interconnect link goes down. The job ends abnormally.

When this option is not specified, the default value that the job ACL defines is adopted.

For a compute node job not on an FX server or a virtual node allocated job, the --net-route specification is ignored, and the communication path does not change when a Tofu interconnect link goes down.

{-o|--out} *pathname*

This option set the path *pathname* where the command outputs the standard output. *pathname* can use the metacharacters as follows:

Metacharacter	Description
%j	Expand the job ID
%J	Expand the sub job ID
%b	Expand the bulk number
%s	Expand the step number
%n	Expand the job name

Note

When some sub jobs have specified the standard output files with the same name, the output content cannot be guaranteed, because there is no file lock among sub jobs.

Do not specify the file name among sub jobs repeatedly.

-p *priority*

This option specifies the priority of submitted jobs.

priority is set as an integer from 0 to 255, 255 is the highest, and 0 is the lowest priority. According to the specified priority, the job priority of user can be requested.

--restart

Automatically re-executes the job if it ended abnormally.

This option cannot be used with option --norestart together.

When this option is specified, if the job is interrupted due to job execution environment stops or error occurs, the job is automatically re-executed.

If neither the option --restart nor --norestart is specified, operations in this case may change with the system settings.

--norestart

Does not automatically re-executes the job, even if it ended abnormally.

This option cannot be used with option --restart together.

When this option is specified, if the job is interrupted due to job execution environment stops or error occurs, the job is not automatically re-executed.

If neither the option --restart nor --norestart is specified, operations in this case may change with the system settings.

--step --sparam "*parameter*[,...]"

This option executes step job.

Step job is a job gathering multiple jobs (sub jobs) together and submit as one job. And there are execution sequence and dependence relationship among these jobs.

If you want to submit a step job, you can specify more than one script at the same time with a single command execution. Please specify multiple scripts by a comma-separated list, or space-separated list when specifying more than one script. If you specify more than one script file, please note the following.

- If an error was detected, it becomes the following operation.
 - If there is a sub job to be a syntax error, all of the sub jobs will not be submitted.
 - In case of other errors (errors in the limit check of the job ACL, etc.), the submitted processing of the sub jobs is suspended at the time of the detected abnormality. The sub jobs later that detect the abnormality will not be submitted.
- If you want to specify a job name for each sub job by the -N|--name option, please specify in each script file.
- If you specify a job name by the argument jnam of --sparam option, it will be submitted as you have set the job name specified by the argument jnam of --sparam option to -N|--name option.
- If you specify a jid in the -z option, job ID's of only the number of sub jobs are output to the standard output.

The parameter of the step job is set by --sparam option.

When --sparam option has not been specified, the parameters are set as the default values.

When --sparam option has been specified, *parameter* is set as follows. When the option specifies multiple values, they will be delimited with a comma (",").

The following displays the setting value of *parameter*.

jid=jobid

If you specify the job ID of the existing step job, you specify the argument jid. This option is used if you want to submit the sub jobs of the second and subsequent step job specify the job ID that has been granted in the first sub job on. If the job ID has not been set, the job will be specified as a sub job of a new step job. When the argument jid and arguments jnam has been specified at the same time, the error occurs (incorrect specified option combination).

jnam=jobname

If you specify the job name (*) of the existing step job, you specify the argument jnam. If you want to submit a step job that each sub jobs of the job name are the same name, you can use this option. If the job name has not been set, the job will be specified as a sub job of a new step job. When the argument jid and arguments jnam has been specified at the same time, the error occurs (incorrect specified option combination).

If you specify a job name of the step job that does not exist, it will be submitted as the sub jobs of the new step job.

*:The job name refers to the job name that specified when you submit first sub job of the step job.

sn=stepno

This parameter sets the step number.

sd=form[:deletetype[:stepno[:...]]]

This parameter sets the dependence relationship. The dependence relationship formula can be set multiple times.

form is set as following formula.

Form	Description
NONE	No dependence relationship If NONE is set, <i>deletetype</i> and <i>stepno</i> are ignored. (The same as the situation where parameter <i>sd</i> has been omitted) none, None can be also accepted as NONE(The other forms are all errors.)
ec == <i>value</i> [,...] pc == <i>value</i> [,...] ec != <i>value</i> [,...] pc != <i>value</i> [,...] ec > <i>value</i> pc > <i>value</i> ec < <i>value</i> pc < <i>value</i> ec >= <i>value</i> pc >= <i>value</i> ec <= <i>value</i> pc <= <i>value</i>	Set the ending condition of the sub job. ec: exit code, the exit status of the sub job shell set by <i>stepno</i> . pc: PJM code, the job exit code of the sub job set by <i>stepno</i> . <i>value</i> can be set as any numeric value. For ==, !=, multiple values can be set using ", ". Example: ec == 1,3,5 -> If ec is one of 1,3,5, the expression is true. ec != 1,3,5 -> If ec is none of 1,3,5, the expression is true. When the execution result of sub job specified by <i>stepno</i> meet the <i>form</i> condition, the sub job will be ended up according to the <i>deletetype</i> .

stepno sets the sub job number of depended sub job.
stepno can set multiple numbers delimiting with colon (":").
 If *stepno* has been omitted, the sub job depends on the result of last executed sub job.
deletetype is set as following.

Deletetype	Description
one	End up only this sub job. Don't end up the sub job which depended on this sub job.
after	End up both this sub job and the sub job which depended on it.
all	End up this sub job and all the sub job afterwards, and cancel the execution of the step job.

When the option --bulk, --mswk or --interact has also been specified, the error occurs (Incorrect specified option combination).

{-s|--stats}

This option outputs the statistical information on submitted jobs to a file.
 This option cannot be used with the -S or --STATS option.
 To output this information to a specified file, use the --spath option.
 The output statistical information is different in FX server and PRIMERGY server.
 For the statistical information that is output, see pjstatinfo(7).

{-S|--STATS}

This option outputs statistical information containing the node information of submitted jobs to a file.
 This option cannot be used with the -s or --stats option.
 To output this information to a specified file, use the --spath option.
 The output statistical information is different in FX server and PRIMERGY server.
 For the statistical information that is output, see pjstatinfo(7).

--spath *pathname*

This option outputs the statistical information of jobs to a file that is specified with *pathname*.
 This option is only effective when used with the -s, --stats, -S, or -STATS option.
pathname can be used metacharacters as follows.

Metacharacter	Description
%j	Expand the job ID
%J	Expand the sub job ID
%b	Expand the bulk number
%s	Expand the step number
%n	Expand the job name

If the --spath option is omitted, the statistical information of jobs is output to the "*keyword*+*jobID*+ .stats" file in the current directory as of the time of job submission.

Keyword is specified by the option -N|--name. If these options have not been specified, the keyword will be set as the script file name.

If the job is submitted from the standard input, the keyword will be set as "STDIN".

The script file name is used as the "keyword". However when the first character of the script file is an numeric character, the "keyword" is created by the script file (J + *script_file_name*).

The number of characters of "keyword" is assumed to be 63 characters in maximums.

When the number of characters of script file name or "J + *script_file_name*" exceeds 63 characters, only first 63 characters are set to the "keyword".

Note

When some sub jobs have specified the Statistical information files with the same name, the output content cannot be guaranteed, because there is no file lock among sub jobs.

Do not specify the file name among sub jobs repeatedly.

--verbose

This option produces verbose output.

This option cannot be used with the -z option.

This option cannot be written in the script file. When the option has been written in the script file, it will be invalid.

--vset "*variable=value*[,...]"

This option defines variables (local variables) which are only valid in specified line of pjsub command option.

In order to refer to the defined variables, add "\$" in front of the variable name, and the variable name is enclosed with parentheses "{}".

This variable becomes valid after the line which this option has been defined.

The maximum size of the variable is 4095 characters including the variable name and defined content, distinguishing the capital letters and lower case letters.

The following displays the description examples in job script.

- Description examples in job scripts.

```
#PJM -L "node=100, elapse=86400"
#PJM --vset "SHOME=/SOURCE/home000/z8000/"
#PJM -I "rank=0-2 ${SHOME}/a.dat %r:./"
```

- Expand the examples above.

```
#PJM -L "node=100, elapse=86400"
#PJM -I "/SOURCE/home000/z8000/a.dat 0:./"
#PJM -I "/SOURCE/home000/z8000/a.dat 1:./"
#PJM -I "/SOURCE/home000/z8000/a.dat 2:./"
```

-w *mode*

This option specifies the location where the response will return to the pjsub command after the job is submitted.

When this option has not been specified, the content specified by pjsub command option will be checked according to the limit values defined by the job ACL, and return (whether success or failure).

The following displays the setting value of *mode*.

nowait

The response returns promptly when the job has been turned to ACCEPT status (temporarily accepted) by job manager.

pjstat command can be used to check whether the job has been submitted successfully.

jobchk

Return when the job has been turned to ACCEPT status (temporarily accepted) and has been checked by job manager. pjstat command can be used to check whether the job has been submitted successfully by processing afterwards.

-x "*variable=value*[,...]"

This option adds the environment variable *variable* set when the job is executing.

variable specifies the environment variable name, and *value* sets the value of the variable.

-X

This option transmits all the environment variables to the compute node.

When the job has been submitted, the automatically transformed environment variables of the node where job was submitted will be transmitted to the compute node with the job itself.

The following displays the transform results of the environment variables of the node where job was submitted.

Before transformation	After transformation
HOME	PJM_O_HOME
LANG	PJM_O_LANG
LOGNAME	PJM_O_LOGNAME
MAIL	PJM_O_MAIL
PATH	PJM_O_PATH
SHELL	PJM_O_SHELL

Before transformation	After transformation
TZ	PJM_O_TZ

When this option has been specified, the other environment variables except the automatically transmitted ones above are also transmitted to the compute node with the job itself. When the job has been submitted, these environment variables are memorized as the environment variables in the same names.

There is an environment variable of OS regulations not succeeded in the job even if optional -X of the pjsub command is specified (LD_LIBRARY_PATH etc.).

Therefore, it is necessary to set the environment variable of OS regulations in the job script.

-z [jid]

This option sets the display content to the standard output when pjsub command is executing.

When the parameter has not been specified, the command will display nothing.

If the parameter is set as *jid*, the command will only display job ID.

When this option has been specified, the error message will be still displayed.

The following displays the difference about the output between the situation where this option has been specified or not.

- -z option is not specified.
Submission successful: "Job *JobID* submitted."
Other situations: Display error message (main reason why the submission failed)
- -z jid is specified.
Submission successful: *JobID*
Other situations: Display error message (main reason why the submission failed)

--script-delimiter *delimiter*

Specify a delimiter (one character) for submit multiple job scripts for step job. Job scripts for step job should be specified one by one as arguments to the pjsub command or separated by delimiter *delimiter*. Mixing of both specification methods is possible.

Delimiter is the one character. And the character can only be single-byte alphabet, numeric character or other symbols. But, hyphen ("-") can not be specified.

If this option is not specified, the comma (",") becomes a delimiter.

To specify a job script containing a comma (","), specify "none" for *delimiter*. This will cause the comma to have no meaning for the delimiter.

This option cannot be written in the script file. When the option has been written, it will be invalid.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following is an execution example of the pjsub command:

[NOTE]

The administrator can change the job statistical information items that are output and the output order.

Therefore, the output job statistical information may differ from the following example.

For a description of each output item, see pjstatsinfo(7).

1. The specification of the -s option of the pjsub command outputs the statistical information on submitted jobs to a file.

```
$ pjsub -L node=2x3x2,elapse=500 -s test.sh
[INFO] PJM 0000 pjsub Job 1006903 submitted.
...
$ cat test.sh.1006903.stats
Job Statistical Information
JOB ID                : 1006903
SUB JOB NUM           : -
START BULKNO          : -
END BULKNO            : -
HOST NAME             : login-node
JOB NAME              : test.sh
```

```

JOB TYPE           : BATCH
JOB MODEL          : NM
USER               : testuser
GROUP              : testgroup
RESOURCE UNIT      : rscunit000
RESOURCE GROUP     : rscgroup000
APRIORITY          : 127
PRIORITY           : 127
SHELL              : /bin/sh
...

```

2. The specification of the -S option of the pjsub command outputs the statistical information including node information on submitted jobs to a file.

In addition to the information output when the -s option is specified, information about each node used by the job ("Node Statistical Information" part) is displayed.

For a description of each output item, see pjstatsinfo(7).

```

$ pjsub -L node=2x3x2,elapsed=500 -S test.sh
[INFO] PJM 0000 pjsub Job 1006904 submitted.
...
$ cat test.sh.1006904.stats
Job Statistical Information
JOB ID           : 1006904
SUB JOB NUM      : -
START BULKNO     : -
END BULKNO       : -
HOST NAME        : login-node
JOB NAME         : test.sh
...
-----
Node Statistical Information
VNODE ID         : -
NODE ID          : 0x01010010
TOFU COORDINATE  : (1,1,1)
NODE COORDINATE  : (1,1,1)
RANK NO          : 0
NODE AVAILABLE   : available
ARCHI INFORMATION : FT
CPU BITMAP (ALLOC) : 0x1
CPU BITMAP (USE)  : 0x1
...
-----
Node Statistical Information
VNODE ID         : -
NODE ID          : 0x01010011
TOFU COORDINATE  : (1,1,2)
NODE COORDINATE  : (1,1,2)
RANK NO          : 1
NODE AVAILABLE   : available
ARCHI INFORMATION : FT
CPU BITMAP (ALLOC) : 0x1
CPU BITMAP (USE)  : 0x1
...

```

[END STATUS]

0: Normal end

1: Abnormal end

[SEE ALSO]

pjdel(1), pjstat(1), pjalter(1), pjstatsinfo(7)

[MESSAGES]

If an error occurs during the execution of the pjsub command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 0001 pjsub Unknown option or no option argument: *opt*.
Try `pjsub --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0002 pjsub Invalid combination: *detail*
Try `pjsub --help` for more information.**

Meaning

Invalid combination of options or arguments for the command displayed in *detail*.

Action

Specify the correct combination of options or arguments.

**[ERR.] PJM 0003 pjsub Unknown option argument: *opt arg*.
Try `pjsub --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0004 pjsub Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0008 pjsub Option arguments must be enclosed in quotation mark.
Script file *filename* line *lineno*.**

Meaning

The double quote or single quote is not a pair at the description of the argument of the pjsub command in the job script. The location of the detected problem is line *lineno* of the job script *filename*.

Action

Describe the double quotes or the single quotes by the pair.

[ERR.] PJM 0009 pjsub Invalid argument: *arg*.

Meaning

The specified argument *arg* is incorrect.

Action

Specify the correct argument.

[ERR.] PJM 0012 pjsub Job *jobid* does not exist.

Meaning

The specified step job (job ID: *jobid*) does not exist. Alternatively, the user is not the owner of the specified step job.

Action

Confirm that the step job exists. Alternatively, check whether the user is the owner of the job.

[ERR.] PJM 0018 pjsub Job name is mismatch: *filename*.

Meaning

The job name specified with the --sparam "jnam=" option for a step job in the job script (file name: *filename*) is different from the job name specified in other job scripts.

Action

Specify the same job name in all job scripts when you specify the job name with the --sparam "jnam=" option for the step job in the job scripts.

[ERR.] PJM 0020 pjsub File open failed: *path (code)*.

Meaning

The specified file path cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and path are correct.

[ERR.] PJM 0021 pjsub Script file name too long.

Meaning

The job script file name is too long.

Action

Specify a job script file name within 4096 characters.

**[ERR.] PJM 0022 pjsub File format error: *filename*.
File line *lineno*.**

Meaning

Line *lineno* of the file *filename* contains a format error.

Action

Write the file in the correct format.

**[ERR.] PJM 0023 pjsub Multiple script files specified.
Try `pjsub --help` for more information.**

Meaning

Multiple job script files were specified.

Action

Do not specify more than one job script file.

[ERR.] PJM 0025 pjsub Line length too long.
Script file *filename* line *lineno*.

Meaning

The maximum number of characters allowed per line in a job script file has been exceeded. The location of the detected problem is line *lineno* of the job script *filename*.

Action

Use up to 4096 characters, including the newline character, per line.

[ERR.] PJM 0026 pjsub Current working directory access failed.

Meaning

Information on the current directory could not be retrieved.

Action

Confirm that the current directory exists on the login node, and execute the command again.

Confirm that the current directory exists on the login node and that you have write permissions, and execute the command again.

[ERR.] PJM 0027 pjsub Current working directory path contains a newline character.

Meaning

The current directory name is incorrect. The directory name contains a linefeed code.

Action

Review the current directory name, and execute the command again. Do not use the linefeed code, backslash +'n', in a directory name.

[ERR.] PJM 0028 pjsub File creation failed: *path*.

Meaning

The specified file *path* cannot be created.

Action

Confirm that the specified file *path* can be created.

[ERR.] PJM 0041 pjsub Getting group information from *gid* failed.

Meaning

The group name could not be obtained based on the group ID *gid*.

Action

Confirm that the group corresponding to the ID is registered in the system.

[ERR.] PJM 0045 pjsub *path* isn't on the global file system.

Meaning

path does not exist on the global file system.

Action

If *path* is "Home Directory", Change the home directory to the global file system.

If *path* is "PJM_LLIO_GFSCACHE=*pathname*" or "PJM_LLIO_SHAREDTMP=*pathname*", specify the global file system in the environment variables PJM_LLIO_GFSCACHE and PJM_LLIO_SHAREDTMP.

[ERR.] PJM 0046 pjsub Option argument is out of range (*low-high*): *option*.

Meaning

A value outside the specifiable range is specified in *option*.

Action

Specify a value in the range from *low* to *high*.

[ERR.] PJM 0047 pjsub Option argument must be multiple of *N*: *option*.

Meaning

The value specified in *option* must be a multiple of *N*.

Action

Specify a multiple of *N*.

[ERR.] PJM 0048 pjsub Option argument must be less than or equal to *N*: *option*.

Meaning

The value specified in *option* must be equal to or less than *N*.

Action

Specify a value equal to or less than *N*.

[ERR.] PJM 0049 pjsub Option argument is invalid. sharedtmp-size + localtmp-size must be less than or equal to *size*.

Meaning

The total of the specified sharedtmp-size and localtmp-size values must be less than *size*.

size is the SSD capacity per node minus the minimum capacity of the second-layer storage cache.

Action

Specify sharedtmp-size and localtmp-size values so that their total is equal to or less than *size*.

[ERR.] PJM 0050 pjsub The option --llo cannot used in this resource unit.

Meaning

The --llo option cannot be specified for use without the layered storage system environment.

Action

Do not specify the --llo option for use without the layered storage system environment.

[ERR.] PJM 0054 pjsub *rscname=value* is less than the lower limit (*limit-value*).

Meaning

The value *value* specified for the resource name *rscname* specified in the -L|--rsc-list option is less than the lower limit value defined by the job ACL function.

The lower limit value for the amounts that can be specified for resources is output to *limit-value*.

Action

Specify a resource limit value that is not below the range defined by the job ACL function.

[ERR.] PJM 0056 pjsub The specified execution time is earlier than the current time: *time*.

Meaning

The time *time* specified for the --at option is earlier than the current time.

Action

Specify a future time for the --at option.

[ERR.] PJM 0057 pjsub *rscname=value* is greater than the upper limit (*limit-value*).

Meaning

The value *value* specified for the resource name *rscname* specified in the -L|--rsc-list option is greater than the upper limit value defined by the job ACL function. Alternatively, it is greater than the amount of resources that can be allocated to a job.

The upper limit value for the amounts that can be specified for resources is output to *limit-value*.

Action

Specify a resource amount that does not exceed the range defined in the job ACL as the limit value for the resource.

Specify the amount of resources (Number of nodes, node shape, etc.) that can be assigned to the job.

For details on the amount of resources that can be allocated to jobs on the FX server, see the table "Size of nodes that can be allocated [FX]" in "Checking resource units and resource groups" in "Chapter 2 Job Operation Procedures" of the manual "Job Operation Software End-User's Guide."

[ERR.] PJM 0058 pjsub *rscname=value* does not exist.

Meaning

- If the resource name *rscname* is *rscunit*
The specified resource unit value does not exist.
- If the resource name *rscname* is *rscgrp*
The specified resource group value does not exist.

Action

Specify the resource unit name or resource group name correctly.

[ERR.] PJM 0059 pjsub *rscname=value* is disabled.

Meaning

- If the resource name *rscname* is *rscunit*
A job currently cannot be submitted to the specified resource unit value.
- If the resource name *rscname* is *rscgrp*
A job currently cannot be submitted to the specified resource group value.

Action

Check the state of the resource unit or resource group with the --rsc option of the pjstat command. Then, specify a resource unit or resource group to which the job can be submitted. If necessary, ask the administrator to change the state of the resource unit or resource group.

[ERR.] PJM 0060 pjsub Step number overflowed.

Meaning

The step number of the step job is outside the prescribed range.

Action

Specify a step number from 0 to 65534 for the step job.

[ERR.] PJM 0063 pjsub The value "m" must be a multiple of the value "n": "*opt=m*", "*opt=n*".

Meaning

The value m, which is specified in the option *opt=m*, must be a multiple of the value n, which is specified in the option *opt=n*.

Action

The value *m* specified in the option *opt=m* must be a multiple of the value *n* specified in the option *opt=n*.

[ERR.] PJM 0068 pjsub System configuration has been changed.

Meaning

The system configuration was changed while a job was being accepted.

Action

Check the resource unit configuration on the system, and then submit the job again.

[ERR.] PJM 0069 pjsub *rscname* in -L option is not specified.

Meaning

- If the resource name *rscname* is *rscunit*
No resource unit name is specified in the -L option.
- If the resource name *rscname* is *rscgrp*
No resource group name is specified in the -L option.

Action

Specify a resource unit name or resource group name in the -L option.

[ERR.] PJM 0070 pjsub No execute permission.

[ERR.] PJM 0070 pjsub No execute permission: *detail*

Meaning

There are two types of messages.

If the message does not display *detail*, pjsub command execution is not permitted. If it displays *detail*, the user does not have the privileges to specify the option or argument.

Action

Display the job ACL function settings with the *pjacl* command. Check whether pjsub command execution is permitted (execute pjsub) for the target resource unit and resource group, and check the privileges to specify the option or argument (execute pjsub(--*xxxx*), where *xxxx* is the option or argument name). If necessary, ask the administrator to add permission.

[ERR.] PJM 0071 pjsub Group not authorized to submit a job: *groupname*.

Meaning

Command execution is not permitted for the permission '*groupname*.'
groupname is displayed in following form.

group (group) : The group name on the operating system.

Action

Display the job ACL settings with the *pjacl* command. Then, confirm the permission for the permission of the pjsub command (permit pjsub) for the target resource unit and resource group to which the job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0072 pjsub Job exceeded the accept limit: *target*.

Meaning

The upper limit on the number of jobs accepted simultaneously has been reached.

At *target*, "ru-accept," "ru-accept-allsubjob," "ru-accept-bulksubjob," and "ru-accept-stepsubjob" indicate batch jobs, and "ru-interact-accept" indicates interactive jobs.

Action

With the `--rscunit`, `--ru`, `--rscgrp`, or `--rg` option of the `pjacl` command, check the number of jobs concurrently accepted for the resource unit to be changed. The number of concurrently accepted jobs is the value of "acceptable job," "acceptable all-subjob," "acceptable bulk-subjob," "acceptable step-subjob," or "acceptable job(interact)" of "limit in rscunit (each users)," "limit in rscgroup (each users)," "limit in rscunit (total users in same group)," "limit in rscgroup (total users in same group)," "limit in rscunit (total all users)" or "limit in rscgroup (total all users)." If necessary, ask the administrator to increase the number of jobs accepted simultaneously.

[ERR.] PJM 0073 pjsub Option value is out of range: *opt*.

Meaning

The value specified by the option *opt* exceeds the range defined by the job ACL function.

Action

With the `pjacl` command, check the upper limit on the limit values that can be specified for the option *opt*. If necessary, ask the administrator to increase the number of accepted commands.

[ERR.] PJM 0074 pjsub Default *rscname* value is illegal: *message*.

Meaning

The combination of default values defined in the job ACL function is invalid for the resource name *rscname*, which can be specified in the `-L|--rsc-list` option.

Action

Contact the administrator.

[ERR.] PJM 0076 pjsub *rscname=value* cannot be specified. The value must be specified as follows: *select-strings*,...

Meaning

The value *value* for the resource name *rscname* specified in the `-L|--rsc-list` option is not among the types that can be specified for the custom resources defined by the job ACL function. *select-strings*,... is a list of types of specifiable custom resources.

Action

Specify a type that is defined by the job ACL function.

[ERR.] PJM 0077 pjsub Too many custom resources specified.

Meaning

You cannot specify 65 or more custom resources in an argument.

Action

Reduce the specified number of custom resources to 64 or less.

[ERR.] PJM 0079 pjsub Arbitrary character string.

Meaning

An error occurred at the job registration exit that was set by the administrator. The contents of the message were set by the administrator.

Action

Referring to the message, find a solution to why acceptance of the job was refused. If you cannot find out the cause from the message, contact the administrator. The administrator is requested to check why acceptance of the job was refused at the job registration exit.

[ERR.] PJM 0084 pjsub Interactive job cannot be submitted except login node.

Meaning

An interactive job is executed on a node other than login node.

Action

Execute interactive jobs on a login node.

[ERR.] PJM 0085 pjsub Interactive job *jobid* was not able to start.

Meaning

The interactive job (job id: *jobid*) could not start.

Action

Execute the command again.

[ERR.] PJM 0089 pjsub Not supported: *message*.

Meaning

The specified function or combination of functions is not supported by the current version of Technical Computing Suite.

Action

By referring to the information displayed in *message*, correct the option specification to avoid the use of an unsupported function or combination of functions, and then submit the job again.

[ERR.] PJM 0090 pjsub PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function and the system state.

[ERR.] PJM 0091 pjsub Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0093 pjsub The operating command cannot be executed except login node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a login or compute cluster management node.

Action

Execute the command on a login or compute cluster management node.

Warning Message

**[WARN] PJM 0024 pjsub Too many arguments in script file.
Script file *filename* line *lineno*.
Ignore the subsequent options: *opts*.**

Meaning

The maximum number of options (64) allowed per line in a job script file has been exceeded. The location of the detected problem is line *lineno* of the job script file *filename*.

Action

The subsequent options that exceed the upper limit are ignored, and processing continues.

[WARN] PJM 0029 pjsub The option is ignored in the use environment: *opt*.

Meaning

The user's environment does not support the option *opt*.

This message is output only when the --verbose option is specified.

Action

The option *opt* is ignored, and processing continues.

Information Message

[INFO] PJM 0000 pjsub Job *jobid* submitted.

Meaning

The job was accepted normally. The job ID is *jobid*.

Action

No action is necessary.

**[INFO] PJM 0080 pjsub Interactive job *jobid* is canceled due to the resource allocation timeout.
The timeout period "t" can be specified by "--sparam wait-time=t".**

Meaning

The interactive job (job ID: *jobid*) was canceled because the computer resources to be allocated to the job could not be determined within a certain period. That was due to insufficient free computer resources or the job ACL limit value.

Action

Check the availability of resources and the job ACL setting value, and submit the job again. You can specify the wait time for the completion of computer resource allocation in the --sparam "wait-time=" option of the pjsub command. The default value is 0 seconds.

[INFO] PJM 0081connected.

Meaning

An interactive job is being prepared. One dot (.) appears every three seconds. 'connected' appears when the preparation is complete.

Action

No action is necessary.

[INFO] PJM 0082 pjsub Interactive job *jobid* started.

Meaning

Execution of an interactive job (job ID: *jobid*) has begun.

Action

No action is necessary.

[INFO] PJM 0083 pjsub Interactive job *jobid* completed.

Meaning

An interactive job (job ID: *jobid*) has been completed.

Action

No action is necessary.

[INFO] PJM 0086 pjsub Interactive job *jobid* canceled.

Meaning

An interactive job (job ID: *jobid*) was canceled.

Action

No action is necessary.

3.1.15 pjwait command

[NAME]

`pjwait` - The job end is waited for.

[SYNOPSIS]

```
/usr/bin/pjwait [-w {waitany|waitnone}] [-z] jobid ...  
/usr/bin/pjwait --help
```

[DESCRIPTION]

The `pjwait` command is a command that waits for the end of the job that specifies it. This command can be executed on the login node and compute cluster management node.

For *jobid*, specify the job ID or sub job ID of a job so that the command waits for the end of the job. You can specify a job ID range.

If the `-w` options are omitted, the `pjwait` command does not return until all the specified jobs end. Once the specified jobs enter the ERROR or HOLD state, the `pjwait` command returns. If the specified jobs have already ended, the `pjwait` command returns immediately. Note that nothing is output for a job in the REJECT state.

The `pjwait` command displays the PJM code, the exit code, and the signal number of the job that specifies it in the form of the following.

jobid PJMcode exitcode signal

The meaning of the PJM code is as follows.

PJM code	Meanings
Numeric value	Code representing the job manager processing result of job execution. For the meaning of the value, see the PC item of <code>pjstat(1)</code> .
-	Step job or bulk job that has ended.
X	Job waiting for one of the following states: ACCEPT, QUEUED, RUNNING-A, RUNNING-P, RUNNING, SUSPEND, SUSPENDED, RESUME, RUNNING-E, RUNOUT
E	Job in the ERROR state
H	Job in the HOLD state

The PJM code of a step or bulk job is determined based on the PJM codes of individual sub jobs. The PJM code is "X" when at least one sub job is waiting. The code is "E" when no sub job is waiting and at least one sub job is in the ERROR state. It is "H" when all sub jobs are in the HOLD state.

The exit code is that of the job script. The signal number identifies the signal whose reception caused the job script to end. With the `-w` option specified such that there is no waiting for job end, the command displays "-" (hyphen) for the exit code together with the signal numbers of jobs that have not ended.

If there is no information on the specified jobs, the command ends normally without displaying anything.

[About specification for a step job, a bulk job, and an interactive job]

- Step job
If you specify the job ID of a step job, the command waits until all the sub jobs end. Then, it displays information on the step job and all the sub jobs. If you specify the sub job IDs of a step job, the command waits until the specified sub jobs end and displays information on only the specified sub jobs. The sub job information is displayed in the order of step number.
- Bulk job
If you specify the job ID of a bulk job, the command waits until all the sub jobs end. Then, it displays information on only the bulk job and does not display any sub job information. If you specify the sub job IDs of a bulk job, the command waits until the specified sub jobs end and displays information on only the specified sub jobs. The sub job information is displayed in the order of bulk number.
If you specify the job ID of a bulk job and a sub job ID of the bulk job, the command displays information on only the bulk job.
- Interactive job
Any information is not displayed when job ID of an interactive job is specified and the pjwait command ends normally.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example: Specify a job ID range of 1 to 100

1-100

Example: Specify a step number range of 1 to 10 for a step job (job ID = 100)

100_1-10

Example: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

101[1-10]

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges

Note the following about specifying multiple job IDs: if the user does not have the privileges for any of the specified jobs, an error message is output because the jobs are explicitly specified. At this time, after outputting information on all of the specified jobs, the pjwait command returns immediately with the exit status of 1 without interrupting job processing.

[OPTIONS]

-w {waitany|waitnone}

This option specifies the *mode* in which to wait for job end.

Either the following is specified. If this option is omitted, the command waits until the end of all specified jobs.

waitany

This command returns when any of the specified jobs ends. If a specified job is a step or bulk job, the command returns when any sub job of the step or bulk job ends. If a specified job has already ended, the command returns immediately.

waitnone

The end of the job that specifies it is not waited for, and this command returns at once. The command checks the execution status of the specified jobs. "X" is the PJM code displayed for a job that has not ended.

-z

This option suppresses output of information on individual jobs to the standard output.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following is a use example of the pjwait command:

1. Multiple jobs are specified.

```
$ pjwait 10 11 12
10 0 1 0
11 1 1 9
12 0 1 0
```

2. Jobs are specified within the range.

If there is no information on a job, nothing is displayed for the job. The following example displays nothing for job IDs 3 to 8 because there is no information on the jobs.

```
$ pjwait 1-9
1 0 1 0
2 1 1 9
9 0 1 0
```

3. The step job is specified.

When job ID of the step job is specified, information on the parents job and all sub jobs is displayed.

```
$ pjwait 13
13 - - -
13_0 0 1 0
13_1 1 1 9
13_2 0 1 0
```

4. The sub job of the step job is specified.

When the sub job ID of a step job is specified, only the sub job is displayed.

```
$ pjwait 13_1
13_1 1 0 0
```

5. The bulk job is specified.

When the job ID of a bulk job is specified, only the bulk job is displayed, and individual sub jobs are not displayed.

```
$ pjwait 14
14 - - -
```

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end (When jobs are specified repeatedly, the command return abnormally even if only one error occurs.)

[SEE ALSO]

pjsub(1), pjdel(1), pjhold(1), pjstat(1)

[MESSAGES]

If an error occurs during the execution of the pjwait command, the following messages is output to the standard error output:

Error Message

[ERR.] PJM 0601 pjwait Unknown option or no option argument: *opt*.
Try `pjwait --help` for more information.

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0603 pjwait Unknown option argument: *opt arg*.
Try `pjwait --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0610 pjwait No jobid(s) specified.
Try `pjwait --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0611 pjwait Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0670 pjwait No execute permission.

Meaning

The user is not permitted to execute the pjwait command.

Action

Display the job ACL settings with the pjacl command, and check whether the pjwait command can be executed (execute pjwait). If necessary, ask the administrator to add permission.

[ERR.] PJM 0671 pjwait No permission: *jobid*.

Meaning

The user is not permitted to operate the target job (job ID: *jobid*).

The operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the pjacl command. Then, confirm the permission (permit pjwait) for the operated objects of the pjwait command for the resource unit and resource group to which the target job is submitted. If necessary, ask the administrator to add permission.

[ERR.] PJM 0690 pjwait PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to check the operating state of the job manager function or the system state.

[ERR.] PJM 0691 pjwait Internal error: *details*.**Meaning**

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. Contact the administrator when the situation does not change in about ten minutes. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 0693 pjwait The operating command cannot be executed except login node and compute cluster management node.**Meaning**

The command cannot be executed on a node other than a login or compute cluster management node.

Action

Execute the command on a login or compute cluster management node.

3.2 Messages in job outputs

Aside from messages for the commands described in "3.1 Commands", other messages might be output at job execution time. This section describes the messages displayed at job execution time.

3.2.1 plexec command

When the mpiexec command is executed in a parallel execution environment, the plexec command, which is an internal command of the parallel execution environment, is called. Therefore, if a problem occurs, messages of the plexec command may be output. The plexec command outputs the following messages to the standard error output:

Error message

[ERR.] PLE 0001 plexec '--np' must be specified.**Meaning**

The --np option is not specified for the internal command 'plexec'.

Action

Review the number of processes specified with the -c|-np|--np|-n|--n option of the mpiexec command.

[ERR.] PLE 0002 plexec PLE service error occurred.(nid=nodeid)(CODE=code)**Meaning**

Connection with the parallel execution environment daemon failed.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0003 plexec cannot create file(filename).(rank=rank)(nid=nodeid)(CODE=code)**Meaning**

The file specified with the --of option of the mpiexec command cannot be created.

Action

Check whether files can be created in the specified directory.

[ERR.] PLE 0005 plexec The specified number of processes is too many.(nid=nodeid)(CODE=code)

Meaning

The number of processes specified with the --c|-np|--np|-n|--n option of the mpiexec command exceeded the maximum specified with the pjsub command.

Action

Review the number of processes specified with the -c|-np|--np|-n|--n option of the mpiexec command.

[ERR.] PLE 0006 plexec cannot get memory.(nid=nodeid)(CODE=code)

Meaning

The node lacks memory resources.

Action

Execute the command again. If the problem persists, contact the administrator. The administrator shall check total memory of the system again.

[ERR.] PLE 0007 plexec A system error occurred.(rank=rank)(nid=nodeid)(CODE=code)

Meaning

A system error occurred.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0008 plexec must be started sequentially.(nid=nodeid)(CODE=code)

Meaning

The mpiexec command was started more than once.

Action

If the --vcoordfile option is not specified, do not issue the mpiexec command repeatedly. To execute multiple mpiexec commands, use the mpiexec command with the --vcoordfile option.

[ERR.] PLE 0009 plexec cannot write file(filename).(rank=rank)(nid=nodeid)(CODE=code)

Meaning

An error occurred in a write operation to the file specified by a --of type option of the mpiexec command.

Action

Check the disk free space. If there is not sufficient free space in the specified output disk, change the output disk to a disk that has sufficient free space. Then execute this command again.

If the problem persists even if there is sufficient free space in the specified output disk, contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0010 plexec cannot open file(filename).(nid=nodeid)(CODE=code)

Meaning

The file specified with the --stdin option of the mpiexec command cannot be opened.

Action

Review the file name specified with the `--stdin` option of the `mpiexec` command.

[ERR.] PLE 0011 plexec The program is not specified.

Meaning

No program is specified.

Action

Specify a program for the `mpiexec` command.

[ERR.] PLE 0012 plexec The executing user is illegal on the node.(*nid=nodeid*)(*CODE=code*)

Meaning

There is an inconsistency in the user information in the node.

Action

Contact the administrator. The administrator shall check whether the user information is consistent on the node *nodeid*.

[ERR.] PLE 0013 plexec The error occurred by setting the directory.(*dir=dirname*)(*nid=nodeid*)(*CODE=code*)

Meaning

A parallel process failed to move to the directory *dirname*.

Action

Contact the administrator. The administrator shall check whether the user has execute permission of the directory *dirname* on the node *nodeid*.

[ERR.] PLE 0014 plexec The process terminated abnormally.(*rank=rank*)(*nid=nodeid*)(*exitstatus=exitstatus*)(*CODE=code1,code2,code3*)

[ERR.] PLE The program that the user specified may be illegal or inaccessible on the node.(*nid=nodeid*)

Meaning

All MPI programs were aborted because the process terminated abnormally. The program that the user specified may be illegal or inaccessible on the node.

Action

See job end code of the job statistical information file.

If the job ended normally, check whether the cause of the abnormal termination resided in the program specified with the `mpiexec` command, or whether the program can be referred on the node (node ID *nodeid*).

If the job was submitted with the option of the memory limitation (`proc-data`, `proc-lockm`, `proc-stak`, `proc-vmem`), the program might have been ended by exceeding the memory limitation.

If memory limitation was not specified, the administrator would set its default values by job ACL function. Check them by using `pjagl` command, and ask the administrator to change these values if necessary.

If the job ended abnormally, take an appropriate action according to job end code.

If the job statistical information file was not output, submit the job with `-s` or `-S` option of `pjsub` command again and investigate the cause.

[ERR.] PLE 0015 plexec cannot execute as sequential job.

Meaning

The command cannot be executed as a sequential job.

Action

To execute an MPI program on a single node (`node=1`) or a single virtual node (`vnode=1`), be sure to specify the `--mpi` option.

[ERR.] PLE 0016 plexec The program arguments is too long**Meaning**

An argument of the specified program is too long.

Action

Check the arguments of the program specified with the mpiexec command.

[ERR.] PLE 0017 plexec The process terminated with the signal.(rank=rank)(nid=nodeid)(sig=signal)**Meaning**

All MPI programs were aborted because the specified program ended with a signal.

Action

See job end code of the job statistical information file.

If the job ended normally, check whether the cause of the signal *signal* resided in the program specified with the mpiexec command.

If the value of *signal* was 9 (SIGKILL), the program might have been terminated by memory limitation of operating system on the node. This indicates that memory resource allocated to the job was insufficient for creating processes of the program.

If the job was submitted with a specified upper limit (node-mem) on memory usage by a node or a specified memory amount (vnode-mem or mem) per virtual node, increase that amount.

If the job was submitted with no memory setting option, the administrator would set its default values by job ACL function. Check them by using pjacl command, and ask the administrator to change these values if necessary.

Note that if the value of vnode-mem or mem is 'unlimited', then the program also might have been terminated by resource competition with other job on the same node. To avoid the lack of memory, select the job execution mode policy SIMPLEX so that the job can occupy the node.

If the job ended abnormally, take an appropriate action according to job end code.

If the job statistical information file was not output, submit the job with -s or -S option of pjsub command again and investigate the cause.

[ERR.] PLE 0018 plexec The program(program filename) cannot be executed.(nid=nodeid)(error=error message)(CODE=code1,code2,code3)**Meaning**

All MPI programs were aborted because the specified program could not be executed.

Action

Check whether the program specified with the mpiexec command can be executed.

[ERR.] PLE 0019 plexec One of MPI processes was aborted.(rank=rank)(nid=nodeid)(CODE=code1,code2,code3)**Meaning**

All MPI programs were aborted because one of MPI processes was aborted.

Action

Check the MPI program.

[ERR.] PLE 0020 plexec The MPI internal error occurred.[The directory (dirname) specified with --tmpdir is illegal.](nid=nodeid)(CODE=code1,code2,code3)**Meaning**

mpiexec command ended by internal error.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0022 plexec has timed out.

Meaning

The internal command plexec terminated because of a timeout. This message is output if interactive job execution cannot start within a certain period after the computer resources to be allocated to the job are determined. The execution start wait time is 15 + (specified value of --sparam "wait-time=" option of pjsub command) seconds. The default value of this option is 0 seconds, so if the option is not specified, the execution start wait time is 15 seconds.

Action

Submit the job again with the --sparam "wait-time=" option to extend the wait time for starting the job by about 1 minute. If the error still occurs even if extending the wait time to about 10 minutes, contact the administrator. The administrator should collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0024 plexec There is no authority to execute the demanded operation.

Meaning

The user does not have the privileges to perform the requested operation.

Action

Confirm the execution privileges for the operation performed by the executed command.

[ERR.] PLE 0030 plexec Duplicated a nodelist.

Meaning

The --vcoordfile and --vnodfile options of the mpiexec command are specified at the same time.

Action

The --vcoordfile and --vnodfile options of the mpiexec command cannot be specified at the same time. Review the options specified for the mpiexec command.

[ERR.] PLE 0037 plexec failed to initialize PSM.

Meaning

Initialization of the system management function failed.

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PLE 0041 plexec The specified prefix(*prefix*) is invalid.

Meaning

There is an error in the specification of the --ofprefix option of the mpiexec command.

Action

Review the value specified with the --ofprefix option of the mpiexec command.

[ERR.] PLE 0042 plexec A virtual coordinate "vcoord" is already used by other plexec.

Meaning

The coordinate *vcoord* node is being used by a process executed by another mpiexec command.

Action

If the MCA parameter "mpi_no_establish_communication" is specified as 1 and the mpiexec command is issued with the --vcoordfile option, then a node whose coordinates match the coordinates in the file specified by the --vcoordfile option is already in use by a process generated by another mpiexec. Review the contents of the file specified by the --vcoordfile option of the mpiexec command or review

the MCA parameter " mpi_no_establish_communication " specification. For more information, see "Development Studio MPI User's Guide".

[ERR.] PLE 0043 plexec The specified virtual coordinate "*vcoord*" is invalid.

Meaning

The specified logical coordinate *vcoord* is invalid.

The specified logical coordinate *vcoord* may be out of the range of the nodes allocated to the job or belong to a dimension incompatible with the node shape.

Action

Review the contents of the file specified with the --vcoordfile option of the mpiexec command.

[ERR.] PLE 0044 plexec The file "*filename*" specified with --vcoordfile does not exist.

Meaning

The file *filename* specified with the --vcoordfile option of the mpiexec command does not exist.

Action

Check the file name specified with the --vcoordfile option of the mpiexec command.

[ERR.] PLE 0045 plexec A virtual node "*vnode*" is already used by other plexec.

Meaning

The virtual node *vnode* is being used by a process executed by another mpiexec command.

Action

A process cannot be created for the same virtual node from multiple mpiexec commands. Review the contents of the file specified with the --vnodfile option of the mpiexec command.

[ERR.] PLE 0046 plexec The specified virtual node "*vnode*" is invalid.

Meaning

The specified virtual node *vnode* is invalid.

The specified virtual node *vnode* may be out of the range of virtual nodes allocated to the job.

Action

Review the contents of the file specified with the --vnodfile option of the mpiexec command.

[ERR.] PLE 0047 plexec The file "*filename*" specified with --vnodfile does not exist.

Meaning

The file *filename* specified with the --vnodfile option of the mpiexec command does not exist.

Action

Review the file name specified with the --vnodfile option of the mpiexec command.

[ERR.] PLE 0048 plexec The number of processes exceed the number of virtual coordinates specified by --vcoordfile.

Meaning

The number of processes specified with the -c|-np|--np|-n|--n option of the mpiexec command exceeds the number of lines of the file specified with the --vcoordfile option.

Action

Make the number of processes specified with the -c|-np|--np|-n|--n option of the mpiexec command equal to or less than the number of lines of the file specified with the --vcoordfile option.

[ERR.] PLE 0049 plexec The number of processes exceed the number of virtual nodes specified by --vnodefile.

Meaning

The number of processes specified with the `-c|-np|--np|-n|--n` option of the `mpiexec` command exceeds the number of lines of the file specified with the `--vnodefile` option.

Action

Make the number of processes specified with the `-c|-np|--np|-n|--n` option of the `mpiexec` command equal to or less than the number of lines of the file specified with the `--vnodefile` option.

[ERR.] PLE 0050 plexec cannot be executed any further.

Meaning

The number of times the `mpiexec` command or `pjaexe` command was executed more than once exceeded the maximum value 128.

Action

Review the job contents, and reduce the number of times the `mpiexec` command or `pjaexe` command is executed more than once to a value that is equal to or less than the maximum value allowed by the Job Operation Software.

[ERR.] PLE 0051 plexec The usernames and uids on all nodes must be defined identically.

Meaning

The user name and the user information regarding the user ID are inconsistent.

Action

The system administrator should confirm that the user information is consistent on each node.

[ERR.] PLE 0052 plexec Invalid job model.

Meaning

There is an error in the job model when executing the `pjsub` command.

The `pjaexe` command can only be used for master-worker job.

Action

When using the `pjaexe` command, specify the `--mswk` option in the `pjsub` command and submit the job again.

[ERR.] PLE 0053 plexec Incompatible version of MPI library.

Meaning

The combination of the executed job and version level of the specified MPI library cannot be executed.

Action

Use the latest version of the MPI library.

[ERR.] PLE 0054 plexec The number of processes exceed the limit on virtual coordinate *vcoord*.

Meaning

The number of processes to be generated on the node at the logical coordinate *vcoord* exceeds the limit on the number of processes that can be generated per the logical coordinate.

The limit on the number of processes that can be generated per the logical coordinate is the value specified in the `--mpi "proc="` option or `--mpi "max-proc-per-node="` option.

Action

Make sure that the number of processes to be generated on the specified logical coordinate *vcoord* does not exceed the limit on the number of processes that can be generated per the logical coordinate.

[ERR.] PLE 0055 plexec The number of processes exceed the limit on virtual node *vnode*.**Meaning**

Multiple processes cannot be generated on the virtual node *vnode* at the same time.

Action

Only one process can be generated on the virtual node. Review the number of processes to be generated on the virtual node.

[ERR.] PLE 0056 plexec The number of unused virtual coordinates are not sufficient.**Meaning**

The number of unused nodes is too small to generate processes.

Action

Confirm each of the following specifications is correct: the number of nodes to be allocated to the job, and the limit on the number of processes that can be generated per node.

If only the number of CPU cores assigned to each process is specified in the `--vcoordfile` option of the `mpiexec` command, the specification order may be the cause of this error.

The Job Operation Software determines the nodes where CPU cores can be reserved depending on the sequence of specified numbers of CPU cores. Some specified sequences may have a node shortage due to an attempt to reserve an unnecessarily large number of nodes. For example, suppose a node group has 16 CPU cores per node. Then, a sequence of 15, 1, 15, and 1 CPU core for 4 processes requires 2 nodes in total to be allocated. In contrast, a sequence of 15, 15, 1, and 1 CPU core requires 3 nodes in total to be allocated.

To avoid this problem, explicitly specify logical node coordinates too so that the processes are placed appropriately.

[ERR.] PLE 0058 plexec The number of CPU cores on a virtual coordinate *vcoord* are not sufficient.**Meaning**

The node at the logical coordinate *vcoord* does not have enough CPU cores to allocate.

Action

Make the total number of CPU cores assigned to processes generated on a node equal to or less than the number of CPU cores installed on the node.

[ERR.] PLE 0060 plexec Format error in `--vcoordfile`.**Meaning**

The file specified in the `--vcoordfile` option of the `mpiexec` command contains a format error.

Action

Review the contents of the file specified in the `--vcoordfile` option of the `mpiexec` command.

For the location and details of the error, see the PLE 0093 information message output at the same time.

[ERR.] PLE 0061 plexec Format error in `--vnodefile`.**Meaning**

The file specified in the `--vnodefile` option of the `mpiexec` command contains a format error.

Action

Review the contents of the file specified in the `--vnodefile` option of the `mpiexec` command.

For the location and details of the error, see the PLE 0093 information message output at the same time.

[ERR.] PLE 0062 plexec Cannot set `numanode_assign_policy`.**Meaning**

Core allocation to the process cannot be implemented under the policy for allocating cores to NUMA nodes.

Action

Review the policy for allocating cores to NUMA nodes, that is specified in the process.
For details on the core allocation policy, see the "Development Studio MPI User's Guide."

[ERR.] PLE 0063 plexec Output file pathname too long.(nid=nodeid)(CODE=code)

Meaning

The file name specified in the standard output or standard error output destination options of the mpiexec command is too long.

Action

Check the file name specified by the mpiexec command.

[ERR.] PLE 0070 plexec The number of processes "n" must be smaller than or equal to the number of virtual nodes "m": "pjsub -L vnode=m", "pjsub --mpi proc=n".

Meaning

The number of processes specified in the option --mpi proc= of the pjsub command must be less than or equal to the number of virtual nodes specified in the option -L vnode=.

Action

In the option --mpi proc= of the pjsub command, specify a number of processes that is less than or equal to the number of virtual nodes specified in the option -L vnode=.

[ERR.] PLE 0080 plexec PMIx service error occurred.(nid=nodeid)(CODE=code1,code2,code3)

Meaning

An error occurred in the service providing the PMIx (Process Management Interface Exascale) used by the MPI library.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

In rare cases, when the job is deleted due to pjdel or pjsig command execution, exceeded job elapsed time and job resource (Memory, etc.) limits, deadline scheduling or node down, the above error message may appear, but should be ignored. To suppress the output of the above message, use the pjsig command with the following signals:.

SIGTERM, SIGHUP, SIGINT, SIGUSR1, SIGUSR2

[ERR.] PLE 0081 plexec pleio service error occurred.(nid=nodeid)(CODE=code1,code2,code3)

Meaning

An error occurred in the service (pleio) for integrating and processing data when outputting standard output and standard error output of MPI programs to a file.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

In rare cases, when the job is deleted due to pjdel or pjsig command execution, exceeded job elapsed time and job resource (Memory, etc.) limits, deadline scheduling or node down, the above error message may appear, but should be ignored. To suppress the output of the above message, use the pjsig command with the following signals:.

SIGTERM, SIGHUP, SIGINT, SIGUSR1, SIGUSR2

[ERR.] PLE 0905 plexec cannot get the job information. (rank=rank)(nid=nodeid)(CODE=code)

Meaning

Job information retrieval failed.

Action

Execute program in the job script. If the program is executed in the job script, contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

Warning message

[WARN] PLE 0601 plexec Cannot set "memory_allocation_policy".(policy=policy)(rank=rank)(nid=nid) (errorno=error_number)

Meaning

The NUMA memory allocation policy cannot be set for the process. The process will execute with the default NUMA memory allocation policy.

Action

Review the NUMA memory allocation policy that is specified in the process.
For details on the NUMA memory allocation policy, see the "Development Studio MPI User's Guide."

[WARN] PLE 0602 plexec "numanode_assign_policy" is simplex, but the job is not simplex.(rank=rank) (nid=nid)

Meaning

For the process, simplex (NUMA node exclusive) was specified as the policy for allocating cores to the NUMA node, but the job is not a node allocated job, and it may not be possible to make the NUMA node exclusive.

Action

Review the job submission method.
For details on the policy for allocating cores to the NUMA node, see the "Development Studio MPI User's Guide."

[WARN] PLE 0603 plexec Invalid "numanode_assign_policy".(policy=policy)(rank=rank)(nid=nid) (error=error message)

Meaning

An invalid value was specified for the environment variable OMPI_MCA_plm_ple_numanode_assign_policy, under the policy for allocating cores to processes. The process will execute with the default core allocation policy.

Action

Review the value of the environment variable OMPI_MCA_plm_ple_numanode_assign_policy.
For details on the core allocation policy and the environment variable OMPI_MCA_plm_ple_numanode_assign_policy, see the "Development Studio MPI User's Guide."

[WARN] PLE 0605 plexec ignore option opt.
or
[WARN] PLE 0605 plexec ignore option opt.(details)

Meaning

The user specified the option *opt* to specify the standard output or standard error output as the output destination of the mpiexec command, but did not have the privileges to specify the option, so it was ignored.
Details may also be output.

details: Destination information for standard output or standard error output

Example: (stdout outputs to *filepath*)

The standard output is output to a file whose file name starts with *filepath*.

"*" in the *filepath* means that each rank number of process has a different name.

For details, see "Development Studio MPI User's Guide."

Action

To enable specification, contact the administrator.

[WARN] PLE 0606 plexec ignore environment variable(*details*).

Meaning

The user specified the environment variable PLE_MPI_STD_EMPTYFILE but did not have the privileges to specify the variable, so it was ignored.

details: PLE_MPI_STD_EMPTYFILE = *Specified value*

Action

To enable specification, contact the administrator.

[WARN] PLE 0607 plexec suffix specified for stdout/stderr can not be converted.(*details*)

Meaning

The set character string for the suffix of the standard output file or standard error output file cannot be converted.

details: Suffix of the standard output file or standard error output file

Action

Contact the administrator.

[WARN] PLE 0608 plexec One of MPI processes was aborted.(*rank=rank*)(*nid=nodeid*) (*CODE=code1,code2,code3*)(*DETAIL=%s*)

or

[WARN] PLE 0608 plexec One of MPI processes was aborted.(*rank=rank*)(*nid=nodeid*) (*CODE=code1,code2,code3*)

Meaning

The MPI program was aborted because one of MPI processes was aborted. Other MPI programs will continue to execute.

Action

Check the MPI program.

[WARN] PLE 0609 plexec The process terminated abnormally.(*rank=rank*)(*nid=nodeid*) (*exitstatus=exitstatus*)(*CODE=code1,code2,code3*)

Meaning

The MPI program was aborted because the process terminated abnormally. Other MPI programs will continue to execute. The program that the user specified may be illegal or inaccessible on the node.

Action

See job end code of the job statistical information file.

If the job ended normally, check whether the cause of the abnormal termination resided in the program specified with the mpiexec command, or whether the program can be referred on the node (node ID *nodeid*).

If the job was submitted with the option of the memory limitation (proc-data, proc-lockm, proc-stak, proc-vmem), the program might have been ended by exceeding the memory limitation.

If memory limitation was not specified, the administrator would set its default values by job ACL function. Check them by using pjacl command, and ask the administrator to change these values if necessary.

If the job ended abnormally, take an appropriate action according to job end code.

If the job statistical information file was not output, submit the job with -s or -S option of pjsub command again and investigate the cause.

[WARN] PLE 0610 plexec The process terminated with the signal.(*rank=rank*)(*nid=nodeid*)(*sig=signal*)

Meaning

The MPI program was aborted because the specified program ended with a signal. Other MPI programs will continue to execute.

Action

See job end code of the job statistical information file.

If the job ended normally, check whether the cause of the signal *signal* resided in the program specified with the `mpiexec` command.

If the value of *signal* was 9 (SIGKILL), the program might have been terminated by memory limitation of operating system on the node. This indicates that memory resource allocated to the job was insufficient for creating processes of the program.

If the job was submitted with a specified upper limit (node-mem) on memory usage by a node or a specified memory amount (vnode-mem or mem) per virtual node, increase that amount.

If the job was submitted with no memory setting option, the administrator would set its default values by job ACL function. Check them by using `pjacl` command, and ask the administrator to change these values if necessary.

Note that if the value of vnode-mem or mem is 'unlimited', then the program also might have been terminated by resource competition with other job on the same node. To avoid the lack of memory, select the job execution mode policy SIMPLEX so that the job can occupy the node.

If the job ended abnormally, take an appropriate action according to job end code.

If the job statistical information file was not output, submit the job with `-s` or `-S` option of `pjsub` command again and investigate the cause.

[WARN] PLE 0611 plexec The program(*program filename*) cannot be executed.(*nid=nodeid*)(*error=error message*)(*CODE=code1,code2,code3*)

Meaning

The MPI program was aborted because the specified program could not be executed. Other MPI programs will continue to execute.

Action

Check whether the program specified with the `mpiexec` command can be executed.

If there is no problem with the execute permissions and the following message is output in *error message*, the number of user processes in the node created by the job may have reached its limit.

execve failed. Resource temporarily unavailable.

For example, this error may occur when a large number of `mpiexec` commands specified VCOORD files are executed in the background. Increase the number of processes that can be generated by using the `proc-crproc` parameter of `{-L|--rsc-list}` option to the `pjsub` command.

If the maximum value does not prevent the same behavior, ask the administrator to review the job ACL function settings.

The administrator should increase the maximum value of the following setting items in the job ACL function so that larger values can be specified in the `proc-crproc` parameter of the `pjsub` command.

Setting item for batch jobs: `joblimit proc-crproc`

Setting item for interactive jobs: `joblimit interact-proc-crproc`

[WARN] PLE 0613 plexec The executing user is illegal on the node.(*nid=nodeid*)(*CODE=code*)

Meaning

Inconsistent user information in node.

Action

Contact the administrator. The administrator is requested to check the user information matches on each node.

[WARN] PLE 0615 plexec The error occurred by setting the directory.(*dir=dirname*)(*nid=nodeid*)(*CODE=code*)

Meaning

Parallel process failed to move to directory *dirname*.

Action

Contact the administrator. The administrator is requested to check the directory *dirname* on each node has execute privileges for the job execution user.

[WARN] PLE 0616 plexec A system error occurred.(*rank=rank*)(*nid=nodeid*)(*CODE=code*)

Meaning

A system error occurred.

Action

Contact the administrator. The administrator shall collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[WARN] PLE 0617 plexec cannot open file(filename).(nid=nodeid)(CODE=code)

Meaning

The file specified by the --stdin option of the mpiexec command cannot be opened.

Action

Review the file name specified with the --stdin option of the mpiexec command.

[WARN] PLE 0618 plexec Output file pathname too long.(rank=rank)(nid=nodeid)(CODE=code)

Meaning

The file name specified in the standard output or standard error output destination options of the mpiexec command is too long.

Action

Check the file name specified by the mpiexec command.

Information message

[INFO] PLE 0090 plexec The process terminated by '~.'

Meaning

The running interactive job ended.

Action

No action is necessary.

[INFO] PLE 0091 plexec ~^Z[suspend tty]

Meaning

A running job has been suspended.

Action

No action is necessary.

[INFO] PLE 0092 plexec Supported escape characters:

~. : terminate processes

~^Z : suspend tty

~? : display the usage of escape characters

~~ : send '~'

Note that escape characters are only effective at the beginning of line.

Meaning

The following escape characters are supported for interactive jobs:

~. : Exits the job shell in execution and ends the interactive job.

~^Z : Suspend the interactive job.

~? : Lists escape characters.

~~ : Transfers the tilde symbol for input for the interactive job.

These characters are valid only when input at the beginning of a line.

Action

No action is necessary.

[INFO] PLE 0093 plexec Format error: *filename:line:message*

Meaning

If the file *filename* contains a format error, this message is output together with an error message to show the location (line) and details (*message*) of the error.

If the file *filename* contains multiple errors, this message is output for each error.

Action

See the detailed message *message* and the error message that was output at the same time, and correct line line in the file *filename*.

[INFO] PLE 0094 plexec The interactive job has received the signal.(sig=*signum*)

Meaning

The interactive job was received the signal *signum*.

Action

If necessary, execute the interactive job again.

Note

When running an interactive job, if you exit the interactive job after entering "Ctrl + c", you will see the above error message before the normal exit message, which should be safely ignored.

3.2.2 Parallel execution environment

Messages about the parallel execution environment may be output during job execution.

Parallel execution environment error messages are output to the standard error output, and information messages such as notification of job re-execution are output to the standard output.

Error message

[ERR.] PLE 9350 PLE Cannot open the file(*filename*) for analyzing options.

Meaning

There may be a problem in the settings of the parallel execution environment.

Action

Contact the administrator. The administrator is requested to check the values of the environment variable PLE_I_MPI_INTEL_OPTIONS and check whether the file *filename* that is set in the environment variable can be referenced. PLE_I_MPI_INTEL_OPTIONS is described in "Appendix E Settings Related to Execution in MPI Processing Systems Other Than Development Studio" in the "Job Operation Software Administrator's Guide for Job Management."

Information message

[INFO] PLE 1700 The job restarts.

Meaning

The job was re-executed.

Action

No action is necessary.

3.2.3 Job management function

Depending on the execution results of a job, a message of the job manager function may be output to the standard error output of the job.



Note

The administrator sets whether to output the job manager function messages described in this section. Therefore, no message may be output even though the relevant event has occurred.

Error message

[ERR.] PJM 7001 Job *jobid* was canceled. A deadline schedule was set on nodes running the job.

Meaning

The job (job ID: *jobid*) was aborted because of the deadline period that was forcibly set by the administrator. (Job end code 7)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

[ERR.] PJM 7002 Job *jobid* was canceled. The job was submitted with no-restart attribute.

Meaning

The job (job ID *jobid*) ended abnormally because, for example, all nodes were restarted. However, the job was canceled because automatic re-execution is disabled. (Job end code 9)

Action

Submit the job again.

[ERR.] PJM 7003 Job *jobid* was canceled. The job exceeded its CPU time limit.

Meaning

The CPU time of the job (job ID: *jobid*) exceeded the upper limit value. (Job end code 10)

Action

Increase the upper limit value of the CPU time for a job, or reduce the CPU time used by the job. Then, submit the job again.

[ERR.] PJM 7004 Job *jobid* was canceled. The job exceeded its elapse time limit.

Meaning

The execution time of the job (job ID *jobid*) exceeded the upper limit value. (Job end code 11)

Action

Increase the upper limit value of the executable time for a job, or make the job end within the executable time. Then, submit the job again.

[ERR.] PJM 7005 Job *jobid* was canceled. The job exceeded its memory limit.

Meaning

The amount of memory used by the job (job ID *jobid*) exceeded the upper limit value. (Job end code 12)

Action

Increase the upper limit value for the amount of memory used by a job, or reduce the amount of memory used by the job. Then, submit the job again.

[ERR.] PJM 7007 Job *jobid* was canceled. The current directory did not exist or have permission.

Meaning

The job (job ID *jobid*) could not access the current directory, standard input file, standard output file, or standard error output file. (Job end code 16)

Action

Confirm that the current directory, standard input file, standard output file, and standard error output file for the job are correctly specified. Then, submit the job again.

[ERR.] PJM 7008 Job *jobid* was canceled. The node running the job was down.

Meaning

A node shut down during job (job ID *jobid*) execution. (Job end code 20)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again. The job is executed on another node that has not shut down..

[ERR.] PJM 7009 Job *jobid* was canceled. The job script execution failed.

Meaning

The job script could not be executed. (Job end code 21)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 7010 Job *jobid* was canceled. An error occurred in ICC.

Meaning

An ICC (Interconnect Controller) error occurred during job (job ID *jobid*) execution on an FX server. (Job end code 22)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again. The job is executed on nodes other than the node where the ICC error occurs.

[ERR.] PJM 7011 Job *jobid* was canceled. The job was killed by OOM killer.

Meaning

OOM killer of the OS terminated the job process. (Job end code 23)

jobid: Job ID

Action

Submit the job again. If the error still occurs, contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 7012 Job *jobid* was canceled. The enhanced CPU statistical information was not able to be gathered.

Meaning

Collection of the enhanced CPU statistical information of a job (job ID *jobid*) executed on an FX server. failed. (Job end code 24)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

[ERR.] PJM 7013 Job *jobid* was canceled. HA failed.

Meaning

Processing to switch nodes occurred during job (job ID *jobid*) execution but failed. (Job end code 25)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

[ERR.] PJM 7014 Job *jobid* was canceled. An error occurred in the prologue or epilogue process.

Meaning

An error occurred during processing by the prologue or epilogue function set by the administrator. (Job end code 26)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to check the processing results of the prologue or epilogue function (job statistical information items PRO EXIT CODE and EPI EXIT CODE), and determine the necessity of taking corrective action for the cause and resubmitting the job.

[ERR.] PJM 7015 Job *jobid* was canceled. An error occurred in the resource management exit process.

Meaning

An error occurred during processing by the resource management exit function set by the administrator. (Job end code 27)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to check the processing results of the resource management exit function, and determine the necessity of taking corrective action for the cause and resubmitting the job.

[ERR.] PJM 7016 Job *jobid* was canceled. The job suspend failed.

Meaning

The suspend process of the job failed, and the job (job ID: *jobid*) was aborted. (Job end code 30)

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

[ERR.] PJM 7018 Job *jobid* was canceled. An error occurred in the job manager.

Meaning

An internal error occurred in the job manager function. (Job end code 100)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

The job entered the ERROR state.

[ERR.] PJM 7019 Job *jobid* was canceled. An error occurred in the job resource management.**Meaning**

An internal error occurred in the job resource management function. (Job end code 140)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

[ERR.] PJM 7021 Job *jobid* was canceled. An error occurred in the llio function.**Meaning**

An internal error occurred in the LLIO function. (Job end code 180)

jobid: Job ID

Action

Contact the administrator. The administrator is requested to collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled, submit the job again.

Warning message

[WARN] PJM 7050 Job *jobid* exceeded its memory limit.**Meaning**

The job (job ID *jobid*) ended normally (job end code 0). However, the operation may not have been performed normally because the amount of memory used by the job exceeded the upper limit value.

Action

Confirm that the job execution results are correct. If you submit the job again, reduce the amount of memory used by the job, or increase the upper limit of memory. To increase the set upper limit value in the job ACL function, ask the administrator.

Information message

[INFO] PJM 7090 Job *jobid* was canceled. User *user* executed pjdcl command.**Meaning**

The job (Job ID: *jobid*) was deleted by the pjdcl command that the user (*user*) had executed.

Action

No action is necessary.

[INFO] PJM 7091 Job *jobid* was canceled. User *user* executed pjhold command.**Meaning**

The job (Job ID: *jobid*) was aborted by the pjhold command that the user (*user*) had executed.

Action

No action is necessary.

3.2.4 Tofu Interconnect [FX]

Jobs may be executed on an FX server with a specification of whether to change the communication path when a Tofu interconnect link goes down. The following messages may be output to the standard error output of those jobs.

Error Message

[ERR.] TOF 0001 Performing the Tofu communication again failed 3 times.(time)(hostname)(rank=rank)(ecode=ecode)

Meaning

Tofu interconnect communication was retried and failed three times.

time: Date and time
hostname: Host name
rank: Rank number
ecode: Error code

Action

Submit the job again. After being resubmitted, the job is allocated to another node.

[ERR.] TOF 0002 There was no available path to perform a Tofu communication again.(time)(hostname)(rank=rank)(ecode=ecode)

Meaning

No communication paths were available for retrying Tofu interconnect communication.

time: Date and time
hostname: Host name
rank: Rank number
ecode: Error code

Action

If automatic re-execution of the job is enabled, it is automatically re-queued. If automatic re-execution of the job is disabled for the job, submit the job again.

After being resubmitted, the job is executed on a node using another path that does not have any links down.

[ERR.] TOF 0003 Performing the Tofu communication again failed because of time-out error.(time)(hostname)(rank=rank)(ecode=ecode)

Meaning

Tofu interconnect communication was retried and failed due to a timeout error.

time: Date and time
hostname: Host name
rank: Rank number
ecode: Error code

Action

Submit the job again. After being resubmitted, the job is allocated to another node.

Information Message

[INFO] TOF 0901 One of Tofu communication was performed again because a Tofu link-down had been detected.(time)(hostname)(rank=rank)(retry=retry)

Meaning

One of the Tofu interconnect communications was retried because a link was detected to be down.

time: Date and time
hostname: Host name
rank: Rank number
retry: Retry count

Action

No action is necessary. Processing continues.

[INFO] TOF 0902 The Tofu library is searching for a path to perform a Tofu communication again.(time)(hostname)(rank=rank)(retry=retry)

Meaning

The Tofu library is searching for a path to retry Tofu interconnect communication.

time: Date and time
hostname: Host name
rank: Rank number
retry: Retry count

Action

No action is necessary. Processing continues.

3.3 Others

This section describes the reference for others.

3.3.1 mpiexec.tcs_intel [PG]

[NAME]

mpiexec.tcs_intel - Intel MPI wrapper command

[SYNOPSIS]

```
/usr/bin/mpiexec.tcs_intel [ arguments of mpiexec.hydra ]
```

[DESCRIPTION]

The mpiexec.tcs_intel(1) command executes mpiexec.hydra, an Intel MPI command, under Technical Computing Suite according to the resource information specified in the pjsub(1) command.

When executing the mpiexec.hydra command, the mpiexec.tcs_intel(1) command automatically supplements the mpiexec.hydra command options according to the option specified in the pjsub(1) command. The supplemented contents are as follows:

- "--mpi proc=num" option specified in the pjsub(1) command
Within the mpiexec.tcs_intel(1) command, the "-n(-np) num" option of the mpiexec.hydra command is supplemented.
- "--mpi max-proc-per-node=num" option specified in the pjsub(1) command
Within the mpiexec.tcs_intel(1) command, the "-perhost num" option of the mpiexec.hydra command is supplemented.

However, if the mpiexec.hydra command option is explicitly specified in the mpiexec.tcs_intel(1) command, the specified option has priority over the option specified in the pjsub(1) command.

[ENVIRONMENT]

PLE_I_MPI_PJPBIND

Set "enable" so that the mpiexec.tcs_intel(1) command uses the pjpbinding(1) command, which automatically binds CPU cores to processes and threads. Otherwise, set "disable".

Operation when this environment variable is not set follows the job operation administrator's setting. When the setting by the job operation administrator is not done, it becomes the same operation as time when "enable" was set.

PLE_I_MPI_PJPBIND_OPT

Set the pjpbind(1) command options to be used within the mpiexec.tcs_intel(1) command.

Operation when this environment variable is not set follows the job operation administrator's setting.

[EXAMPLE]

The following examples show job scripts, each for a different purpose.

Example 1: Single Program Multiple Data

Job executed with the following values: number of nodes = 16, number of processes = 64, and maximum number of processes per node = 4

```
#PJM -L "node=16"
#PJM --mpi "proc=64"
#PJM --mpi "max-proc-per-node=4"
export PATH=${PATH}:< PATH of mpiexec.hydra >
mpiexec.tcs_intel ./a.out
```

Example 2: Multiple Program Multiple Data

Job executed with the following values: number of nodes = 16, number of processes = 32 per program, and maximum number of processes per node = 4

```
#PJM -L "node=16"
#PJM --mpi "max-proc-per-node=4"
export PATH=${PATH}:< PATH of mpiexec.hydra >
mpiexec.tcs_intel -n 32 ./a.out : -n 32 ./b.out
```

[END STATUS]

The returned value is the end status of the executed mpiexec.hydra command.

[NOTE]

The PATH environment variable in the job must include the installation directory of the mpiexec.hydra command.

3.3.2 pjstatsinfo

[NAME]

Job statistical information items

[DESCRIPTION]

PJSTATSINFO describes the contents of the job statistical information output by the pjstat, pjsub, or pmdumpjobinfo command. Their columns in a table are described below.

- Item:
Entries and descriptions of job statistical information items
- Item name:
Names that identify items. Administrators use these names when setting the output contents of job statistical information. Items shown with a hyphen (-) for this name are not output by the pmdumpjobinfo command.
- Data type:
Types of item values. These data types are used by administrators when they set the output contents of job statistical information.
- pjstat/.stats file:
Item names of the job statistical information output either by the pjstat command with the -s/-S option specified or to the .stats file generated by the pjsub command with the -s/-S option specified. Items shown with a hyphen (-) for this item name are neither output by pjstat nor output to the .stats file.

Job statistical information (Output items for each job)

The following data is output as job statistical information.

Item	Item name	Data type	pjstat/.stats file
Record name The output of "JI" indicates a job information record.	record	string	-
Job ID	jid	uint32	-
Step number For jobs that are not sub jobs of a step job, pmdumpjobinfo outputs "".	stepno	uint16	-
Bulk number For jobs that are not sub jobs of a bulk job, pmdumpjobinfo outputs "".	bulkno	uint32	-
Job ID or sub job ID The pjstat command and the .stats file show the output contents. For a sub job ID, the output ID is in the <i>job ID[bulk number]</i> or <i>job ID_step number</i> format.	-	-	JOB ID
Number of completed sub jobs of a bulk or step job This item is used for output the "SUB JOB NUM" of the pjstat command.	esnum	uint32	-
Total number of sub jobs of a bulk or step job For sub jobs of a bulk or step job and for normal jobs, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	snum	int32	-
Number of sub jobs of a bulk or step job The pjstat command and the .stats file show the output contents. " <i>number of completed sub jobs/total number of sub jobs</i> " For sub jobs of a bulk or step job and for normal jobs, "-" is output.	-	-	SUB JOB NUM
Starting bulk number For other than bulk job summary information, it is as follows. - pjstat/.stats: "-" - pmdumpjobinfo: ""	sbn	uint32	START BULKNO
Ending bulk number For other than bulk job summary information, it is as follows. - pjstat/.stats: "-" - pmdumpjobinfo: ""	ebn	uint32	END BULKNO
Host name of a node executing the pjsb command For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	hostname	string	HOST NAME
Job name The name specified in the -N or --name option of the pjsb command is output. If the pjsb command specifies neither the -N nor --name option, the file name of the job script is output. For jobs submitted from the standard input, "STDIN" is output.	jnam	string	JOB NAME

Item	Item name	Data type	pjstat/.stats file
For step job summary information, the job name that specified when you submit first sub job of the step job is output.			
Job type <ul style="list-style-type: none"> - pjstat/.stats BATCH (Batch job) INTERACT (Interactive job) - pmdumpjobinfo 1 (Batch job) 2 (Interactive job) 	jtyp	uint32	JOB TYPE
Job model <ul style="list-style-type: none"> - pjstat/.stats NM (Normal job) BU (Bulk job) ST (Step job) MW (Master worker job) - pmdumpjobinfo 1 (Normal job) 2 (Bulk job) 3 (Step job) 5 (Master worker job) 	jmdl	uint32	JOB MODEL
Compute node type One of the following numerical values is output according to the compute node type: 1 (PRIMERGY server) 2 (FX server)	nodetype	uint32	-
Job execution user name For the step job summary information, the name of the user who submitted the first sub job of the step job is output.	usr	string	USER
Job execution group name For step job summary information, the group name that specified when you submit first sub job of the step job is output.	grp	string	GROUP
Job execution user ID For the step job summary information, the user ID who submitted the first sub job of the step job is output.	uid	uint32	-
Job execution group ID For step job summary information, the group ID that specified when you submit first sub job of the step job is output.	gid	uint32	-
Resource unit name for an executed job If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the Resource unit name that specified when you submit first sub job of the step job is output.	rscu	string	RESOURCE UNIT
Resource group name for an executed job If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the Resource group name that specified when you submit first sub job of the step job is output.	rscg	string	RESOURCE GROUP

Item	Item name	Data type	pjstat/.stats file
Resource unit name for a submitted job For step job summary information, "" is output.	arscu	string	-
Resource group name for a submitted job For step job summary information, "" is output.	arscg	string	-
Job priority in a resource unit The lowest priority is 0, and the highest is 255. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	apriority	int16	APRIORITY
Job priority The lowest priority is 0, and the highest is 255. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	pri	int16	PRIORITY
Absolute path name of the shell used to execute job scripts For step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: "- "	sh	string	SHELL
Whether to send e-mail at each state - pjstat/.stats The following characters are output side by side: b (Send e-mail at start of processing) e (Send e-mail at end of processing) r (Send e-mail at job restart) s or S (Send statistical information by e-mail. Add node information if "S") - (Do not send e-mail) If e-mail is sent in all of the states, "bers" or "berS" is output. If e-mail is sent only at the start of processing and at job restart, "b-r-" is output. For sub jobs of a bulk job, "----" is output. For step job summary information, the mail transmission designation specified when submitting the first sub job of the step job is output. - pmdumpjobinfo The output is a logical OR of the following numerical values: 1 (Processing start) 2 (Processing end) 4 (Job restart) 8 (Statistical information without node information) 16 (Statistical information with node information)	mail	int32	MAIL SEND FLAG
E-mail address The address specified by the --mail-list option of the pjsub command is output.	adr	string	MAIL ADDRES

Item	Item name	Data type	pjstat/.stats file
Step job dependency expression For step job summary information and for jobs that are not step jobs, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: ""	sde	string	STEP DEPENDENCY EXP
umask value of the user submitting the job For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	mask	int32	FILE MASK
Path name of the standard output file of jobs For bulk/step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: ""	std	string	STANDARD OUT FILE
Path name of the standard error output file of jobs For bulk/step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: ""	stde	string	STANDARD ERR FILE
Path name of the .stats file that is output when the pjsub command specifies the -s/-S option	infop	string	INFORMATION FILE
Path name of the job submission directory For step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: "- "	pjsubdir	string	PJSUB DIRECTORY
Character string specified in the --fs option of the pjsub command For step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: ""	fs	string	FILE SYSTEM NAME
Character string specified in the --appname option of the pjsub command For step job summary information, the following is output: - pjstat/.stats: "" - pmdumpjobinfo: ""	appname	string	APPLICATION NAME
Rules on assigning ranks for created processes One of the following numerical values is output according to the specified rules on assigning ranks for created processes: 0 (Rules on assigning ranks for created processes is not specified) 1 (rank-map-bynode) 2 (rank-map-bychip) For the following, "" is output: - Node allocated jobs executed on a PRIMERGY server compute node - Step job summary information	rankm	int32	-
Number of the placement of the ranks for created processes If the rules on assigning ranks for created processes is rank-map-by-chip, the <i>n</i> value specified in the --mpi rank-map-by-chip= <i>n</i> option of the pjsub command is output. For the following, "" is output:	rankmnm	int32	-

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - Jobs executed on an FX server compute node - Node allocated jobs executed on a PRIMERGY server compute node - Step job summary information 			
Virtual node placement policy The following numerical values are output: 0 (PACK) 1 (UNPACK) 2 (Absolutely PACK) 3 (Absolutely UNPACK) For jobs executed on an FX server compute node and for step job summary information, "" is output.	vpol	uint64	-
Number of placed virtual nodes If the virtual node placement policy is UNPACK or Absolutely UNPACK, the <i>n</i> value specified in the -P vn-policy=unpack= <i>n</i> or -P vn-policy=abs-unpack= <i>n</i> option of the pjsub command is output. If no virtual node placement policy is specified, 1 is output. For jobs executed on an FX server compute node and for step job summary information, "" is output.	vnodeargunum	uint64	-
Execution mode policy One of the following numerical values is output according to the execution mode policy: 0 (SHARE) 1 (SIMPLEX) For jobs executed on an FX server compute node, "" is output. For node allocated jobs executed on a PRIMERGY server compute node, "" is output. For step job summary information, "" is output.	epol	uint64	-
NUMA policy One of the following numerical values is output according to the NUMA policy: 0 (PACK) 1 (UNPACK) For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	numapol	uint64	NUMA POLICY
Job execution start time specified in the --at option of the pjsub command <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - - pmdumpjobinfo: Number of seconds since January 1, 1970 For jobs without a job execution start time specified in the --at option of the pjsub command, interactive job and for step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	specifiedsdt	time	SPECIFIED JOB START DATE
Whether the job automatic re-execution are suppressed	autorestart	uint8	AUTO RESTART

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats no (Do not automatically re-execute jobs) yes (Automatically re-execute jobs) - pmdumpjobinfo 0 (Do not automatically re-execute jobs) 1 (Automatically re-execute jobs) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Upper limit on the number of processes generated by an MPI program</p> <p>The number of processes specified in the --mpi proc option of the pjsub command is output. If --mpi proc option is omitted, the number of virtual nodes or nodes is output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	mpiprocr	int32	MPI PROC (REQUIRE)
<p>Upper limit on the number of processes generated by an MPI program</p> <p>The number of processes specified in the --mpi proc option of the pjsub command is output.</p> <p>If --mpi proc option is omitted, the following is</p> <ul style="list-style-type: none"> - For jobs executed on an FX server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of specified by the --mpi shape option. If --mpi max-proc-per-node option is omitted, the number of specified by the --mpi shape option is output. - For jobs executed on an PRIMERGY server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of virtual nodes or nodes. If --mpi maxproc-per-node option is omitted, the number of virtual nodes or nodes is output. <p>For the summary information on a bulk job or a step job, the total of all sub jobs is output.</p>	mpiproca	int32	MPI PROC (ALLOC)
<p>Number of processes (static processes) generated at the time of MPI program startup</p> <p>The number of processes specified in the --mpi proc option of the pjsub command is output. If --mpi proc option is omitted, the number of virtual nodes or nodes is output. For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	mpistaticprocr	int32	MPI STATIC PROC (REQUIRE)
<p>Number of processes (static processes) generated at the time of MPI program startup</p> <p>The number of processes specified in the --mpi proc option of the pjsub command is output.</p> <p>If --mpi proc option is omitted, the following is</p>	mpistaticproca	int32	MPI STATIC PROC (ALLOC)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - For jobs executed on an FX server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of specified by the --mpi shape option. If --mpi max-proc-per-node option is omitted, the number of specified by the --mpi shape option is output. - For jobs executed on an PRIMERGY server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of virtual nodes or nodes. If --mpi maxproc-per-node option is omitted, the number of virtual nodes or nodes is output. <p>For the summary information on a bulk job or a step job, the total of all sub jobs is output.</p>			
<p>CPU usage time limit for each process</p> <p>If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. If there is no limit, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: unlimited - pmdumpjobinfo: 18446744073709551615 (=0xffffffffffffffff) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pcl	uint64	PROC CPU LIMIT
<p>Core file size limit for each process</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pcfl	uint64	PROC COREFILE LIMIT
<p>Limit on the number of processes</p> <p>If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. If there is no limit, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: unlimited - pmdumpjobinfo: 18446744073709551615 (=0xffffffffffffffff) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pcpl	uint64	PROC CREATE PROCESS LIMIT
<p>Data segment limit for each process</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. 	pdl	uint64	PROC DATA LIMIT

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Lock memory size limit for each process</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1 (val2)</i> The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	prml	uint64	PROC LOCKED MEMORY LIMIT
<p>POSIX message queue size limit for each process</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1 (val2)</i> The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pmql	uint64	PROC MESSAGE QUEUE LIMIT
<p>File descriptor limit for each process</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. If there is no limit, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: unlimited - pmdumpjobinfo: 18446744073709551615 (=0xffffffffffffffff) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pofl	uint64	PROC OPEN FILES LIMIT
<p>Limit on the number of signals for each process</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. If there is no limit, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: unlimited - pmdumpjobinfo: 18446744073709551615 (=0xffffffffffffffff) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ppsl	uint64	PROC PENDING SIGNAL LIMIT

Item	Item name	Data type	pjstat/.stats file
Process file size limit <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ppl	uint64	PROC PERMFILE LIMIT
Stack segment limit for each process <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	psl	uint64	PROC STACK LIMIT
Virtual memory size limit for each process <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pvml	uint64	PROC VIRTUAL MEMORY LIMIT
Comment <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "" - pmdumpjobinfo: "" 	cmt	string	COMMENT
Information added by the administrator <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "" - pmdumpjobinfo: "" 	supplementaryinfo	string	SUPPLEMENTARY INFORMATION
Job submission option flag	pjsuboptionflag	uint64	-

Item	Item name	Data type	pjstat/.stats file
<p>The output is a logical OR of the following values according to the option specified in the pjsub command:</p> <p>0x0000000000000001 (--at)</p> <p>0x0000000000000002 (--bulk)</p> <p>0x0000000000000004 (--dir-prefix)</p> <p>0x0000000000000080 (--comment)</p> <p>0x0000000000000200 (-e)</p> <p>0x0000000000000800 (--gid)</p> <p>0x0000000000001000 (--gname)</p> <p>0x0000000000008000 (--interact)</p> <p>0x00000000000010000 (-j)</p> <p>0x00000000000020000 (--rsc-list)</p> <p>0x00000000000040000 (-m)</p> <p>0x00000000000080000 (--mail-list)</p> <p>0x000000000000100000 (--mpi)</p> <p>0x000000000000400000 (--name)</p> <p>0x0000000000001000000 (--norestart)</p> <p>0x0000000000004000000 (-o)</p> <p>0x0000000000002000000 (-p)</p> <p>0x0000000000004000000 (--restart)</p> <p>0x0000000000008000000 (-s)</p> <p>0x00000000000010000000 (-S)</p> <p>0x00000000000040000000 (--sparam)</p> <p>0x00000000000080000000 (--spath)</p> <p>0x000000000000100000000 (--step)</p> <p>0x000020000000000000 (--vset)</p> <p>0x000040000000000000 (-w)</p> <p>0x000080000000000000 (-x)</p> <p>0x000100000000000000 (-X)</p> <p>0x002000000000000000 (--reason)</p> <p>0x004000000000000000 (--fs)</p> <p>0x008000000000000000 (--appname)</p> <p>0x010000000000000000 (-P--vn-policy)</p> <p>0x020000000000000000 (-P--exec-policy)</p> <p>0x040000000000000000 (-P)</p> <p>For step job summary information, "" is output.</p>			
<p>Extended job submission option flag</p> <p>The output is a logical OR of the following value according to the option specified in the pjsub command:</p> <p>0x0000000000000001 (--net-route dynamic)</p> <p>0x0000000000000002 (--llo)</p> <p>For step job summary information, "" is output.</p>	pjsuboptionexflag	uint64	-
<p>-L option argument flag of the pjsub command</p> <p>The output is a logical OR of the following values according to the argument specified in the -L option of the pjsub command:</p> <p>0x0000000000000001 (node)</p> <p>0x0000000000000004 (elapsed)</p> <p>0x0000000000000008 (node-mem)</p> <p>0x0000000000000020 (rscunit)</p> <p>0x0000000000000040 (rscgrp)</p> <p>0x0000000000000100 (proc-core)</p> <p>0x0000000000000200 (proc-cpu)</p> <p>0x0000000000000400 (proc-crproc)</p> <p>0x0000000000000800 (proc-data)</p> <p>0x0000000000001000 (proc-lockm)</p>	largflag	uint64	-

Item	Item name	Data type	pjstat/.stats file
0x0000000000002000 (proc-msgq) 0x0000000000004000 (proc-openfd) 0x0000000000008000 (proc-psig) 0x0000000000010000 (proc-filesz) 0x0000000000020000 (proc-stack) 0x0000000000040000 (proc-vmem) 0x0000000000080000 (vnode) 0x0000000000100000 (vnode-core) 0x0000000000200000 (core-mem) 0x0000000000400000 (vnode-mem) 0x0000000000800000 (custom resource) For step job summary information, "" is output.			
--mpi option argument flag of the pjsb command The output is a logical OR of the following values according to the argument specified in the --mpi option of the pjsb command: 0x0000000000000001 (shape) 0x0000000000000002 (proc) 0x0000000000000004 (rank-map-bynode) 0x0000000000000008 (rank-map-bychip) 0x0000000000000010 (rank-map-hostfile) 0x0000000000000020 (assign-online-node) For step job summary information, "" is output.	mpiargflag	uint64	-
--step option argument flag of the pjsb command The output is a logical OR of the following values according to the argument specified in the --step option of the pjsb command: 0x0000000000000001 (jid) 0x0000000000000002 (sd) 0x0000000000000004 (sn) 0x0000000000000010 (jnam) For step job summary information, "" is output.	stepargflag	uint64	-
-P option argument flag of the pjsb command The output is a logical OR of the following values according to the argument specified in the -P option of the pjsb command: 0x0000000000000001 (vn-policy) 0x0000000000000002 (exec-policy) For step job summary information, "" is output.	pargflag	uint64	-
--llo option argument flag of the pjsb command The output is a logical OR of the following values according to the argument specified in the --llo option of the pjsb command: 0x0000000000000001 (sharedtmp-size) 0x0000000000000002 (localtmp-size) 0x0000000000000004 (cn-cached-write-size) 0x0000000000000008 (cn-cache-size) 0x0000000000000010 (stripe-count) 0x0000000000000020 (stripe-size) 0x0000000000000040 (auto-readahead) 0x0000000000000080 (async-close) 0x0000000000000100 (cn-read-cache) 0x0000000000000200 (sio-read-cache) 0x0000000000000400 (uncompleted-fileinfo-path) 0x0000000000000800 (perf) 0x0000000000001000 (perf-path) For step job summary information, "" is output.	lloargflag	uint64	-

Item	Item name	Data type	pjstat/.stats file
Number of compute cluster management node or PJM service starts The output number of restarts from waiting for execution to exit of job for compute cluster management node or PJM service. The count is reset to 1 when the number exceeds 255.	servicestartflag	uint8	-
Number of job retries The number of times that the job manager function requests the job resource management function for the execution of jobs (The first request is not counted). For bulk/step job summary information, the total for sub jobs is output.	rnum	uint16	RETRY NUM
Number of job re-execution times The number of times that the job was re-executed is output. For bulk/step job summary information, the total for sub jobs is output.	rerunnum	uint32	-
State before a job transition to the state shown in the STATE item <ul style="list-style-type: none"> - pjstat/.stats The following character strings are output: ACC (ACCEPT: Accepted job submission) ERR (ERROR: Fixed state due to error) HLD (HOLD: Fixed state due to user) QUE (QUEUED: Waiting for job execution) RNA (RUNNING-A: Acquiring resources required for job execution) RNE (RUNNING-E: Processing epilogue) RNO (RUNOUT: Waiting for completion of job end processing) RNP (RUNNING-P: Processing prologue) RUN (RUNNING: Executing job) SPP (SUSPEND: Suspending) SPD (SUSPENDED: Already suspended) RSM (RESUME: Resuming) - pmdumpjobinfo The following numerical values are output: 1 (ACCEPT: Accepted job submission) 2 (QUEUED: Waiting for job execution) 5 (RUNNING-A: Acquiring resources required for job execution) 6 (RUNNING: Executing job) 7 (RUNOUT: Waiting for completion of job end processing) 12 (HOLD: Fixed state due to user) 13 (ERROR: Fixed state due to error) 17 (RUNNING-P: Processing prologue) 19 (SUSPEND: Suspending) 20 (SUSPENDED: Already suspended) 21 (RESUME: Resuming) For step job summary information, the following is output: - pjstat: If there are no sub jobs being executed or no sub jobs being executed, the state before transition of the sub job to be executed next will be output. "-" is output for summary information of completed step jobs.	lst	int16	LAST STATE

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - .stats: "-" - pmdumpjobinfo: "" 			
Job state <ul style="list-style-type: none"> - pjstat/.stats The following character strings are output: ACC (ACCEPT: Accepted job submission) CCL (CANCEL: Terminated due to job execution cancellation) ERR (ERROR: Fixed state due to error) EXT (EXIT: Completed job end processing) HLD (HOLD: Fixed state due to user) QUE (QUEUED: Waiting for job execution) RNA (RUNNING-A: Acquiring resources required for job execution) RNE (RUNNING-E: Processing epilogue) RNO (RUNOUT: Waiting for completion of job end processing) RNP (RUNNING-P: Processing prologue) RUN (RUNNING: Executing job) SPP (SUSPEND: Suspending) SPD (SUSPENDED: Already suspended) RSM (RESUME: Resuming) - pmdumpjobinfo output The following numerical values are output: 2 (QUEUED: Waiting for job execution) 9 (EXIT: Completed job end processing) 11 (CANCEL: Terminated due to job execution cancellation) 12 (HOLD: Fixed state) 13 (ERROR: Fixed state due to error) <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat: If there is no sub job being executed or no sub job being executed, the status of the sub job to be executed next will be output. "-" is output for summary information of completed step jobs. - .stats: "-" - pmdumpjobinfo: "" 	st	int16	STATE
Whether a backfill concerning job scheduling occurred <p>The following numerical values are output:</p> <p>0 (No backfill occurred) 1 (Backfill occurred)</p> <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	backfill	uint8	BACKFILL
Job submission time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For step job summary information, the acceptance time of the first submitted sub job is output.</p>	adt	time	ACCEPT DATE
Last queuing time	qdt	time	QUEUED DATE

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
Scheduling start time The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	schedsdt	timespec	-
Transition time to the RUNNING-A state The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	radt	time	-
Prologue execution start time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	psdt	timespec	PRO START DATE
Transition time to the RUNNING-P state The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	rpdt	time	-
Prologue execution end time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1,1970 <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	pedt	timespec	PRO END DATE
Job execution start time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: Sub job start time - pmdumpjobinfo: "-" 	sdt	timespec	JOB START DATE
Transition time to the RUNNING state	rdt	time	-

Item	Item name	Data type	pjstat/.stats file
The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.			
Job execution end time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	edt	timespec	JOB END DATE
Epilogue execution start time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	esdt	timespec	EPI START DATE
Transition time to the RUNNING-E state The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	redt	time	-
Epilogue execution end time <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	eedt	timespec	EPI END DATE
Transition time to the RUNOUT state The number of seconds since January 1 is output. For bulk/step job summary information, "" is output.	rodt	time	-
Transition time to the EXIT or CANCEL state <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	exc	time	EXIT DATE
Transition time to the HOLD or ERROR state The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	hedt	time	-
Initial job execution start time	fsdt	time	-

Item	Item name	Data type	pjstat/.stats file
Time when the job started for the first time. The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.			
Job deletion request time The time of receiving the job deletion request by the pjdel command is output. <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	deldtr	time	JOB DELETE DATE (REQUIRE)
Job deletion time The time when the job was deleted from the compute node is output by job deletion request. <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	deldt	time	JOB DELETE DATE
Preceding sub job end time Only for the sub jobs preceded by a sub job in a step job, the following values are output: <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD HH:MM:SS</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>HH</i>: Hour, <i>MM</i>: Minute, <i>SS</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For the other sub jobs of the step job, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	apsjedt	time	ALL PREC SUBJOB EXIT DATE
Cumulative RUNNING-A time (seconds) The output cumulative value is the amount of time that a job is in the RUNNING-A state before entering another state. The value of the previous time is added when the job is re-executed. For bulk/step job summary information, the total for sub jobs is output.	tratm	uint64	-
Cumulative RUNNING time (seconds) The output cumulative value is the amount of time that a job is in the RUNNING state before entering another state. The value of the previous time is added when the job is re-executed. For bulk/step job summary information, the total for sub jobs is output.	trtm	uint64	-
Cumulative wait time (seconds)	twtm	uint64	-

Item	Item name	Data type	pjstat/.stats file
<p>The output cumulative value is the amount of time that a job is in the QUEUED state. The value of the previous time is added when the job is re-executed.</p> <p>For bulk/step job summary information, the total for sub jobs is output.</p>			
<p>Cumulative hold time (seconds)</p> <p>The output cumulative value is the amount of time that a job is in the HOLD state before being released.</p> <p>For bulk/step job summary information, the total for sub jobs is output.</p>	thldtm	uint64	HOLD TIME
<p>Name of the user who placed this job in the HOLD state</p> <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	lhusr	string	LAST HOLD USER
<p>Number of times that the job entered the HOLD state</p> <p>For bulk/step job summary information, the total for sub jobs is output.</p>	holnm	uint32	HOLD NUM
<p>Execution trigger of the job resource manager exit script prealloc</p> <p>The following values are output:</p> <ul style="list-style-type: none"> 0 (Not executed) 1 (Job start) 15 (pmresume command execution) <p>For bulk/step job summary information, "" is output.</p>	prealloctrigger	uint8	-
<p>Exit code of the job resource manager exit script prealloc</p> <p>The following values are output:</p> <ul style="list-style-type: none"> 0 (Normal end) 1 (Instruction to place job in error state) 2 (Instruction to execute job again) 3 (Instruction to place job in HOLD state) 4 (Instruction to delete job) <p>If an error occurred, one of the following values is output:</p> <ul style="list-style-type: none"> 102 (Execution failure of job resource manager exit script) 255 (Error other than above) <p>For bulk/step job summary information, "" is output.</p>	preallocec	uint32	-
<p>Start time of the job resource manager exit script prealloc</p> <p>The number of seconds since January 1, 1970 is output.</p> <p>For bulk/step job summary information, "" is output.</p>	preallocsdt	timespec	-
<p>End time of the job resource manager exit script prealloc</p> <p>The number of seconds since January 1, 1970 is output.</p> <p>For bulk/step job summary information, "" is output.</p>	preallocedt	timespec	-
<p>Execution trigger of the job resource manager exit script predel</p> <p>The following values are output:</p> <ul style="list-style-type: none"> 0 (Not executed) 3 (pjdel command execution) 4 (pjhold command execution) 5 (Deletion request from job manager or job scheduler function) 6 (Compute node error) 8 (Excessive time elapsed) 9 (Excessive memory usage) 	predeltrigger	uint8	-

Item	Item name	Data type	pjstat/.stats file
13 (OOM Killer operation) 16 (Error detection trigger in job execution environment [McKernel only]) For bulk/step job summary information, "" is output.			
Exit code of the job resource manager exit script predel The following values are output: 0 (Normal end) 1 (Instruction to place job in error state) 2 (Instruction to execute job again) 3 (Instruction to place job in HOLD state) 4 (Instruction to delete job) If an error occurred, one of the following values is output: 102 (Execution failure of job resource manager exit script) 255 (Error other than above) For bulk/step job summary information, "" is output.	predelec	uint32	-
Start time of the job resource manager exit script predel The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	predelsdt	timespec	-
End time of the job resource manager exit script predel The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	predeledt	timespec	-
Execution trigger of the job resource manager exit script postfree The following values are output: 0 (Not executed) 2 (Job end) 14 (pmsuspend command execution) For bulk/step job summary information, "" is output.	postfreetrigger	uint8	-
Exit code of the job resource manager exit script postfree The following values are output: 0 (Normal end) 1 (Instruction to place job in error state) 2 (Instruction to execute job again) 3 (Instruction to place job in HOLD state) 4 (Instruction to delete job) If an error occurred, one of the following values is output: 102 (Execution failure of job resource manager exit script) 255 (Error other than above) For bulk/step job summary information, "" is output.	postfreeec	uint32	-
Start time of the job resource manager exit script postfree The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	postfreesdt	timespec	-
End time of the job resource manager exit script postfree The number of seconds since January 1, 1970 is output. For bulk/step job summary information, "" is output.	postfreeedt	timespec	-
Time when the job resource manager exit function collected data from each node - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i> : Year, <i>MM</i> : Month, <i>DD</i> : Day, <i>hh</i> : Hour, <i>mm</i> : Minute, <i>ss</i> : Second)	prmdt	timespec	PRM DATE

Item	Item name	Data type	pjstat/.stats file
<p>- pmdumpjobinfo: Number of seconds since January 1, 1970</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>If the occurrence of a communication path ICC error or node going down affected the job results, the node ID of that node is output.</p> <p>For cases where the node had no effect and for bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	affectednid	uint32	AFFECTED NODE ID
<p>Job script exit code</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	ec	int32	EXIT CODE
<p>Signal number when the job script ended with the receipt of a signal</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	sn	int32	SIGNAL NO
<p>Job end code (PJM code) indicating the job manager processing result from job execution</p> <p>The following values are output:</p> <p>0 (Normal job end)</p> <p>1 (CANCEL due to pjdel command)</p> <p>2 (REJECT due to refusal to accept job)</p> <p>3 (Execution rejected by gate check)</p> <p>4 (HOLD due to pjhold command)</p> <p>6 (CANCEL due to evaluation of step job dependency expression)</p> <p>7 (CANCEL due to forced specification of deadline schedule)</p> <p>8 (CANCEL due to gate check)</p> <p>9 (EXIT when job rebuilt because "not re-executable" specified)</p> <p>11 (Job execution timeout due to elapsed time limit violation)</p> <p>12 (Forced termination due to memory usage limit violation)</p> <p>16 (Termination due to no accessibility to current directory or standard input/standard output/standard error output file)</p> <p>18 (Job end due to execution of subsequent job or start of deadline schedule, when job has been running for longer than minimum executable time)</p> <p>20 (Node down)</p> <p>21 (Shell execution failure)</p> <p>22 (ICC error)</p> <p>23 (Termination due to OOM Killer operation)</p> <p>25 (HA failure)</p> <p>26 (Error in prologue or epilogue process)</p> <p>27 (Error in job resource manager exit process)</p> <p>28 (Job execution environment error)</p> <p>29 (Invalid job execution environment specified)</p> <p>30 (Abort due to suspend or resume process failure)</p> <p>100 (Internal error in job manager)</p> <p>120 (Internal error in job scheduler)</p> <p>140 (Internal error in job resource management)</p> <p>160 (Internal error in Tofu library)</p> <p>180 (Internal error in layered storage system)</p>	pc	uint32	PJM CODE

Item	Item name	Data type	pjstat/.stats file
For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""			
Message corresponding to job processing results The meanings of output messages are as follows. ANOTHER JOB STARTED Already running for longer than the minimum time for the elapsed time limit, the job ended due to the start of another job. <i>MM/DD hh:mm DELAY</i> The scheduled execution start time of the job is later than the specified execution start time. DEADLINE SCHEDULE STARTED Already running for longer than the minimum time for the elapsed time limit, the job ended due to the start of the deadline schedule. ELAPSE LIMIT EXCEEDED The elapsed time exceeded the limit. FILE IO ERROR The current directory at the user job submission time was not accessible. GATE CHECK The job manager exit function canceled the job. IMPOSSIBLE SCHED Scheduling could not be done. INSUFF CPU The number of physical CPUs is insufficient. INSUFF MEMORY The amount of physical memory is insufficient. INSUFF NODE The number of physical nodes is insufficient. INSUFF <i>customrsc</i> The custom resource defined with the resource name <i>customrsc</i> is insufficient. INTERNAL ERROR Internal error. INVALID HOSTFILE The host file specified in the rank-map-hostfile parameter of the pjsub command is invalid. LIMIT OVER MEMORY The memory limit was exceeded during job execution. LOST COMM All-to-all communication among parallel processes is not guaranteed. NO CURRENT DIR The current directory or standard input/standard output/standard error output file at the user job submission time is not accessible. NOT EXIST <i>customrsc</i> No custom resource is defined with the resource name <i>customrsc</i> .	ermsg	string	REASON

Item	Item name	Data type	pjstat/.stats file
<p>RESUME FAIL The job failed to resume.</p> <p>RSCGRP NOT EXIST The resource group does not exist, or the number of nodes allocated to the resource group is 0.</p> <p>RSCGRP STOP The resource group is stopped.</p> <p>RSCUNIT NOT EXIST The resource unit does not exist.</p> <p>RSCUNIT STOP The resource unit is stopped.</p> <p>RUNLIMIT EXCEED The number of simultaneous job executions exceeded the limit.</p> <p>SUSPEND FAIL The job failed to suspend.</p> <p>USELIMIT EXCEED The job manager is waiting for execution because of the limit on simultaneous node or CPU core usage.</p> <p>USER NOT EXIST The job execution user does not exist in the system.</p> <p>WAIT SCHED The job is exempt from scheduling because the limit on the number of jobs subject to scheduling has been reached.</p> <p>Other character strings Message specified in the --reason option of the pjdel, pjhold, or pmsuspend command. Or a message set by the administrator on the job manager exit function, the job scheduler exit function, or the job resource management exit function. The display format for any message in the --reason option of the pjhold or pmsuspend command is "<i>command execution user name:message</i>". If the --reason option is not specified, "<i>command execution user name:</i>" is displayed.) For bulk/step job summary information, "" is output:</p>			
<p>Prologue script exit code The following values are output: 0 (Normal end) 1 (Instruction to place job in error state) 2 (Instruction to execute job again) 3 (Instruction to place job in HOLD state) 4 (Instruction to delete job) 101 (Forced termination of prologue process because of cancellation or hold of job) 102 (Prologue script execution failure) 103 (Forced termination of prologue process because of OOM Killer operation) 104 (Forced termination of prologue process because of hardware error) 105 (Forced termination of prologue process because node went down) 106 (Forced termination of prologue process because of memory amount limit violation)</p>	pec	uint32	PRO EXIT CODE

Item	Item name	Data type	pjstat/.stats file
108 (Forced termination of prologue process because of elapsed time limit violation) 109 (Forced termination of prologue process because of current directory access failure) 110 (Forced termination due to job execution environment error) 255 (Error other than above) For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""			
Epilogue script exit code The following values are output: 0 (Normal end) 1 (Instruction to place job in error state) 2 (Instruction to execute job again) 3 (Instruction to place job in HOLD state) 4 (Instruction to delete job) 101 (Forced termination of epilogue process because of cancellation or hold of job) 102 (Epilogue script execution failure) 103 (Forced termination of epilogue process because of OOM Killer operation) 104 (Forced termination of epilogue process because of hardware error) 105 (Forced termination of epilogue process because node went down) 106 (Forced termination of epilogue process because of memory amount limit violation) 108 (Forced termination of epilogue process because of elapsed time limit violation) 109 (Forced termination of epilogue process because of current directory access failure) 110 (Forced termination due to job execution environment error) 255 (Error other than above) For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	eec	uint32	EPI EXIT CODE
Total node down time (seconds) The total time that a compute node was down in a master-worker job is output. For bulk/step job summary information, the total for sub jobs is output.	tndt	uint64	NODE DOWN TIME (TOTAL)
Starting time (seconds) of the job execution environment The time taken for the start processing of the job execution environment is output. The value is the total of Node/Virtual node statistical information(jebt). For bulk/step job summary information, the total for sub jobs is output.	jebt	uint64	-
Ending time (seconds) of the job execution environment The time taken for the end processing of the job execution environment is output. The value is the total of Node/Virtual node statistical information(jest).	jest	uint64	-

Item	Item name	Data type	pjstat/.stats file
For bulk/step job summary information, the total for sub jobs is output.			
Start time of the last suspend <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For jobs that were never suspended and for bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	suspendd	time	SUSPEND DATE
Start time of the last resume <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For jobs that were never resumed and for bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	resume	time	RESUME DATE
Cumulative suspend time This is the cumulative time of the periods when the job was suspended. <i>DD HH.MM.SS (s)</i> (<i>DD</i> : Day, <i>HH</i> : Hour, <i>MM</i> : Minute, <i>SS</i> : Second. <i>s</i> : Expressed as number of seconds) <i>DD</i> is omitted when the time is 24 hours or less. If the job has been re-executed, this is the cumulative time since the beginning of the re-execution. For jobs without a suspend period, 0 is output. For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	suspend	time	SUSPEND TIME
Number of suspends If the job has been re-executed, this is the cumulative count since the beginning of the re-execution. For jobs that were never suspended, 0 is output. For bulk/step job summary information, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	suspendnum	int32	SUSPEND NUM
Specification method for elapsed time The following values are output: 1: Specify the elapsed time limit (pjsub -L elapse= <i>limit</i>) 2: Specify the elapsed time limit as a range (pjsub -L elapse= <i>min-max</i>) For step job summary information, "" is output.	elptm	uint8	-
Elapsed time limit specified at job submission	elpl	uint64	ELAPSE TIME (LIMIT)

Item	Item name	Data type	pjstat/.stats file
<p>If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output.</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>DD hh:mm:ss (s)</i> (<i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second. <i>s</i>: Expressed as number of seconds) <i>DD</i> is omitted when the time is 24 hours or less. If the elapsed time limit is specified as a range, the output has the <i>min-max</i> (<i>min-max</i>) format. (The value in parentheses expresses the limit as a number of seconds.) - pmdumpjobinfo: Value in units of seconds 18446744073709551615 (=0xffffffffffffff) means no limit (unlimited). For jobs where the elapsed time limit is specified as a range (pjsub -L elapse=<i>min-max</i>), the upper limit value (max) of elapsed time limit value is output in seconds. <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Upper limit value of the elapsed time limit specified as a range</p> <p>For jobs where the elapsed time limit is specified as a range (pjsub -L elapse=<i>min-max</i>), the upper limit value (<i>max</i>) of elapsed time limit value is output in seconds. 18446744073709551615 (=0xffffffffffffff) means no limit (unlimited). For jobs where the elapsed time limit is not specified as a range (pjsub -L elapse=<i>limit</i>), pmdumpjobinfo outputs "". For step job summary information, pmdumpjobinfo outputs "".</p>	maxelpl	uint64	-
<p>Lower limit value of the elapsed time limit specified as a range</p> <p>For jobs where the elapsed time limit is specified as a range (pjsub -L elapse=<i>min-max</i>), the lower limit value (<i>min</i>) of elapsed time limit value is output in seconds. 18446744073709551615 (=0xffffffffffffff) means no limit (unlimited). For jobs where the elapsed time limit is not specified as a range (pjsub -L elapse=<i>limit</i>), pmdumpjobinfo outputs "". For step job summary information, pmdumpjobinfo outputs "".</p>	minelpl	uint64	-
<p>Elapsed time from job execution start to end</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>DD hh:mm:ss (s)</i> (<i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second. <i>s</i>: Expressed as number of seconds) <i>DD</i> is omitted when the time is 24 hours or less. Values of less than a second are rounded up. - pmdumpjobinfo: Value in units of seconds <p>For bulk/step job summary information, the total for sub jobs is output.</p>	elp	timespec	ELAPSE TIME (USE)
<p>Total user CPU time (milliseconds) spent on job execution</p> <p>The value is the total of Node/Virtual node statistical information(uctmut). For bulk/step job summary information, the total for sub jobs is output.</p>	uctmut	uint64	USER CPU TIME (USE)

Item	Item name	Data type	pjstat/.stats file
Total system CPU time (milliseconds) spent on job The value is the total of Node/Virtual node statistical information(sctmut). For bulk/step job summary information, the total for sub jobs is output.	sctmut	uint64	SYSTEM CPU TIME (USE)
Total user CPU time and system CPU time (milliseconds) The value is the total of Node/Virtual node statistical information(usctmut). For bulk/step job summary information, the total for sub jobs is output.	usctmut	uint64	CPU TIME (TOTAL)
Physical memory amount limit for each node <ul style="list-style-type: none"> - pjstat/.stats: <i>val1 (val2)</i> The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	mszl	uint64	MEMORY SIZE (LIMIT)
Allocated physical memory amount for each node <ul style="list-style-type: none"> - pjstat/.stats: <i>val1 (val2)</i> The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	msza	uint64	MEMORY SIZE (ALLOC)
Maximum physical memory usage for each node Outputs the amount of memory used by the node that uses the most physical memory. The value is the maximum of Node/Virtual node statistical information(mmszu). <ul style="list-style-type: none"> - pjstat/.stats: <i>val1 (val2)</i> The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes For bulk job summary information, the maximum value of the sub job is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	mmszu	uint64	MAX MEMORY SIZE (USE)

Item	Item name	Data type	pjstat/.stats file
Physical memory amount limit for each virtual node <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffff) means no limit (unlimited). <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	vmszl	uint64	VNODE MEMORY SIZE (LIMIT)
Allocated physical memory amount for each virtual node <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffff) means no limit (unlimited). <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	vmsza	uint64	VNODE MEMORY SIZE (ALLOC)
Maximum physical memory usage for each virtual node <p>Outputs the amount of memory used by the virtual node that uses the most physical memory.</p> <p>The value is the maximum of Node/Virtual node statistical information(mmszu).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the maximum value of the sub job is output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	vmszu	uint64	VNODE MAX MEMORY SIZE (USE)
Number of cores for each virtual node <p>For step job summary information, pmdumpjobinfo outputs "".</p>	vcnum	uint32	-
Number of requested CPU cores <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	cnumr	uint32	CPU NUM (REQUIRE)
Number of CPU cores actually allocated	cnumat	uint32	CPU NUM (ALLOC)

Item	Item name	Data type	pjstat/.stats file
<p>The value is the total of Node/Virtual node statistical information(cnuma).</p> <p>For bulk/step job summary information, the total for sub jobs is output.</p> <p>For step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>Number of CPU cores actually used by the job</p> <p>The value is the total of Node/Virtual node statistical information(cnumu).</p> <p>For bulk job summary information, the total for sub jobs is output</p> <p>For step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	cnumut	uint32	CPU NUM (USE)
<p>Number of assistant cores actually allocated to the job</p> <p>The value is the total of Node/Virtual node statistical information(anuma).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p> <p>This item is not output by default. To output the item, the administrator needs to configure it.</p>	anuma	uint32	ASSISTANT CORE NUM (ALLOC)
<p>Number of assistant cores used by the job</p> <p>The value is the total of Node/Virtual node statistical information(acnumut).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p> <p>This item is not output by default. To output the item, the administrator needs to configure it.</p>	acnumut	uint32	ASSISTANT CORE NUM (USE)
<p>Total user CPU time (milliseconds) for assistant cores</p> <p>The value is the total of Node/Virtual node statistical information(auctmut).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p> <p>This item is not output by default. To output the item, the administrator needs to configure it.</p>	auctmut	uint64	ASSISTANT CORE USER CPU TIME (USE)
<p>Total system CPU time (milliseconds) for assistant cores</p> <p>The value is the total of Node/Virtual node statistical information(asctmut).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p>	asctmut	uint64	ASSISTANT CORE SYSTEM CPU TIME (USE)

Item	Item name	Data type	pjstat/.stats file
<p>- pmdumpjobinfo: ""</p> <p>This item is not output by default. To output the item, the administrator needs to configure it.</p>			
<p>Maximum memory usage by assistant cores</p> <p>The value is the total of Node/Virtual node statistical information(ammszu).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2²⁰ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the maximum value of the sub job is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" <p>This item is not output by default. To output the item, the administrator needs to configure it.</p> <p>This item can be acquired only in McKernel mode.</p>	ammszu	uint64	ASSISTANT CORE MAX MEMORY SIZE (USE)
<p>Receive data size for Tofu user communication</p> <p>This is the receive data size for user-level communication via Tofu. The value is the total of Node/Virtual node statistical information(tucrb).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	tucrb	int64	TOFU USER COMM RECV BYTE
<p>Send data size for Tofu user communication</p> <p>This is the send data size for user-level communication via Tofu. The value is the total of Node/Virtual node statistical information(tucsb).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	tucsb	int64	TOFU USER COMM SEND BYTE
<p>Receive data size for Tofu system communication</p>	tscrb	int64	TOFU SYSTEM COMM RECV BYTE

Item	Item name	Data type	pjstat/.stats file
<p>This is the receive data size for system communication via Tofu. The value is the total of Node/Virtual node statistical information(tscrb).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Send data size for Tofu system communication</p> <p>This is the send data size for system communication via Tofu. The value is the total of Node/Virtual node statistical information(tscsb).</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	tscsb	int64	TOFU SYSTEM COMM SEND BYTE
<p>Number of Fujitsu profiler use times</p> <p>The value is the total of Node/Virtual node statistical information(fjprofiler).</p> <p>For bulk/step job summary information, the total value for sub jobs is output.</p>	fjprofiler	int64	FJ PROFILER
<p>Number of starts of a program using the sector cache driver/library</p> <p>The value is the total of Node/Virtual node statistical information(sectocache).</p> <p>For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	sectocache	int64	SECTOR CACHE
<p>Number of starts of a program using the inter-core hardware barrier driver/library</p> <p>The value is the total of Node/Virtual node statistical information(intranodebarrier).</p> <p>For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	intranodebarrier	int64	INTRA NODE BARRIER
CMG numbers used by the job	cmgno	int32_array	-

Item	Item name	Data type	pjstat/.stats file
CMG numbers are output, separated by spaces. For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, "" is output.			
Average power consumption of the compute core group for each CMG (Estimate) The power consumption of the compute core group for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(avgpccoc). <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpccoc	double_array	AVG POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)
Maximum power consumption of the compute core group for each CMG (Estimate) From the periodically collected power consumption (W) of the compute core group for each CMG, the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxpccoc). <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpccoc	double_array	MAX POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)
Minimum power consumption of the compute core group for each CMG (Estimate) From the periodically collected power consumption (W) of the compute core group for each CMG, the minimum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(minpccoc). <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. 	minpccoc	double_array	MIN POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Power consumption of the compute core group for each CMG (Estimate)</p> <p>The output for the power consumption (Wh) of the compute core group for each CMG from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(ecocc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecocc	double_array	ENERGY CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)
<p>Average power consumption of the L2 cache for each CMG (Estimate)</p> <p>The power consumption of the L2 cache for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avgpcolc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avgpcolc	double_array	AVG POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Maximum power consumption of the L2 cache for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the L2 cache for each CMG, the maximum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(maxpcolc).</p>	maxpcolc	double_array	MAX POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats: This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Minimum power consumption of the L2 cache for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the L2 cache for each CMG, the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcolc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minpcolc	double_array	MIN POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Power consumption of the L2 cache for each CMG (Estimate)</p> <p>The output for the power consumption (Wh) of the L2 cache for each CMG from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(ecolc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecolc	double_array	ENERGY CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Average power consumption of the memory for each CMG (Estimate)</p> <p>The power consumption of the memory for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p>	avgpcomc	double_array	AVG POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<p>The value is the total of Node/Virtual node statistical information(avgpcomc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Maximum power consumption of the memory for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the memory for each CMG, the maximum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(maxpcomc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxpcomc	double_array	MAX POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)
<p>Minimum power consumption of the memory for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the memory for each CMG, the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcomc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minpcomc	double_array	MIN POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)
<p>Memory power consumption for each CMG (Estimate)</p> <p>The output for the memory power consumption (Wh) foreach CMG from job start to job end is shown to up to six decimal places.</p>	ecomc	double_array	ENERGY CONSUMPTION OF

Item	Item name	Data type	pjstat/.stats file
<p>The value is the total of Node/Virtual node statistical information(ecomc).</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			MEM/CMG(<i>num</i>) (IDEAL)
<p>Average Tofu power consumption (Estimate)</p> <p>Tofu power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avGPCOT).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avGPCOT	double	AVG POWER CONSUMPTION OF TOFU (IDEAL)
<p>Maximum Tofu power consumption (Estimate)</p> <p>From the periodically collected Tofu power consumption (W), the maximum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(maxPCOT).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxPCOT	double	MAX POWER CONSUMPTION OF TOFU (IDEAL)
<p>Minimum Tofu power consumption (Estimate)</p> <p>From the periodically collected Tofu power consumption (W), the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minPCOT).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minPCOT	double	MIN POWER CONSUMPTION OF TOFU (IDEAL)
<p>Tofu power consumption (Estimate)</p> <p>The output for the Tofu power consumption (Wh) from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(ecot).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p>	ecot	double	ENERGY CONSUMPTION OF TOFU (IDEAL)

Item	Item name	Data type	pjstat/.stats file
- pjstat/.stats: "-" - pmdumpjobinfo: ""			
Average power consumption at the periphery inside the CPU (Estimate) The power consumption at the periphery inside the CPU from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(avGPCOP). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avGPCOP	double	AVG POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Maximum power consumption at the periphery inside the CPU (Estimate) From the periodically collected power consumption (W) at the periphery inside the CPU, the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxPCOP). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxPCOP	double	MAX POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Minimum power consumption at the periphery inside the CPU (Estimate) From the periodically collected power consumption (W) at the periphery inside the CPU, the minimum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(minPCOP). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minPCOP	double	MIN POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Power consumption at the periphery inside the CPU (Estimate) The output for the power consumption (Wh) at the periphery inside the CPU from job start to job end is shown to up to six decimal places. The value is the total of Node/Virtual node statistical information(ecOP). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecOP	double	ENERGY CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Average power consumption of the optical module (Estimate) The power consumption of the optical module from job start to job end is divided by the job execution time. The resulting value (W) of	avGPCOM	double	AVG POWER CONSUMPTION OF

Item	Item name	Data type	pjstat/.stats file
<p>up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avgpcoom).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			OPTICAL MODULE (IDEAL)
<p>Maximum power consumption of the optical module (Estimate)</p> <p>From the periodically collected power consumption (W) of the optical module, the maximum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(maxpcoom).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	maxpcoom	double	MAX POWER CONSUMPTION OF OPTICAL MODULE (IDEAL)
<p>Minimum power consumption of the optical module (Estimate)</p> <p>From the periodically collected power consumption (W) of the optical module, the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcoom).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	minpcoom	double	MIN POWER CONSUMPTION OF OPTICAL MODULE (IDEAL)
<p>Power consumption of the optical module (Estimate)</p> <p>The output for the power consumption (Wh) of the optical module from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(ecoom).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	ecoom	double	ENERGY CONSUMPTION OF OPTICAL MODULE (IDEAL)
<p>Average PCI-E power consumption (Estimate)</p> <p>The PCI-E power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avgpcop).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	avgpcop	double	AVG POWER CONSUMPTION OF PCI-E (IDEAL)

Item	Item name	Data type	pjstat/.stats file
Maximum PCI-E power consumption (Estimate) From the periodically collected PCI-E power consumption (W), the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxpcop). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcop	double	MAX POWER CONSUMPTION OF PCI-E (IDEAL)
Minimum PCI-E power consumption (Estimate) From the periodically collected PCI-E power consumption (W), the minimum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(minpcop). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minpcop	double	MIN POWER CONSUMPTION OF PCI-E (IDEAL)
PCI-E power consumption (Estimate) The output for the PCI-E power consumption (Wh) from job start to job end is shown to up to six decimal places. The value is the total of Node/Virtual node statistical information(ecop). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecop	double	ENERGY CONSUMPTION OF PCI-E (IDEAL)
Average node power consumption (Estimate) The node power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(avgpcon). This value does not take into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value is the same when same job is executed on different nodes. There may be a difference of +/- 40% from average node power consumption (actual measurement). For bulk job summary information, the total for sub jobs is output. For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpcon	double	AVG POWER CONSUMPTION OF NODE (IDEAL)
Maximum node power consumption (Estimate) From the periodically collected node power consumption (W), the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxpcon).	maxpcon	double	MAX POWER CONSUMPTION OF NODE (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<p>This value does not take into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value is the same when same job is executed on different nodes. There may be a difference of +/- 40% from maximum node power consumption (actual measurement).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>Minimum node power consumption (Estimate)</p> <p>From the periodically collected node power consumption (W), the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcon).</p> <p>This value does not take into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value is the same when same job is executed on different nodes. There may be a difference of +/- 40% from minimum node power consumption (actual measurement).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	minpcon	double	MIN POWER CONSUMPTION OF NODE (IDEAL)
<p>Node power consumption (Estimate)</p> <p>The output for the node power consumption (Wh) from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(econ).</p> <p>This value does not take into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value is the same when same job is executed on different nodes. There may be a difference of +/- 40% from node power consumption (actual measurement).</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	econ	double	ENERGY CONSUMPTION OF NODE (IDEAL)
<p>Average node power consumption (actual measurement)</p> <p>The node power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avgpconm).</p> <p>This value takes into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value also varies due to individual differences and data patterns to be processed even when same job is executed on different nodes. There may be a difference</p>	avgpconm	double	AVG POWER CONSUMPTION OF NODE (MEASURED)

Item	Item name	Data type	pjstat/.stats file
of +/- 40% from average node power consumption (Estimated). For bulk/step job summary information, the total for sub jobs is output.			
Maximum node power consumption (actual measurement) From the periodically collected node power consumption (W), the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxpconm). This value takes into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value also varies due to individual differences and data patterns to be processed even when same job is executed on different nodes. There may be a difference of +/- 40% from maximum node power consumption (Estimated). For bulk/step job summary information, the total for sub jobs is output.	maxpconm	double	MAX POWER CONSUMPTION OF NODE (MEASURED)
Minimum node power consumption (actual measurement) From the periodically collected node power consumption (W), the minimum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(minpconm). This value takes into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value also varies due to individual differences and data patterns to be processed even when same job is executed on different nodes. There may be a difference of +/- 40% from minimum node power consumption (Estimated). For bulk/step job summary information, the total for sub jobs is output.	minpconm	double	MIN POWER CONSUMPTION OF NODE (MEASURED)
Node power consumption (actual measurement) The output for the node power consumption (Wh) from job start to job end is shown to up to six decimal places. The value is the total of Node/Virtual node statistical information(econm). This value takes into account variations in power consumption due to difference of computing and I/O node devices which are assistant core, PCI Express and so on. This value also varies due to individual differences and data patterns to be processed even when same job is executed on different nodes. There may be a difference of +/- 40% from node power consumption (Estimated). For bulk/step job summary information, the total for sub jobs is output.	econm	double	ENERGY CONSUMPTION OF NODE (MEASURED)
Power knob usage information The output binary bit pattern indicates whether the Power API is used and whether the power knob is operated. The value is a logical OR of Node/Virtual node statistical information(uiopa). 0: Power API not used, and power knob not operated 1: Power API used, and power knob not operated 11: Power API used, and power knob operated For bulk/step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:	uiopa	uint8	UTILIZATION INFO OF POWER API

Item	Item name	Data type	pjstat/.stats file
- pjstat/.stats: "-" - pmdumpjobinfo: ""			
PKG number PKG numbers of CPUs used by the job are output, separated by spaces. For bulk/step job summary information and for jobs executed on an FX server compute node, "" is output.	pkgno	int32_array	-
Average CPU power consumption for each package The CPU power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(avgcocpkg). - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgcocpkg	double_array	AVG POWER CONSUMPTION OF CPU/PKG(<i>num</i>)
Maximum CPU power consumption for each package From the periodically collected CPU power consumption (W) for each package, the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxcocpkg). - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxcocpkg	double_array	MAX POWER CONSUMPTION OF CPC/PKG(<i>num</i>)
Minimum CPU power consumption for each package From the periodically collected CPU power consumption (W) for each package, the minimum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(mincocpkg). - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.	mincocpkg	double_array	MIN POWER CONSUMPTION OF CPU/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<p>- pmdumpjobinfo: As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-" - pmdumpjobinfo: ""</p>			
<p>CPU power consumption for each package</p> <p>The output for the CPU power consumption (Wh) for each package from job start to job end is shown to up to six decimal places. The value is the total of Node/Virtual node statistical information(ecocpkg).</p> <p>- pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p> <p>- pmdumpjobinfo As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-" - pmdumpjobinfo: ""</p>	ecocpkg	double_array	ENERGY CONSUMPTION OF CPU/PKG(<i>num</i>)
<p>Average memory power consumption for each package</p> <p>The memory power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(avgpcompkg).</p> <p>- pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p> <p>- pmdumpjobinfo As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-" - pmdumpjobinfo: ""</p>	avgpcompkg	double_array	AVG POWER CONSUMPTION OF MEM/PKG(<i>num</i>)
<p>Maximum memory power consumption for each package</p> <p>From the periodically collected memory power consumption (W) for each package, the maximum value of up to six decimal places is output. The value is the total of Node/Virtual node statistical information(maxpcompkg).</p> <p>- pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p>	maxpcompkg	double_array	MAX POWER CONSUMPTION OF MEM/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Minimum memory power consumption for each package</p> <p>From the periodically collected memory power consumption (W) for each package, the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcmpkg).</p> <ul style="list-style-type: none"> - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minpcmpkg	double_array	MIN POWER CONSUMPTION OF MEM/PKG(<i>num</i>)
<p>Memory power consumption for each package</p> <p>The output for the memory power consumption (Wh) for each package from job start to job end is shown to up to six decimal places.</p> <p>The value is the total of Node/Virtual node statistical information(ecompkg).</p> <ul style="list-style-type: none"> - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecompkg	double_array	ENERGY CONSUMPTION OF MEM/PKG(<i>num</i>)
<p>Average PP0 power consumption for each package</p> <p>The PP0 power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(avGPCOP0pkg).</p> <ul style="list-style-type: none"> - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. 	avGPCOP0pkg	double_array	AVG POWER CONSUMPTION OF PP0/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<p>- pmdumpjobinfo</p> <p>As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>Maximum PP0 power consumption for each package</p> <p>From the periodically collected PP0 power consumption (W) for each package, the maximum value of the maximum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(maxpcop0pkg).</p> <p>- pjstat/.stats</p> <p>This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p> <p>- pmdumpjobinfo</p> <p>As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	maxpcop0pkg	double_array	MAX POWER CONSUMPTION OF PP0/PKG(<i>num</i>)
<p>Minimum PP0 power consumption for each package</p> <p>From the periodically collected PP0 power consumption (W) for each package, the minimum value of up to six decimal places is output.</p> <p>The value is the total of Node/Virtual node statistical information(minpcop0pkg).</p> <p>- pjstat/.stats</p> <p>This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p> <p>- pmdumpjobinfo</p> <p>As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	minpcop0pkg	double_array	MIN POWER CONSUMPTION OF PP0/PKG(<i>num</i>)
<p>PP0 power consumption for each package</p> <p>The output for the PP0 power consumption (Wh) for each package from job start to job end is shown to up to six decimal places</p> <p>The value is the total of Node/Virtual node statistical information(ecop0pkg).</p> <p>- pjstat/.stats</p> <p>This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number.</p>	ecop0pkg	double_array	ENERGY CONSUMPTION OF PP0/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<p>- pmdumpjobinfo</p> <p>As many values as the number of packages are output, separated by spaces.</p> <p>For bulk/step job summary information and for jobs executed on an FX server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>Power information acquisition status</p> <p>The output is a logical OR of the following numerical values:</p> <p>0x1: Power information is not affected by another node-sharing job.</p> <p>0x2: Power information is affected by another node-sharing job.</p> <p>0x4: Acquisition of information from some nodes failed.</p> <p>0x8: Information was not acquired from some nodes because of a preceding job.</p> <p>The value is a logical OR of Node/Virtual node statistical information(pcs).</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	pcs	uint16	POWER CONSUMPTION STATE
<p>Power measurement start time</p> <p>The value is the earliest time of Node/Virtual node statistical information(pcmsdt).</p> <p>- pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second)</p> <p>- pmdumpjobinfo: Number of seconds since January 1, 1970</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	pcmsdt	timespec	POWER CONSUMPTION MEASURE START DATE
<p>Power measurement end time</p> <p>The value is the latest time of Node/Virtual node statistical information(pcmedt).</p> <p>- pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second)</p> <p>- pmdumpjobinfo: Number of seconds since January 1, 1970</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	pcmedt	timespec	POWER CONSUMPTION MEASURE END DATE
<p>The time when the power consumption exceeds the threshold set by the power limit function</p> <p>When excess power does occur, the time when the power consumption exceeds the threshold set by the power limit function is output.</p> <p>The value is the earliest time of Node/Virtual node statistical information(pcdt).</p>	pcdt	timespec	POWER CAPPING DATE

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats : <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>When excess power does not occur or for bulk job summary information, and for step job summary information and for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats : "-" - pmdumpjobinfo : "" <p>This item is not output by default of the job operation software. To output the item, the administrator must set the job statistics information in the papjmstats.conf file.</p>			
<p>Size (bytes) of the shared temporary area of first-layer storage</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	llioszr	uint64	LLIO SHARED TMP SIZE (REQUIRE)
<p>Size (bytes) of the node temporary area of first-layer storage</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	llioszr	uint64	LLIO LOCAL TMP SIZE (REQUIRE)
<p>Readahead for first-layer storage and second-layer storage</p> <p>0: Disable automatic readahead. 1: Enable automatic readahead.</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	llioautorh	int32	LLIO AUTO READAHEAD
<p>Operation when closing files on first-layer storage and second-layer storage</p> <p>0: Synchronous close 1: Asynchronous close</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	llioasyncl	int32	LLIO ASYNC CLOSE
<p>Threshold value (bytes) of whether or not to cache data during writing to first-layer storage</p> <p>When writing data to first-layer storage, if the write size is equal to or less than this value, the data is not immediately written to the storage.</p> <p>Instead, it is temporarily stored in compute node cache.</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p>	llocnwsz	uint64	LLIO CN CACHED WRITE SIZE

Item	Item name	Data type	pjstat/.stats file
- pjstat/.stats: "-" - pmdumpjobinfo: ""			
Size (bytes) of the compute node cache managed by first-layer storage (LLIO) For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	llicncsz	uint64	LLIO CN CACHE SIZE
Operation of whether or not to cache data on a compute node when a job reads a file from first-layer storage 0: Do not cache. 1: Cache. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	llicnrc	int32	LLIO CN READ CACHE
Operation of whether or not to cache files in first-layer storage when the files are read from second-layer storage to a compute node 0: Do not cache. 1: Cache If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: "".	lliosiorc	int32	LLIO SIO READ CACHE
Number of stripes per file when files are distributed across first-layer storage If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lliostrpec	uint64	LLIO STRIPE COUNT
Size (bytes) of stripes when files are distributed across first-layer storage If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lliostrpsz	uint64	LLIO STRIPE SIZE
Output file path for the names of unexported files on second-layer storage If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For bulk/step job summary information, "" is output.	llioucpath	string	LLIO UNCOMPLETED FILEINFO PATH
Total size (bytes) of unexported files on second-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	llioucfsz	uint64	LLIO UNCOMPLETED FILEINFO SIZE
LLIO performance information file path	llioppath	string	LLIO PERF PATH

Item	Item name	Data type	pjstat/.stats file
<p>Depending on the combination of the --llio perf and perf-path options of the pjsb command, the following is output:</p> <ul style="list-style-type: none"> - --llio perf <i>current directory path+script file name +.job ID+.llio_perf</i> - --llio perf-path=<i>path</i> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" - --llio perf,perf-path=<i>path</i> <i>path</i> is the path specified in the --llio perf-path option of the pjsb command. <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. For bulk/step job summary information, "" is output.</p>			
<p>Cache hit ratio (%) of second-layer storage cache</p> <p>The cache hit ratio is: number of cache hits/(number of cache hits + number of cache misses). The value is rounded up to the nearest million.</p> <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	2chr	double	2NDLAYERCACHE CACHE HIT RATIO
<p>Minimum available capacity (bytes) of second-layer storage cache</p> <p>For bulk job summary information, the minimum value of the sub job is output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	2mfs	uint64	2NDLAYERCACHE MIN FREE SPACE
<p>Maximum usage (bytes) of the shared temporary area of first-layer storage</p> <p>For bulk job summary information, the maximum value of the sub job is output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	smus	uint64	SHAREDTMP MAX USE SPACE
<p>Maximum usage (bytes) of the node temporary area of first-layer storage</p> <p>The output value shows the node with the greatest usage of the node temporary area among the compute nodes allocated to the job.</p> <p>For bulk job summary information, the maximum value of the sub job is output.</p> <p>For step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	lmus	uint64	LOCALTMP MAX USE SPACE
<p>Lowest cache hit ratio (%) of compute node cache that is second-layer storage cache</p> <p>The output value shows the node with the lowest cache hit ratio for compute node cache among the compute nodes allocated to the job.</p> <p>The value is rounded up to the nearest million.</p> <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	2ncmchr	double	2NDLAYERCACHE NODE CACHE MIN CACHE HIT RATIO

Item	Item name	Data type	pjstat/.stats file
Lowest cache hit ratio (%) of compute node cache for the shared temporary area of first-layer storage The output value shows the node with the lowest cache ratio for compute node cache among the compute nodes allocated to the job. The value is rounded up to the nearest million. For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	sncmchr	double	SHAREDTMP NODE CACHE MIN CACHE HIT RATIO
Lowest cache hit ratio (%) of compute node cache for the node temporary area of first-layer storage The output value shows the node with the lowest cache ratio for compute node cache among the compute nodes allocated to the job. The value is rounded up to the nearest million. For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lncmchr	double	LCOALTMP NODE CACHE MIN CACHE HIT RATIO
Available capacity (bytes) of the compute node cache with the lowest cache hit ratio For bulk/step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	mncmrfs	uint64	MIN NODE CACHE HIT RATIO FREE SPACE
read traffic (bytes) of second-layer storage cache For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	2rta	uint64	2NDLAYERCACHE READ TRANSFER AMOUNT
write traffic (bytes) of second-layer storage cache For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	2wta	uint64	2NDLAYERCACHE WRITE TRANSFER AMOUNT
read traffic (bytes) in the shared temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	srta	uint64	SHAREDTMP READ TRANSFER AMOUNT
write traffic (bytes) in the shared temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	swta	uint64	SHAREDTMP WRITE TRANSFER AMOUNT
read traffic (bytes) in the node temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lrta	uint64	LOCALTMP READ TRANSFER AMOUNT

Item	Item name	Data type	pjstat/.stats file
write traffic (bytes) in the node temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lwta	uint64	LOCALTMP WRITE TRANSFER AMOUNT
Number of open times of second-layer storage cache For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	2open	uint64	2NDLAYERCACHE OPEN COUNT
Number of statfs times of second-layer storage cache For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	2statfs	uint64	2NDLAYERCACHE STATFS COUNT
Number of sync times of second-layer storage cache For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	2sync	uint64	2NDLAYERCACHE SYNC COUNT
Number of open times in the shared temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	sopenc	uint64	SHAREDTMP OPEN COUNT
Number of open times in the node temporary area of first-layer storage For bulk job summary information, the total for sub jobs is output. For step job summary information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	lopenc	uint64	LOCALTMP OPEN COUNT
Start time of the LLIO initialization process - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i> : Year, <i>MM</i> : Month, <i>DD</i> : Day, <i>hh</i> : Hour, <i>mm</i> : Minute, <i>ss</i> : Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, "" is output.	lloinitst	timespec	-
End time of the LLIO initialization process - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i> : Year, <i>MM</i> : Month, <i>DD</i> : Day, <i>hh</i> : Hour, <i>mm</i> : Minute, <i>ss</i> : Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 For bulk/step job summary information, "" is output.	lloinitet	timespec	-
Start time of waiting for cache flushing to complete in second-layer storage	2flst	timespec	2NDLAYERCACHE FLASH START TIME

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>End time of waiting for cache flushing to complete in second-layer storage</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	2flet	timespec	2NDLAYERCACHE FLASH END TIME
<p>Start time of LLIO end processing</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, "" is output.</p>	llofinist	timespec	-
<p>End time of LLIO end processing</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD HH:MM:SS</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>HH</i>: Hour, <i>MM</i>: Minute, <i>SS</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>For bulk/step job summary information, "" is output.</p>	llofiniet	timespec	-
<p>Number of requested virtual nodes</p> <p>If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output.</p> <p>For jobs on an FX server compute node, node allocated jobs on a PRIMERGY server compute node, and step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	vnumr	uint32	VNODE NUM (REQUIRE)
<p>Number of allocated virtual nodes</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For jobs on an FX server compute node, node allocated jobs on a PRIMERGY server compute node, and step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	vnuma	uint32	VNODE NUM (ALLOC)
<p>Number of virtual nodes used</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For jobs on an FX server compute node, node allocated jobs on a PRIMERGY server compute node, and step job summary</p>	vnumu	uint32	VNODE NUM (USE)

Item	Item name	Data type	pjstat/.stats file
information, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""			
Node allocation type 0: Torus mode (node-exclusive allocation) 1: Mesh mode (node-exclusive allocation) 2: Node-sharing allocation (not supported) 3: Non-contiguous mode (node-exclusive allocation) For step job summary information and for jobs on a PRIMERGY server compute node, "" is output.	alloctype	uint32	-
Number of requested nodes For step job summary information, "" is output.	nnumr	uint32	-
X coordinate on the shape of requested nodes For step job summary information, "" is output.	nnumr_x	uint32	-
Y coordinate on the shape of requested nodes For step job summary information, "" is output.	nnumr_y	uint32	-
Z coordinate on the shape of requested nodes For step job summary information, "" is output.	nnumr_z	uint32	-
Limit on node allocation The output is a logical OR of the following numerical values: 0x0: No limit 0x1: strict specified 0x2: strict-io specified 0x4: io-exclusive specified For step job summary information, "" is output.	nalimit	int32	-
Number and shape of requested nodes The pjstat command and the .stats file show the output contents. - For a node-exclusive job The number and shape of requested nodes are output. [3-dimensional shape] <i>N</i> . <i>X</i> <i>Y</i> <i>Z</i> [:mesh :noncont] [:strict :strict-io] [:io-exclusive] [2-dimensional shape] <i>N</i> . <i>X</i> <i>Y</i> [:mesh :noncont] [:io-exclusive] [1-dimensional shape] <i>N</i> [:mesh :noncont] [:io-exclusive] <i>N</i> is the number of nodes. <i>X</i> , <i>Y</i> , and <i>Z</i> are node coordinates. ":mesh" represents mesh mode. ":noncont" represents non-contiguous mode. If neither ":mesh" nor ":noncont" is displayed, the mode is torus mode. - For a node-sharing job (not supported) Number of requested virtual nodes - For a job on a PRIMERGY server compute node Number of requested nodes	-	-	NODE NUM (REQUIRE)

Item	Item name	Data type	pjstat/.stats file
<p>If the number of requested nodes is not specified at the job submission time, "-" is output.</p> <p>If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output.</p> <p>For step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>However, for the summary information for a step job that mixes both a sub job executed on a FX server compute node and a sub job executed on a PRIMERGY server compute node, "-" is output to the .stats file.</p>			
<p>Number of allocated nodes</p> <p>For bulk job summary information, the total for sub jobs is output.</p> <p>For step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	nnuma	uint32	-
<p>X coordinate on the shape of allocated nodes</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	nnuma_x	uint32	-
<p>Y coordinate on the shape of allocated nodes</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	nnuma_y	uint32	-
<p>Z coordinate on the shape of allocated nodes</p> <p>For bulk/step job summary information, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	nnuma_z	uint32	-
<p>Number and shape of allocated nodes</p> <p>The pjstat command and the .stats file show the output contents.</p> <ul style="list-style-type: none"> - For a node-exclusive job Shape and number of allocated nodes $N \cdot X \times Y \times Z$ [:strict :strict-io][:io-exclusive] (torus mode) $N \cdot X \times Y \times Z$:mesh (mesh mode) N:noncont (non-contiguous mode) N is the number of nodes. X, Y, and Z are the respective coordinate values (including unavailable nodes). - For a node-sharing job (not supported) Shape and number of allocated nodes $N \cdot X \times Y \times Z$ - For a node allocated job executed on a PRIMERGY server compute node N is the number of allocated nodes <p>For step job summary information and for virtual node allocated jobs on a PRIMERGY server compute node, "-" is output.</p> <p>For bulk job summary information, the total for sub jobs is output.</p>	-	-	NODE NUM (ALLOC)
<p>Number of nodes used</p> <ul style="list-style-type: none"> - For a node-exclusive job Number of allocated nodes actually used by the job 	nnumu	uint32	NODE NUM (USE)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - For a node-sharing job (not supported) Number of allocated virtual nodes actually used by the job - For a job on a PRIMERGY server compute node Number of allocated nodes actually used by the job <p>For bulk job summary information, the total for sub jobs is output. Step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Number of unavailable nodes</p> <ul style="list-style-type: none"> - For a node-exclusive job Number of allocated nodes that are unavailable - For a node-sharing job (not supported) Number of allocated virtual nodes that are unavailable - For a job on a PRIMERGY server compute node Number of allocated nodes that are unavailable <p>For bulk job summary information, the total for sub jobs is output. Step job summary information, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	nnumv	uint32	NODE NUM (UNUSED)
<p>Node ID list of nodes used</p> <p>Node IDs (hexadecimal) are output, separated by spaces. For bulk/step job summary information, "-" is output. This item is not output by the pmdumpjobinfo command.</p>	nidlu	string	NODE ID (USE)
<p>Tofu coordinate list of nodes used</p> <p>The Tofu coordinates (X, Y, Z) of the nodes are output, separated by spaces. If any of the nodes belong to the same Tofu unit, the same Tofu coordinates are output multiple times. For jobs on a PRIMERGY server compute node and for bulk/step job summary information, "-" is output. This item is not output by the pmdumpjobinfo command.</p>	tofulu	string	TOFU COORDINATE (USE)
<p>Node IDs of allocated nodes, and their corresponding rank numbers</p> <p>Node IDs and rank numbers are output in the <i>nodeID(rank number)</i> format, separated by spaces. For bulk/step job summary information, "-" is output. This item is not output by the pmdumpjobinfo command.</p>	rank	string	NODE ID (RANK)
<p>Amount of requested custom resources</p> <p>This item is output as many times as the number of custom resource definitions. If nothing is specified in the argument of the pjsb command, <DEFAULT> appears at the end of the output. Power knob settings are output as the following custom resource: Frequency: freq Memory access limit: throttling_state Instruction issue limit: issue_state SIMD width: simd_state FLA+B, FLA: fl_pipe_state EXA+B, EXA: ex_pipe_stat</p>	CR-STR- <i>custom resource name</i> -req CR-NUM- <i>custom resource name</i> -req	string uint64	<i>custom resource name</i> (REQUIRE)

Item	Item name	Data type	pjstat/.stats file
Amount of allocated custom resources This item is output as many times as the number of custom resource definitions.	CR-STR- <i>custom resource name</i> - alloc CR-NUM- <i>custom resource name</i> -alloc	string uint64	<i>custom resource name</i> (ALLOC)
Amount of custom resources used This item is output as many times as the number of custom resource definitions. If the amount of use is not registered in a hook of the job operation management function, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	CR-STR- <i>custom resource name</i> -use CR-NUM- <i>custom resource name</i> -use	string uint64	<i>custom resource name</i> (USE)

Node/Virtual node statistical information (Output items for each node/virtual node)

The following data is output as node/virtual node statistical information.

Item	Item name	Data type	pjstat/.stats file
Record name The output of "JN" indicates a node/virtual node statistical information record.	record	string	-
Virtual node ID The output virtual node ID is a decimal number. For jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	vnid	string	VNODE ID
Node ID (hexadecimal)	nid	string	NODE ID
Tofu coordinates x,y,z of a node For jobs executed on a PRIMERGY server compute node, "" is output to the pmdumpjobinfo output.	tofu_x tofu_y tofu_z	uint32 uint32 uint32	-
Tofu coordinates (x,y,z) of a node The pjstat command and the .stats file show the output contents. For jobs executed on a PRIMERGY server compute node, "-" is output.	-	-	TOFU COORDINATE
Coordinates x,y,z of the node used For jobs executed on a PRIMERGY server compute node, "" is output to the pmdumpjobinfo output.	node_x node_y node_z	uint32 uint32 uint32	-
Coordinates (X,Y,Z) of the node used The pjstat command and the .stats file show the output contents. For jobs executed on a PRIMERGY server compute node, "-" is output.	-	-	NODE COORDINATE
Rank number of the job executed on a node or virtual node If there are multiple rank numbers, they are output separated by spaces.	rankno	string	RANK NO
Node state - pjstat/.stats available: Available	nodest	int32	NODE AVAILABLE

Item	Item name	Data type	pjstat/.stats file
- pmdumpjobinfo 1: Available			
Node machine type PG: PRIMERGY server FT: FX server	machine	string	ARCHI INFORMATION
Character string representing, in bitmap format (hexadecimal notation), the CPU ID of an allocated CPU Bits corresponding to the allocated CPU ID are set in order starting with the lowest.	cabitmap	string	CPU BITMAP (ALLOC)
Character string representing, in bitmap format (hexadecimal notation), the CPU ID of the CPU actually used Bits corresponding to the used CPU ID are set in order starting with the lowest.	cubitmap	string	CPU BITMAP (USE)
Upper limit on the number of processes generated by an MPI program The number of processes specified in the --mpi proc option of the pjsub command is output. If --mpi proc option is omitted, the following is - For jobs executed on an FX server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of specified by the --mpi shape option. If --mpi max-proc-per-node option is omitted, the number of specified by the --mpi shape option is output. - For jobs executed on an PRIMERGY server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of virtual nodes or nodes. If --mpi maxproc-per-node option is omitted, the number of virtual nodes or nodes is output. For the summary information on a bulk job or a step job, the total of all sub jobs is output.	mpiproca	int32	MPI PROC (ALLOC)
Number of processes (static processes) generated at the time of MPI program startup The number of processes specified in the --mpi proc option of the pjsub command is output. If --mpi proc option is omitted, the following is - For jobs executed on an FX server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of specified by the --mpi shape option. If --mpi max-proc-per-node option is omitted, the number of specified by the --mpi shape option is output. - For jobs executed on an PRIMERGY server compute node outputs the number of processes specified in the --mpi max-proc-per-node option multiplied by the number of virtual nodes or nodes. If --mpi max-proc-per-node option is omitted, the number of virtual nodes or nodes is output.	mpistaticproca	int32	MPI STATIC PROC (ALLOC)

Item	Item name	Data type	pjstat/.stats file
For the summary information on a bulk job or a step job, the total of all sub jobs is output.			
Starting time (seconds) of the job execution environment The time taken for the start processing of the job execution environment is output.	jebt	uint64	-
Ending time (seconds) of the job execution environment The time taken for the end processing of the job execution environment is output.	jest	uint64	-
Total user CPU time (milliseconds) spent on job execution	uctmut	uint64	USER CPU TIME (USE)
Total system CPU time (milliseconds) spent on job execution	sctmut	uint64	SYSTEM CPU TIME (USE)
Total user CPU time and system CPU time (milliseconds)	usctmut	uint64	CPU TIME (TOTAL)
Amount of memory allocated to the node <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2^{20} bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. If there is no limit, unlimited is displayed. - pmdumpjobinfo: Value in units of bytes 18446744073709551615 (=0xffffffffffffffff) means no limit (unlimited). 	msza	uint64	MEMORY SIZE (ALLOC)
Maximum physical memory usage <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2^{20} bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes 	mmszu	uint64	MAX MEMORY SIZE (USE)
Number of CPU cores actually allocated in the node	cnuma	uint32	CPU NUM (ALLOC)
Number of CPU cores actually used by the job	cnumu	uint32	CPU NUM (USE)
Number of actually allocated assistant cores For jobs executed on a PRIMERGY server compute node, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" This item is not output by default. To output the item, the administrator needs to configure it.	anuma	uint32	ASSISTANT CORE NUM (ALLOC)
Number of assistant cores used by the job For jobs executed on a PRIMERGY server compute node, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" This item is not output by default. To output the item, the administrator needs to configure it.	acnumut	uint32	ASSISTANT CORE NUM (USE)
Total user CPU time (milliseconds) for assistant cores For jobs executed on a PRIMERGY server compute node, the following is output: <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	auctmut	uint64	ASSISTANT CORE USER CPU TIME (USE)

Item	Item name	Data type	pjstat/.stats file
This item is not output by default. To output the item, the administrator needs to configure it.			
Total system CPU time (milliseconds) for assistant cores For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: "" This item is not output by default. To output the item, the administrator needs to configure it.	asctmut	uint64	ASSISTANT CORE SYSTEM CPU TIME (USE)
Maximum memory usage by assistant cores - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MiB (2^{20} bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: "" This item is not output by default. To output the item, the administrator needs to configure it. This item can be acquired only in McKernel mode.	ammszu	uint64	ASSISTANT CORE MAX MEMORY SIZE (USE)
Number of processes in progress This item is output only by the pjstat command. For running jobs, the number of processes running in the job is output. If the job is not running, 0 is output.	procnum	int32	PROC NUM
Number of execution cycles For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf1	int64	PERF COUNT 1
Number of floating-point instruction operations 1 For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf2	int64	PERF COUNT 2
Number of floating-point instruction operations 2 For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf3	int64	PERF COUNT 3
Number of memory read requests For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf4	int64	PERF COUNT 4
Number of memory write requests	perf5	int64	PERF COUNT 5

Item	Item name	Data type	pjstat/.stats file
For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""			
Sleep cycle For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf6	int64	PERF COUNT 6
Reserved item for future expansion For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf7	int64	PERF COUNT 7
Reserved item for future expansion For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf8	int64	PERF COUNT 8
Reserved item for future expansion For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	perf9	int64	PERF COUNT 9
Receive data size for Tofu user communication This is the receive data size for user-level communication via Tofu. - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10 ⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	tucrb	int64	TOFU USER COMM RECV BYTE
Send data size for Tofu user communication This is the send data size for user-level communication via Tofu. - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10 ⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	tucsb	int64	TOFU USER COMM SEND BYTE
Receive data size for Tofu system communication This is the receive data size for system communication via Tofu.	tscrb	int64	TOFU SYSTEM COMM RECV BYTE

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Send data size for Tofu system communication</p> <p>This is the send data size for system communication via Tofu.</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>val1</i> (<i>val2</i>) The unit for <i>val1</i> is MB (10⁶ bytes). The value is rounded up to the nearest ten. The unit for the <i>val2</i> value is bytes. - pmdumpjobinfo: Value in units of bytes <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	tscsb	int64	TOFU SYSTEM COMM SEND BYTE
<p>Number of Fujitsu profiler use times</p>	fjprofiler	int64	FJ PROFILER
<p>Number of starts of a program using the sector cache driver/library</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	sectorcache	int64	SECTOR CACHE
<p>Number of starts of a program using the inter-core hardware barrier driver/library</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	intranodebarrier	int64	INTRA NODE BARRIER
<p>CMGs numbers used by the job</p> <p>CMG numbers are output, separated by spaces. For jobs executed on a PRIMERGY server compute node, "" is output.</p>	cmgno	int32_array	-
<p>Average power consumption of the compute core group for each CMG (Estimate)</p> <p>The power consumption of the compute core group for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p>	avgpcocc	double_array	AVG POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>			
<p>Maximum power consumption of the compute core group for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the compute core group for each CMG, the maximum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	maxpcocc	double_array	MAX POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)
<p>Minimum power consumption of the compute core group for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the compute core group for each CMG, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	minpcocc	double_array	MIN POWER CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)
<p>Power consumption of the compute core group for each CMG (Estimate)</p> <p>The output for the power consumption (Wh) of the compute core group for each CMG from job start to job end is shown to up to six decimal places.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <p>- pjstat/.stats: "-"</p> <p>- pmdumpjobinfo: ""</p>	ecocc	double_array	ENERGY CONSUMPTION OF CORES/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<p>Average power consumption of the L2 cache for each CMG (Estimate)</p> <p>The power consumption of the L2 cache for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "- " - pmdumpjobinfo: "" 	avgpcolc	double_array	AVG POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Maximum power consumption of the L2 cache for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the L2 cache for each CMG, the maximum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "- " - pmdumpjobinfo: "" 	maxpcolc	double_array	MAX POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Minimum power consumption of the L2 cache for each CMG (Estimate)</p> <p>From the periodically collected power consumption (W) of the L2 cache for each CMG, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "- " - pmdumpjobinfo: "" 	minpcolc	double_array	MIN POWER CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)
<p>Power consumption of the L2 cache for each CMG (Estimate)</p>	ecolc	double_array	ENERGY CONSUMPTION OF L2CACHE/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<p>The output for the power consumption (Wh) of the L2 cache for each CMG from job start to job end is shown to up to six decimal places.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Average memory power consumption for each CMG (Estimate)</p> <p>The memory power consumption for each CMG from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avgpcomc	double_array	AVG POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)
<p>Maximum memory power consumption for each CMG (Estimate)</p> <p>From the periodically collected memory power consumption (W) for each CMG, the maximum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. - pmdumpjobinfo As many values as the number of CMGs used by the job are output, separated by spaces. <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxpcomc	double_array	MAX POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)
<p>Minimum memory power consumption for each CMG (Estimate)</p> <p>From the periodically collected memory power consumption (W) for each CMG, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number. 	minpcomc	double_array	MIN POWER CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pmdumpjobinfo <p>As many values as the number of CMGs used by the job are output, separated by spaces.</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Memory power consumption for each CMG (Estimate)</p> <p>The output for the memory power consumption (Wh) for each CMG from job start to job end is shown to up to six decimal places.</p> <ul style="list-style-type: none"> - pjstat/.stats <p>This item appears as many times as the number of CMGs used by the job. <i>num</i> in the item name is the CMG number.</p> <ul style="list-style-type: none"> - pmdumpjobinfo <p>As many values as the number of CMGs used by the job are output, separated by spaces.</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecomc	double_array	ENERGY CONSUMPTION OF MEM/CMG(<i>num</i>) (IDEAL)
<p>Average Tofu power consumption (Estimate)</p> <p>The Tofu power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avgpcot	double	AVG POWER CONSUMPTION OF TOFU (IDEAL)
<p>Maximum Tofu power consumption (Estimate)</p> <p>From the periodically collected Tofu power consumption (W), the maximum value of up to six decimal places is output.</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxpcot	double	MAX POWER CONSUMPTION OF TOFU (IDEAL)
<p>Minimum Tofu power consumption (Estimate)</p> <p>From the periodically collected Tofu power consumption (W), the minimum value of up to six decimal places is output.</p> <p>For jobs executed on a PRIMERGY server, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minpcot	double	MIN POWER CONSUMPTION OF TOFU (IDEAL)
<p>Tofu power consumption (Estimate)</p> <p>The output for the Tofu power consumption (Wh) from job start to job end is shown to up to six decimal places.</p> <p>For jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecot	double	ENERGY CONSUMPTION OF TOFU (IDEAL)

Item	Item name	Data type	pjstat/.stats file
Average power consumption at the periphery inside the CPU (Estimate) The power consumption at the periphery inside the CPU from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpcocp	double	AVG POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Maximum power consumption at the periphery inside the CPU (Estimate) From the periodically collected power consumption (W) at the periphery inside the CPU, the maximum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcocp	double	MAX POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Minimum power consumption at the periphery inside the CPU (Estimate) From the periodically collected power consumption (W) at the periphery inside the CPU, the minimum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minpcocp	double	MIN POWER CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Power consumption at the periphery inside the CPU (Estimate) The output for the power consumption (Wh) at the periphery inside the CPU from job start to job end is shown to up to six decimal places. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecocp	double	ENERGY CONSUMPTION OF CPU PERIPHERALS (IDEAL)
Average power consumption of the optical module (Estimate) The power consumption of the optical module from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpcoom	double	AVG POWER CONSUMPTION OF OPTICAL MODULE (IDEAL)
Maximum power consumption of the optical module (Estimate) From the periodically collected power consumption (W) of the optical module, the maximum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcoom	double	MAX POWER CONSUMPTION OF OPTICAL MODULE (IDEAL)

Item	Item name	Data type	pjstat/.stats file
Minimum power consumption of the optical module (Estimate) From the periodically collected power consumption (W) of the optical module, the minimum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minpcoom	double	MIN POWER CONSUMPTION OF OPTICAL MODULE (IDEAL)
Power consumption of the optical module (Estimate) The output for the power consumption (Wh) of the optical module from job start to job end is shown to up to six decimal places. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecoom	double	ENERGY CONSUMPTION OF OPTICAL MODULE (IDEAL)
Average PCI-E power consumption (Estimate) The PCI-E power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpcop	double	AVG POWER CONSUMPTION OF PCI-E (IDEAL)
Maximum PCI-E power consumption (Estimate) From the periodically collected PCI-E power consumption (W), the maximum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcop	double	MAX POWER CONSUMPTION OF PCI-E (IDEAL)
Minimum PCI-E power consumption (Estimate) From the periodically collected PCI-E power consumption (W), the minimum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minpcop	double	MIN POWER CONSUMPTION OF PCI-E (IDEAL)
PCI-E power consumption (Estimate) The output for the PCI-E power consumption (Wh) from job start to job end is shown to up to six decimal places. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecop	double	ENERGY CONSUMPTION OF PCI-E (IDEAL)
Average node power consumption (Estimate) The node power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output:	avgpcon	double	AVG POWER CONSUMPTION OF NODE (IDEAL)

Item	Item name	Data type	pjstat/.stats file
- pjstat/.stats: "-" - pmdumpjobinfo: ""			
Maximum node power consumption (Estimate) From the periodically collected node power consumption (W), the maximum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcon	double	MAX POWER CONSUMPTION OF NODE (IDEAL)
Minimum node power consumption (Estimate) From the periodically collected node power consumption (W), the minimum value of up to six decimal places is output. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	minpcon	double	MIN POWER CONSUMPTION OF NODE (IDEAL)
Node power consumption (Estimate) The output for the node power consumption (Wh) from job start to job end is shown to up to six decimal places. For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	econ	double	ENERGY CONSUMPTION OF NODE (IDEAL)
Average node power consumption (actual measurement) The node power consumption from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.	avgpconm	double	AVG POWER CONSUMPTION OF NODE (MEASURED)
Maximum node power consumption (actual measurement) From the periodically collected node power consumption (W), the maximum value of up to six decimal places is output.	maxpconm	double	MAX POWER CONSUMPTION OF NODE (MEASURED)
Minimum node power consumption (actual measurement) From the periodically collected node power consumption (W), the minimum value of up to six decimal places is output.	minpconm	double	MIN POWER CONSUMPTION OF NODE (MEASURED)
Node power consumption (actual measurement) The output for the node power consumption (Wh) from job start to job end is shown to up to six decimal places.	econm	double	ENERGY CONSUMPTION OF NODE (MEASURED)
Power knob usage information The output binary bit pattern indicates whether the Power API is used and whether the power knob is operated. 0: Power API not used, and power knob not operated 1: Power API used, and power knob not operated 11: Power API used, and power knob operated For jobs executed on a PRIMERGY server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	uiopa	uint8	UTILIZATION INFO OF POWER API
PKG number	pkgno	int32_array	-

Item	Item name	Data type	pjstat/.stats file
<p>PKG numbers of CPUs used by the job are output, separated by spaces.</p> <p>For jobs executed on an FX server compute node, "" is output.</p>			
<p>Average CPU power consumption for each package</p> <p>The CPU power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avgpcocpkg	double_array	AVG POWER CONSUMPTION OF CPU/PKG(<i>num</i>)
<p>Maximum CPU power consumption for each package</p> <p>From the periodically collected CPU power consumption (W) for each package, the maximum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxpcocpkg	double_array	MAX POWER CONSUMPTION OF CPC/PKG(<i>num</i>)
<p>Minimum CPU power consumption for each package</p> <p>From the periodically collected CPU power consumption (W) for each package, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	minpcocpkg	double_array	MIN POWER CONSUMPTION OF CPU/PKG(<i>num</i>)
<p>CPU power consumption for each package</p> <p>The output for the CPU power consumption (Wh) for each package from job start to job end is shown to up to six decimal places.</p>	ecocpkg	double_array	ENERGY CONSUMPTION OF CPU/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>Average memory power consumption for each package</p> <p>The memory power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	avgpcompkg	double_array	AVG POWER CONSUMPTION OF MEM/PKG(<i>num</i>)
<p>Maximum memory power consumption for each package</p> <p>From the periodically collected memory power consumption (W) for each package, the maximum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	maxpcompkg	double_array	MAX POWER CONSUMPTION OF MEM/PKG(<i>num</i>)
<p>Minimum memory power consumption for each package</p> <p>From the periodically collected memory power consumption (W) for each package, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. 	minpcompkg	double_array	MIN POWER CONSUMPTION OF MEM/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
For jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""			
Memory power consumption for each package The output for the memory power consumption (Wh) for each package from job start to job end is shown to up to six decimal places. - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. For jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	ecompkg	double_array	ENERGY CONSUMPTION OF MEM/PKG(<i>num</i>)
Average PP0 power consumption for each package The PP0 power consumption for each package from job start to job end is divided by the job execution time. The resulting value (W) of up to six decimal places is output. - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. For jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	avgpcop0pkg	double_array	AVG POWER CONSUMPTION OF PP0/PKG(<i>num</i>)
Maximum PP0 power consumption for each package From the periodically collected PP0 power consumption (W) for each package, the maximum value of up to six decimal places is output. - pjstat/.stats This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo As many values as the number of packages are output, separated by spaces. For jobs executed on an FX server compute node, the following is output: - pjstat/.stats: "-" - pmdumpjobinfo: ""	maxpcop0pkg	double_array	MAX POWER CONSUMPTION OF PP0/PKG(<i>num</i>)
Minimum PP0 power consumption for each package	minpcop0pkg	double_array	MIN POWER CONSUMPTION OF PP0/PKG(<i>num</i>)

Item	Item name	Data type	pjstat/.stats file
<p>From the periodically collected PP0 power consumption (W) for each package, the minimum value of up to six decimal places is output.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 			
<p>PP0 power consumption for each package</p> <p>The output for the PP0 power consumption (Wh) for each package from job start to job end is shown to up to six decimal places.</p> <ul style="list-style-type: none"> - pjstat/.stats: This item is output as many times as the number of packages. <i>num</i> in the item name is the PKG number. - pmdumpjobinfo: As many values as the number of packages are output, separated by spaces. <p>For jobs executed on an FX server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	ecop0pkg	double_array	ENERGY CONSUMPTION OF PP0/PKG(<i>num</i>)
<p>Power information acquisition status</p> <p>The output is a logical OR of the following numerical values: 0x1: Power information is not affected by another node-sharing job. 0x2: Power information is affected by another node-sharing job. 0x4: Acquisition of information from some nodes failed. 0x8: Information was not acquired from some nodes because of a preceding job.</p>	pcs	uint16	POWER CONSUMPTION STATE
<p>Power measurement start time</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 	pcmsdt	timespec	POWER CONSUMPTION MEASURE START DATE
<p>Power measurement end time</p> <ul style="list-style-type: none"> - pjstat/.stats: <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 	pcmedt	timespec	POWER CONSUMPTION MEASURE END DATE
<p>The time when the power consumption exceeds the threshold set by the power limit function</p>	pcdt	timespec	POWER CAPPING DATE

Item	Item name	Data type	pjstat/.stats file
<p>When excess power does occur, the time when the power consumption exceeds the threshold set by the power limit function is output.</p> <ul style="list-style-type: none"> - pjstat/.stats : <i>YYYY/MM/DD hh:mm:ss</i> (<i>YYYY</i>: Year, <i>MM</i>: Month, <i>DD</i>: Day, <i>hh</i>: Hour, <i>mm</i>: Minute, <i>ss</i>: Second) - pmdumpjobinfo: Number of seconds since January 1, 1970 <p>When excess power does not occur or for jobs executed on a PRIMERGY server compute node, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats : "-" - pmdumpjobinfo : "" <p>This item is not output by default of the job operation software. To output the item, the administrator must set the job statistics information in the papjmstats.conf file.</p>			
read traffic (bytes) of second-layer storage cache	2rta	uint64	2NDLAYERCACHE READ TRANSFER AMOUNT
write traffic (bytes) of second-layer storage cache	2wta	uint64	2NDLAYERCACHE WRITE TRANSFER AMOUNT TIME
read time (seconds) of second-layer storage cache Values of less than a second are rounded up.	2rtt	uint64	2NDLAYERCACHE READ TRANSFER
write time (seconds) of second-layer storage cache Values of less than a second are rounded up.	2wtt	uint64	2NDLAYERCACHE WRITE TRANSFER TIME
read traffic (bytes) in the shared temporary area of first-layer storage	srta	uint64	SHAREDTMP READ TRANSFER AMOUNT
write traffic (bytes) in the shared temporary area of first-layer storage	swta	uint64	SHAREDTMP WRITE TRANSFER AMOUNT
read time (seconds) of the shared temporary area of first-layer storage Values of less than a second are rounded up.	srtt	uint64	SHAREDTMP READ TRANSFER TIME
write time (seconds) of the shared temporary area of first-layer storage Values of less than a second are rounded up.	swtt	uint64	SHAREDTMP WRITE TRANSFER TIME
read traffic (bytes) in the node temporary area of first-layer storage	lrta	uint64	LOCALTMP READ TRANSFER AMOUNT
write traffic (bytes) in the node temporary area of first-layer storage	lwta	uint64	LOCALTMP WRITE TRANSFER AMOUNT
read time (seconds) of the node temporary area of first-layer storage Values of less than a second are rounded up.	lrtt	uint64	LOCALTMP READ TRANSFER TIME
write time (seconds) of the node temporary area of first-layer storage Values of less than a second are rounded up.	lwtt	uint64	LOCALTMP WRITE TRANSFER TIME
Amount of requested custom resources	CR-STR- <i>custom resource name</i> -req	string uint64	<i>custom resource name</i> (REQUIRE)

Item	Item name	Data type	pjstat/.stats file
<p>This item is output as many times as the number of custom resource definitions for each node.</p> <p>If nothing is specified in the argument of the pjsub command, <DEFAULT> appears at the end of the output.</p>	CR-NUM- <i>custom resource name</i> -req		
<p>Amount of allocated custom resources</p> <p>This item is output as many times as the number of custom resource definitions for each node.</p>	CR-STR- <i>custom resource name</i> -alloc CR-NUM- <i>custom resource name</i> -alloc	string uint64	<i>custom resource name</i> (ALLOC)
<p>Amount of custom resources used</p> <p>This item is output as many times as the number of custom resource definitions for each node.</p> <p>If the amount of use is not registered in a hook of the job operation management function, the following is output:</p> <ul style="list-style-type: none"> - pjstat/.stats: "-" - pmdumpjobinfo: "" 	CR-STR- <i>custom resource name</i> -use CR-NUM- <i>custom resource name</i> -use	string uint64	<i>custom resource name</i> (USE)

[SEE ALSO]

pjstat(1), pjsub(1), papjmstats.conf(5), papjmstatsadm(8), pmdumpjobinfo(8)

Chapter 4 Command and Configuration File Reference for Administrator

This chapter provides commands and configuration file references for administrators. Command messages are also described for commands.

4.1 Commands

This section provides a command reference and command messages for administrators. Commands are listed in alphabetical order.

4.1.1 pachecker command

[NAME]

pachecker - Command that checks the software environment

[SYNOPSIS]

```
/usr/sbin/pachecker [ -c clstname ]  
                    [ scope options ]  
                    [ --type TYPE[,...] ]  
                    [ --check directory ]  
                    [ -d directory ]  
  
/usr/sbin/pachecker --help  
scope options(exclusive):  
    --nodegrp nodegid[,...]  
    --bootgrp bootgid[,...]  
    -n nodeid[,...]  
    -f filename
```

[DESCRIPTION]

The pachecker command is a command for the system administrators to check the software environment. By default, it checks the Technical Computing Suite packages and Technical Computing Suite settings. It checks following items,

- Confirmation that packages of the Technical Computing Suite applies
- Confirmation of setting reflection of the Technical Computing Suite
- In addition to this check, the system administrator can check following items by writing the items to check in a definition file (referred to below as "check definition file"), placing the file in the specified location, and then executing the command.
- Confirmation that packages applies
- Confirmation of presence of file
- Confirmation of comparative result of content of file
- Confirmation of execution result of arbitrary command

The system administrator creates the check definition file.

This command checks every node in the target range regardless of the power status of the node.

This command can be executed on the active system management node.

[OPTIONS]

-c *clstname*

Specify the name of the cluster to check.

If this option is not specified, the set cluster name in the environment variable PXMYCLST is used.

--nodegrp *nodegid*[,...]

Specify a node group ID in *nodegid* to check only that specific node group in the cluster.

Write *nodegid* in hexadecimal.

You can specify multiple node group IDs by delimiting them with a comma (e.g., 0x01,0x02,0x03). You can specify a range of IDs by using a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid*[,...]

Specify a boot group ID in *bootgid* to check only that specific boot group in the cluster.

Write *bootgid* in hexadecimal.

You can specify multiple boot group IDs by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103). You can specify a range of IDs by using a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-n *nodeid*[,...]

Specify a node ID in *nodeid* to check only that specific node in the cluster.

Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A). You can specify a range of IDs by using a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-f *filename*

This option executes processing to check the node IDs written in the file specified by filename.

One node ID per line is written in hexadecimal in the file filename.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

(Example)

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--type *TYPE*[,...]

Specify this option to limit the types of check. You can specify the following types.

- pkg : Executes a status check of only the application of the Technical Computing Suite package.
- conf : Executes a status check of only the reflection of Technical Computing Suite settings.
- plugin : Executes only the check items arbitrarily defined by the system administrator.

If this option is not specified, all the types are assumed specified.

--check *directory*

Specifies a directory for the check definition file.

You can specify an absolute or relative path.

If this option is not specified, the directory /etc/opt/FJSVtcs/plugin/psm/check is assumed to be specified.

When specifying this option and the --type option, be sure to specify the confirmation type plugin.

-d *directory*

Specify a directory as the output destination of the execution result file.

You can specify an absolute or relative path.

If this option is not specified, the current directory is the output destination.

--help

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[NOTES]

Be careful about the following when specifying options.

The --nodegrp, --bootgrp, -n, and -f options cannot be specified at the same time.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can specify only one cluster name in the environment variable PXMYCLST.

An example is shown below.

Example

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The following end codes are returned.

0: All the execution results are OK.

1: The execution results have at least one FAILED or DIFF result.

2: The check definition file format is incorrect.

3: An error such as an incorrect option has occurred.

Check definition file

[NAME]

Check definition file - Configuration file for defining arbitrary check items

[SYNOPSIS]

```
/etc/opt/FJSVtcs/plugin/psm/check/file name or  
arbitrary directory/file name  
* Arbitrary file name
```

[DESCRIPTION]

The check definition file is a configuration file for defining arbitrary check items.

The file is written in yaml format.

Descriptions in the check definition file are divided into section lines and item lines. A section line is written in the "section name:" format. A colon (:) must be added to the end of the section name.

Write only "-" on the line placed between a section line and an item line.

Relative to the "-" line, an item line must be indented by at least one single-byte space.

Since an item line is written in the "key:value" format, it needs a single-byte space (a tab cannot be used) after the colon (:).

If there are multiple items, add a "-" line and an item line.

Comments can be inserted anywhere in the file. A comment begins with the # character and ends at the end of the line.

Place the check definition file in the /etc/opt/FJSVtcs/plugin/psm/check directory or the arbitrary directory on the active system management node. You can see the check items by executing the pachecker command.

If you want to use a check definition file located in the arbitrary directory, the --check option is required.

According to the check targets, the following four sections are available. You can define multiple instances of these sections and mix the sections.

Package section

FileExist section

FileMatch section

Command section

Each section contains the following common items:

NodeType

Model

Action

Package

This section is provided to check the differences between package versions.

Set the following items:

Package

NodeType

Model

Action

Package

Specify the name of the package to check. This item is required.

You can specify multiple package names by delimiting them with a comma (e.g., glibc,glibc-common).

If you want, you can specify the architecture that package corresponds by appending the architecture after a period delimiter(e.g.,glibc.i686).

NodeType

Specify the node type to check. You can specify multiple node types by delimiting them with a comma.

Be sure to specify either NodeType or Model.

The following table shows the target node types. You can also specify the node types of a multiuse node.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
CN	Compute node
LN	Login node
BIO	Boot I/O node
GIO	Global I/O node
SIO	Storage I/O node
SCM	Storage cluster management node
MGS	MGS node
OSS	OSS node
MDS	MDS node
ALL	All node type

Model

Specify the model to check. You can specify multiple models by delimiting them with a comma.

Be sure to specify either NodeType or Model.

You can specify the following models and arbitrary models.

Abbreviation	Model name
FT	FX server
PG	PRIMERGY server

Action

This action outputs the specified message together with execution results when the judgment from the execution results is FAILED.

FileExist

This section is used to check whether a file exists.

The check tries to find a file whose file path, permission, owner, or owner group matches the specified value within the specified target range. The permission, owner, and owner group can be omitted.

Set the following items:

Name

Path

Permission

User

Group

NodeType

Model

Action

Name

Specify a check item name. This item is required.

The maximum length of the character string is 32 characters.

The specified check item name cannot be the same as an existing one in the check definition file.

Path

Specify the absolute path to the target file. This item is required.

Permission

Specify a permission for the target file. This item can be omitted.

Specify a 3- or 4-digit octal value (e.g., 755).

If this item is omitted, the permission is not checked.

User

Specify a UID for the owner of the target file. This item can be omitted.

If this item is omitted, the owner is not checked.

Group

Specify a GID for the owner group of the target file. This item can be omitted.

If this item is omitted, the owner group is not checked.

NodeType

This is the same as NodeType in the Package section.

Model

This is the same as Model in the Package section.

Action

This is the same as Action in the Package section.

FileMatch

This section is used to check the differences between files in the results of a comparison of their contents.

Specify a file on the comparison source node to check for any file matching that file within the target range.

If the type of the target file is binary, compare check sum values instead of comparing the file contents.

This function automatically judges whether a file is binary.

Set the following items:

Name

SrcNodeID

SrcPath

DstPath

NodeType

Model

Action

Name

Specify a check item name. This item is required.

The maximum length of the character string is 32 characters.

The specified check item name cannot be the same as an existing one in the check definition file.

SrcNodeID

Specify the node ID of the comparison source node. This item is required.

You can specify only one node.

SrcPath

Specify the absolute path to the comparison source file. This item is required.

Multiple paths cannot be specified.

DstPath

Specify the absolute path to the comparison destination file. This item is required.

Multiple paths cannot be specified.

NodeType

This is the same as NodeType in the Package section.

Model

This is the same as Model in the Package section.

Action

This is the same as Action in the Package section.

Command

This section is used to check the execution results of an arbitrary command. Besides commands, scripts created by the system administrator can be executed.

Set the following items:

Name

Type

ExecCmd

Expected

NodeType

Model

Action

Name

Specify a check item name. This item is required.

The maximum length of the character string is 32 characters.

The specified check item name cannot be the same as an existing one in the check definition file.

Type

Specify a method for checking an arbitrary command.

Specify ReturnCode to check return values. Specify Diff to check differences.

Be sure to specify either one of the methods.

ExecCmd

Specify the command to execute, or specify a script created by the system administrator. This item is required.

You can also specify command lines with options and pipes. (They need to be enclosed in double quotation marks.)

For a script, the system administrator must have distributed the script to the target node in advance.

The following environment variable is set to the execution process of ExecCmd.

Environment	Description
PSM_CLSTNAME	Cluster name of own node
PSM_NODEID	Node ID of own node

Environment	Description
PSM_NODETYPE	Node type of own node
PSM_MODEL	Model of own node
PSM_MNGNET	IP address of the management network of own node
PSM_CTRLNET	IP address

of the control network of own node

Expected

Specify an expected return value for an executed command. This item can be omitted.

You can specify only numeric characters.

If this item is omitted, 0 is the expected value. If type is Diff, the written value is not used.

NodeType

This is the same as NodeType in the Package section.

Model

This is the same as Model in the Package section.

Action

This is the same as Action in the Package section.

[EXAMPLE]

Examples of the check definition file are shown below.

This example checks differences in the rsh package version on a boot I/O node (BIO) and bash on a node that is a PG model.

```
Package:
-
    Package: rsh
    NodeType: BIO
-
    Package: bash
    Model: PG
```

This example checks for an /etc/modprobe.d/libmlx4.conf file that has a permission of 644 and an owner UID of 0 (root) on the compute cluster management node (CCM) and login node (LN).

```
FileExist:
-
    Name: libmlx4.conf
    Path: /etc/modprobe.d/libmlx4.conf
    NodeType: CCM,LN
    Permission: 644
    User: 0
```

This example checks whether the /etc/hosts files on all nodes are the same as /etc/hosts on 0x00000001.

```
FileMatch:
-
    Name: hosts
    SrcNodeID: 0x00000001
    SrcPath: /etc/hosts
    DstPath: /etc/hosts
    NodeType: ALL
    Action: "pmscatter /etc/hosts"
```

This example checks whether the return value is 0 by executing "/usr/sbin/ip a |grep ib0 |grep UP" from a global I/O node (GIO).

```
Command:
-
    Name: ibup
    Type: ReturnCode
    ExecCmd: "/usr/sbin/ip a |grep ib0 |grep UP"
    Expected: 0
    NodeType: GIO
```

[MESSAGES]

If an error occurs during the execution of the pachecker command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 4101 pachecker Result of the check items contains "FAILED" or "DIFF".

Meaning

The check results include FAILED or DIFF.

Action

Review the set state and target range of the target node, and execute the command again.

[ERR.] PSM 4102 pacheker There is no file in the check definition directory.

Meaning

No file exists in the check definition directory.

Action

Place a check definition file in the check definition directory /etc/opt/FJSVtcs/plugin/psm/check or the directory specified by the --check option, and execute the command again.

[ERR.] PSM 4103 pachekcer Failed writing to the file.(*filename*).

Meaning

Writing to the *filename* file failed.

filename: Output file

Action

Confirm that the file can be write and created, and execute the command again.

[ERR.] PSM 4104 pachecker Format of definition file is incorrect.(*details*)

Meaning

The format of the check definition file is incorrect.

details: Cause of the error

Action

Review the format of the check definition file, and execute the command again.

Warning Message

[WARN] PSM 4111 pachecker Failed to execute command.(cmd=*cmd*, node=*node*)

Meaning

A node failed to execute a batch operation command for a check.

cmd: Batch operation command name

node: Node name

Action

Check the state of the node that failed to execute the batch operation command.

[WARN] PSM 4112 pachecker SrcNodeID does not exist.(Name=*name* SrcNodeID=*srcnodeid*)

Meaning

The SrcNodeID in the check definition file does not exist in the cluster to be checked.

name: Check item name

srcnodeid: The node ID of the comparison source node

Action

Check the SrcNodeID in the FileMatch section of the check definition file. Even if SrcNodeID does not exist in the cluster to be checked, the software environment is checked.

Information Message

[INFO] PSM 4181 pachecker Result of the check items are OK.

Meaning

All check results are OK.

Action

No action is necessary.

**[INFO] PSM 4182 pachecker The execution result file was output.
*filename***

Meaning

The execution result file *filename* was output.

Action

No action is necessary.

4.1.2 paclone command

[NAME]

paclone - Command that creates or restores a disk image

[SYNOPSIS]

```
/usr/sbin/paclone <sub-command>  
[ --help ]
```

sub-command:
{backup, restore, status}

[DESCRIPTION]

The `paclone` command is a command for administrators to create and restore disk images and check their status. This command can be executed only on the active system management node.

[OPTIONS]

`--help`

This option displays how to use this command.

[SUBCOMMAND]

`backup`

The function is to create a disk image of the specified node. The subcommand simultaneously creates an `image.info` file that contains disk image information.

[SYNOPSIS]

```
paclone backup { -m hostname | -c clstname { -n nodeid | --bootgrp bootgid } }
               -d [{hostname|ip address}:]dirname
               [ -k diskname[,...] | -p partition[,...] | -i by-id[,...] | -u uuid[,...] ]
               [ --force ]
               [ --fsck { yes | no | noexec } ]
               [ --power { reboot | off } ]
               [ --kparam kparam ]
paclone backup --help
```

[OPTIONS]

`-m hostname | --hostname hostname`

Specify the host name of the node that is the backup source.

If the specified host name does not exist, the subcommand ends with an error message. Multiple host names cannot be specified.

`-c clstname`

Specify the cluster name of the node that is the backup source.

The environment variable `PXMYCLST` cannot be used.

`-n nodeid`

Specify the node ID of the node that is the backup source. If the specified node ID does not exist, the subcommand ends with an error message. Multiple node IDs cannot be specified.

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option.

`--bootgrp bootgid`

Specify the boot group ID of the BoB that is the backup source. If the specified boot group ID does not exist, the subcommand ends with an error message. Multiple boot group IDs cannot be specified.

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option.

`-d [{hostname|ip address}:]dirname`

This option creates a disk image in the directory specified by *dirname*. You can specify a directory on any node.

The format is `<host name>:<directory>`.

You can specify the host name of a node that is in the node information registered by the installation function.

You can specify an IP address instead of a host name.

The "`<host name>`:" part can be omitted.

If omitted, a directory on the command execution node (active system management node) is assumed specified.

The specified path to the directory must be an absolute path. This option must be specified.

(Example) Specifying `/work/backup/imagedir` on server1

```
-d server1:/work/backup/imagedir
```

If the `-m` option is specified together with this option, the subcommand creates a subdirectory with a name in the `hostname_YYYYM-MDDhhmmss` format under *dirname*, and then a disk image in the subdirectory.

If the `-n` option is specified together with this option, the subcommand creates a subdirectory with a name in the

cluster_nodeid_YYYYMMDDhhmmss format under *dirname*, and then a disk image in the subdirectory.

If the `--bootgrp` option is specified together with this option, the subcommand creates a subdirectory with a name in the cluster_bootgid_YYYYMMDDhhmmss format under *dirname*, and then a disk image in the subdirectory.

`-k diskname[,...] | --disk diskname[,...]`

Specify the device name of the source disk for creating a disk image.

You can specify multiple device names by delimiting them with a comma.

(Example) `-k sda,sdb`

`-p partition[,...] | --part partition[,...]`

Specify the device name of the source partition for creating a disk image.

You can specify multiple device names by delimiting them with a comma.

(Example) `-p sda1,sda2`

`-i by-id[,...] | --by-id by-id[,...]`

Specify the by-id of the source disk or partition for creating a disk image. Mixing of disks and partitions can not be specified.

You can specify multiple by-ids by delimiting them with a comma.

(Example) Specifying by-id names (sda1,sda2)

`-i scsi-3500000e01ee1ddc0-part1,scsi-3500000e01ee1ddc0-part2`

`-u uuid[,...] | --uuid uuid[,...]`

Specify the UUID of the source partition for creating a disk image.

You can specify multiple UUIDs by delimiting them with a comma.

(Example) Specifying UUIDs (sda1,sda2)

`-u 8e39c712-f01b-409e-b52f-fec32b53d58c,b135db24-9a38-4ce4-8073-ca15ec416344`

`--fsck { yes | no | noexec }`

Specify the operation of the file system before the disk image is created. You can specify the following operations.

yes:

The operation for a fat/ext file system is to execute the `fsck -y` command.

The `fsck -y` command response is "yes" to all repair inquiries that may arise during a file system check.

The operation for an xfs file system is to execute the `xfs_repair` command to perform repairs.

no:

The operation for a fat/ext file system is to execute the `fsck -n` command.

The `fsck -n` command response is "no" to all repair inquiries that may arise during a file system check.

A file system error causes processing to terminate abnormally without performing a backup.

The operation for an xfs file system is to execute the `xfs_repair -n` command to end processing without performing repairs.

noexec:

The operation is to not check the file system.

If this option is not specified, "no" is assumed specified.

`--power { reboot | off }`

Specify the operation of the target node after the disk image is created. You can specify the following operations.

reboot:

This operation restarts the target node.

off:

This operation powers off the target node.

If this option is not specified, "reboot" is assumed specified.

`--force`

This option performs a backup without a yes/no prompt asking whether to execute it.

`--kparam kparam`

Specify a kernel parameter to pass to the backup and restore kernels.

`--help`

This option displays how to use this subcommand.

[SUBCOMMAND]

restore

The function is to restore a disk image, created by the backup function, on the specified node.

The subcommand creates a partition on the restore destination node based on the partition configuration in the `image.info` file created at the backup time. Then, it restores the disk image.

With this function, you can use a plugin (referred to below as a "restore plugin") that executes arbitrary post-restore processing on the target node.

[SYNOPSIS]

```
pacclone restore { -m hostname[,...]
                  | -c clstname { -n nodeid[,...] | --bootgrp bootgid[,...] | -f filename }
                  | -f filename }
                  -d [{hostname|ip address}:]dirname
                  [ -p partition[,...] ]
                  [ --force ]
                  [ --power { reboot | off } ]
                  [ --kparam kparam ]
                  [ --plugin pluginlist [ --clear-plugin ] ]
pacclone restore --help
```

[OPTIONS]

`-m hostname[,...] | --hostname hostname[,...]`

Specify the host name of the node that is the restore target.

If the specified host name does not exist, the subcommand ends with an error message.

You can specify multiple host names by delimiting them with a comma (e.g., `host01,host02,host03`).

`-c clstname`

Specify the cluster name of the node that is the restore target.

The environment variable `PXMYCLST` cannot be used.

`-n nodeid[,...]`

Specify the node ID of the node that is the restore target.

If the specified node ID does not exist, the subcommand ends with an error message.

When specifying this option, be sure to specify a cluster name in `clstname` in the `-c` option.

You can specify multiple node IDs by delimiting them with a comma (e.g., `0xFFFF0003,0xFFFF0004,0xFFFF0005`).

You can specify a range of IDs by using a hyphen (e.g., `0xFFFF0003-0xFFFF0005`).

If the range specified with a hyphen includes a node ID that does not exist, the node ID is ignored and processing continues.

`--bootgrp bootgid[,...]`

Specify the boot group ID of the BoB that is the restore target.

If the specified boot group ID does not exist, the subcommand ends with an error message.

When specifying this option, be sure to specify a cluster name in `clstname` in the `-c` option.

You can specify multiple boot group IDs by delimiting them with a comma (e.g., `0x0101,0x0102,0x0103`).

You can specify a range of IDs by using a hyphen (e.g., `0x0101-0x0103`).

If the range specified with a hyphen includes a boot group ID that does not exist, the boot group ID is ignored and processing continues.

`-d [{hostname|ip address}:]dirname`

This option restores a disk image from directly under the directory specified by `dirname` to the destination node. You can specify a directory on any node.

The format is `<host name>:<directory>`.

You can specify the host name of a node that is in the node information registered by the installation function.
You can specify an IP address instead of a host name.
The "<host name>:" part can be omitted.
If omitted, a directory on the command execution node (active system management node) is assumed specified.
The specified path to the directory must be an absolute path. This option must be specified.

-f *filename*

Write the restore targets in the file specified by *filename*.
If the specified target node does not exist, the subcommand ends with an error message.
With the -c option specified, you can write node IDs or boot group IDs in the file *filename*.
With the -c option not specified, you can write host names in the file *filename*.
Write one node ID or node group ID in hexadecimal per line.
Write one host name per line.
Lines with the comment symbol (#) written at the beginning are ignored.

(Example) Node IDs

```
0xFFFF0003
0xFFFF0004
0xFFFF0005
```

(Example) Boot group IDs

```
0x0101
0x0102
0x0104
```

(Example) Host names

```
host01
host02
host03
```

-p *partition*[,...] | --part *partition*[,...]

Specify the partition whose disk image is the restore target.
If the specified disk image partition does not exist, an error occurs.
If this option is not specified, the restore target is all the disk images in the directory specified by the -d option.
If this option is specified, no partitioning operations are performed on the restore destination node.

--power { reboot | off }

Specify the operation of the target node after restore. You can specify the following operations.

reboot:

This operation restarts the target node.

off:

This operation powers off the target node.

If this option is not specified, "reboot" is assumed specified.

--force

This option executes restore without a yes/no prompt asking whether to execute it.

--kparam *kparam*

Specify a kernel parameter to pass to the restore kernel.

--plugin *pluginlist*

Specify a file that contains a list of plugin files, in *pluginlist*. The files will be executed as restore plugins.
By splitting them up on separate lines, you can specify multiple plugin files in the file.
Specify *pluginlist* with an absolute or relative path (e.g., --plugin /tmp/list).
Using an absolute or relative path, specify one plugin file per line in *pluginlist*.
Lines beginning with a comment, blank lines, blanks at the beginning or end of lines, and tabs in *pluginlist* are ignored. Inline comments cannot be used.
Multiple files cannot be specified with the same file name.

Plugin files are executed in the order listed in *pluginlist*. Multiple *pluginlist* files cannot be specified.

Executed restore plugins are placed in /root/paclone_plugin on the restore destination node.

If the restore plugin already exists in /root/paclone_plugin, that existing file is deleted, and then the new restore plugin is placed there.

--clear-plugin

Specify this option not to place restore plugins in /root/paclone_plugin on the restore destination node after execution of the restore plugins.

--help

This option displays how to use this subcommand.

[SUBCOMMAND]

status

This subcommand displays the progress of backup or restore on the backup or restore destination nodes, and the success or failure of processing.

[SYNOPSIS]

```
paclone status show
paclone status clear [ -s status | -m hostname[...] | -f hostfile ]
paclone status --help
```

[ACTIONS]

show

This action displays the progress of backup or restore on the backup or restore destination nodes, and the success or failure of processing.

clear

This action clears a node from the backup or restore completed state.

The only information that can be cleared is information about nodes where backup or restore has not begun or has completed successfully.

Information about nodes where backup or restore is in progress or has failed is not cleared.

To clear information about nodes where backup or restore is in progress or has failed, specify the --hostname option or -f option.

[OPTIONS]

-m hostname[,...] | --hostname hostname[,...]

Specify the name of the host targeted for clearing.

If status information is not registered for the specified host name, the host name is ignored and processing continues. You can specify multiple host names by delimiting them with a comma.

You can specify this option only when the clear option is specified.

-s status | --status status

Specify the STATUS targeted for clearing in status.

All node information of the specified STATUS is no longer displayed.

You can specify NotStarted or Completed for the STATUS.

If Completed is specified, only node information in the Succeeded STATE is cleared.

You can specify this option only when the clear option is specified.

-f hostfile

Specify the file that contains the host names of the nodes targeted for clearing.

Write one node per line.

If this option is specified, information is cleared regardless of the STATUS or STATE.

If status information is not registered for the specified host name, the host name is ignored and processing continues.

You can specify this option only when the clear option is specified.

(Example)

host001
host002
host003

--help

This option displays how to use this subcommand.

[NOTE]

- The NFS publishing setting must be implemented in advance for the directory specified by the -d option. It is also necessary to grant read and write permissions to the root user.
- To specify a host name, the installation function must have already completed registration of its node information.

[When the backup subcommand is specified]

- The -m and -c options cannot be specified at the same time.
- The -k, -k, -i, and -u options cannot be specified at the same time.

[When the restore subcommand is specified]

- The -m and -c options cannot be specified at the same time.
- The restore plugins are executed from the current directory, which is /root.
- If the restore plugin end code is 0, it is treated as a success. If the code is not 0, it is treated as a failure.
- Restore processing neither terminates abnormally nor stops according to restore plugin execution results.

[EXIT STATUS]

The following exit status are returned.

0: Normal end

1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the paclone command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "[4.2.2 Common command messages of the install function.](#)"

Error Message

[ERR.] INST 3022 paclone Cannot backup the disk or partition.(details)

Meaning

The disk or partition cannot be backed up.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 3023 paclone Cannot restore the disk image.(details)

Meaning

The disk image cannot be restored.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 3024 paclone The specified node information is not registered.(*details*)

Meaning

The specified node is not registered in node information.

details: Cause of the error

Action

Register the node information, and execute the command again.

[ERR.] INST 3025 paclone Command Failed.(*details*).

Meaning

The command failed.

details: Cause of the error

Action

Remove the cause, referring to the contents of the message output immediately after the error, and execute the command again.

[ERR.] INST 3026 paclone No data available.(*details*)

Meaning

No applicable information, such as execution target nodes, exists.

details: Cause of the error

Action

Review the option arguments, and execute the command again.

[ERR.] INST 3027 paclone System management function is not installed.

Meaning

The system management function is not installed on the active system management node.

Action

If the system management function is not installed, backups and restores that specify the node ID or boot group ID are not allowed. Specify the host name of the target node, and execute the command again.

[ERR.] INST 3028 Failed to get host information.(*details*)

Meaning

Host information cannot be obtained.

details: Cause of the error

Action

Register host information, and execute the command again.

[ERR.] INST 3029 paclone Failed to distribute or replicate process plugin-ins.(*details*)

Meaning

The plugin cannot be distributed or replicated.

details: IP address of the relay node that failed to distribute and replicate the plugin.

Action

If the relay node is not started, start it and execute the command again. If the relay node is started, there may be a reason such as a communication error or space exhaustion of the distribution destination. Confirm that the status of the relay node, remove the cause, and execute the command again.

[ERR.] INST 3030 paclone Power operation error.(*details*)

Meaning

The power operation for the target node failed.

details: Cause of the error

Action

Confirm the operability of the power of the target node, and execute the command again.

[ERR.] INST 3031 paclone Not found plug-in file.(*filename*)

Meaning

The plugin file *filename* was not found.

Action

Specify an existing file, and execute the command again.

[ERR.] INST 3032 paclone Invalid plug-in file name.(*details*)

Meaning

The plugin file name does not meet naming conventions.

details: Cause of the error

Action

In accordance with naming conventions, define a file name with up to 63 characters, including single-byte alphanumeric characters, the hyphen, the underscore, and the period. However, the first character cannot be a hyphen.

Information Message

[INFO] INST 3201 paclone Command succeeded.(*details*).

Meaning

The command completed successfully.

details: Function name (backup or restore)

Action

No action is necessary.

4.1.3 paclstmgr command

[NAME]

paclstmgr - Manages the incorporation of nodes into operation

[SYNOPSIS]

```
/usr/sbin/paclstmgr [-c clstname] --disable [--ic]
                    scope options
                    [--nodetype nodetype[,...]]
```

```

        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [-k]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --hard-mainte [--force]
scope options
        [--nodetype nodetype[,...]]
        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [-k]
        [-P num[:model][,num[:model],...]]
        [--interval num[:model][,num[:model],...]]
        [--interval-plugin filename]
        [--interval-option filename]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --soft-mainte [--force]
scope options
        [--nodetype nodetype[,...]]
        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [-k]
        [--no-service-stop | --service-stop]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --recover [--disable [--ic]]
scope options
        [--nodetype nodetype[,...]]
        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [--no-restart | --service-restart | --warm-reboot | --cold-reboot]
        [--force-all]
        [-P num[:model][,num[:model],...]]
        [--interval num[:model][,num[:model],...]]
        [--interval-plugin filename]
        [--interval-option filename]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --enable
scope options
        [--nodetype nodetype[,...]]
        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --service {stop | start | restart}
scope options
        [--nodetype nodetype[,...]]
        [--excludetype nodetype[,...]]
        [--model model[,...]]
        [--plugin-pre plugindir]
        [--plugin-post plugindir]
/usr/sbin/paclstmgr [-c clstname] --set-failover {enable | disable | list}
        {-n nodeid[,...] | -f filename}
/usr/sbin/paclstmgr [-c clstname] --failover nodeid [--force]
/usr/sbin/paclstmgr [-c clstname] --failback nodeid [--force]
/usr/sbin/paclstmgr --help

scope options(exclusive):
        -a
        --nodegrp nodegid[,...]
        --bootgrp bootgid[,...]

```



```

-n nodeid[,...] [--giogrp | --siogrp]
--cmu nodeid[,...] [--giogrp | --siogrp]
--coord xcoord,ycoord
--rscunit rscuname[,...]
-f filename

```

[DESCRIPTION]

`pacstmgr` is a command for system administrators, which prohibits the assignment of new jobs to compute nodes and switches nodes so that the nodes cannot be used during hardware or software maintenance.

This command can be executed only on the active system management node.

When this command is executed, power control may be exercised such as powering off or restarting a node. In these cases of power control, you can control the number of concurrently executed jobs by specifying the `-P` option, `--interval` option, `--interval-plugin` option, or `--interval-option` option.

If the `--hard-mainte` option is specified, power-off is executed by the `papwrctl` command.

Under the following conditions, a restart is executed by the `papwrctl` command. However, if the `-a` option is specified, the `pastop` and `pastart` commands are used for the restart.

- The `--recover` option and `--warm-reboot` option are specified.
- The `--recover` option and `--cold-reboot` option are specified.

If the `--recover` option is specified and the nodes are incorporated from a state other than the software maintenance state (SoftMaintenance), power-on is executed by the `papwrctl` command.

[OPTIONS]

`-c clstname`

This option operates on the cluster specified by *clstname*.

If this option is not specified, the cluster name set by the environment variable `PXMYCLST` is used.

`-n nodeid[,...]`

If you want to narrow down the execution targets to a specific node in the cluster, specify the node ID in *nodeid*. Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., `0x01010010,0x01010012,0x0101002A`). You also can specify

a range with a hyphen (e.g., `0x01010010-0x0101002A`). If the specified node ID is incorrect or does not exist in the cluster, the node ID is ignored and processing continues.

If the `--set-failover` option is specified, processing includes the node paired with the specified node in a redundant configuration.

`--giogrp`

If you want to narrow down the execution targets to a specific GIO group in the cluster, specify this option as a sub option of the `-n` option or `--cmu` option.

The unit of operation for FX server is a GIO group, and nodes in the same GIO group as the node specified by the `-n` option or `--cmu` option are gathered in the execution targets.

For nodes that are not FX server, the `--giogrp` option specification is ignored.

`--siogrp`

If you want to narrow down the execution targets to a specific SIO group in the cluster, specify this option as a sub option of the `-n` option or `--cmu` option.

The unit of operation for FX server is an SIO group, and nodes in the same SIO group as the node specified by the `-n` option or `--cmu` option are gathered in the execution targets.

For nodes that are not FX server, the `--siogrp` option specification is ignored.

`--nodegrp nodegid[,...]`

If you want to narrow down the displayed objects to a specific node group in the cluster, specify the node group ID in *nodegid*. Write *nodegid* in hexadecimal.

You can specify multiple node GIDs by delimiting them with a comma (e.g., `0x01,0x02,0x03`). You also can specify a range with a hyphen (e.g., `0x01-0x03`).

If the specified node group ID is incorrect or does not exist in the cluster, the node group is ignored and processing continues.

--bootgrp *bootgid*[,...]

If you want to narrow down the displayed objects to a specific boot group in the cluster, specify the boot group ID in *bootgid*. Write *bootgid* in hexadecimal.

You can specify multiple boot GIDs by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103). You also can specify a range with a hyphen (e.g., 0x0101-0x0103).

If the specified boot group ID is incorrect or does not exist in the cluster, the boot group is ignored and processing continues.

--cmu *nodeid*[,...]

If you want a node in the CMU of the FX server in the cluster to be included in the execution targets, specify the node ID in *nodeid*. Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

You also can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

If the specified node ID is incorrect or does not exist in the cluster, the node ID is ignored and processing continues.

The unit of operation is a CMU, and nodes in the same CMU as the specified node are gathered in the execution targets.

If the specified node ID is not for a FX server, the command runs in the same way as when the -n option is specified.

--coord *xcoord*,*ycoord*

If you want to narrow down the controlled objects by using the coordinates of an FX server in a specific Tofu logical configuration in the cluster, specify the coordinates in *xcoord* and *ycoord*. With the coordinates, you can specify not only a specific location but also a range with a hyphen and all for the entire coordinate axis.

[Example]

--coord 4,1-5: The operated object that have an X-coordinate of 4 and a Y-coordinate from 1 to 5.

--coord 4,all: The operated object that have an X-coordinate of 4.

--coord all,5: The operated object that have a Y-coordinate of 5.

If the specified coordinate is incorrect or does not exist in the cluster, the coordinate is ignored and processing continues.

--rscunit *rscuname*[,...]

If you want to narrow down the displayed objects to a specific resource unit in the cluster, specify the resource unit name in *rscuname*.

You can specify multiple resource unit names by delimiting them with a comma (e.g., unitA,unitB).

If the specified resource unit name is incorrect or does not exist in the cluster, the resource unit is ignored and processing continues.

-a

This option targets all the nodes in a cluster, except the system management node and multiuse node.

-f *filename*

This option executes the process for node IDs written in the file that is specified by *filename*.

In the *filename* file, write each node ID in hexadecimal on one line.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

0x01010010

0x01010011

0x0101002A

--nodetype *nodetype*[,...]

If you want to narrow down the execution targets to only a specific node type in the cluster, specify the node type in *nodetype*.

You can specify multiple node types by delimiting them with a comma (e.g., CMM,CCS).

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
CN	Compute node

Abbreviation	Node type
LN	Login node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

If the specified node type does not exist in the cluster, the node type is ignored and processing continues.

If the nodes of the specified type are in a redundant configuration, operations are performed on both the active and standby nodes.

--excludetype nodetype[,...]

If you want to exclude a specific node type in the cluster from the execution targets, specify the node type in *nodetype*. You can specify multiple node types by delimiting them with a comma (e.g., BIO,GIO).

Even if a node type is specified as an execution target by the --nodetype option, specification with this option to exclude it has priority.

The same node types that can be specified are the same as for the --nodetype option.

--model model[,...]

If you want to narrow down the execution targets to only a specific model in the cluster, specify the model in *model*.

You can specify multiple models by delimiting them with a comma (e.g., FT,PG).

You can specify the following models and any models defined in the cluster.

Abbreviation	Model name
PG	PRIMERGY server
FT	FX server

If the specified model does not exist in the cluster, processing terminates.

--hard-mainte

This option specifies hardware maintenance.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--soft-mainte

This option specifies software maintenance.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

When specifying this option, also specify the --no-service-stop option or --service-stop option. If they are omitted, processing continues as though the --service-stop option were specified.

--no-service-stop

Specify this option to perform software maintenance without stopping the Job Operation Software service.

When specifying this option, specify the --no-restart option for the processing by the --recover option to incorporate nodes into operation from maintenance mode.

--service-stop

Specify this option to stop the Job Operation Software service to perform software maintenance.

When specifying this option, specify an option other than the --no-restart option for the processing by the --recover option to incorporate nodes into operation from maintenance mode.

--disable [--ic]

This option prevents new jobs from being assigned to compute nodes.

You can also prevent that assignment of new jobs that use interconnects mounted on the target compute nodes as communication paths, by specifying the --ic option.

--service

Specify this option to perform Job Operation Software service operations such as in recovery processing after trouble occurs.

If start is the specified argument, the service starts. If stop is specified, the service stops. If restart is specified, the service restarts.

--set-failover

This option enable or disables automatic failover and displays setting information.

Nodes with a node type and redundant configuration in the target scope are configured and their setting status is displayed.

If enable is the specified argument, failover is enabled. If disable is specified, failover is disabled. If list is specified, the failover setting status is displayed.

The --failover option of this command specifies manual failover, which is executed without being subject to the disable setting.

Also, automatic failover is disabled for the following kinds of nodes regardless of the setting of this option. One kind is nodes placed in software maintenance mode by specification of the --soft-mainte option. The other kind is nodes placed in hardware maintenance mode by specification of the --hard-mainte option.

--failover *nodeid*

This option generates a failover and switches the active node. *nodeid* specifies the node ID of the active node that is the switching destination.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

You cannot specify a node ID that is incorrect, does not belong to a redundant node, or does not exist in the cluster.

--fallback *nodeid*

This option switches back to a node. *nodeid* specifies the node ID of the node that is the switchback destination.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

You cannot specify a node ID that is incorrect, does not belong to a redundant node, or does not exist in the cluster.

-k

This option immediately re-executes the following jobs for a transition to maintenance mode:

running jobs on compute nodes, and jobs using interconnects mounted on compute nodes as communication paths. If the target node is not a compute node, this option is ignored.

--recover [--disable [--ic]]

This option incorporates nodes in maintenance mode (HardMaintenance or SoftMaintenance state) into operation.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

If the target nodes are in a mixed state, you cannot incorporate them into operation.

A node with a hardware failure or other problem cannot be incorporated into operation directly. You need to first change it to maintenance mode and then specify this option to incorporate it into operation.

This command changes target compute nodes to Manual state not to allocate new jobs, if it is executed with --disable option. If --disable option is specified, the ICC route of the compute node is incorporated into operation.

Specify the --ic option to suppress the use of interconnects mounted on the target compute nodes as job communication paths.

--no-restart

This option incorporates nodes into operation from maintenance mode without restarting the Job Operation Software service.

The option is valid only at recovery from software maintenance mode.

--service-restart

This option restarts the Job Operation Software service to incorporate nodes into operation from maintenance mode.

The option is valid only at recovery from software maintenance mode.

--warm-reboot

This option executes a high-speed restart without a hardware reset to incorporate nodes into operation from maintenance mode.
The option is valid only at recovery from software maintenance mode.
If the specified nodes are not FX server, the command runs in the same way as when the --cold-reboot option is specified.

--cold-reboot

This option restarts nodes to incorporate them into operation from maintenance mode.
The option is valid only at recovery from software maintenance mode.

--force-all

Specify this option to incorporate a mix of nodes in software maintenance mode and nodes in hardware maintenance mode into operation.

-P num[:model][,num[:model],...] | --concurrency num[:model][,num[:model],...]

You can specify the number of targets for simultaneous power control in num.
You can specify a value ranging from 1 to the total number of nodes in the specified cluster. However, the specified value may exceed the number of nodes matching the option-specified conditions. In that case, the command runs as though the specified maximum number of nodes is the number of nodes matching the option-specified conditions.
If this option is not specified, the value is determined according to the papwrctl command specifications.
Also, you can vary the specified value by model by specifying each model in model.
If the specifications have both an overall value specified and a specified value for each model, the specified value for each model has priority.
If no model is specified, all models are targets.
This option is passed to the papwrctl command, pastart command, and pastop command in cases where power control is exercised with the --hard-mainte and --recover options specified.

--interval num[:model][,num[:model],...]

You can specify a power control interval in milliseconds for individual maintenance targets.
You can specify a value ranging from 0 to 86400000.
If this option is not specified, the value is determined according to the papwrctl command specifications.
Also, you can vary the specified value by model by specifying each model in model. If the specifications have both an overall value specified and a specified value for each model, the specified value for each model has priority.
If multiple values are specified for a model, an error occurs.
This option is passed to the papwrctl command, pastart command, and pastop command in cases where power control is exercised with the --hard-mainte and --recover options specified.

--interval-plugin filename

Specify the path of the plugin command executed at the power control interval. The specified command is executed before the initial control processing and at every interval.
If the plugin command does not exist, this command terminates abnormally. When the plugin command return code is other than 0, this command continues processing.
This option is passed to the papwrctl command, pastart command, and pastop command in cases where power control is exercised with the --hard-mainte and --recover options specified.

--interval-option filename

Specify the path to the file for changing the following at the power control interval: the number of targets for simultaneous power control, and the power control interval for individual control targets.
After the plugin command is executed, the specified file is read before the initial control processing and at every interval.
You can write the -P and --interval options in the file. Other options, even when written, are ignored.
This option is passed to the papwrctl command, pastart command, and pastop command in cases where power control is exercised with the --hard-mainte and --recover options specified.

--plugin-pre plugin

Specify the path to the directory where the plugin command that is executed before node switching to the operating state has been placed.
You can specify this option only when the --enable, --disable, --hard-mainte, --soft-mainte, --recover, or --service option is specified.
The pacstmgr command waits for the return of the plugin command to continue processing. If the plugin command does not exist or terminates abnormally (the return code is other than 0), the pacstmgr command terminates abnormally.

--plugin-post plugin

Specify the path to the directory where the plugin command that is executed after node switching to the operating state has been placed. You can specify this option only when the --enable, --disable, --hard-mainte, --soft-mainte, --recover, or --service option is specified. The pacstmgr command waits for the return of the plugin command to continue processing. If the plugin command does not exist or terminates abnormally (the return code is other than 0), the pacstmgr command terminates abnormally.

--enable

This option incorporates nodes in Manual or AbnormalEnd state.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--help

This option displays information on how to use this command.
If this option is specified, all the arguments and other options are ignored.

[NOTE]

Keep the following in mind when specifying options or arguments.

- You cannot specify the -a, --nodegrp, --bootgrp, -n, --cmu, --coord, --rscunit, and -f options at same time.
- You cannot specify the --hard-mainte, --soft-mainte, --disable, --set-failover, --failover, --failback, --recover, --service, and --enable options at same time.
- You cannot specify the --warm-reboot and --cmu options at the same time.
- If --failover or --failback option is specified, it's cannot be canceled by "Ctrl + C".

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option. Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

[papwradm\(8\)](#)

[MESSAGES]

If an error occurs during the execution of the pacstmgr command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 4001 paclstmgr The status of nodes is not Disable.

Meaning

The status of the target node is not Disable.

Action

Review the status of the target node, and execute the command again.

[ERR.] PSM 4002 paclstmgr Cannot recover by mixed status.

Meaning

Recovery is not possible because the status is mixed.

Action

Set the status of nodes in the specified range such that they are all the same.

[ERR.] PSM 4003 paclstmgr The system monitoring function is not active.

Meaning

The system monitoring function is not working, and processing cannot continue.

Action

Execute the command again on the active system management node.

[ERR.] PSM 4004 paclstmgr Nodes outside the redundant target is specified.

Meaning

The specified node is not switched by the paclstmgr command.

The node that can switch by the paclstmgr command is as follows.

- System management node
- Compute cluster management node
- Compute cluster sub management node
- Storage cluster management node
- MGS node
- MDS node
- OSS node

Action

Review the specified node, and execute the command again.

[ERR.] PSM 4005 paclstmgr Specified node is not active.

Meaning

The specified node is not the active node.

Action

Specify the active node.

[ERR.] PSM 4006 paclstmgr Node that is not redundant is specified.

Meaning

The specified node is not the redundant configuration node.

Action

Review the specified node, and execute the command again.

[ERR.] PSM 4008 paclstmgr Nodes outside the target is specified switched back.

Meaning

The specified node is a node that cannot switch back.

The node that can switch back is as follows.

- MGS node
- MDS node
- OSS node

Action

Review the specified node, and execute the command again.

[ERR.] PSM 4009 paclstmgr Target node is not ready to be switched back.

Meaning

The specified node is not ready to switch back.

Action

1. If you didn't specify the switched node, switch back it after switching the target node.
2. If you specified the switched node, confirm it is status that can switch back. After that, switch back it.
3. If it is status to which the node cannot be switched back due to factors other than 1 and 2, collect investigation data about the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] PSM 4010 paclstmgr Node to be the failover pair is not in one side.

Meaning

The status is not one-sided operation on the node of the pair of the redundant configuration.

Action

Confirm the status is one-sided operation on the node of the pair of the redundant configuration. If status is not one-sided operation on it, collect investigation data about the system management node according to the "Job Operation Software Administrator's Guide for Maintenance". Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] PSM 4011 paclstmgr Cannot access specified directory. (*directory*)

Meaning

The specified directory is not accessible.

directory: Specified directory

Action

Confirm that the path and privileges to the specified directory are correct, remove the cause, and execute the command again.

[ERR.] PSM 4012 paclstmgr Plugin is not found in specified directory. (*directory*)

Meaning

The specified directory does not have the plugin.

directory: Specified directory

Action

Specify the directory containing the plugin, and execute the command again.

[ERR.] PSM 4013 paclstmgr File permission of plugin is invalid. (*filename*)

Meaning

The plugin privileges are incorrect.

filename: Plugin file

Action

Set 0550 or 0700 for the permission of the plugin.

[ERR.] PSM 4014 paclstmgr Failover failed.(*details*)

Meaning

Failover failed.

details: Cause of the failover failure (Example: Situation where the standby node cannot be promoted to be the active node)

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 4016 paclstmgr The status of related nodes is not Disable.

Meaning

The status of nodes related to the specified node is not Disable.

Action

Isolate nodes related to the specified node from job operation.

Alternatively, add the -k option to the paclstmgr command, and execute the command again.

[ERR.] PSM 4017 paclstmgr Recover operation is not executed, because of related node is running.

Meaning

Recovery failed because a control instruction involved in stopping the boot I/O node or storage I/O node was issued while the related node was running.

Action

Specify all the nodes in the boot group, and execute the command again.

[ERR.] PSM 4018 paclstmgr The ARCH_STATUS of nodes is not ICC_Disable or ICC_Stopped.

Meaning

ARCH STATUS of the target node is not ICC_Disable or ICC_Stopped.

Action

This is because the ICC is available for the target compute node. Use the --disable --ic option of the paclstmgr command to prevent the interconnect on the target compute node from being used as a communication path for jobs, and then execute the command again.

Warning Message

[WARN] PSM 4051 paclstmgr Execute init to panic the node specified with *nodeid*.**Meaning**

If a failover is performed when the FEFSSR service is not active, papwrctl init is executed on the specified node (*nodeid*).

Action

No action is necessary.

[WARN] PSM 4052 paclstmgr There are nodes that failed in the operation.(*filename*)**Meaning**

There is a node that is inoperable.

filename: Error file (list of IDs of inoperable nodes)

Action

Remove the cause, referring to the error file *filename*, and execute the command again for the inoperable node.

[WARN] PSM 4053 paclstmgr Active node is specified.**Meaning**

The specified node is active.

Action

Check again whether job operation is affected. If there is no problem, continue processing.

Information Message

[INFO] PSM 4081 paclstmgr Abnormal network status of standby SMM stopped failover process.**Meaning**

Failover was not executed because of trouble in the network of the standby system management node (SMM).

Action

No action is necessary.

[INFO] PSM 4082 paclstmgr Failure of failover makes the active node active again.**Meaning**

Failover failed. The active node is promoted again.

Action

No action is necessary.

[INFO] PSM 4083 paclstmgr Failover can't be done, because target node has no failover pair node.**Meaning**

Failover was not executed because there is no failover pair.

Action

No action is necessary.

[INFO] PSM 4084 paclstmgr Failover process starts now. *nodeid1* is going to failover to *nodeid2*.**Meaning**

Failover started. The node ID *nodeid1* was switched to *nodeid2*.

Action

No action is necessary.

[INFO] PSM 4085 paclstmgr Failover process finished normally. *nodeid1* was failovered to *nodeid2*.

Meaning

Failover has been completed. The node ID *nodeid1* was switched to *nodeid2*.

Action

No action is necessary.

[INFO] PSM 4086 paclstmgr No change of the status occurred.

Meaning

The status has not changed.

Action

No action is necessary.

4.1.4 paconsole command

[NAME]

paconsole - Makes console connection.

[SYNOPSIS]

```
/usr/sbin/paconsole [ -c clstname ] -n nodeid  
/usr/sbin/paconsole --help
```

[DESCRIPTION]

paconsole is a command for system administrators, which makes console connection the specified nodes.
This command can be executed on the system management node.

[OPTIONS]

-c *clstname*

This option connects to the console of the cluster specified in *clstname*.

If no cluster name is specified, the cluster name set by the environment variable PXMYCLST is specified.

-n *nodeid*

If you want to narrow down the displayed objects to a specific node in the cluster, specify a node ID in *nodeid*.

Write *nodeid* in hexadecimal.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

Example

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[NOTES]

"#." is inputted when escaping from OS console.

[MESSAGES]

If an error occurs during the execution of the paconsole command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 5301 paconsole Connection to the console failed.

Meaning

The console connection failed.

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] PSM 5302 paconsole This model is not supported. (Model: *model*)

Meaning

Console connection to the model of the specified node is not supported.

Action

Nodes of models other than the "FT" and "PG" models cannot be connected to the console.

Information Message

[INFO] PSM 5381 paconsole Connection to the console closed.

Meaning

Connection to the console has been completed.

Action

No action is necessary.

4.1.5 padeadline command

[NAME]

padeadline - Shows the setting and setting of schedule at deadline.

[SYNOPSIS]

```
/usr/sbin/padeadline [-c clstname] [--ph]
                        [--force]
                        {-a|--nodegrp nodegid[,...]|--bootgrp bootgid[,...]
                        |{--rscunit|--ru} rscuname[,...]|{--rscunit|--ru} rscuname {--rscgrp|--rg}
                        rscgname[,...]
                        |--cmu nodeid[,...]
                        |-n nodeid[,...]|--coord xcoord,ycoord|-f nodelist}
                        --period start={YYYYmmdd[hhmm]|now}[ ,end={YYYYmmdd[hhmm]|unlimited}]
                        [--enforce] [--ic]
/usr/sbin/padeadline [-c clstname] [--pg]
                        [--force]
                        {-a|--nodegrp nodegid[,...]
                        |{--rscunit|--ru} rscuname[,...]|{--rscunit|--ru} rscuname {--rscgrp|--rg}
                        rscgname[,...]
                        |-n nodeid[,...]|-f nodelist}
                        --period start={YYYYmmdd[hhmm]|now}[ ,end={YYYYmmdd[hhmm]|unlimited}]
                        [--enforce] [--ic]
/usr/sbin/padeadline [-c clstname] [--force] --cancel N[,...]
/usr/sbin/padeadline [-c clstname] [--ph|--pg] [{--rscunit|--ru} rscuname[,...]] [{--rscgrp|--rg}
                        rscgname[,...]] --show
/usr/sbin/padeadline --help
```

[DESCRIPTION]

padeadline command is a command for the manager to display setting/release of the deadline when the job is not allocated schedule and a set content for the period when the resource specified in the option was specified.

When an FX server compute node and a PRIMERGY server compute node are included in the same cluster, it is necessary to specify by optional --ph option(FX server) or --pg option(PRIMERGY server) which of the FX server or PRIMERGY server compute node is to be the target.

This command can be executed on the system management node and the compute cluster management node.

[OPTIONS]

-c *clstname*

The schedule of the deadline of the cluster *clstname* is set, and displayed.

When *clstname* is set to environment variable PXMYCLST, this option is omissible.

When this option is specified, this option is disregarded in execution by the compute cluster management node.

--ph

In the environment that an FX server compute node and a PRIMERGY server compute node are included in the same cluster, when the deadline schedule is set to FX server, it is necessary to specify this option.

The specification of this option is unnecessary in FX server single environment.

--pg

In the environment that an FX server compute node and a PRIMERGY server compute node are included in the same cluster, when the deadline schedule is set to PRIMERGY server, it is necessary to specify this option.

The specification of this option is unnecessary in PRIMERGY server single environment.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

-a

The deadline schedule is set to all the compute nodes in the cluster.

--nodegrp *nodegid*[,...]

The deadline schedule is set to all the compute nodes of the node group *nodegid*.

nodegid is described by the hexadecimal number. *nodegid* can specify two or more comma districts by switching off (0x01,0x05,0x010). Moreover, *nodegid* can specify the range by the hyphen (0x02-0x50).

The character string length that can be specified for *nodegid* is 256 bytes or less.

--bootgrp *bootgid*[,...]

Only FX server is effective.

The deadline schedule is set to all the compute nodes of the boot group *bootgid*.

bootgid is described by the hexadecimal number. *bootgid* can specify two or more comma districts by switching off (0x0101,0x0102,0x0103). Moreover, *bootgid* can specify the range by the hyphen (0x0102-0x0105).

The character string length that can be specified for *bootgid* is 256 bytes or less.

{--rscunit|--ru} *rscuname*[,...]

The deadline schedule is set to all the compute nodes of the resource unit *rscuname*.

rscuname can specify two or more comma districts by switching off (unit0,unit1).

The character string length that can be specified for *rscuname* is 256 bytes or less.

{--rscgrp|--rg} *rscgname*[,...]

The deadline schedule is set to all the compute nodes of the resource group *rscgname*.

rscgname can specify two or more comma districts by switching off (group0,group1).

The character string length that can be specified for *rscgname* is 256 bytes or less.

Only the resource group which is set by the shape specification or node ID specification can be specified for *rscgname*.

If you specify the resource group which is set by node number specification or raito specification, the command will fail.

After setting the deadline schedule, even if you change the range of the resource group, the target compute nodes of the deadline schedule are compute nodes included in the resource group at the time of execution padeadline command.

--cmu *nodeid*[,...]

Only FX server is effective. The deadline schedule is set to all compute nodes of the same CMU as *nodeid*.

nodeid can specify two or more comma districts by switching off (0x01010010,0x01010012,0x01010013). Moreover, *nodeid* can specify the range by the hyphen (0x01010010-0x0101002A).

The character string length that can be specified for *nodeid* is 256 bytes or less.

All specified compute nodes must belong to one of the resource units.

-n *nodeid*[,...]

The deadline schedule is set to the compute node *nodeid*.

nodeid can specify two or more comma districts by switching off (0x01010010,0x01010012,0x01010013). Moreover, *nodeid* can specify the range by the hyphen (0x01010010-0x01010015).

The character string length that can be specified for *nodeid* is 256 bytes or less.

All specified compute nodes must belong to one of the resource units.

--coord *xcoord*,*ycoord*

Only FX server is effective.

The deadline schedule is set to all the compute nodes within the range of Tofu coordinates *xcoord* and *ycoord*.

The order of specifying *xcoord* and *ycoord* assumes fixation, and specifies x coordinates and y coordinates sequentially. z coordinates are objects always all.

example)

It sets it to the node of Y coordinates "1-5" of x axis "4" for "--coord 4,1-5".

It sets it to all nodes of x axis "4" for "--coord 4 and all".

It sets it to all nodes of y axis "5" for "--coord all and 5".

-f *filename*

The object resource of the deadline schedule is specified with file *filename*.

Node ID is described in the file by one node a line.

The deadline schedule is set to the specified node.

The character string length that can be specified for *filename* is 256 bytes or less.

example)

```
#cat /work/user1/list1
0x01010010
0x01010020
:
# padeadline -f /work/user1/list1
```

--period start={YYYYMMDD[hhmm]|now}[.end={YYYYMMDD[hhmm]|unlimited}]

Beginning date YYYYMMDD and end date YYYYMMDD of the deadline schedule are specified.

The specification at the beginning date is indispensable. The specification at the end date can be omitted.

It is considered that "Unlimited" is specified when the specification at the end date is omitted.

When unlimited is specified at the end date, the deadline concerned schedule maintains the state at the deadline until being canceled.

When now is specified at the beginning date, time now is set.

"00" is set at the second of the beginning date when "hhmm" (season) is specified and "59" is set at the second of the end date

"00:00:00" is set at the second of the season of the beginning date when the specification of "hhmm" is omitted and "23:59:59" is set at the second of the season of the end date.

--enforce

The deadline schedule is compulsorily set even if there is a job under execution when the influence is when the deadline schedule is set.

--ic

This option suppresses use of interconnect mounted on the specified resource as a communication path of the job.

Specify the --ic option to perform hardware maintenance with interconnect stopped.

--cancel *N* [,...]

The deadline schedule is canceled.

Schedule number *N* of the deadline given when setting it is specified at the canceled deadline.

Schedule number of deadline can specify two or more comma districts by switching off. Moreover, Schedule number of deadline can specify the range by the hyphen.

The maximum number of character strings that can be specified for *N* is 255 characters.

--show

The set deadline schedule is displayed.

It is displayed in order with early start time of the deadline schedule, and the ending deadline schedule is not displayed.

The output item is shown as follows.

Output item and meaning

item	meaning	
NO	Schedule number of deadline	
TYPE	Method of specifying resource when setting it at deadline schedule. The meaning of the character string output with TYPE is the following.	
	Output character string	Method of specifying object resource
	a	-a
	ng	--nodegrp
	bg	--bootgrp
	ru	--rscunit or --ru
	rg	--rscgrp or --rg
	cmu	--cmu
	n	-n
	cd	--coord
	f	-f
START	Beginning date of schedule at deadline	
END	End date of schedule at deadline	

item	meaning
TARGET	Resource specified with TYPE (If the use of interconnect mounted on the specified resource is also suppressed, "InterConnect" appears along with the resource information.)

{--rscunit|--ru} *rscuname*[,...]

This option displays the deadline schedule that affect the node that is included in the specified resource unit name *rscuname*.
rscuname can specify two or more comma districts by switching off (unit0,unit1).
When this option is omitted, all of the deadline schedule are displayed.

{--rscgrp|--rg} *rscgname*[,...]

This option displays the deadline schedule that affect the node that is included in the specified resource group name *rscgname*.
rscgname can specify two or more comma districts by switching off (group0,group1).
When this option is omitted, all of the deadline schedule are displayed.

--help

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[REMARKS]

- Set the deadline schedule newly after deleting the set deadline schedule when you change the resource and the time of the set deadline schedule.
- When --force option has been omitted, use enter the directly for inquiries as to whether it can be implemented, not use pipes or redirects.

[MESSAGES]

If an error occurs during the execution of the padeadline command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 6201 padeadline Unknown option *opt*.
Try `padeadline --help` for more information.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 6202 padeadline Combination of option is illegal.
Try `padeadline --help` for more information.

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

**[ERR.] PJM 6203 padeadline Unknown option argument *arg*.
Try `padeadline --help` for more information.**

Meaning

For the option, the specified argument *arg* is unknown, or no argument is specified.

Action

Specify the correct argument.

[ERR.] PJM 6204 padeadline Argument error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 6205 padeadline Invalid node ID is specified.

Meaning

The specified node ID is invalid.

Action

Check whether the node ID and options specified together as a combination are the correct combination.

[ERR.] PJM 6212 padeadline Deadline number *no* does not exist.

Meaning

The specified deadline number *no* does not exist.

Action

Confirm the set deadline schedule.

[ERR.] PJM 6220 padeadline Cannot open the file (*path*) : *code*.

Meaning

The specified file *path* cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 6230 padeadline Read error (*path*): *code*.

Meaning

Reading of the target resource file *path* of the deadline schedule failed.

code: Internal code for maintenance

Action

The file with the description of the target resource of the deadline schedule may be damaged. Confirm the file contents.

[ERR.] PJM 6231 padeadline Seek error (*path*): *code*.

Meaning

Locating the specified file *path* failed.

code: Internal code for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6236 padeadline Resource group type is invalid: *rscgname*.

Meaning

A deadline schedule cannot be set for the resource group *rscgname* with node resources defined based on the number of nodes or a ratio of the number of nodes.

Action

If you want to set a deadline schedule for a resource group with node resources defined based on the number of nodes or a ratio of the number of nodes, set the schedule for the resource unit to which the resource group belongs.

[ERR.] PJM 6260 padeadline Bad alloc schedule.

Meaning

A deadline cannot be set because a job is running.

Action

Specify the --enforce option or change the deadline time, and then execute the command again.

[ERR.] PJM 6262 padeadline PJM daemon is not ready.

Meaning

The job manager function is not ready to execute the specified operation.

Action

Wait a moment, and execute the command again.

[ERR.] PJM 6263 padeadline Reserved job exist.

Meaning

A deadline schedule cannot be set because there are jobs with a planned execution start time that is guaranteed during the specified deadline schedule.

Action

Set the deadline schedule again by specifying the --enforce option in the padeadline command.
This reschedules the job and resets the planned execution start time.

[ERR.] PJM 6272 padeadline Out of limit.

Meaning

The upper limit on the number of deadline settings has been reached.

Action

The upper limit on the number of deadline settings has been reached.

[ERR.] PJM 6290 padeadline PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 6291 padeadline Internal error: *details*.

Meaning

An internal error was detected.

details: detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6293 padeadline This command can be executed on SMM/CCM.

Meaning

Can only be executed on the system management node or compute cluster management node.

Action

Execute the command on the system management node or compute cluster management node.

[ERR.] PJM 6295 padeadline No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

Warning Message

[WARN] PJM 6234 padeadline Illegal data.

Meaning

Invalid data was detected.

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

Information Message

[INFO] PJM 6200 padeadline Deadline-schedule *N* operation.

Meaning

- If operation is scheduled
Successfully set deadline *N*.
- If *operation* is cancelled
Removed setting for deadline *N*.

Action

No action is necessary.

4.1.6 padumpmgr command

[NAME]

padumpmgr - Batch manages dumps

[SYNOPSIS]

```
/usr/sbin/padumpmgr --list [-c clstname[,...]] [ scope options ]
                        [ [-E | --expand] -v ]
                        [ --time YYYYMMDD[hhmmss][-{YYYYMMDD[hhmmss]|now}] ]
                        [ --data [--delimiter "delimiter"] ]

/usr/sbin/padumpmgr --get [-c clstname] [ scope options ]
                        -d dirname
                        [ --dumpid dumpid[,...]
                        | --time YYYYMMDD[hhmmss][-{YYYYMMDD[hhmmss]|now}] ]
                        [ --force ] [ --erase ]

/usr/sbin/padumpmgr --del [-c clstname] [ scope options ]
                        [ --dumpid dumpid[,...]
                        | --time YYYYMMDD[hhmmss][-{YYYYMMDD[hhmmss]|now}] ]
                        [ --force ]

/usr/sbin/padumpmgr --refresh [-c clstname] [ --force ]

/usr/sbin/padumpmgr --help

scope options(exclusive):
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    -n nodeid[,...]
    -f filename
    --rscunit rscuname[,...]
```

[DESCRIPTION]

The padumpmgr command batch-processes dumps for the nodes of the specified cluster.

This command can be executed only on the active system management node.

[OPTIONS]

--list

This option displays a dump list.

--get

This option transfers a dump to the system management node.

The files of fjsnap, summary and vmlinux in PRIMERGY server is gathered, if these files exist in the dump directory.

--del

This option deletes a dump of the PRIMERGY server.

The files of fjsnap, summary and vmlinux in PRIMERGY server is deleted, if these files exist in the same directory with the memory dump.

When specified for FX server, this option is ignored.

--refresh

This option updates the dump list.

The update does not cover the FX server. (The system monitoring function automatically updates that information.)

-c *clstname*

This option performs the dump operation for the cluster specified by *clstname*.

Only if the --list option is specified, you can specify multiple cluster names by delimiting them with a comma (e.g., cluster1,cluster2).

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

-v

This option displays dump information about the nodes specified in the -c option, the --nodegrp option, or another option. This information is displayed on one line for each node.

You can use this option only if the --list option is specified.

If this option is used in combination with either the -n or -f option, it is ignored.

-E | --expand

This option expands the hierarchy of the dump list to show one more level.

If -E option is specified in the -c option at the cluster specification time, the display is in units of node groups.

With the -E option specified in the --nodegrp option, this option displays a boot group in a node group.

You can use this option only if the --list option is specified.

If this option is used in combination with either the -n or -f option, it is ignored.

--dumpid *dumpid*

If this option is specified, the dump specified by *dumpid* is the target.

You can specify multiple dump targets by delimiting them with a comma.

If this option is omitted, all dumps are targets.

You can use this option only if the --get or --del option is specified.

--time *YYYYMMDD[hhmmss][-{ YYYYMMDD[hhmmss]|now}]*

If this option is specified, the existing dumps within the period specified by *YYYYMMDD[hhmmss][-{ YYYYMMDD[hhmmss]|now}]* are targets.

If the date that follows "-" is omitted, the dumps of the date specified in front of "-" are targets.

If this option is omitted, all dumps are targets.

-d *dirname*

This option uses *dirname* to specify the dump transfer destination directory.

You can use this option only if the --get option is specified. If the --get option is specified, this option is required.

--nodegrp *nodegid[,...]*

If you want to narrow down the controlled objects to a specific node group in the cluster, specify the node group ID in *nodegid*.

Write *nodegid* in hexadecimal.

You can specify multiple node groups by delimiting them with a comma (e.g., 0x01,0x02,0x03).

Also, you can specify a range with a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid[,...]*

If you want to narrow down the controlled objects to a specific boot group in the cluster, specify a boot group ID in *bootgid*.

Write *bootgid* in hexadecimal.

You can specify multiple boot groups by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103).

Also, you can specify a range with a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-n *nodeid[,...]*

If you want to narrow down the controlled objects to a specific node in the cluster, specify a node ID in *nodeid*.

Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

Also, you can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-f *filename*

This option executes a process for the node ID written in the file specified by *filename*.

For *filename*, write a single node ID in hexadecimal on one line.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit *rscuname*[,...]

If you want to narrow down the controlled objects to a specific resource unit in the cluster, specify the resource unit name in *rscuname*.

You can indicate multiple resource units by delimiting them with a comma (e.g., unitA,unitB).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST

--force

This option forcibly executes a dump operation without displaying a yes/no prompt asking whether to execute the operation.

You can use this option only if the --get, --del, or --refresh option is specified.

--data

This option displays the fields of the settings delimited with a delimiter (",").

You can use this option only if the --list option is specified.

--delimiter "*delimiter*"

In the display format of the --data option, this option changes the delimiter to that specified by *delimiter*.

You can use this option only if the --list option is specified.

--erase

After the dumps are transferred, they are deleted from the target PRIMERGY server.

The files of fjsnap, summary and vmlinux in PRIMERGY server is deleted, if these files exist in the same directory with the memory dump.

You can use this option only if the --get option is specified.

When specified for FX server, this option is ignored.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[NOTE]

Keep the following in mind when specifying options.

- You cannot specify the --nodegrp, --bootgrp, -n, -f, and --rscunit options at the same time.
- You cannot specify the --dumpid, --time options at the same time.

[EXIT STATUS]

The exit status is returned as follows:

- 0: Normal end
- 1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the padumpmgr command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "[4.2.1 Common command messages of the system management function.](#)"

Error Message

[ERR.] PSM 5401 padumpmgr The system monitoring function is not active.

Meaning

The system monitoring function is not working, and processing cannot continue.

Action

Execute the command again on the active system management node.

Warning Message

[WARN] PSM 5451 padumpmgr Failed in forwarding the dump.(details)

Meaning

Forwarding the dump file is failed.

details: Target node ID, etc. where transfer failed

Action

Review the forwarding directory, and execute the command again, as required.

[WARN] PSM 5452 padumpmgr Failed in deleting the dump.(details)

Meaning

Deleting the dump file is failed.

details: Target node ID, etc. where deletion failed

Action

Review the dump file list, and execute the command again, as required.

[WARN] PSM 5453 padumpmgr There is not becoming empty in dump ID.

Meaning

An ID cannot be assigned because it would exceed the upper limit on the number of manageable dumps.

Action

Review the dump file list, and unnecessary dump files are deleted.

[WARN] PSM 5454 padumpmgr Failed in the update of the dump list.

Meaning

The node failed in the dump file list update.
The target node may not have started.

Action

Confirm the operating states of the target node by using the pashowclst command. After starting the target node, execute the command again.

[WARN] PSM 5455 padumpmgr Dump file cannot be deleted.([details](#))

Meaning

Dump files on the FX server cannot be deleted.
details: Node IDs, etc. outside the deletion range

Action

No action is necessary.

Information Message

[INFO] PSM 5481 padumpmgr No such dump.([details](#))

Meaning

There is no target dump to be processed.
details: Cluster name, etc. where no dump is subject to processing

Action

No action is necessary.

[INFO] PSM 5482 padumpmgr Not a target dump.

Meaning

There is no target dump to be processed.

Action

No action is necessary.

4.1.7 paeventadm command

[NAME]

paeventadm - Performs operations on the log monitoring condition configuration file.

[SYNOPSIS]

```
/usr/sbin/paeventadm --set [ --force ]  
/usr/sbin/paeventadm --set [ -c clstname ] -f filename [ --force ]  
/usr/sbin/paeventadm --show  
/usr/sbin/paeventadm --help
```

[DESCRIPTION]

The paeventadm command reflects the monitoring conditions defined in the log monitoring condition definition file of paevent.conf. This command can be executed only on the active system management node.

[OPTIONS]

--set

This option sets the written contents in the log monitoring condition definition file.

If the --force option is not specified, a message appears with a yes/no prompt asking whether to execute the command.

If this setting fails with any node, an error file is output to the current directory.

--show

This option displays the current set contents.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

-c *clstname*

The name of a cluster that is an execution target is specified in *clstname*.

If you do not set this option and the environment variable PXMYCLST and specify the --set option when executing the paeventadm command, the setting is executed for all clusters.

-f *filename*

This option processes the node IDs written in the file specified in filename.

Write one node ID in hexadecimal per line in filename.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify the name of a specific cluster in the environment variable PXMYCLST.

--help

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

The cluster name can be specified for the environment variable PXMYCLST instead of -c option.

The specification of PXMYCLST is given to priority more than -c option.

PXMYCLST can specify only one cluster.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[NOTE]

You cannot specify the --set, --show options at the same time.

[EXIT STATUS]

The exit status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

paevent.conf(5)

[MESSAGES]

If an error occurs during the execution of the `paeventadm` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

Error Message

[ERR.] PSM 6301 paeventadm Process control failed.(*details*)

Meaning

An error occurred in the execution of a child process.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 6305 pxeventadm Failed to send files.(*details*)

Meaning

File transfer to the target node failed.

details: Cause of the error

Action

Confirm the file contents of *details*, and execute the command again.

[ERR.] PSM 6306 paeventadm Regular expression error.(*details*)

Meaning

The regular expression is incorrect.

details: Character string with the error

Action

Review the definitions of the log monitoring configuration file.

4.1.8 paledctl command [FX]

[NAME]

`paledctl` - Location LED control command for the FX server

[SYNOPSIS]

```
/usr/sbin/paledctl [ -c clstname ]  
                  --bootgrp bootgid[,...]  
                  [ -v ]  
                  { on | off | show }  
/usr/sbin/paledctl --help
```

[DESCRIPTION]

paledctl command is a command for system administrators to control location LEDs of the FX server.
It can control and display the status of location LEDs of BoB.
This command can be executed on the system management node.

[ARGUMENTS]

on

This argument turns on the location LEDs.

off

This argument turns off the location LEDs.

show

This argument displays the location LED states.
The following states are displayed.

OFF: Off
ON: On
UNKNOWN: Unknown

[OPTIONS]

-c *clstname*

The location LEDs of the cluster specified in *clstname* are controlled.
If no cluster name is specified, the set cluster name in the environment variable PXMYCLST is specified.

--bootgrp *bootgid*[,...]

Specify the boot group ID of a BoB in *bootgid* to control the location LEDs.
Write *bootgid* in hexadecimal.
You can specify multiple boot group IDs by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103).
You can specify a range of IDs by using a hyphen (e.g., 0x0101-0x0103).
When specifying this option, be sure to specify a cluster name in *clstname* in the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-v

This option displays all lines, including those where RESULT is OK.
If this option is not specified, only the lines where RESULT is NG are displayed.
If the show argument is specified, this option is ignored.

--help

This option displays information on how to use this command.
If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option.
Specification with the -c option has priority over that with the environment variable PXMYCLST.
You can set only one cluster name with the environment variable PXMYCLST.
An example is shown below:
Example

```
export PXMYCLST=cluster1
```

EXIT STATUS

The end status is returned as follows:

- 0: Normal end
- 1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the `paledctl` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

Error Message

[ERR.] PSM 2401 paledctl LED control failed.(*details*)

Meaning

LED control failed.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

Information Message

[INFO] PSM 2481 paledctl LED control was executed.

Meaning

LED control was executed.

Action

No action is necessary.

4.1.9 pamonitor_notice command

[NAME]

`pamonitor_notice` - Command for notifying the monitoring function that a node is down

[SYNOPSIS]

```
/usr/sbin/pamonitor_notice -c clstname
                           -n nodeid[,...]
                           --detail detail
/usr/sbin/pamonitor_notice --help
```

[DESCRIPTION]

The `pamonitor_notice` command is a command for system administrators to notify the monitoring function that a node is down. The specified node is treated as a node that has gone down and is isolated from job operations. A dump is automatically collected. You can view the specified detailed information with the `pashowclst` command.

[OPTIONS]

-c *clstname*

Notifications are provided about any nodes being down in the cluster specified in *clstname*. Only the local cluster can be specified to provide this notification for node types other than the system management node.

-n *nodeid*[,...]

Specify the ID of a node in *nodeid* to provide any notifications about the node being down. Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010011,0x0101002A).

You can specify a range of IDs by using a hyphen (e.g., 0x01010010-0x0101002A).

--detail *detail*

Specify detailed information, such as the reason for a node going down.

Use up to 63 letters using half-width alphanumeric characters, half-width spaces, and the following symbols.

_ - . , () / ! ?

You can view the specified detailed information by executing the `pashowclst` command with the `--detail` option specified.

The detailed information is cleared at the following times. One is when the relevant node is reincorporated into job operations (at the timing of the transition to Running in the STATUS field of the `pashowclst` command). The other is when the active or standby system management node is restarted.

--help

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[EXIT STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the `pamonitor_notice` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "[4.2.1 Common command messages of the system management function.](#)"

Error Message

[ERR.] PSM 3401 pamonitor_notice Notice failed.([details](#))

Meaning

Notification processing failed.

details: Cause of the error

Action

Error notification processing failed for a node, as indicated in *details*.

If the upper-level node managing the error notification destination node stopped, execute the command when the upper-level node has been started.

Information Message

[INFO] PSM 3482 pamonitor_notice Notice completed.

Meaning

Notification processing has completed.

Action

No action is necessary.

4.1.10 pamonitoradm command

[NAME]

pamonitoradm - Configures system monitoring.

[SYNOPSIS]

```
/usr/sbin/pamonitoradm --set [ --force ]  
/usr/sbin/pamonitoradm --set -c clstname -f errodelist [ --force ]  
/usr/sbin/pamonitoradm --show  
/usr/sbin/pamonitoradm --help
```

[DESCRIPTION]

The pamonitoradm command executes the process that specified by the options, to the system monitoring configuration file pamonitor.conf which is used by the system monitoring function.

This command can be executed on the active system management node.

[OPTIONS]

--set

This option sets the content set to the system monitoring configuration file, and displays a set result (success or failure). If you execute the command without defining a configuration file, an error is returned. If the --force option is not specified, a message appears with a yes/no prompt asking whether to execute the command.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--show

This option displays the current set contents.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

-c *clstname*

This option is specified together with the -f option. It specifies the name of the cluster to be reflecting settings.

-f *errodelist*

This option sets the content of the system monitoring configuration file to a specific node.

Because node IDs are output to the error list file to the home directory when there is a node that failed in the reflection for the reasons for the stop of the power supply etc., it is possible to execute it again by specifying the file for the argument.

Describe one node ID in the specified file every one line as follows.

The line from which comment (#) is described at the head of line is ignored.

This command continues processing ignoring the lines where character strings other than node ID and node ID that does not exist are described.

[Example]

```
0x03010001  
0x03020001  
0x03030001
```

[EXIT STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end

[SEE ALSO]

pamonitor.conf(5)

[MESSAGES]

There are no messages specific to the pamonitoradm command. Only common command messages are output. For details on the common command messages, see "[4.2.1 Common command messages of the system management function.](#)"

4.1.11 pamopluginadm command

[NAME]

pamopluginadm - Configures system monitoring plugin

[SYNOPSIS]

```
/usr/sbin/pamopluginadm --set [ --force ]  
/usr/sbin/pamopluginadm --set -c clstname -f errodelist [ --force ]  
/usr/sbin/pamopluginadm --show  
/usr/sbin/pamopluginadm --help
```

[DESCRIPTION]

The pamopluginadm command executes the process specified by an option in the system monitoring plugin configuration file pamoplugin.conf, which is used by the system monitoring plugin function.

This command can be executed on the active system management node.

[OPTION]

--set

This option sets the content set to the system monitoring plugin configuration file, and displays a set result (success or failure). If you execute the command without defining a configuration file, an error is returned. If the --force option is not specified, a message appears with a yes/no prompt asking whether to execute the command.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--show

This option displays the current set contents.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

-c *clstname*

This option is specified together with the -f option. It specifies the name of the cluster to be reflecting settings.

-f *errodelist*

This option sets the content of the system monitoring plugin configuration file to a specific node.

Because node IDs are output to the error list file to the home directory when there is a node that failed in the reflection for the reasons for the stop of the power supply etc., it is possible to execute it again by specifying the file for the argument.

Describe one node ID in the specified file every one line as follows.

The line from which comment (#) is described at the head of line is ignored.

This command continues processing ignoring the lines where character strings other than node ID and node ID that does not exist are

described.
[Example]

```
0x03010001
0x03020001
0x03030001
```

[EXIT STATUS]

The following end codes are returned.
0: Normal end
1: Abnormal end

[SEE ALSO]

pamoplugin.conf(5)

[MESSAGES]

There are no messages specific to the pamopluginadm command. Only common command messages are output. For details on the common command messages, see "[4.2.1 Common command messages of the system management function.](#)"

4.1.12 papjmadm command

[NAME]

papjmadm - Shows the setting and content of a configuration file for job operations of an entire cluster.

[SYNOPSIS]

```
/usr/sbin/papjmadm --set [-c clstname] [--force]
/usr/sbin/papjmadm --show [-c clstname]
/usr/sbin/papjmadm --check [-c clstname] [--file filename]
/usr/sbin/papjmadm --help
```

[DESCRIPTION]

papjmadm is a command for cluster administrators to configure or update a configuration file for job operations of an entire cluster, and display its configuration.
This command can be executed only on the system management node.

[OPTIONS]

-c *clstname*

Configure, update or show a configuration file which is for job operations for cluster *clstname*.
If *clstname* is set in environment variable PXYCLST, this option can be omitted.

--set

The content of a current configuration file is set and shows the result of the configuration (either success or failure).
If this command is executed without the --force option set, a message always appears with a yes/no prompt asking whether to execute it.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--show

This option shows currently set contents in the form of a configuration file. For a default setting that is not explicitly set, "# default" must be added to the end of the line.

--check

This option verifies the format of a current configuration file, compares the file to the actual configuration, and displays the result of these checks.

If the --file option is set, this option will verify the format of a specified file, compare the file to the actual configuration, and display the result of these checks.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

An example of --set option is as follows:

```
$ papjmadm --set
[WARNING]
papjmadm is going to set config of cluster1
Do you really want to continue (y/n)? quit
Do you really want to continue (y/n)? y
```

[END CODE]

0: Normal end

1: Abnormal end

[END STATUS]

papjm.conf(5), pmpjmadm(8)

[MESSAGES]

If an error occurs during the execution of the papjmadm command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 6001 papjmadm Unknown option *opt*.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 6002 papjmadm Combination of option is illegal.

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

[ERR.] PJM 6003 papjmadm Unknown option argument *arg*. Try `papjmadm --help` for more information.

Meaning

For the option, the specified argument *arg* is unknown, or no argument is specified.

Action

Specify the correct argument.

[ERR.] PJM 6004 papjmadm Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

If the message displays a cluster name in *arg*, confirm that the ClusterName setting in the papjm.conf file is correct.

[ERR.] PJM 6020 papjmadm Cannot open the file (*path*) : *code*.

Meaning

The specified file path cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 6021 papjmadm File pathname too long.

Meaning

The file name is too long.

Action

Specify a file name within 4096 characters.

[ERR.] PJM 6022 papjmadm File format error: *details*.

Meaning

The format of coding in the configuration file is incorrect.

details

invalid section detected

The section name is incorrect.

invalid item data detected

The item name is incorrect, or there are duplicate lines.

no section detected or no section

The file does not contain any section.

itemcnt

In JobSelectPolicy or the JobEvaluation section, two or more values are defined in the same item in the same section.

ERROR:line = *line-number submessage*

In the line shown with the line number, there is an error.

Action

Write the file in the correct format.

When "ERROR:line = *line-number submessage*" is output in details, remove the cause of the error referring to information on the submessage. After that, write the file in the correct format.

[ERR.] PJM 6029 papjmadm Cannot get the file : *code*.

Meaning

The configuration file for the specified index does not exist.

code: Internal code for maintenance

Action

Review the argument for the option, and execute the command again.

[ERR.] PJM 6050 papjmadm Operation failed : *code*.

Meaning

Command processing failed.

code: Internal code for maintenance

Action

If a pmscatter or pmexe command message is output together with this message, see the corrective action of the respective message. If command execution fails even after taking the above corrective action, or if no pmscatter or pmexe command message is output, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6090 papjmadm PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 6091 papjmadm Internal error: *details*.

Meaning

An internal error was detected.

details: detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6093 papjmadm This command can be executed on SMM.

Meaning

Can only be executed on the system management node.

Action

Execute the command on the system management node.

[ERR.] PJM 6095 papjmadm No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

Information Message

[INFO] PJM 6000 papjmadm Operation completed.

Meaning

Command processing ended normally.

Action

No action is necessary.

4.1.13 papjmstatsadm command

[NAME]

papjmstatsadm - Command that operates the statistical information configuration file for the cluster administrator

[SYNOPSIS]

```
/usr/sbin/papjmstatsadm --set [-c clstname] [--force]
/usr/sbin/papjmstatsadm --show [-c clstname]
/usr/sbin/papjmstatsadm --check [-c clstname] [--file filename]
/usr/sbin/papjmstatsadm --help
```

[DESCRIPTION]

Using options, this command operates the configuration file for the cluster administrator.
This command can be used only on the system management node.

[OPTIONS]

-c *clstname*

This option sets, updates, or displays the configuration file related to job operation management of the *clstname* cluster.
You can omit this option by setting a cluster name in the environment variable PXMYCLST.

--set

This option sets the contents of the current configuration file and displays the result (success or failure) of this setting.
If the --force option is not specified, a yes/no prompt asking whether to execute the command always appear when the command is executed.

--force

This option forcibly executes the instruction without a yes/no prompt asking whether to execute it.

--show

This option displays the current settings in the format of the configuration file.

--check

This option checks the format of the contents of the current configuration file and displays the results.
If the --file option is specified, this option checks the format of the contents of the specified file and displays the results.

--help

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following examples show uses of the papjmstatsadm command.

- Example of specifying the --set option to reflect settings

```
$ papjmstatsadm --set
[WARNING]
papjmstatsadm is going to set config
Do you really want to continue (y/n)? quit
Do you really want to continue (y/n)? y
```

[END CODE]

0: Normal end
1: Abnormal end

[REMARKS]

If you modify the Command or Item section that affect what `pjstat` commands display in `papjmstats.conf` file and run the `--set` option, the information of jobs in the EXIT/CANCEL/REJECT state is removed.

[SEE ALSO]

`papjmstats.conf(5)`

[MESSAGES]

If an error occurs during the execution of the `papjmstatsadm` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 6301 papjmstatsadm Unknown option *opt*.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 6302 papjmstatsadm Combination of option is illegal.

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

[ERR.] PJM 6303 papjmstatsadm Unknown option argument *arg*. Try ``papjmstatsadm --help`` for more information.

Meaning

For the option, the specified argument *arg* is unknown, or no argument is specified.

Action

Specify the correct argument.

[ERR.] PJM 6304 papjmstatsadm Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

If the message displays a cluster name in *arg*, confirm that the ClusterName setting in the `papjm.conf` file is correct.

[ERR.] PJM 6320 papjmstatsadm Cannot open the file (*path*) : *code*.

Meaning

The specified file path cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 6321 papjmstatsadm File pathname too long.

Meaning

The file name is too long.

Action

Specify a file name within 4096 characters.

[ERR.] PJM 6322 papjmstatsadm File format error: *details*.

Meaning

The format of coding in the configuration file is incorrect.

details:

invalid section detected

The section name is incorrect.

invalid item data detected

The item name is incorrect, or there are duplicate lines.

no section detected, no section or Cluster section is not found.

The file does not contain any section

The value of PATH is duplicated.

The Record section has multiple definitions of the PATH item in the same section.

The length of characters per line exceeds the maximum value.

The length of one line is too long. Write a line with up to 511 characters.

section was duplicated.

The same cluster section contains multiple File and Command sections.

required item is not specified.

An item that cannot be omitted is not written in the Item section.

ERROR:line = *line-number submessage*

The line indicated by the line number has an error.

Action

Write the file in the correct format.

If "ERROR:line = *line-number submessage*" is output in *details*, remove the cause of the error, referring to information in the submessage. Then, write the file in the correct format.

[ERR.] PJM 6323 papjmstatsadm Cannot access the file specified in PATH (*path*).

Meaning

The job statistical information file specified in the PATH item of the Record section cannot be accessed.

path: path name

Action

Check whether output to the job statistical information file specified on compute cluster management node is possible, remove the cause, and re-execute processing. If the output directory of the statistical information file does not exist, create it and re-execute.

[ERR.] PJM 6329 papjmstatsadm Cannot get the file : *code*.**Meaning**

The configuration file for the specified index does not exist.

code: Internal code for maintenance

Action

Review the argument for the option, and execute the command again.

[ERR.] PJM 6350 papjmstatsadm Operation failed : *code*.**Meaning**

Command processing failed.

code: Internal code for maintenance

Action

If a pmscatter or pmexe command message is output together with this message, see the corrective action of the respective message. If command execution fails even after taking the above corrective action, or if no pmscatter or pmexe command message is output, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6390 papjmstatsadm PJM daemon is not present.**Meaning**

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 6391 papjmstatsadm Internal error: *details*.**Meaning**

An internal error was detected.

details: detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6393 papjmstatsadm This command can be executed on SMM.**Meaning**

Can only be executed on the system management node.

Action

Execute the command on the system management node.

[ERR.] PJM 6395 papjmstatsadm No cluster specified on SMM.**Meaning**

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

Information Message

[INFO] PJM 6300 papjstatsadm Operation completed.

Meaning

Command processing ended normally.

Action

No action is necessary.

4.1.14 papwradm command

[NAME]

papwradm - Reflects the system control settings.

[SYNOPSIS]

```
/usr/sbin/papwradm --set [ --force ]  
/usr/sbin/papwradm --show  
/usr/sbin/papwradm --help
```

[DESCRIPTION]

papwradm is a command for system administrators, which executes the process specified by an option in the system control function settings written in the system control configuration file papwr.conf.

You can set the following in the system control configuration file:

- IPMI credentials for power control of PRIMERGY server
- The default value, upper limit, and lower limit information for number of concurrent executions and an interval

This command can be executed on the active system management node.

[OPTIONS]

--set

This option sets the contents written in the system control configuration file.

If the --force option is not specified, a message appears with a yes/no prompt asking whether to execute the command.

--show

This option displays the current set contents. The content is displayed by the configuration file format.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

papwr.conf(5)

[MESSAGES]

There are no messages specific to the `papwrctl` command. Only common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

4.1.15 papwrctl command

[NAME]

`papwrctl` - Controls a single power unit.

[SYNOPSIS]

```
/usr/sbin/papwrctl [ -c clstname ]
                    [ --nodetype nodetype[,...] ]
                    [ --excludetype nodetype[,...] ]
                    [ --model model[,...] ]
                    scope options
                    [ -w ]
                    [ -v ]
                    [ --force ]
                    [ --force-all ]
                    [ --jobnowait ]
                    [ -P num[:model][,num[:model],...] ]
                    [ --interval num[:model][,num[:model],...] ]
                    [ --interval-plugin filename ]
                    [ --interval-option filename ]
                    [ --boottype { pxe | disk } ]
                    { on | off | force-off | cold-reboot | warm-reboot | reset | cycle | init | status
                      | all-status }
/usr/sbin/papwrctl --help

scope options(exclusive):
    -a
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    -n nodeid[,...] [ --giogrp | --siogrp ]
    --cmu nodeid[,...] [ --giogrp | --siogrp ]
    --coord xcoord, ycoord
    --rscunit rscuname[,...]
    -f filename
```

[DESCRIPTION]

`papwrctl` is a command for system administrators which controls the power for the specified nodes.

This command can be executed only on the active system management node.

Nodes excluded from power control by this command and nodes with batch power-on or power-off in progress cannot be started without the force execution option.

When arguments except "status" and "all-status" are specified for the argument of this command, the active system management node is excluded from the target.

By default (`--jobnowait` is not specified), if the "off" is specified in this command while a job runs on the target nodes, the command waits for the job to end.

If PRIMERGY server or arbitrary models is specified for the power control target in this command, a control instruction is issued for each node. If the FX server is specified for the power control target in the command, the unit for control instructions to the FX server varies depending on the option-specified conditions.

- If the range of the specified FX server is 1 BoB or more, a control instruction is issued for each BoB.
- If the range of the specified FX server is less than 1 BoB, a control instruction is issued for each node.

[ARGUMENTS]

on

This argument turns on the power of the specified node. If the node power is already on, the power-on is considered successfully executed.

off

This argument shuts down the specified node and then turns off the power. If the node power is already off, the power-off is considered successfully executed.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

In cases with the FX server controlled in units of BoB, the power to a BoB is turned off.

force-off

This argument powers off the specified node regardless of the status of the OS. Shutdown processing is not executed before the power-off.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

In cases with the FX server controlled in units of BoB, the power to a BoB is turned off.

cold-reboot

This argument restarts the node specified without stopping the power supply.

You cannot specify this argument for the FX server.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

warm-reboot

This argument executes a high-speed restart without a hardware reset on the specified node.

For nodes that are not FX server, the operation is the same as with cold-reboot.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

reset

This argument resets hardware of the specified node regardless of the status of the OS.

For the FX server, you can perform this operation only by specifying nodes individually or by specifying a node list.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

cycle

This argument forcibly turns off the power and then turns on the power of the specified node. Shut-down processing is not executed before power-off.

You cannot specify this argument for the FX server.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

init

This argument generates an INIT interrupt. If the OS has a dump function configured, the function collects a dump by this.

For the FX server, you can perform this operation only by specifying nodes individually or by specifying a node list.

If this argument is specified, a message appears with a yes/no prompt asking whether to execute.

status

This argument displays the power supply status of nodes.

all-status

This argument displays the power supply status of nodes, the power supply status of node's Tofu paths (ICC), the power supply status of CMU, and the power supply status of BoB.

[OPTIONS]

-c *clstname*

This option uses *clstname* to specify the name of the target cluster for power control. If no cluster name is specified, the cluster name set by the environment variable `PXMYCLST` is specified.

-a

This option makes all nodes in the cluster subject to control, except the local node.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--nodegrp *nodegid*[,...]

If you want to narrow down the controlled objects to a specific node group in the cluster, specify the node group ID in *nodegid*.

Write *nodegid* in hexadecimal. You can specify multiple node groups by delimiting them with a comma (e.g., 0x01,0x02,0x03). Also, you can specify a range with a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid*[,...]

If you want to narrow down the controlled objects to a specific boot group in the cluster, specify a boot group ID in *bootgid*.

Write *bootgid* in hexadecimal. You can specify multiple boot groups by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103). Also, you can specify a range with a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-n *nodeid*[,...]

If you want to narrow down the controlled objects to a specific node in the cluster, specify a node ID in *nodeid*. Write *nodeid* in hexadecimal. You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A). Also, you can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--giogrp

If you want to narrow down the control targets to only a specific GIO group in the cluster, specify this option as a sub option of the -n option or --cmu option.

The unit of control for FX server is a GIO group, and nodes in the same GIO group as the node specified by the -n option or --cmu option are gathered in the power control targets.

For PRIMERGY server, the --giogrp option specification is ignored.

--siogrp

If you want to narrow down the control targets to only a specific SIO group in the cluster, specify this option as a sub option of the -n option or --cmu option.

The unit of control for FX server is an SIO group, and nodes in the same SIO group as the node specified by the -n option or --cmu option are gathered in the power control targets.

For PRIMERGY server, the --siogrp option specification is ignored.

--cmu *nodeid*[,...]

If you want to narrow down the control targets to only CMUs in the cluster, specify a node ID in *nodeid*.

Write *nodeid* in hexadecimal. You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A). Also, you can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

If a FX server node ID is specified, the unit of control is a CMU, and two nodes in the same CMU as the specified node are gathered in the power control targets.

If a node ID of the PRIMERGY server is specified, the command runs in the same way as when the -n option is specified.

--coord *xcoord,ycoord*

If you want to narrow down the controlled objects by using the coordinates of an FX server node in a specific Tofu logical configuration in the cluster, specify the coordinates in *xcoord,ycoord*. As for coordinates, you can specify not only a specific position but also a range specification by a hyphen and "all" indicating the entire coordinate axis.

[Example]

--coord 4,1-5 = Stops nodes that have an X-coordinate of 4 and a Y-coordinate from 1 to 5.

--coord 4,all = Stops all nodes that have an X-coordinate of 4.

--coord all,5 = Stops all nodes that have a Y-coordinate of 5.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

If the FX server is excluded from the controlled objects narrowed down by the --model or --nodetype option, the command ends with an error.

--rscunit *rscuname*[,...]

If you want to narrow down the controlled objects to a specific resource unit in the cluster, specify the resource unit name in *rscuname*. You can indicate multiple resource units by delimiting them with a comma (e.g., unitA,unitB).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-f *filename*

This option executes a process for the node ID written in the file specified by *filename*.

For filename, write a single node ID in hexadecimal on one line.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--nodetype *nodetype*[,...]

If you want to narrow down the controlled objects to a specific node type in the cluster, specify the node type in *nodetype*. You can specify multiple node types by separating them with a comma (e.g., CCM,CCS).

The target nodes for power control when this option is specified becomes it within the range specified with scope options.

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
CN	Compute node
LN	Login node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

If the specified node type does not exist in the cluster, the node type is ignored and processing continues.

If the specified node type is a redundant configuration, both active and standby nodes is power controlled.

When SMM is specified, only the standby system management node is targeted for power control.

--excludetype *nodetype*[,...]

If you want to exclude a specific node type in the cluster from the control targets, specify the node type in *nodetype*. You can specify multiple node types by delimiting them with a comma (e.g., BIO,GIO).

Even if a node type is specified as a control target by --nodetype option, specification of the node type with this option has priority. The node types that can be specified are the same as for the --nodetype option.

--model *model*[,...]

If you want to limit the controlled objects to a specific model in the cluster, specify the model in *model*.

You can specify multiple models by delimiting them with a comma (e.g., FT,PG).

The controlled system when this option is specified becomes it within the range specified with scope options.

You can specify the following models and arbitrary models defined in the cluster.

Abbreviation	Model name
FT	FX server
PG	PRIMERGY server

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

If the specified model does not exist in the cluster, processing terminates.

-w

Use this option to wait for services to start or stop at the node start or stop time.

During the power-on sequence, the command with this option specified waits for services to start. During the power-off sequence, no response to the command is returned until power-off is confirmed.

This option is effective only when the on, off, or force-off argument is specified. An error is returned in cases with other arguments.

If the option is not specified, the response is returned after a power operation instruction.

This option doesn't wait user definition services.

-v

This option displays the progress of power control.

If the -w option is specified, this option displays the progress of status confirmation.

--force

This option forcibly executes power control without displaying a yes/no prompt asking whether to execute this command.

--force-all

This option executes the requested operation until completion. Even if an error occurs during command processing, the command is not terminated.

If the option is not specified and an error occurs during command processing, processing ends at that point.

The specified power operation is forcibly executed even while normal start is not guaranteed, so be careful when specifying this option.

- The specified power operation is forcibly executed even while cluster power control is in progress.
- The specified power operation is forcibly executed even nodes excluded from power control.
- Even if startup on each model fails, the other models are controlled.
- Even with the --force-all option specified, the active system management node cannot be specified as a power control target.
- Since the power is turned on even when the BoB is in a mixed power states, the Tofu port does not link up and may not start normally.

--jobnowait

This option controls the power without waiting for jobs to end.

If the option is not specified, the command waits for jobs to end.

-P *num[:model][,num[:model],...]* | --concurrency *num[:model][,num[:model],...]*

As many control instructions as specified in the -P option are issued to the control targets at the interval specified in the --interval option, in step-by-step control of the control targets.

Specify the number of targets and the model for control instructions in *num[:model][,num[:model],...]*. For *num*, specify the number of

control targets for control instructions at each interval specified in the `--interval` option. If you want to specify the number of control targets by model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 1 to the total number of nodes in the specified cluster.
- For *model*, you can specify any model that exists in the specified cluster.
- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.
- If you specify a combination of the `-f`, `-n`, `--nodatype`, and `--excludetype` options for an FX server, power operation is applied on a per-node basis.

If this option is not specified, the unit of 30 nodes is set for PRIMERGY server and arbitrary models. For the FX server, if the range of the FX server specified in the control targets is 1 BoB or more, the unit of 24 BoBs is set. If the range of the FX server specified in the control targets is less than 1 BoB, the unit of 16 nodes is set. However, if the number of nodes in the cluster is smaller than the specified value, the number of nodes in the cluster is set as the value. If a value is set in the Default item of the Concurrency section in the `papwradm` command, this specified value is applied.

The example shown below specifies the number of concurrent executions in units of 10 BoBs for cases where the FX server is specified with a range of 1 BoB or more.

[Example]

`-P 10:FT`

`--interval num[:model][,num[:model],...]`

As many control instructions as specified in the `-P` option are issued to the control targets at the interval specified in the `--interval` option, in step-by-step control of the control targets.

Specify the interval and model for control instructions in `num[:model][,num[:model],...]`. For *num*, specify the time in milliseconds to wait after the control instruction is given to the number of control targets specified with the `-P` option. If you want to specify the time for each model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 0 to 86400000.
- For *model*, you can specify any model that exists in the specified cluster.
- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.

If this option is not specified, 0 is specified for all the models of the control targets. If a value is set in the Default item of the Interval section in the `papwradm` command, this specified value is applied.

The example shown below specifies 3,000 milliseconds for the control instruction interval for the FX server and 1,000 milliseconds for the control instruction interval for others.

[Example]

`--interval 1000,3000:FT`

`--interval-plugin filename`

Specify the path of the plugin command executed at an interval. The specified command is executed before the first control processing and at every interval.

If the plugin command does not exist, or if the plugin command return code is other than 0, the `papwrctl` command terminates abnormally.

`--interval-option filename`

Specify the number of targets for concurrent power control at an interval, and specify the path to the file for changing the power control interval for individual control targets. After the plugin command is executed, the specified file is read before the initial control processing and at every interval. You can write the `-P` and `--interval` options in the file. Other options, even when written, are ignored.

`--boottype { pxe | disk }`

Specify the boot type for power-on.

You can specify either of the following values. If this option is not specified, disk is assumed specified.

Specifiable value	Description
pxe	PXE boot
disk	Boot from the default boot device

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[NOTE]

Keep the following in mind when specifying options or arguments.

- You cannot specify the -a, --nodegrp, --bootgrp, -n, --cmu, --coord, --rscunit, and -f options at the same time.
- You cannot specify the on, off, force-off, cold-reboot, warm-reboot, reset, cycle, init, status and all-status arguments at the same time.
- You cannot specify the warm-reboot and --cmu options at the same time.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

Example

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

[papwr.conf\(5\)](#)

[MESSAGES]

If an error occurs during the execution of the papwrctl command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 2001 papwrctl Power control failed.(details)

Meaning

Power control failed.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 2002 papwrctl Node starting process was timed out.(*nodeid*)

Meaning

Node start processing in BoB with the node ID *nodeid* has timed out.

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 2003 papwrctl Node stopping process was timed out.(*nodeid*)

Meaning

Node stop processing in BoB with the node ID *nodeid* has timed out.

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 2004 papwrctl Power control failed for invalid state.(*details*)

Meaning

Power control failed because it was disabled while the node was powered on or off.

details: Cause of the error

Action

After the completion of node power-on or power-off, execute the power control command again.

[ERR.] PSM 2005 papwrctl Power control failed by under system configuration information change.(*id*)

Meaning

Power control failed because it was disabled while system configuration information was being changed.

id: ID of the controlled node or boot group

Action

After the completion of the system configuration information change, execute the power control command again.

[ERR.] PSM 2006 papwrctl Boot process failed.(*details*)

Meaning

Start processing failed.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 2007 papwrctl BIO node power control is not executed, because of related compute node is running.(*details*)

Meaning

The control failed because a control instruction involved in stopping the boot I/O node was issued while the compute node was running.

details: Cause of the error

Action

Issue the control instruction while the compute node is stopped.

[ERR.] PSM 2008 papwrctl Concurrency option setting value exceeds the upper limit value.(upper limit=num)

Meaning

The specified number of concurrent executions exceeds the set upper limit value.

num: Upper limit on the number of concurrent executions

Action

Specify a number of concurrent executions that is less than or equal to the upper limit value, or issue a control instruction with the --force-all option specified.

[ERR.] PSM 2009 papwrctl Concurrency option setting value exceeds the lower limit value.(lower limit=num)

Meaning

The specified number of concurrent executions is less than the set lower limit value.

num: Lower limit on the number of concurrent executions

Action

Specify a number of concurrent executions that is equal to or greater than the lower limit value, or issue a control instruction with the --force-all option specified.

[ERR.] PSM 2010 papwrctl Interval option setting value exceeds the upper limit value.(upper limit=num)

Meaning

The specified interval exceeds the set upper limit value.

num: Upper limit on the interval

Action

Specify an interval less than or equal to the upper limit value, or issue a control instruction with the --force-all option specified.

[ERR.] PSM 2011 papwrctl Interval option setting value exceeds the lower limit value.(lower limit=num)

Meaning

The specified interval is less than the set lower limit value.

num: Lower limit on the interval

Action

Specify an interval equal to or greater than the lower limit value, or issue a control instruction with the --force-all option specified.

[ERR.] PSM 2012 papwrctl Interval plugin command abnormally finished.(details)

Meaning

Control processing was interrupted because the interval plugin command ended abnormally. The return value of the interval plugin command is output in *details*.

details: Return value of the interval plugin command

Action

Investigate the cause of the abnormal end of the interval plugin command.

[ERR.] PSM 2013 papwrctl Boot type setting failed.(*details*)

Meaning

Setting of the boot type failed.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PSM 2014 papwrctl Power control plugin command path is not set.(*details*)

Meaning

The plugin was not set because the set plugin information was invalid.

details: Cause of the error

Action

Using the control function configuration file papwr.conf and the control function configuration command papwradm, correctly set the path information for the power control plugin command.

[ERR.] PSM 2015 papwrctl Warm reboot failed.(*details*)

Meaning

Fast restart processing failed.

details: Cause of the error

Action

Issue the fast restart instruction while the node is running.

[ERR.] PSM 2016 papwrctl Power control plugin for this model is not set.(*details*)

Meaning

The power control plugin command is not set for the model described in *details*.

details: Model name

Action

Using the control function configuration file papwr.conf and the control function configuration command papwradm, set the power control plugin command.

[ERR.] PSM 2017 papwrctl Mixed BoB power status.

Meaning

Power on operation is not possible with mixed BoB power states.

Action

Confirm the BoB power states are not mixed and execute the command again.

[ERR.] PSM 2018 papwrctl Check BoB power status failed.

Meaning

BoB power status check failed.

Action

Confirm you can connect to the BMC that manages the BoB.

Warning Message

[WARN] PSM 2051 papwrctl Cannot get the job information. Retrying...

Meaning

Acquisition of job information failed. The system is retrying.

Action

No action is necessary because the system is retrying job information acquisition until successful.

If you want to avoid the waiting due to job end confirmation, such as because service on the compute cluster management node is stopped, execute the command again with the --jobnowait option.

[WARN] PSM 2052 papwrctl Node power on is not executed, because of related BIO node is not running. (nodeid)

Meaning

The node did not start because the boot I/O (BIO) node inside BoB is not running.

nodeid: Node ID of the target node

Action

After starting the boot I/O node inside BoB, start this node.

Information Message

[INFO] PSM 2081 papwrctl All power control was executed.

Meaning

Power control was executed.

Action

No action is necessary.

[INFO] PSM 2082 papwrctl Failed and retry.(details)

Meaning

Power control failed. The command is retried.

details: Cause of the error

Action

No action is necessary.

[INFO] PSM 2083 papwrctl Excluded from the power control.(nodeid)

Meaning

No power control because the node is excluded from power control.

nodeid: Node ID of the node excluded from power control

Action

No action is necessary.

If you want to force it to start, execute the papwrctl command with the --force-all option specified.

4.1.16 papwrmgradm command

[NAME]

`papwrmgradm` - Command that operates the configuration file for the power management function

[SYNOPSIS]

```
/usr/sbin/papwrmgradm --set [--force]
/usr/sbin/papwrmgradm --set -c clstname [-f nodefile] [--force]
/usr/sbin/papwrmgradm --show
/usr/sbin/papwrmgradm --help
```

[DESCRIPTION]

The `papwrmgradm` command is to define and show a definition of the configuration file (`papwrm.conf`) for the power management function. This command can be executed on the active system management node only by the root user. If any node failed to deploy the configuration, the failed node and its details are displayed in `/var/log/FJSVtcs/pwr/papwrmgradm` directory. Depending on the reason for the failure, there are two types of output:

- `papwrmgradm_compute_cluster_name_failed`
Output from the `pmscatter` command when a file that is located in the `/etc/opt/FJSVtcs/pwr` directory is distributed to each node and the distribution fails on some nodes. The output in the file is the node ID of the node that failed to deploy and its details.
- `papwrmgradm_compute_cluster_name_excluded`
Output is generated when files placed under the `/etc/opt/FJSVtcs/pwr` directory are distributed to each node with the `pmscatter` command, and any nodes are excluded from distribution. The output in the file is the node ID of the node that is excluded from the distribution and its details.

[OPTIONS]

`--set`

This option sets the contents of the current configuration file and displays the results of this setting. Before reflecting the settings, the command displays a (yes/no) prompt asking to confirm the execution.

`--force`

This option forcibly executes the instruction without displaying a (yes/no) prompt asking to confirm the execution before reflecting the settings.

`-c clstname`

This option sets the contents of the configuration file only for the compute cluster specified in *clstname*.

`-f nodefile`

This option sets the contents of the configuration file only for the nodes written in the file specified in *nodefile*.

Write node IDs (example: 0x01010001) in the *nodefile* file.

Write one ID per line.

When specifying this option, you need to also specify the `-c` option.

When distribution of settings fails on any node, use this option for resetting.

In resetting, not change the setting contents.

`--show`

This option displays the current settings in the format of the configuration file. To display default settings that have not been explicitly set, add "# default" to the end of the line.

`--help`

This option displays how to use this command.

[EXIT STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end

[SEE ALSO]

papwrm.conf(5)

[MESSAGES]

If an error occurs during the execution of the papwrmgradm command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PWRM 0100 papwrmgradm This command is already running

Meaning

Execution of the papwrmgradm command is in progress.

Action

Confirm that no other instances of the papwrmgradm command are being executed. Then, execute the command again.

[ERR.] PWRM 0101 papwrmgradm This command cannot be executed on this node

Meaning

An attempt was made to execute the command on a node other than the system management node.

Action

Execute the command on the system management node.

[ERR.] PWRM 0102 papwrmgradm Permission denied: papwrmgradm

Meaning

The user does not have the privileges to execute the papwrmgradm command.

Action

Check the user who executed the command. If the correct user executed the command, grant system or cluster administrator privileges to the user to execute the command again.

[ERR.] PWRM 0107 papwrmgradm Invalid option: *opt*

Meaning

The specified command option *opt* is invalid.

Action

Specify a valid option.

[ERR.] PWRM 0108 papwrmgradm Invalid option argument: *details*

Meaning

The setting value for the argument *arg* of the command option *opt* is invalid.

The setting value for the argument of the command option is invalid. Alternatively, the file specified in the argument of the command option contains an invalid setting value.

details: Invalid command option and argument, or invalid setting value in the file

Action

Specify a valid setting value.

[ERR.] PWRM 0109 papwrmgradm Internal error: *details*

Meaning

An internal error occurred.

details: Detailed information used for maintenance

Action

Collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PWRM 0141 papwrmgradm File not found: *fname*

Meaning

The file *fname* does not exist.

Action

Confirm that the file exists.

[ERR.] PWRM 0142 papwrmgradm File permission denied: *fname*

Meaning

The user does not have access rights to the specified file *fname*.

Action

Grant file and directory access rights to the user.

[ERR.] PWRM 0143 papwrmgradm File cannot be read: *fname*

Meaning

The specified file *fname* cannot be read.

Action

Specify a file that can be read.

[ERR.] PWRM 0144 papwrmgradm pmscatter command failed: *details*

Meaning

Distribution of the configuration file failed.

details: Detailed information for maintenance

Action

Using the pashowclst command, check whether the system management, compute cluster management, and compute nodes are running. If any of these nodes is not running, redistribute the configuration file in the following operations.

1. The configuration file could not be distributed to a node (called "non-distribution node" below). Obtain information about the node from `"/var/log/FJSVtcs/pwr/papwrmgradm/papwrmgradm_compute cluster name_failed"` of the active system management node.
2. If the non-distribution node is the system management node or compute cluster management node, restart the node. Specify the node in the `-c` and `-f` options of the `papwrmgradm` command. Execute the command to distribute the configuration file again.
3. If the non-distribution node is a compute node undergoing maintenance, specify the node in the `-c` and `-f` options of the `papwrmgradm` command. Execute the command to distribute the configuration file again.
4. If the non-distribution node is a compute node powered off by the auto power control function for compute nodes, specify the `--disable` option in the `pacstmgr` command. Execute the command to isolate the node from operation. Specify the `on` option in the `papwrctl` command, and start the target node. Then, specify the node in the `-c` and `-f` options of the `papwrmgradm` command.

Execute the command to distribute the configuration file again. After the redistribution, specify the --enable option in the paclstmgr command, and execute the command to incorporate the target node into operation.

If the configuration file cannot be distributed again even after you take the above actions, collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PWRM 0161 papwrmgradm Invalid item name: *details*

Meaning

A setting item name is wrong.

details: Detailed information pointing out what is wrong

Action

Correct the configuration file, and execute the command again.

[ERR.] PWRM 0162 papwrmgradm Invalid parameter: *details*

Meaning

A setting value is invalid.

details: Detailed information pointing out what is invalid

Action

Review the setting value, and execute the command again. In addition, if a file name is output in detail, check the file exists and whether it is readable.

[ERR.] PWRM 0163 papwrmgradm Invalid section name: *details*

Meaning

A section name is invalid.

details: Detailed information pointing out what is invalid

Action

Review the section name, and execute the command again.

Information Message

[INFO] PWRM 0110 papwrmgradm The processing of the configuration file was completed

Meaning

Reflection of the configuration file has completed.

Action

No action is necessary.

4.1.17 parender command

[NAME]

parender - Setup file creation support command

[SYNOPSIS]

```
/usr/sbin/parender [ -c clstname ]  
                  [ --dstfile filename ]  
                  filename  
/usr/sbin/parender --help
```

[DESCRIPTION]

The `parenter` command expands the keywords enclosed in double curly brackets (`{{ { }}`) in a file into information on the command execution node and saves the information.

If an error occurs partway through processing, the file is not rewritten.

[ARGUMENTS]

filename

Specify the name of a file in *filename* to expand keywords in the file.

[OPTIONS]

`-c clstname`

Information on the cluster specified in *clstname* is expanded.

This option must be specified for execution of the command on a node that belongs to multiple clusters.

If this option is not specified, the set cluster name in the environment variable `PXMYCLST` is used.

`--dstfile filename`

Specify the name of the write destination file in *filename*.

If this option is omitted, the original file is overwritten.

If the write destination file already exists, it is overwritten.

`--help`

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[KEYWORDS]

The keywords subject to this command are shown below.

`PX_MYNODE.CLSTNAME`

This keyword expands into the cluster name of the command execution node.

`PX_MYNODE.NODETYPE`

This keyword expands into the node type of the command execution node.

If the node has a multiple node types, the keyword expands into comma-separated values (e.g., `CN,BIO`).

`PX_MYNODE.NODEID`

This keyword expands into the node ID of the command execution node.

`PX_MYNODE.MNG_NET`

This keyword expands into the management network IP address of the command execution node.

If the command execution node is FX server but it's not boot I/O node, the keyword expands into a Tofu network IP address.

`PX_MYNODE.CNODE_NUM`

This keyword expands into a node number within the BoB of the command execution node.

The node number is padded with zeros to three digits (e.g., `012`).

This keyword cannot be specified for nodes other than FX server.

`PX_SAMEBGRP.nodetype[num].NODETYPE`

This keyword expands into the *num*th node type name of the node type *nodetype* belonging to the same boot group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node has a dual-use configuration, the keyword expands into comma-separated values (e.g., `CN,BIO`).

This keyword cannot be specified for nodes that do not belong to a boot group.

PX_SAMEBGRP.*nodetype*[*num*].NODEID

This keyword expands into the *num*th node ID of the node type *nodetype* belonging to the same boot group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

This keyword cannot be specified for nodes that do not belong to a boot group.

PX_SAMEBGRP.*nodetype*[*num*].MNG_NET

This keyword expands into the *num*th management network IP address of the node type *nodetype* belonging to the same boot group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node is FX server but it's not boot I/O node, the keyword expands into a Tofu network IP address.

This keyword cannot be specified for nodes that do not belong to a boot group.

PX_SAMEBGRP.*nodetype*[*num*].CNODE_NUM

This keyword expands into a node number within the *num*th BoB of the node type *nodetype* belonging to the same boot group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

The node number is padded with zeros to three digits (e.g., 012).

This keyword cannot be specified for nodes other than FX server.

This keyword cannot be specified for nodes that do not belong to a boot group.

PX_SAMENGRP.*nodetype*[*num*].NODETYPE

This keyword expands into the *num*th node type name of the node type *nodetype* belonging to the same node group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node has a dual-use configuration, the keyword expands into comma-separated values (e.g., CN,BIO).

This keyword cannot be specified for nodes that do not belong to a node group.

PX_SAMENGRP.*nodetype*[*num*].NODEID

This keyword expands into the *num*th node ID of the node type *nodetype* belonging to the same node group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

This keyword cannot be specified for nodes that do not belong to a node group.

PX_SAMENGRP.*nodetype*[*num*].MNG_NET

This keyword expands into the *num*th management network IP address of the node type *nodetype* belonging to the same node group and the same cluster as the command execution node.

num is a numeric value that is assigned in the order from the smallest to the largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node is FX server but it's not boot I/O node, the keyword expands into a Tofu network IP address.

This keyword cannot be specified for nodes that do not belong to a node group.

PX_SAMENGRP.*nodetype*[*num*].CNODE_NUM

This keyword expands into a node number within the *num*th BoB of the node type *nodetype* belonging to the same node group and the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

The node number is padded with zeros to three digits (e.g., 012).

This keyword cannot be specified for nodes other than FX server.

This keyword cannot be specified for nodes that do not belong to a node group.

PX_SAMECLST.*nodetype*[*num*].NODETYPE

This keyword expands into the *num*th node type name of the node type name *nodetype* belonging to the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node has a dual-use configuration, the keyword expands into multiple comma-separated node types (e.g., CN,BIO).

PX_SAMECLST.*nodetype*[*num*].NODEID

This keyword expands into the *num*th node ID of the node type name *nodetype* belonging to the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

PX_SAMECLST.*nodetype*[*num*].MNG_NET

This keyword expands into the *num*th node ID of the node type name *nodetype* belonging to the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

If the node is FX server but it's not boot I/O node, the keyword expands into a Tofu network IP address.

PX_SAMECLST.*nodetype*[*num*].CNODE_NUM

This keyword expands into a node number within the *num*th BoB of the node type *nodetype* belonging to the same cluster as the command execution node.

num is a numerical value that is assigned in order from smallest to largest node ID (minimum: 0 maximum: specified number of nodes by the *nodetype* - 1).

This node number is padded with zeros to three digits (e.g., 012).

This keyword cannot be specified for nodes other than FX server.

ENV.*variable*

This keyword expands the value of the environment variable *variable* of the command execution node.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can specify only one cluster name in the environment variable PXMYCLST.

An example is shown below.

[Example] export PXMYCLST=cluster1

[EXIT STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the parent command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

Error Message

[ERR.] PSM 6001 parender UTF-8 codec cannot decode the file.

Meaning

Decoding of the specified file failed.

Action

Specify a file in UTF-8 format.

[ERR.] PSM 6002 parender Invalid variables are included.(*details*)

Meaning

The file contains an invalid variable.

details: Invalid variable

Action

Remove the cause, referring to the information output immediately after the error, and execute the command again.

4.1.18 parscadm command

[NAME]

parscadm - Command that operates configuration file for the cluster manager.

[SYNOPSIS]

```
/usr/sbin/parscadm --set [--force]
/usr/sbin/parscadm --set -c clstname -f nodelist [--force]
/usr/sbin/parscadm --show
/usr/sbin/parscadm --check [--file filename]
/usr/sbin/parscadm --help
```

[DESCRIPTION]

The parscadm command is a command to operate the configuration file for the cluster manager by the option. This command is for system administrator and cluster administrator. It can be used only in the system management node.

[OPTIONS]

--set

The content of the configuration file is reflected in all nodes in the cluster.

When executed without the --force option specified, the command inquires whether to continue execution (y/n).

If reflection of the configuration file fails, the node ID and the reason for the failure are output to an error file.

This command is executed only on the nodes in the "os-running" or "on" power state. For nodes not included in the execution targets, the node ID and the reason for the exclusion are output to the error file.

The error file is output to \$HOME/parscadm_<cluster name>.<command execution date and time>.<process ID>_failed.

--force

The command is compulsorily executed without hearing right or wrong of execution.

--show

The content of the configuration file set now is displayed.

When the default value exists in the item not described in the configuration file, an item concerned is displayed.

When a past configuration file is displayed, number (*idx*) of the history is specified.

--check

The comparison confirmation of a confirmation of the format and an actual composition is checked against the content of a configuration file now at the time of being in the place of default, and the result is displayed. Moreover, the file specified for *filename* was able to be targeted.

When you check the comparison of a past configuration file, number (*idx*) of the history is specified.

--file *filename*

The file name to be checked is specified.

-c *clstname*

This option is specified with the -f option, and specify the cluster name that reflects the settings.

-f *odelist*

To describe the node ID for which you want to reflect the settings on specified nodelist file.

Ignore and continue processing that nonexistent node ID line in which a character string other than node ID is described.

EXAMPLE

```
0x01010010
0x01010011
0x01010012
```

--help

The format on this command are displayed.

When this option is specified, the argument and other options are all disregarded.

[EXAMPLE]

The content of the configuration file is reflected in all nodes in the cluster.

```
# parscadm --set
[WARNING]
Do you really want to continue (y/n)? y
[INFO] PRM 1002 parscadm The processing of the configuration file was completed
```

[END STATUS]

The end status is returned as follows:

0: Normal end

Other than 0: Abnormal end

[SEE ALSO]

pmrscadm(8), parsc.conf(5)

[MESSAGES]

If an error occurs during the execution of the parscadm command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PRM 1201 parscadm Duplicated option: *opt*

Meaning

The option *opt* is specified more than once.

opt: Option name

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 1202 parscadm Option error**Meaning**

The specified option is invalid.

Action

Review the option, and execute the command again.

[ERR.] PRM 1203 parscadm Invalid cluster name: *clstname***Meaning**

Incorrect cluster name *clstname* specified.

clstname: Cluster name

Action

Specify the correct cluster name, and execute the command again.

[ERR.] PRM 1204 parscadm Not a file: *fname***Meaning**

The file specified by *fname* is not a file.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 1208 parscadm Not found file: *fname***Meaning**

The file specified by *fname* does not exist.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 1209 parscadm Cannot be used on this node**Meaning**

The parscadm command cannot be used on this node.

Action

Execute the command on the system management node.

[ERR.] PRM 1212 parscadm The cluster other than compute cluster is specified: *clstname***Meaning**

The specified cluster name *clstname* is not for a compute cluster.

Action

Specify the cluster name of a compute cluster, and execute the command again.

[ERR.] PRM 1213 parscadm There are no contents of the file: *fname***Meaning**

The file contains nothing.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 1214 parscadm Read conffile failed: *fname*(*reason*)

Meaning

The configuration file specified by *fname* contains a syntax error.

fname: File name

reason: Error reason

Action

Take the corrective action described in *reason*, and execute the command again.

[ERR.] PRM 1215 parscadm item name *itemname* is invalid

Meaning

The item name or section name specified by *itemname* is incorrect.

itemname: Item name or section name

Action

Specify the correct item name or section name, and execute the command again.

[ERR.] PRM 1216 parscadm *itemname* is duplicate

Meaning

The item name or section name specified by *itemname* has been specified repeatedly.

itemname: Item name or section name

Action

Delete either of the repeated items or sections, and execute the command again.

[ERR.] PRM 1217 parscadm *itemname* is not found

Meaning

The *itemname* item must be specified but has not been specified.

itemname: Item name

Action

Add the *itemname* item, and execute the command again.

[ERR.] PRM 1218 parscadm item value *itemname* is invalid

Meaning

The value specified by *itemname* cannot be set.

itemname: Item name

Action

Correct the item value specified by *itemname*, and execute the command again.

[ERR.] PRM 1219 parscadm No adjustment of the value between items(*itemname1* *itemname2*)

Meaning

The values specified by *itemname1* and *itemname2* are not consistent with each other.

itemname1: Compared item name

itemname2: Compared item name

Action

Correct the item values specified by *itemname1* and *itemname2*, and execute the command again.

[ERR.] PRM 1225 parscadm Connection error: *clstname*

Meaning

Operation of the configuration file failed because communication was not possible with the active compute cluster management node in the cluster *clstname*.

Action

Check the state of the compute cluster management node in the cluster *clstname* by using the `pashowclst` command. If the error occurs even though the compute cluster management node is in the normal state, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PRM 1227 parscadm Setting of job memory failed: *filename*

Meaning

The memory for jobs cannot be set for the BoB described in the file.

filename: File name

Action

Failed to set memory for jobs for target boot group. Execute the `parscadm` command again with the `--set` option. If the problem persists, contact a custom engineer (CE) to verify that the BMC configuration is correct.

After the problem is resolved, set the *odelist* specified with the `-f` option of the `parscadm` command to the node IDs of all nodes in BoB that failed to be set and execute the command again.

[ERR.] PRM 1240 parscadm Because the job operation has not stopped, the configuration file cannot be reflected.

Meaning

The configuration file cannot be reflected because job operations have not stopped.

Action

Stop job operations and make sure that there are no running jobs in the target cluster. Then, execute the command again. For details on how to reflect the configuration file, see "Reflecting and referencing the `parsc.conf` file" in "Chapter 3 Job Operation Management Function Settings" in "Job Operation Software Administrator's Guide for Job Management."

[ERR.] PRM 1250 parscadm Internal error (*details*)

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 1260 parscadm No memory (*details*)

Meaning

Memory acquisition failed.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 1599 parscadm Internal error (*details*)

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

Warning Message

[WARN] PRM 1101 parscadm item name *itemname* is invalid

Meaning

The item name or section name specified by *itemname* is incorrect.

itemname: Item name or section name

Action

Specify the correct item name or section name, and execute the command again.

[WARN] PRM 1102 parscadm *itemname* is duplicate

Meaning

The item name or section name specified by *itemname* has been specified repeatedly.

itemname: Item name or section name

Action

Delete either of the repeated items or sections, and execute the command again.

[WARN] PRM 1103 parscadm *itemname* is not found

Meaning

The *itemname* item must be specified but has not been specified.

itemname: Item name

Action

Add the *itemname* item, and execute the command again.

[WARN] PRM 1104 parscadm item value *itemname* is invalid

Meaning

The value specified by *itemname* cannot be set.

itemname: Item name

Action

Correct the item value specified by *itemname*, and execute the command again.

[WARN] PRM 1105 parscadm No adjustment of the value between items(*itemname1* *itemname2*)

Meaning

The values specified by *itemname1* and *itemname2* are not consistent with each other.

itemname1: Compared item name

itemname2: Compared item name

Action

Correct the item values specified by *itemname1* and *itemname2*, and execute the command again.

[WARN] PRM 1109 parscadm There are failed nodes in reflection of configuration file: *fname*

Meaning

The configuration file cannot be reflected on the node described in *fname*.

fname: File name

Action

After starting the node described in *fname*, try again to reflect the configuration file.

[WARN] PRM 1111 parscadm There are nodes that failed in reflection of configuration file(Internal error): *fname*

Meaning

The configuration file cannot be reflected to the node described in *fname* due to an internal error.

fname: File name

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

Information Message

[INFO] PRM 1001 parscadm You need to reboot nodes to reflect the setting.

Meaning

Restart the node to reflect the configuration file.

Action

Restart the compute cluster management node, the compute cluster sub management node, and all compute nodes of the cluster to be configured. For details on how to restart, see "Shifting to the Maintenance Mode and Its Recovery" in "Details of the System Management Function" in "Job Operation Software Administrator's Guide System Administration."

[INFO] PRM 1002 parscadm The processing of the configuration file was completed

Meaning

parscadm command ended normally.

Action

No action is necessary.

[INFO] PRM 1003 parscadm Configuration file processing not executed

Meaning

The `parscadm` command was not executed.

Action

No action is necessary.

4.1.19 paserviceadm command

[NAME]

`paserviceadm` - Configures user definition service

[SYNOPSIS]

```
/usr/sbin/paserviceadm --set [ --force ]  
/usr/sbin/paserviceadm --set -c clstname -f errnodelist [ --force ]  
/usr/sbin/paserviceadm --show  
/usr/sbin/paserviceadm --help
```

[DESCRIPTION]

The `paserviceadm` command executes the process specified by an option in the user definition service configuration file `paservice.conf`, which is used by the system monitoring function.

This command can be executed on the active system management node.

[OPTION]

--set

This option sets the content set to the user definition service configuration file, and displays a set result (success or failure). If you execute the command without defining a configuration file, an error is returned. If the `--force` option is not specified, a message appears with a yes/no prompt asking whether to execute the command.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--show

This option displays the current set contents.

--help

This option displays information on how to use this command and its version information.

If this option is specified, all the arguments and other options are ignored.

-c *clstname*

This option is specified together with the `-f` option. It specifies the name of the cluster to be reflecting settings.

-f *errnodelist*

This option sets the content of the user definition service configuration file to a specific node.

Because node IDs are output to the error list file to the home directory when there is a node that failed in the reflection for the reasons for the stop of the power supply etc., it is possible to execute it again by specifying the file for the argument.

Describe one node ID in the specified file every one line as follows.

The line from which comment (#) is described at the head of line is ignored.

This command continues processing ignoring the lines where character strings other than node ID and node ID that does not exist are described.

(Example)

```
0x03010001  
0x03020001  
0x03030001
```

[EXIT STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end

[SEE ALSO]

paservice.conf(5)

[MESSAGES]

There are no messages specific to the paserviceadm command. Only common command messages are output. For details on the common command messages, see "[4.2.1 Common command messages of the system management function.](#)"

4.1.20 pashowclst command

[NAME]

pashowclst - Displays the status or configuration information of a cluster.

[SYNOPSIS]

To display the node status:

[Cluster specification display]

/usr/sbin/pashowclst

```
[ {-E|--expand} | -v ] [--nodetype nodetype[,...]]
[-c clstname[,...]]
{ [-d [STOPPED,ERROR,DISABLE][-t [time]]] |
  [-D [-t [time]]] |
  [--filter item=data[,...] [item=data[,...]]] |
  [--hard-mainte] | [--soft-mainte] }
```

[node-group/boot-group/CMU/coordinate/resource-unit specification display]

/usr/sbin/pashowclst

```
[ {-E|--expand} | -v ] [--nodetype nodetype[,...]]
[[ -c clstname ]
{ --nodegrp nodegid[,...] | --bootgrp bootgid[,...] |
  --cmu nodeid[,...] [--giogrp|--siogrp] | --coord xcoord,ycoord |
  --rscunit [rscuname[,...]] } ]
{ [-d [STOPPED,ERROR,DISABLE][-t [time]]] |
  [-D [-t [time]]] |
  [--filter item=data[,...] [item=data[,...]]] |
  [--hard-mainte] | [--soft-mainte] }
```

[Detailed status change reason display]

/usr/sbin/pashowclst -v --detail

```
[--nodetype nodetype[,...]]
[[ -c clstname ]
{ --nodegrp nodegid[,...] | --bootgrp bootgid[,...] |
  --cmu nodeid[,...] [--giogrp|--siogrp] | --coord xcoord,ycoord |
  --rscunit [rscuname[,...]] } ]
{ [-d [STOPPED,ERROR,DISABLE][-t [time]]] |
  [-D [-t [time]]] |
  [--filter item=data[,...] [item=data[,...]]] }
```

[Node specification display]

/usr/sbin/pashowclst -c clstname -n nodeid[,...][--giogrp|--siogrp][--detail]

Redundancy configuration state display:

/usr/sbin/pashowclst -m [--filegrp]

```
[ {-E|--expand} ] [--nodetype nodetype[,...]]
[ [-c clstname] [--nodegrp nodegid[,...]] ]
```

```
[Node specification display]
/usr/sbin/pashowclst -m -c clstname -n nodeid[,...]

To display configuration information:
[Cluster specification display]
/usr/sbin/pashowclst -l
    [ {-E |--expand} | { -v [--nodetype nodetype[,...]]} ]
    [-c clstname[,...]]

[node-group/boot-group/CMU/coordinate/resource-unit specification display]
/usr/sbin/pashowclst -l
    [ {-E |--expand} | { -v [--nodetype nodetype[,...]]} ]
    [[-c clstname]
    { --nodegrp nodegid[,...] | --bootgrp bootgid[,...] |
      --cmu nodeid[,...] | --coord xcoord,ycoord |
      --rscunit [rscuname[,...]] } ]

[Node specification display]
/usr/sbin/pashowclst -l -c clstname -n nodeid[,...][--giogrp|--siogrp]

Node status, service status, and configuration information display:
/usr/sbin/pashowclst -a
    [-c clstname[,...]]
    [[-d [STOPPED,ERROR,DISABLE]] | [-D]]

Display process option:
/usr/sbin/pashowclst [--data [--delimiter "delimiter"]]

To display help:
/usr/sbin/pashowclst --help
```

[DESCRIPTION]

pashowclst is a command for system administrators, which displays the status or configuration information of nodes in a cluster. This command can be executed on the system management node.

[OPTIONS]

-E|--expand

This option expands the hierarchy to show one more level, which is the display unit.
 If -E option is specified when the cluster specified in -c option, the display is in units of node groups.
 When -E option is specified with --nodegrp option, the display is boot groups in the node group.
 When -E option is specified with --bootgrp option, the display is nodes in the boot group.
 You can use this option only when the -c option is specified.

-v

This option displays information about the nodes specified in the -c option, the --nodegrp option, or another option.
 This information is displayed on one line for each node.
 This option also displays the status of services managed by the monitoring mechanism. The status is displayed in the "service name (operation status indicator)" format.
 The following tables show the meanings of the displayed service names and the operation status indicators.
 If the user definition services are configured, it is similarly displayed.

Service name	Meaning
PJM	Job manager function
MRD, SRD, or NRD	Job resource management function
PLE	Parallel execution environment
FEFS	Layered file system

Service name	Meaning
FEFSSR	Layered file system I/O
PWRD	Power management
IC	High-speed interconnect

(Note) Service names are displayed only for the nodes where the service is enabled.

Operation status indicator	Meaning
b	Service has not started
s	Service is being initialized
o	Service is running
d	Service stop processing in progress
x	Service stopped or error occurred
a	Layered file system I/O error
!	Degraded
w	Standby system startup completed. System can be switched back
*	Operating on one side
f	Failover in progress
?	Unknown

--detail

Specify this option together with the -v or -n option to display the details of the cause of a node status change.

-t [*time*]

When -v option is specified, specify this option together with the -d or -D option to display information on the elapsed time from the last node status update.

You can specify the target elapsed time in minutes as an argument in *time*. In that case, the command displays only the nodes with an elapsed time exceeding the specified time.

The elapsed time is displayed in "*hh:mm:ss*" format.

[Example]

12:34:56

--nodetype *nodetype*[,...]

If you want to narrow down the displayed objects to a specific node type in the cluster, specify the node type in *nodetype*.

You can specify multiple node types by delimiting them with a comma. (e.g., CCM,CCS)

If *nodetype* does not exist in the cluster, the node type is ignored and processing continues.

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node

Abbreviation	Node type
LN	Login node
SCM	Storage cluster management node
OSS	OSS node
MDS	MDS node
MGS	MGS node

-c *clstname*[,...]

This option displays the states and configuration information for the nodes of the cluster specified by *clstname*.

You can specify multiple cluster names by delimiting them with a comma (e.g., cluster1,cluster2).

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodegrp *nodegid*[,...]

If you want to narrow down the displayed objects to a specific node group in the cluster, specify the node group ID in *nodegid*. Write *nodegid* in hexadecimal. You can specify multiple node groups by delimiting them with a comma (e.g., 0x01,0x02,0x03).

Also, you can specify a range with a hyphen (e.g., 0x01-0x03).

If the specified node group does not exist in the cluster, the node group is ignored and processing continues.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid*[,...]

If you want to narrow down the displayed objects to a specific boot group in the cluster, specify the boot group ID in *bootgid*. Write *bootgid* in hexadecimal. You can specify multiple boot groups by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103).

Also, you can specify a range with a hyphen (e.g., 0x0101-0x0103).

If the specified boot group does not exist in the cluster, the boot group is ignored and processing continues.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--cmu *nodeid*[,...]

If you want to narrow down the display targets to nodes in the CMU of the FX server in the cluster, specify the node ID in *nodeid*. Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

Also, you can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

If the specified node ID does not exist in the cluster, the node ID is ignored and processing continues.

The unit of display is a node, and the nodes in the same CMU as the specified node are gathered in the display targets.

If the specified node ID is not for a FX server, the command runs in the same way as when the -n option is specified.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--coord *xcoord*,*ycoord*

If you want to narrow down the displayed targets to FX server nodes of the specific coordinates in the Tofu logical configuration in the cluster, specify the coordinates in *xcoord*,*ycoord*.

With the coordinates, you can specify not only a specific location but also a range with a hyphen and "all" for the entire coordinate axis.

[Example]

--coord 4,1-5 = The X-coordinate 4 Y-coordinates It displays the node from 1 to 5.

--coord 4,all = The X-coordinate It displays all of 4 nodes.

--coord all,5 = The Y-coordinate It displays all of 5 nodes.

If the specified coordinate does not exist in the cluster, the coordinate is ignored and processing continues.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit [*rscuname*[,...]]

If you want to narrow down the displayed objects to a specific resource unit in the cluster, specify the resource unit name in *rscuname*.

You can specify multiple resource unit names by delimiting them with a comma (e.g., unitA,unitB).

If the specified resource unit name is incorrect or does not exist in the cluster, the resource unit is ignored and processing continues. When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-d [STOPPED,ERROR,DISABLE]

This option displays information about nodes whose status is not RUNNING.

If STOPPED, ERROR, or DISABLE is set for the argument, this option displays only the information about the nodes in the specified status.

-D

This option displays information about nodes whose status is not RUNNING and nodes including degraded service. Also, when -D option is specified, the nodes which the user definition services is not running are displayed.

--filter *item=data[,...]* [*item=data[,...]*]

This option extracts and displays the nodes in the state specified in *data* of the item specified in *item*. You can specify multiple *data* entries by delimiting them with a comma.

[Example]

```
--filter STATUS=Stopped,SoftError REASON=OSInit,NodeDown
```

You can specify the following for *item*.

STATUS

For *data*, specify the status which displayed in STATUS field when -v option is specified

You can specify the following for *data*.

data	Description of status displayed
Running	The job operation software service has started, and the node is operational.
Stopped	The node (or the OS) has not been started. Jobs cannot be executed.
Init	Node startup processing is in progress. Jobs cannot be executed.
Stopping	Node stop processing is in progress. Jobs cannot be executed.
SoftError	Software and services is abnormal. Jobs cannot be executed.
PreDisable	Processing to restrict node usage is in progress. Existing jobs waits to end. New jobs cannot be assigned.
Disable	The node cannot be used. Jobs cannot be executed.
Unknown	The status is unknown. The monitoring function of the job operation software has not yet started.

REASON

For *data*, specify the status which displayed in REASON field when -v option is specified

You can specify the following for *data*.

data	[Description of status displayed]
DeadlineSchedule	Deadline scheduling is in progress.
SrvDown	An error of a job operation-related service was detected.
NodeDown	An error was detected in the alive monitoring of the node.
Manual	The server cannot be used manually.
SoftMaintenance	Software is under maintenance.
ResetRequest	ResetRequest has occurred.
ReservedAlarm	ReservedAlarm has occurred.
HardMaintenance	Hardware is under maintenance.

data	[Description of status displayed]
AbnormalEnd	The node has not stopped normally.
RouterFatal	RouterFatal has occurred.
PortRouterFatal	PortRouterFatal has occurred.
PortReservedAlarm	PortReservedAlarm has occurred.
Alarm	Alarm has occurred.

--hard-mainte

This option displays the setting status of hardware maintenance of the node.

--soft-mainte

This option displays the setting status of software maintenance of the node.

-n *nodeid*[,...]

If you want to narrow down the displayed objects to a specific node in the cluster, specify the node ID in *nodeid*.

Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

Also, you can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

If does not exist in the cluster, the node ID is ignored and processing continues.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--giogrp

If you want to narrow down the display targets to only a specific GIO group in the cluster, specify this option as a sub option of the -n option or --cmu option.

The unit of display for FX server nodes is a GIO group, and nodes in the same GIO group as the node specified in the -n option or --cmu option are gathered in the display targets.

For nodes that are not FX server nodes, the --giogrp option specification is ignored.

--siogrp

If you want to narrow down the display targets to only a specific SIO group in the cluster, specify this option as a sub option of the -n option or --cmu option.

The unit of display for FX server nodes is an SIO group, and nodes in the same SIO group as the node specified in the -n option or --cmu option are gathered in the display targets.

For nodes that are not FX server nodes, the --siogrp option specification is ignored.

-l

This option displays the configuration information of the nodes of a cluster.

If a specific cluster name is not specified in the -c option or the environment variable PXMYCLST, this option displays all the cluster information managed by the execution node.

However, you can specify only one cluster with the environment variable PXMYCLST.

-m

This option displays information on each redundant node, which is either the active or standby node, in a cluster.

If this option is used with the -E option, specify the --nodegrp option together. If the --nodegrp option is not specified in this case, the function of -E is ignored.

--filegrp

Information on MGS, MDS, and OSS node is added and displayed in -m option.

--data

This option displays the fields of the settings delimited with a delimiter (.).

You can use this option in all options except the --help option.

--delimiter "delimiter"

In the display format of the **--data** option, this option changes the delimiter to that specified by *delimiter*.
You can use this option in all options except the **--help** option.

-a

This option displays the status and configuration information of nodes and services with regard to the nodes in a cluster.
The displayed information covers all the clusters managed by the execution node, unless a specific cluster name is specified by the **-c** option or the environment variable **PXMYCLST**.
However, you can specify only one cluster with the environment variable **PXMYCLST**.

--help

This option displays information on how to use this command.
If this option is specified, all the arguments and other options are ignored.

[NOTE]

Keep the following in mind when specifying options.

- You cannot specify the **--nodegrp**, **--bootgrp**, **-n**, **--cmu**, **--coord**, and **--rscunit** options at the same time.
- You cannot specify the **-E**, **-v**, and **-n** options at the same time.
- You cannot specify the **-E**, **--cmu**, **--coord**, and **--rscunit** options at the same time.
- You cannot specify the **-v**, **--hard-mainte**, and **--soft-mainte** options at the same time.
- You cannot specify the **-d**, **-D**, **--filter**, **--hard-mainte**, and **--soft-mainte** options at the same time.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable **PXMYCLST** instead of with the **-c** option.
Specification with the **-c** option has priority over that with the environment variable **PXMYCLST**.
You can set only one cluster name with the environment variable **PXMYCLST**.
An example is shown below.

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The exit status is returned as follows:

- 0: Normal end
- 1: Abnormal end and When the node to be displayed doesn't exist.

[MESSAGES]

If an error occurs during the execution of the **pashowclst** command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

.....
Not only the following messages but also common command messages are output. For details on the common command messages, see
"4.2.1 Common command messages of the system management function."
.....

Error Message

[ERR.] PSM 3001 pashowclst Resource unit is not set.

Meaning

The resource unit is not set.

Action

Set the resource unit configuration.

Information Message

[\[INFO\] PSM 3081 pashowclst No data available.\(details\)](#)

Meaning

There is no applicable information.

Action

Review the option arguments.

4.1.21 pasnap command

[NAME]

pasnap - Runs a batch process to collect component files

[SYNOPSIS]

```
/usr/sbin/pasnap [ -c clstname ]
                    scope_options
                    [ --nodetype nodetype[,...] ]
                    [ --excludetype nodetype[,...] ]
                    [ --component component[,...] ]
                    [ -P num:group[,...] ]
                    [ --direct ]
                    [ -d dirname ]
                    [ --snaptype {log|conf} ]
                    --time YYYYMMDD[hhmmss][-{YYYYMMDD[hhmmss]|now} ]
/usr/sbin/pasnap --help

scope options (exclusive):
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    --rscunit rscuname[,...]
    -n nodeid[,...]
    -f filename
    -a
    -n nodeid --case troublename
    --jobid jobid
```

[DESCRIPTION]

pasnap is a command for administrators who run a batch process to collect the component files for the nodes of the specified cluster. This command can be executed only on the system management nodes and multiuse nodes. When executed, the command creates a snap directory by default in the current directory to collect the component files (log files, configuration files, or command output) that are set for each Job Operation Software component. When executed, the command creates an execution result file, error file, and execution exclusion file in the snap directory.

Collected materials	File name
Execution result file	pasnap_operation name_(target name_)result
Error file	pasnap_operation name_(target name_)failed
Execution exclusion file	pasnap_operation name_(target name_)excluded

The following table outlines the above *operation name*.

Operation name	Operation description
exec_snap	Create investigation materials on each node
gather	Collect investigation materials on each node
exec_rmsnap	Delete investigation materials on each node
scatter	Distribute definition files to each node

The following table outlines the above *target name*.

Target name	Description
bio	Processing on boot I/O nodes
nodelist	Node list file on boot I/O nodes
troublefile	Trouble definition file

[OPTIONS]

-c *clstname*

This option uses *clstname* to specify the cluster name of a collected object.

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodegrp *nodegid*[,...]

This option uses *nodegid* to specify a node group ID to set only a specific node group in a cluster as an executed object.

nodegid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x01,0x02,0x03).

You can also specify a range with a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid*[,...]

This option uses *bootgid* to specify a boot group ID to set only a specific boot group in a cluster as an executed object.

bootgid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x0101,0x0102,0x0103).

You can also specify a range with a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-n *nodeid*[,...]

This option uses *nodeid* to specify a node ID to set only a specific node in a cluster as a collected object.

nodeid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

You can also specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

A master node ID cannot be specified.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

When specifying this option together with the --case option, you can specify only one *nodeid*.

-f *filename*

This option executes a process for a node ID written in the file that is specified with filename.

Each of the node IDs is written on one line of the node list by the hexadecimal number.

[Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-a

This option uses to specify for all the nodes in a cluster as a collected object.

If --nodetype is specified, all nodes of the specified node types are targeted.

If --nodetype is omitted, all nodes in the specified cluster are targeted.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit *rscuname*[,...]

This option uses *rscuname* to specify a resource unit name to set only a specific resource unit in a cluster as an executed object.

You can specify multiple resource unit name by separating them with a comma (e.g., unitA,unitB).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--nodetype *nodetype*[,...]

If some node type of cluster is target of collection node, please specify the node type to *nodetype*.

Multiple node type delimited by "," can be specified (e.g., CCM,CCS).

Following node types can be specified. Also, type of the multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
OSS	OSS node
MDS	MDS node
MGS	MGS node

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--excludetype *nodetype*[,...]

If some node type is not target of collection node, please specify the *nodetype*.

It can be specified multiple node type delimited by "," (ex.CCM,CCS). Even if the nodes is specified for the execution target by the --nodetype, the exclusion specification of this option is given to priority.

The node types that can be specified are the same as for the --nodetype option.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--component *component*[,...]

Specify component as the component type to collect component files.

You can specify multiple types by delimiting them with a comma (e.g., INST,PSM).

If the specified type is incorrect or no existing nodes are of that type, the type is ignored and processing continues.

If this option is omitted, all component investigation data but OS is target.

Specify ALL, if you want to collect all component investigation data including OS.

You can specify any of the following component.

Abbreviation	Component
INST	Installation function
PSM	System management function
PWRM	Power management function
PLE	Parallel execution environment
PRM	Resource management function
PJM	Job manager function
FEFS	Fujitsu Exabyte File System
OS	OS
ALL	Covers all above components

-P *num:group[,num:group]* | --concurrency *num*":group[,num:group]

The number of nodes processed at one time is specified in *num:group*.

If this option is not specified, all the nodes that are the processing targets are processed in parallel.

For *num*, specify the parallel processing quantity. For *group*, specify the parallel processing quantity you want to change in the unit of a group.

You can specify any of the following values.

Unit	Parallel processing quantity
nodegrp	The minimum value is 1. The maximum value is the number of node groups that exist in the specified cluster.
bootgrp	The minimum value is 1. The maximum value is the maximum number of boot groups that exist in the specified cluster.
node	The minimum value is 1. The maximum value is the maximum number of nodes that exist in the specified cluster.

--direct

With this option specified, the command logs in to the material collection source node to collect detailed materials.

This option enables the collection of information from nodes other than the boot I/O node of the FX server.

If component files for the FX server are collected without this option specified, the command output among the component files are not collected.

-d *dirname*

The directory *dirname* which stores the extracted data is specified.

A snap directory is created to the directory subordinate who specified, and a material the component files are stored.

If this option is not specified, snap directory is created the current directory.

--snaptype {log|conf}

Specify the type of materials to collect.

If this option is not specified, all material types are collected.

You can specify either of the following types.

Material type	Description
log	Log files required for investigation
conf	Configuration files and command output required for investigation

--time *YYYYMMDD[hhmmss]*[-{ *YYYYMMDD[hhmmss]*|now}]

Specify the start time and end time for collected logs.

If the date specification after "-" is omitted from this specified option, logs on the specified date are collected.

With this option specified, the command collects only the log files among the component files.

--jobid *jobid*

Specify a job ID in *jobid* or a sub job ID in *jobid* to collect component files.

With this option specified, the command collects log files from the compute nodes that executed the jobs with the specified *jobid*. The period of the log files is from the start to end of job execution.

To use this option, you need to configure the job manager function to save the job information.

You can use this option only on the system management node.

Examples of specification are shown below.

Example 1) Job ID specified (e.g., 123456789)

The targets are all the nodes used in the specified job.

Use following method for step jobs or bulk jobs.

Example 2) Job ID + step number specified (e.g., 123456789_1)

The targets are all the nodes used in the specified step job.

Example 3) Job ID + bulk number specified (e.g., 123456789[0])

The targets are the nodes used in the specified bulk job.

--case *troublename*

Specify a trouble event name in *troublename*.

With this option specified, the command collects the necessary component files according to the trouble event.

Specify one of the trouble event names written in the trouble definition file *patrouble.conf*.

When specifying this option, be sure to specify the target node in *nodeid* in the *-n* option.

If the *-n* option specifies a node other than the boot I/O node of the FX server, the command logs in to the material collection source node to collect detailed materials, in the same way as when the *--direct* option is specified.

--help

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

The cluster name can be specified for the environment variable *PXMYCLST* instead of *-c* option.

The specification of *PXMYCLST* is given to priority more than *-c* option.

PXMYCLST can specify only one cluster.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[NOTE]

- Component files of target component is not collected on the node which doesn't exist the target component.
- You cannot specify the *--nodegrp*, *--bootgrp*, *-n*, *-f*, *--rscunit*, *--jobid* and *-a* options at the same time.
- If node types that don't exist in the cluster are specified in *--nodetype* or *--excludetype*, specified node types are ignored.
- If the node ID specified in the *-n* option does not exist in the cluster, processing ignores that node.
- If you attempt to run *pasnap* from multiple different nodes simultaneously on the same node, the first *pasnap* attempt will fail.
- If the specified node has a large size of component files (log files), the *--time* option takes a long time to extract the log output at the specified time.
Do not specify the *--time* option if you want to reduce the time it takes to collect all log files.

[SEE ALSO]

patrouble.conf(5)

[MESSAGES]

If an error occurs during the execution of the `pasnap` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

Error Message

[ERR.] PSM 6201 pasnap Invalid case name.(*troublename*)
See the config file (/etc/opt/FJSVtcs/patrouble.conf).

Meaning

The specified trouble event name *troublename* is invalid.

Action

Review the specified trouble event name *troublename*, and execute the command again.

[ERR.] PSM 6202 pasnap Invalid snaptype.(*snaptpe*)

Meaning

The specified material type *snaptpe* is invalid.

Action

Review the specified material type *snaptpe*, and execute the command again.

[ERR.] PSM 6203 pasnap Invalid time.(*time*)

Meaning

The specified time *time* is invalid.

Action

Review the specified time *time*, and execute the command again.

[ERR.] PSM 6204 pasnap Cannot get the job information.

Meaning

Failed to get job information.

Action

Confirm that you can obtain the job statistical information, and execute the command again.

4.1.22 pastart command

[NAME]

`pastart` - Executes a batch power-on operation.

[SYNOPSIS]

```
/usr/sbin/pastart [ [ -c clstname[,...] ]
                  { --nodegrp nodegid[,...] } ]
                  [ -w ]
                  [ -v ]
                  [ --force-all ]
                  [ -P num[:model][,num[:model],...] ]
                  [ --interval num[:model][,num[:model],...] ]
                  [ --timeout commandtimeout ]
                  [ --ccstimeout ccstimeout ]
                  [ --interval-plugin filename ]
                  [ --interval-option filename ]

/usr/sbin/pastart --help
```

[DESCRIPTION]

pastart is a command for administrators, which powers on the specified clusters and node groups.

This command can be executed only on the active system management node.

Multiuse nodes in the compute cluster or the storage cluster is excluded from the target of batch power on. Therefore, node does not be turned power on.

Moreover, it makes an error of this command when two or more clusters including a multiuse cluster are specified.

[OPTIONS]

-c *clstname*[,...]

This option turns on the power of the cluster specified by *clstname*. You can specify multiple cluster names by delimiting them with a comma (e.g., *cluster1,cluster2*). If no cluster name is specified, the cluster name set by the environment variable *PXMYCLST* is specified.

--nodegrp *nodegid*[,...]

If you want to power on only a specific node group in the cluster, specify the node group ID in *nodegid*. Write *nodegid* in hexadecimal. You can specify multiple node groups by delimiting them with a comma (e.g., *0x01,0x02,0x03*). Also, you can specify a range with a hyphen (e.g., *0x01-0x03*).

When specifying this option, be sure to specify a cluster name in *clstname* with the **-c** option, or specify a specific cluster name in the environment variable *PXMYCLST*.

If this option is specified, compute cluster sub management nodes in specific node group are powered on also.

-w

This option specifies waiting for services to start. If the option is specified, no response to the command is returned until service-up is confirmed.

If the option is not specified, the response is returned after a power-on instruction.

-v

This option displays the progress of power control.

If the **-w** option is set, this option displays the progress of start confirmation.

--force-all

This option executes the requested operation until completion. Even if an error occurs during command processing, the command is not terminated.

If the option is not specified and an error occurs during command processing, processing ends at that point.

-P *num[:model]*[,*num[:model]*,...] | --concurrency *num[:model]*[,*num[:model]*,...]

As many control instructions as specified in the **-P** option are issued to the control targets at the interval specified in the **--interval** option, in step-by-step control of the control targets.

Specify the number of targets and the model for control instructions in *num[:model]*[,*num[:model]*,...]. For *num*, specify the number of control targets for control instructions at each interval specified in the **--interval** option. If you want to specify the number of control targets by model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 1 to the total number of nodes in the specified cluster.
- For *model*, you can specify any model that exists in the specified cluster.
- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.

If this option is not specified, the unit of 30 nodes is set for PRIMERGY server and arbitrary models. And the unit of 24 BoB is set for FX server. However, if the number of nodes in the cluster is smaller than the specified value, the number of nodes in the cluster is set as the value. If a value is set in the Default item of the Concurrency section in the `papwradm` command, this specified value is applied. The example shown below specifies the number of concurrent executions in units of 10 BoBs.

[Example]

```
-P 10:FT
```

```
--interval num[:model][,num[:model],...]
```

As many control instructions as specified in the `-P` option are issued to the control targets at the interval specified in the `--interval` option, in step-by-step control of the control targets.

Specify the interval and model for control instructions in *num[:model][,num[:model],...]*. For *num*, specify the time in milliseconds to wait after the control instruction is given to the number of control targets specified with the `-P` option. If you want to specify the time for each model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 0 to 86400000.
- For *model*, you can specify any model that exists in the specified cluster.
- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.

If this option is not specified, 1000 is specified for all the models of the control targets. If a value is set in the Default item of the Interval section in the `papwradm` command, this specified value is applied.

The example shown below specifies 3,000 milliseconds for the control instruction interval for the FX server and 1,000 milliseconds for the control instruction interval for others.

[Example]

```
--interval 1000,3000:FT
```

```
--timeout commandtimeout
```

Specify a timeout in minutes for end of command processing. The specifiable values are 0 to 1440.

If the command timeout occurs, the command is forcibly terminated. The control instruction isn't given after the timeout period.

If this option is not specified or 0 is specified for this option, `pastart` command waits indefinitely.

[Example]

```
--timeout 600
```

```
--ccstimeout ccstimeout
```

Specify the startup waiting time of the compute cluster sub management node in minutes. The specifiable values are 0 to 1440.

The waiting time in minutes to judge timeout when the compute cluster sub management node does not start normally can be specified for this option.

The compute cluster sub management nodes which was judged timeout will not be powered on.

If this option is not specified, startup waiting time is 20 minutes.

If 0 is specified for this option, `pastart` command waits indefinitely.

```
--interval-plugin filename
```

Specify the path of the plugin command executed at an interval. The specified command is executed before the first control processing and at every interval.

If the plugin command does not exist, or if the plugin command return code is other than 0, the `pastart` command terminates abnormally.

```
--interval-option filename
```

Specify the number of targets for concurrent power control at an interval, and specify the path to the file for changing the power control interval for individual control targets. After the plugin command is executed, the specified file is read before the initial control

processing and at every interval.

You can write the -P and --interval options in the file. Other options, even when written, are ignored.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

[papwr.conf\(5\)](#)

[MESSAGES]

If an error occurs during the execution of the pastart command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 2101 pastart Command is already running.(*clstname*)

Meaning

The cluster *clstname* is under power control.

Action

When the cluster is under power control, other power controls are unavailable. After the completion of cluster power control or after cluster power control is stopped, execute the command again.

[ERR.] PSM 2102 pastart Power control failed.(*details*)

Meaning

Power control failed.

details: Cause of the error

Action

Collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 2103 pastart Power control for CCS node was timed out. So, related node group does not boot. (nodegid)

Meaning

The node group with the node group ID *nodegid* did not start because the wait timeout period for power control completion of the compute cluster sub management node (CCS) elapsed during the start process of the whole system.

Action

After the compute cluster sub management node (CCS) starts, start the node group again.

[ERR.] PSM 2104 pastart Storage cluster is not running.

Meaning

The storage cluster required for the compute cluster was not started first.

Action

Confirm the node configuration. Then, after starting the storage cluster required for the compute cluster, start the compute cluster again.

[ERR.] PSM 2105 pastart Batch power control containing multi use cluster cannot be performed.

Meaning

A multiuse cluster is included in the multiple clusters specified.

Action

Multiple clusters cannot be controlled when they include a multiuse cluster. To execute control of the multiuse cluster, specify it alone.

Warning Message

[WARN] PSM 2151 pastart Pre or post plugin script abnormally finished.(details)

Meaning

The plugin script ended abnormally.

details: Cause of the error

Action

Check the cause of the abnormal end of the plugin script.

[WARN] PSM 2152 pastart Power control for CCS was timed out.

Meaning

The wait timeout period for power control completion of the compute cluster sub management node (CCS) elapsed during the start process of the whole system. The start process of the whole system continues.

Action

Take action as follows according to the state of the compute cluster sub management node (CCS).

- If "Power ON (on)" is shown for the power of the compute cluster sub management node (CCS) and the OS has not successfully completed startup

After power-on, the start process may fail before OS startup or incorporation into operation.

Collect the dump file of the compute cluster sub management node (CCS) according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk. If a panic occurred during

the start process and a dump file was collected, contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk without collecting the dump.

- If "Power ON (on)" is shown for the power of the compute cluster sub management node (CCS) and the OS has successfully completed startup

After the OS starts, a service may fail to start.

Collect system information about the compute cluster sub management node according to the "Job Operation Software Administrator's Guide for Maintenance," and contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk.

Information Message

[INFO] PSM 2181 pastart All power control was executed.

Meaning

Power control was executed.

Action

No action is necessary.

4.1.23 pastop command

[NAME]

pastop - Executes a batch power-off operation

[SYNOPSIS]

```
/usr/sbin/pastop [ [ -c clstname[,...] ]
                  { --nodegrp nodegid[,...] } ]
                  [ -w ]
                  [ -v ]
                  [ --force ]
                  [ --force-all ]
                  [ --emergency ]
                  [ --jobnowait ]
                  [ -P num[:model][,num[:model],...] ]
                  [ --interval num[:model][,num[:model],...] ]
                  [ --timeout commandtimeout ]
                  [ --shutdown shutdowntimeout { none | force-off | init } [ model ] ]
                  [ --interval-plugin filename ]
                  [ --interval-option filename ]
/usr/sbin/pastop --help
```

[DESCRIPTION]

pastop is a command for system administrators, which powers off the specified clusters and node groups.

If the storage cluster is shared with other clusters, power-off is not executed unless all of the sharing clusters have stopped operating.

This command can be executed only on the active system management node.

Before the command is executed, a message appears with a yes/no prompt asking whether to execute the command.

Multiuse nodes in the compute cluster or the storage cluster is excluded from the target of batch power off. Therefore, node does not be turned power off.

Moreover, it makes an error of this command when two or more clusters including a multiuse cluster are specified.

[OPTIONS]

-c *clstname* [,...]

This option turns off the power of the cluster specified by *clstname*. You can specify multiple cluster names by delimiting them with a comma (e.g., cluster1,cluster2). If no cluster name is specified, the cluster name set by the environment variable PXMYCLST is specified.

--nodegrp *nodegid*[,...]

If you want to power off only a specific node group in the cluster, specify the node group ID in *nodegid*. Write *nodegid* in hexadecimal. You can specify multiple node groups by delimiting them with a comma (e.g., 0x01,0x02,0x03). Also, you can specify a range with a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* with the **-c** option, or specify a specific cluster name in the environment variable PXYCLST.

-w

This option specifies waiting for power-off. If the option is specified, no response to the command is returned until power-off is confirmed.

If this option is not specified, the response is returned after a power-off instruction.

-v

This option displays the progress of power control.

If the **-w** option is specified, this option displays the progress of stop confirmation.

--force

This option forcibly executes power control without displaying a yes/no prompt asking whether to execute control.

--force-all

This option executes the requested operation until completion. Even if an error occurs during command processing, the command is not terminated.

If the option is not specified and an error occurs during command processing, processing ends at that point.

--emergency

This option, when specified, forcibly executes emergency power-off (force-off) even if a batch power-on or power-off operation is in progress.

--jobnowait

This option controls the power without waiting for jobs to end.

If the option is not specified (default), the command waits for jobs to end before controlling the power.

-P *num[:model]*[,*num[:model]*,...] | --concurrency *num[:model]*[,*num[:model]*,...]

As many control instructions as specified in the **-P** option are issued to the control targets at the interval specified in the **--interval** option, in step-by-step control of the control targets.

Specify the number of targets and the model for control instructions in *num[:model]*[,*num[:model]*,...]. For *num*, specify the number of control targets for control instructions at each interval specified in the **--interval** option. If you want to specify the number of control targets by model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 1 to the total number of nodes in the specified cluster.

- For *model*, you can specify any model that exists in the specified cluster.

- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.

If this option is not specified, the unit of 30 nodes is set for PRIMERGY server and arbitrary models. And the unit of 24 BoB is set for FX server. However, if the number of nodes in the cluster is smaller than the specified value, the number of nodes in the cluster is set as the value. If a value is set in the Default item of the Concurrency section in the papwadm command, this specified value is applied. The example shown below specifies the number of concurrent executions in units of 10 BoBs.

(Example)

-P 10:FT

--interval *num[:model]*[,*num[:model]*,...]

As many control instructions as specified in the **-P** option are issued to the control targets at the interval specified in the **--interval** option, in step-by-step control of the control targets.

Specify the interval and model for control instructions in *num[:model]*[,*num[:model]*,...]. For *num*, specify the time in milliseconds to wait after the control instruction is given to the number of control targets specified with the **-P** option. If you want to specify the time

for each model, specify both *model* and *num*. If you specify *num* without specifying *model*, all the models in the specified range for the control targets are targeted.

- For *num*, you can specify a value ranging from 0 to 86400000.
- For *model*, you can specify any model that exists in the specified cluster.
- If you specify *num* without specifying *model* and you also specify both *num* and *model*, the value from specifying both *num* and *model* has priority.

If this option is not specified, 0 is specified for all the models of the control targets. If a value is set in the Default item of the Interval section in the `papwradm` command, this specified value is applied.

The example shown below specifies 3,000 milliseconds for the control instruction interval for the FX server and 1,000 milliseconds for the control instruction interval for others.

(Example)

```
--interval 1000,3000:FT
```

`--timeout commandtimeout`

Specify the timeout time until the command ends in minutes. You can specify a value ranging from 0 to 1440.

If the timeout is reached, the command is forcibly terminated. Control instructions after the timeout are not given.

If this option is not specified or 0 is specified in this option, this command waits indefinitely.

(Example)

```
--timeout 600
```

`--shutdown shutdowntimeout { none | force-off | init } [model]`

This option uses *shutdowntimeout* to specify an OS shutdown timeout period in minutes starting from an OS shutdown instruction. At the end of this period, the operation is considered to have timed out. The specifiable values are 0 to 1440.

If 0 is specified in *shutdowntimeout*, there are no timeouts.

The option also uses `none | force-off | init` to specify the action to take for a timeout.

none: Take no action.

force-off: Forcibly execute power-off.

init: Generate an INIT interrupt.

You can specify a shutdown timeout for only a specific model by specifying the model in *model*. If *model* is not specified, all the models are subject to the shutdown timeout.

When the model that does not exist in the specified cluster is specified, `pastop` becomes an error.

If this option is not specified, the timeout time is 30 minutes.

`none` is the default action for a shutdown timeout. If no action is specified, the default is used.

You cannot specify `none`, `force-off`, and `init` for this option at the same time.

`--interval-plugin filename`

Specify the path of the plugin command executed at an interval. The specified command is executed before the first control processing and at every interval.

If the plugin command does not exist, or if the plugin command return code is other than 0, the `pastop` command terminates abnormally.

`--interval-option filename`

Specify the number of targets for concurrent power control at an interval, and specify the path to the file for changing the power control interval for individual control targets. After the plugin command is executed, the specified file is read before the initial control processing and at every interval.

You can write the `-P` and `--interval` options in the file. Other options, even when written, are ignored.

`--help`

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name with the environment variable PXMYCLST instead of with the -c option. Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can set only one cluster name with the environment variable PXMYCLST.

An example is shown below:

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[SEE ALSO]

[papwr.conf\(5\)](#)

[MESSAGES]

If an error occurs during the execution of the pastop command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 2201 pastop Command is already running.(*clstname*)

Meaning

The cluster *clstname* is under power control.

Action

When the cluster is under power control, other power controls are unavailable. After the completion of cluster power control, execute the command again. Alternatively, execute the command with the --force-all option.

[ERR.] PSM 2202 pastop Power control failed.(*details*)

Meaning

Power control failed.

details: Cause of the error

Action

Collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 2203 pastop Associated compute cluster is not stopped.

Meaning

The compute cluster requiring the storage cluster has not stopped.

Action

Confirm the node configuration. Then, after stopping the compute cluster requiring the storage cluster, stop the storage cluster again.

[ERR.] PSM 2204 pastop Batch power control containing multi use cluster cannot be performed.

Meaning

A multiuse cluster is included in the multiple clusters specified.

Action

Multiple clusters cannot be controlled when they include a multiuse cluster. To execute control of the multiuse cluster, specify it alone.

Warning Message

[WARN] PSM 2251 pastop Pre or post plugin script abnormally finished.([details](#))

Meaning

The plugin script ended abnormally.

details: Cause of the error

Action

Check the cause of the abnormal end of the plugin script.

[WARN] PSM 2252 pastop Shutdown processing was timed out.

Meaning

The specified timeout time was controlled because an argument other than none is specified in the --shutdown option.

Action

The command to stop the power supply is running. After waiting a moment, use the pashowclst command to confirm that the specified model *model* has stopped. If model is not specified, confirm that all models have stopped.

Information Message

[INFO] PSM 2281 pastop All power control was executed.

Meaning

Power control was executed.

Action

No action is necessary.

4.1.24 pasyspwr command

[NAME]

pasyspwr - Command that outputs arbitrary information on internal system power consumption

[SYNOPSIS]

```
[Latest power consumption information of all devices]
PXMYCLST=clstname /usr/sbin/pasyspwr
    [ --average | --momentary ]
    [ --last ]
    [ -v ]
    [ --data [ --delimiter delimiter ] ]
```



```
[Latest power consumption information of specified devices]
/usr/sbin/pasyspwr scope options
    [ --average | --momentary ]
    [ --last ]
    [ -v ]
    [ --data [ --delimiter delimiter ] ]

[Past power consumption information of specified devices]
/usr/sbin/pasyspwr --trace scope options
    [ --average ] [ --momentary ]
    [ --time YYYYMMDDhhmm[ss][-{YYYYMMDDhhmm[ss]|now}] ]
    [ --data [ --delimiter delimiter ] ]

/usr/sbin/pasyspwr --help

scope options
    -c clstname
    -n {nodeid[,...] | nodeid-nodeid | all} | --nodelist file
    --extdev {extdevname[,...] | all} | --extdevlist file
    --pwrgrp {pwrgrpname[,...] | all} | --pwrgrplist file

    The scope options can be combined as follows:
        -c --extdev
        -c --extdevlist
    extdevname is a name defined in /etc/opt/FJSVtcs/pwr/ptdev. pwrgrpname is a name defined in
    papwr.conf(8)
```

[DESCRIPTION]

The pasyspwr command is a command for administrators to output system power consumption information. It can output the following power consumption information.

- Latest power consumption information of all devices (without the *scope options* and the --trace option)
 The command outputs total power consumption of all compute nodes and all external devices in the compute cluster specified by the environment variable PXMYCLST. If the environment variable PXMYCLST is not specified, the command fails by error.
 The command without the -v option outputs a total power consumption value.
 The command with the -v option outputs individual power consumption information in addition to the total value.
 Latest power consumption information of specified devices (with the *scope options*, without the --trace option)
 The command outputs information on the most recently measured power consumption of the compute nodes, the external devices, or individual devices composing the power group specified in the scope options.
 The command without the -v option outputs a total power consumption value.
 The command with the -v option outputs individual power consumption information in addition to the total value.
 Past power consumption information of specified devices (with the *scope options*, without the --trace option)
 The command outputs information on the past power consumption of the compute nodes, the external devices, or individual devices composing the power group specified in the scope options.
 If no *scope options* are specified, the command fails by error.
 The command without the --time option outputs power consumption information for all periods saved in the system power database.
 The command with the --time option outputs information on power consumption measured at the specified time or during the specified period.
 The command without the --time option outputs power consumption information for the latest 10 minutes.
 The command with the --time option outputs power consumption information for the specified time or period.

This command can be executed on the active system management node only by the root user.

This command outputs the following items.

Item name	Description
TOTAL_PWR	Sum of power consumption of the specified compute nodes and external equipment (unit: W)
PWR_TYPE	Type of power consumption used for calculating the sum MOM: All are momentary power consumption. AVE: All are average power consumption.

Item name	Description
	MIX: Average power consumption and momentary power consumption are mixed. ERR: Sum acquisition error
DATE	Output date and time

The following describes the displayed items and their meaning:

Item name	Description
NODE/DEV	Node ID of compute node or external equipment name
TYPE	Type of power consumption MOM: Momentary power consumption AVE: Average power consumption
PWR	Power consumption (W)
M_DATE	Measurement time
B_DATE	Reference time
SER_NO	Serial number. For compute nodes, a value that increments by 1 every minute is output.

[OPTIONS]

-c *clstname*

This option outputs power consumption information about the compute cluster specified in *clstname*. You can specify only one compute cluster. Nothing other than a compute cluster can be specified. You can omit this option by setting the environment variable PXYMYCLST.

If the -n option or --nodelist option is specified, the command outputs power consumption information of the specified compute nodes. This option can be specified with the --extdev option or --extdevlist option.

This option can not be specified with the --pwrgrp option or --pwrgrpulist option.

-n {*nodeid*[,...] | *nodeid-nodeid* | all}

This option outputs power consumption information about each compute node specified in *nodeid*.

Specify *nodeid* in hexadecimal. You can specify multiple nodeids by delimiting them with a comma (,) (example: 0x01010AF1,0x01010AF2).

You can specify a range with IDs separated by a hyphen (-) (example: 0x01010AF1-0x01010AF9).

If all is specified, all compute nodes are targeted.

If a duplicate *nodeid* is specified, the duplicate is ignored.

A non-existent *nodeid* cannot be specified.

This option cannot be specified together with the --nodelist, --pwrgrp, or --pwrgrpulist option.

When specifying this option, be sure to specify a compute cluster name in *clstname* in the -c option, or specify a compute cluster name in the environment variable PXYMYCLST.

The number of characters specified for this option cannot be 1,000 or more.

--nodelist *file*

This option outputs power consumption information about the compute nodes written in the file (file full path) specified in *file*. The number of characters specified for *file* cannot be 256 or more. A non-existent or empty file cannot be specified.

Write node IDs in the file (example: 0x010100AF1).

Write one node ID per line.

To specify a range of compute nodes, write node IDs separated by a hyphen (-) (example: 0x010100AF1-0x01010AF9) in the file.

If a duplicate node ID is written, the duplicate is ignored.

A non-existent node ID cannot be specified.

This option cannot be specified together with the -n, --pwrgrp, or --pwrgrpulist option.

When specifying this option, be sure to specify a compute cluster name in *clstname* in the -c option, or specify a compute cluster name in the environment variable PXYMYCLST.

--extdev {*extdevname*[,...] | all}

This option outputs power consumption information about each external device specified in *extdevname*.

Choose from the names defined in `/etc/opt/FJSVtcs/pwrwm/extdev` to specify *extdevname*. You can specify multiple *extdevnames* by delimiting them with a comma (example: `facility1,facility2`).

If all is specified, all external devices are targeted.

If a duplicate *extdevname* is specified, the duplicate is ignored.

A non-existent *extdevname* cannot be specified.

This option can be specified together with the `-c` option.

It cannot be specified together with `--extdevlist`, `--pwrgrp`, or `--pwrgrp1` option.

The number of characters specified for this option cannot be 1,000 or more.

--extdevlist *file*

This option outputs power consumption information about the external devices written in the file (file full path) specified in *file*. The number of characters specified for *file* cannot be 256 or more. A non-existent or empty file cannot be specified.

Write external device names in the file (example: `facility1`).

Write one device name per line.

If a duplicate external device name is written, the duplicate is ignored. A non-existent external device name cannot be written.

This option can be specified together with the `-c` option.

It cannot be specified together with the `--extdev`, `--pwrgrp`, or `--pwrgrp1` option.

--pwrgrp {*pwrgrpname*[,...] | all}

This option outputs power consumption information about individual devices composing the power group specified in *pwrgrpname*.

Choose from the names defined in `papwrwm.conf(5)` to specify *pwrgrpname*.

You can specify multiple *pwrgrpnames* by delimiting them with a comma (,) (example: `pwrgrp1,pwrgrp2`).

If all is specified, all of the power group is targeted.

If a duplicate *pwrgrpname* is specified, the duplicate is ignored.

A non-existent *pwrgrpname* cannot be specified.

This option cannot be specified together with other scope options.

The number of characters specified for this option cannot be 1,000 or more.

--pwrgrp1 *file*

This option outputs power consumption information about individual devices composing the power group written in the file (file full path) specified in *file*. The number of characters specified for *file* cannot be 256 or more. A non-existent or empty file cannot be specified.

Write a power group in the file (example: `grp1`).

Write one power group name per line.

If a duplicate power group name is written, the duplicate is ignored. A non-existent power group name cannot be written.

This option cannot be specified together with other *scope options*.

--average

If the `--trace` option is not specified, this option preferentially uses average power consumption to calculate a total power consumption value. If neither the `--average` nor `--momentary` option is specified, the calculation preferentially uses average power consumption.

If the `--trace` option is specified, this option outputs average power consumption.

If average power consumption is not measured, the command does not output them.

If neither the `--average` nor `--momentary` option is specified, both average power consumption and momentary power consumption are output.

--momentary

If the `--trace` is not specified, this option preferentially uses momentary power consumption to calculate a total power consumption value. If neither the `--average` nor `--momentary` option is specified, the calculation preferentially uses average power consumption.

If the `--trace` option is specified, this option outputs momentary power consumption.

If momentary power consumption is not measured, the command does not output them.

If neither the `--average` nor `--momentary` option is specified, both average power consumption and momentary power consumption are output.

--last

This option uses the most recently measured power consumption to calculate a total value, even if the measurement time exceeds the allowable range.

-v

This option outputs line-by-line information on the power consumption of compute nodes and external devices. The information has been used to calculate a total power consumption value.

--data

This option outputs each fields separated with a delimiter (comma ","). You can change the delimiter by the --delimiter option.

--delimiter *delimiter*

This option changes the delimiter in the output format of the --data option to the delimiter specified in *delimiter*, and outputs data.

--trace

This option outputs power consumption information saved in the system power database. If the --time option is not specified, power consumption information for the latest 10 minutes is output. If this option is specified, individual power consumption information instead of a total value is output.

If this option (without the --data option) is specified, only up to 10 characters can be output for extdevname and pwrgrpname. Set extdevname and pwrgrpname to 10 characters or less, or specify with the --data option.

--time *YYYYMMDDhhmm[ss][-{YYYYMMDDhhmm[ss]}now]*

For the compute nodes and external devices specified by scope options, this option outputs the average power consumption and/or momentary power consumption at the closest measurement time (but within one hour) earlier than the specified time. If ss is not specified, ss=00.

You can specify the period for output by adding a time with a preceding hyphen (-) to the end of the option. If "now" instead of a time is specified, power consumption information up to the latest is output.

If the type of the external device is ENE (integral power consumption), the average power consumption is calculated from the integral power consumption up to 1 minute before the specified time.

--help

This option displays how to use this command.

[Calculation of a total power consumption value]

When calculating a total power consumption value, specify the --average option to preferentially use average power consumption or the --momentary option to preferentially use momentary power consumption. If neither option is specified, the calculation preferentially uses average power consumption.

The measurement time of power consumption used to calculate the total value varies with each compute node or external device. For this reason, you can set an allowable range of variation in measurement time until the total value is calculated. For details on how to set the allowable range, see papwrn.conf(8). If the allowable range is exceeded, power consumption outside the range is not included in the calculation of the total value.

Even when average power consumption has priority, momentary power consumption is used to calculate the total value if average power consumption was not but momentary power consumption was measured within the allowable range. Conversely, even when momentary power consumption has priority, average power consumption is used to calculate the total value if momentary power consumption was not but average power consumption was measured within the allowable range. If neither average power consumption nor momentary power consumption was measured within the allowable range, the total value cannot be calculated.

Even when the measurement time exceeds the allowable range, however, if the --last option is specified, the most recently measured power consumption is used to calculate the total value.

Kinds of the power consumption to calculate the total value is output as the item PWR_TYPE.

Kinds of the power consumption to calculate the total value	Descripton
AVE	Summed only average power consumption
MON	Summed only momentary power consumption
MIX	Summed combination of average power consumption and momentary power consumption
ERR	Can not calculate the total value.

Even if the kind is ERR, exit status of this command is 0.

[ENVIRONMENT VARIABLE]

PXMYCLST

To output latest power consumption information of all devices, you must set the target compute cluster name to the environment variable PXMYCLST. You can specify only one compute cluster name.

To output latest/past power consumption information of specified devices, you can set the compute cluster name to the environment variable PXMYCLST instead of specifying the -c option. You can specify only one compute cluster name.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

An example is shown below.

[Example] export PXMYCLST=cluster1

[EXIT STATUS]

The following end codes are returned.

0 : Normal end

1 : Abnormal end

[SEE ALSO]

papwrm.conf(5), papwrmgradm(8)

[MESSAGES]

If an error occurs during execution of the pasyspwr command, one of the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PWRM 1101 pasyspwr This command cannot be executed on this node

Meaning

An attempt was made to execute the command on a node other than the system management node.

Action

Execute the command on the system management node.

[ERR.] PWRM 1102 pasyspwr Permission denied: pasyspwr

Meaning

The user does not have the privileges to execute the pasyspwr command.

Action

Check the user who executed the command.

If the correct user executed the command, grant system or cluster administrator privileges to the user to execute the command again.

[ERR.] PWRM 1103 pasyspwr Invalid cluster name: *clstname*

Meaning

The specified cluster name *clstname* is invalid.

Action

Specify a valid cluster name.

[ERR.] PWRM 1105 pasyspwr Invalid scope option: *details*

Meaning

The specified scope option is invalid.

details: Detailed information pointing out what is invalid

Action

Specify the correct scope option.

Note

For details on the scope option, see "[OPTIONS]" in "4.1.24 pasyspwr command."

[ERR.] PWRM 1107 pasyspwr Invalid option: *opt*

Meaning

The specified command option *opt* is invalid.

Action

Specify a valid option.

[ERR.] PWRM 1108 pasyspwr Invalid option argument: *opt arg*

Meaning

The setting value for the argument *arg* of the command option *opt* is invalid.

Action

Specify a valid setting value.

[ERR.] PWRM 1109 pasyspwr Internal error: *details*

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

Only the cluster names of compute clusters can be specified for command options. Check whether the specified option is correct. If it is correct, collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PWRM 1141 pasyspwr File not found: *file*

Meaning

The file *file* does not exist.

Action

Specify the correct file.

[ERR.] PWRM 1142 pasyspwr File permission denied: *file*

Meaning

The user does not have privileges for the file *file*.

Action

Specify the correct file.

[ERR.] PWRM 1143 pasyspwr cannot be read file: *file*

Meaning

The file *file* cannot be read.

Action

Specify the correct file.

[ERR.] PWRM 1161 pasyspwr *name* is not found

Meaning

The node ID, external device name, or power group name specified by *name* does not exist

Action

Specify the correct node ID, external device, or power group name.

Information Message

[INFO] PWRM 1150 pasyspwr System power is not available

Meaning

There is no system power information.

Action

Confirm that the SystemPower section is set in the papwrn.conf file. If not, set the SystemPower section in the papwrn.conf file, execute the papwrngradm command --set option, wait at least one minute, and then execute the pasyspwr command again.

4.1.25 pmalter command

[NAME]

pmalter - Changes the parameters of a submitted job.

[SYNOPSIS]

```
/usr/sbin/pmalter [-c clstname]
                  [{-L|--rsc-list} {elapse|rscunit|ru|rscgrp|rg}=value[,...]]
                  [-p priority]
                  [{-P|--apriority} priority]
                  jobid ...
/usr/sbin/pmalter --help
```

[DESCRIPTION]

The pmalter command changes the following parameters of a submitted job:

- Job elapsed time limit ({-L|--rsc-list} *elapse*=*elapselimit*)
- Name of the resource unit executing the job ({-L|--rsc-list} {*rscunit*|*ru*}=*rscuname*)
- Name of the resource group executing the job ({-L|--rsc-list} {*rscgrp*|*rg*}=*rscgname*)
- Job Priority among the same user (-p *priority*)
- Job priority within a resource unit ({-P|--apriority} *priority*)

The pmalter command can be executed on the system management node and compute cluster management node.

The job ID, a sub job ID of a step job or a sub job ID of a bulk job is specified for *jobid*.

If the job ID of a bulk job or step job is specified, all the sub jobs in the job are targets (except the ended sub jobs in the step job).

You can change the parameters, except resource unit names, of a sub job of a step job by specifying its sub job ID.

When specifying the sub job ID of a bulk job, only the elapsed time limit value can be changed.

The pmalter command cannot be used for interactive jobs.

The pmalter command cannot be used for jobs that specify the minimum value of elapsed time (example: specifying only minimum value 10-, minimum value and maximum value 10-100).

For each parameter, the states of jobs and sub jobs that pmalter command can be accepted are listed as follows. If the targets are not in any of these states, the pmalter command ends with an error.

- Job elapsed time limit

You can change parameters with the `pmlalter` command only if the target jobs and sub jobs are in the QUEUED, RUNNING, HOLD, or ERROR state.

Changing the elapsed time limit value in the RUNNING state is possible only when the function to change the elapsed time limit value for jobs in the RUNNING state is enabled and the jobs are on the FX server compute node.

- Name of the resource unit executing the job
- Name of the resource group executing the job
- Job Priority among the same user
- Job priority within a resource unit

You can change parameters with the `pmlalter` command only if the target jobs and sub jobs are in the QUEUED, HOLD, or ERROR state.

If you have changed the parameters of a job in the QUEUED state, the job is rescheduled.

The specified parameter values are checked according to the job ACL function settings, and an error occurs in the following cases.

- The job priority exceeds the specifiable upper limit.
- The owner of the job is not permitted to use the specified resource unit or resource group.

If the `pmlalter` command failed, the parameter and the status of the job does not change.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using "-" (hyphen). Similarly, you can also specify a range of sub job IDs.

Example 1: Specify a job ID range of 1 to 100

1-100

Example 2: Specify a step number range of 1 to 10 for a step job (job ID = 100)

100_1-10

Example 3: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 100)

100[1-10]

When multiple job IDs (sub job IDs) or a range of job IDs (sub job IDs) are specified, the processing is not aborted, and all target jobs (sub jobs) are processed, even if the error occurs in a certain job (sub job).

- Multiple specification
Because the job IDs (sub job IDs) are specified explicitly, the error message is displayed when the error occurs.
- Range specification
Because the error message is not displayed, it is necessary to confirm the job attributes by the `pjstat` command.

[OPTIONS]

-c *clstname*

This option specifies the target cluster name *clstname*. If the -c option is omitted, the cluster name is the value of the environment variable `PXMYCLST`. If even the environment variable `PXMYCLST` has not been set, an error occurs.

When this option is specified, this option is disregarded in execution by the compute cluster management node.

The option specification method conforms to that of the `pjsub` command.

{-L|--rsc-list} "resource=*value*[,resource=*value*...]"

This option specifies the upper limit value of the resources required by jobs.

The numerical values or character strings that can be specified in *value* for each *resource* are different.

The resources shown below can be specified for *resource*. When specifying multiple values or strings, delimit them with a comma (",").

elapse=elapsedtimelimit

This changes the job elapsed time limit to *elapsedtimelimit* seconds. The minimum value is 1 second, and the maximum value is unlimited.

It is not possible to specify the minimum value specification of elapsed timelimit (example: specify minimum value only 10-, minimum value and maximum value specified 10-100).

If an acceptance message is output after executing `pmalter` command for the job in the `RUNNING` state, it does not necessarily mean that the job elapsed time limit value change succeeded. Please confirm whether the job elapsed time limit was correctly changed in the `RUNNING` state by the `pjstat` command.

Changing the elapsed time limit value in the `RUNNING` state is possible only when the function to change the elapsed time limit value for jobs in the `RUNNING` state is enabled and the jobs are on the FX server compute node.

`{rscunit|ru}=rscuname`

This changes the resource unit where the job is submitted to *rscuname*. Permission to submit the job to the changed resource unit is necessary.

The resource unit name can be set with 63 bytes at most.

When a resource group with `rscgrp=rscgname` is not specified, this also changes a resource group to the default resource group of the resource unit *rscuname* according to the job ACL function settings.

It is not allowed to change from a resource unit with FX server to a resource unit with PRIMERGY server. Also, it is not allowed to change from a resource unit with PRIMERGY server to a resource unit with FX server.

`{rscgrp|rg}=rscgname`

This changes the resource group where the job is submitted to *rscgname*. Permission to submit the job to the changed resource group is necessary.

The resource group name can be set with 63 bytes at most.

`-p priority`

This option sets the job priority to *priority*. You can specify a value from 0 to 255 for *priority*. The lowest priority is 0, and the highest priority is 255.

The job priority among the same user is determined according to the specified priority.

`{-P | --apriority} priority`

This option sets the job priority within a resource unit to *priority*. You can specify a value from 0 to 255 for *priority*. The lowest priority is 0, and the highest priority is 255.

The job priority within a resource unit is determined according to the specified priority.

`--help`

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[END STATUS]

0: Normal end

1: Abnormal end (When multiple jobs are specified, 1 is returned even if only one of them ends with an error.)

[SEE ALSO]

`pjalter(1)`, `pjstat(1)`, `pjsub(1)`

[MESSAGES]

If an error occurs during the execution of the `pmalter` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 1001 pmalter Unknown option or no option argument: *opt*.
Try `pmalter --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 1003 pmlter Unknown option argument: *opt arg*.
Try `pmlter --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 1004 pmlter Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 1006 pmlter No parameter specified.
Try `pmlter --help` for more information.**

Meaning

No parameter to be changed is specified.

Action

Specify the parameter to be changed.

**[ERR.] PJM 1010 pmlter No jobid(s) specified.
Try `pmlter --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 1011 pmlter Invalid jobid: *jobid*.

Meaning

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 1012 pmlter Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job exists.

[ERR.] PJM 1013 pmlter Job *jobid* is not in the state that pmlter can be accepted.

Meaning

The command cannot be executed in the state of the job specified by the job (job ID *jobid*).

Action

To change any of the following parameters, execute the command when the state of the job is QUEUED, HOLD, or ERROR:

- Resource unit for executing a job
- Resource group for executing a job
- Priority among jobs of the same user
- Priority in resource unit

To change any of the following parameters, execute the command when the state of the job is QUEUED, RUNNING, HOLD, or ERROR. In the RUNNING state, you can change the elapsed time limit value only under both of the following conditions. First, the function to change the elapsed time limit values of jobs in the RUNNING state is enabled. Second, the job is on an FX server compute node.

- Elapsed time limit value

[ERR.] PJM 1014 pmalter Job *jobid* is interactive job. pmalter cannot be executed for interactive jobs.

Meaning

The job specified by job ID *jobid* is interactive job. The pmalter command cannot be executed for interactive jobs.

Action

Specify any job other than an interactive job.

[ERR.] PJM 1015 pmalter It is unable to change resource by specifying sub job ID.

Meaning

You cannot specify a sub job ID for a step job when changing the resource unit name.

You cannot specify a sub job ID for a bulk job when changing the resource unit name or the resource group name.

Action

Review the specified job ID.

[ERR.] PJM 1051 pmalter The specified time is too short: *limittime*, *jobid*.

Meaning

Resource information specified for a job (job ID *jobid*) with the -L|--rsc-list option is incorrect.

- If *limittime* is elapse

The specified elapsed time limit value is shorter than the execution time that has already elapsed.

Action

Specify a limit value higher than the execution time that has already elapsed.

[ERR.] PJM 1052 pmalter Job *jobid* parameter is being changed.

Meaning

The command cannot be executed because a parameter change is being processed for the job specified with the job ID *jobid*.

Action

Wait a moment, and execute the command again.

[ERR.] PJM 1054 pmalter *rscname=value* that have been specified in the job *jobid* is less than the lower limit (*limit-value*).

Meaning

When a job with the job ID *jobid* was submitted, the value *value* for the resource name *rscname* specified in the -L|--rsc-list option was lower than limit-value, the lower limit value defined by the job ACL function of a changed resource unit or resource group.

Action

Specify a resource limit value that is not below the range defined by the job ACL function of the changed resource unit or resource group. Alternatively, specify a resource unit or resource group whose resource limit value is not below the range defined by the job ACL function.

[ERR.] PJM 1057 pmlter *rscname=value* that have been specified in the job *jobid* is greater than the upper limit (*limit-value*).

Meaning

When a job with the job ID *jobid* was submitted, the value *value* for the resource name *rscname* specified in the -L|--rsc-list option exceeded limit-value. *limit-value* is either the upper limit value defined by the job ACL function of a changed resource unit or resource group, or the amount of resources that can be allocated to a job.

Action

Specify a resource amount that does not exceed the range defined in the job ACL as the limit value for the resource. Or, specify a resource unit or resource group whose resource limits do not exceed the range defined in the job ACL feature and whose resource amount (Number of nodes, node shape, etc.) can be assigned to the job.

For details on the amount of resources that can be allocated to jobs on the FX server, see the table "Size of nodes that can be allocated [FX]" in "Checking resource units and resource groups" in "Chapter 2 Job Operation Procedures" of the manual "Job Operation Software End-User's Guide."

[ERR.] PJM 1058 pmlter *rscname=value* does not exist.

Meaning

- If the resource name *rscname* is *rscunit*
The specified resource unit *value* does not exist.
- If the resource name *rscname* is *rscgrp*
The specified resource group *value* does not exist.

Action

Specify an existing resource unit name or resource group name.

[ERR.] PJM 1059 pmlter *rscname=value* is disabled.

Meaning

- If the resource name *rscname* is *rscunit*
A job currently cannot be submitted to the specified resource unit *value*.
- If the resource name *rscname* is *rscgrp*
A job currently cannot be submitted to the specified resource group *value*.

Action

Check the state of the resource unit or resource group with the --rsc option of the *pjstat* command. Then, specify a resource unit or resource group to which the job can be submitted. If necessary, ask the administrator to change the state of the resource unit or resource group.

[ERR.] PJM 1061 pmlter The architecture of compute nodes in the resource unit is different before and after the change.

Meaning

The resource unit for FX servers cannot be changed to the resource unit for PRIMERGY servers. Likewise, the resource unit for PRIMERGY servers cannot be changed to the resource unit for FX servers.

Action

Specify a resource unit that has the same architecture.

[ERR.] PJM 1070 pmalter No execute permission: *jobid*.

Meaning

Changing the job parameter is failed because there is no authority to submit a job *jobid* in the resource unit or the resource group that is specified.

Action

Display the job ACL settings for the job submission user with the `pjacl` command. Then, confirm the job submission authority (execute `pjsub` and execute `pjsub(--xxxx)`: *xxxx* is the option name) for the resource unit or the resource group that is specified.

[ERR.] PJM 1071 pmalter No permission: *jobid*.

Meaning

The user does not have sufficient permission to operate the job *jobid*.

Changing of the parameter is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the `pjacl` command. Then, confirm the permission (permit `pmalter`) for the operated objects of the `pmalter` command for the resource unit and resource group to which the target job is submitted.

[ERR.] PJM 1072 pmalter Job exceeded the accept limit: *target, jobid*.

Meaning

The following problem was encountered in an attempt made to change the resource unit or group for executing the job specified with the job ID *jobid*: The maximum number of jobs that can be concurrently accepted for the change-to resource unit or group is reached.

The value of *target* indicates the following:

- `ru-accept`: Number of batch jobs that can be concurrently accepted for a resource unit
- `ru-interact-accept`: Number of interactive jobs that can be concurrently accepted for a resource unit
- `rg-accept`: Number of batch jobs that can be concurrently accepted for a resource group
- `rg-interact-accept`: Number of interactive jobs that can be concurrently accepted for a resource group
- `ru-accept-allsubjob`: Number of sub jobs and normal jobs (batch jobs) for bulk and step jobs that can be concurrently accepted for a resource unit
- `ru-accept-bulksubjob`: Number of bulk sub jobs that can be concurrently accepted for a resource unit
- `ru-accept-stepsubjob`: Number of step sub jobs that can be concurrently accepted for a resource unit
- `rg-accept-allsubjob`: Number of sub jobs and normal jobs (batch jobs) for bulk and step jobs that can be concurrently accepted for a resource group
- `rg-accept-bulksubjob`: Number of bulk sub jobs that can be concurrently accepted for a resource group
- `rg-accept-stepsubjob`: Number of step sub jobs that can be concurrently accepted for a resource group

Action

Using the `pjacl` command with the `--rscunit` or `--rscgrp` option, confirm the maximum number of jobs that can be concurrently accepted for the change-to resource unit or group.

The maximum number of jobs that can be concurrently accepted for a resource unit is the "acceptablejob" or "acceptable job(interact)", "acceptable all-subjob", "acceptable bulk-subjob", or "acceptable step-subjob" value of the item "limit in rscunit (each users)," "limit in rscunit (total users in same group)," or "limit in rscunit (total all users)."

The maximum number of jobs that can be concurrently accepted for a resource group is the "acceptablejob" or "acceptable job(interact)", "acceptable all-subjob", "acceptable bulk-subjob", or "acceptable step-subjob" value of the item "limit in rscgroup (each users)," "limit in rscgroup (total users in same group)," or "limit in rscgroup (total all users)."

[ERR.] PJM 1073 pmalter Option value is out of range: *target*, *jobid*.**Meaning**

The option value *target* specified for job *jobid* exceeds the effective range defined by job ACL.

Action

Display the job ACL settings, and then confirm the definition value of the specified option.

[ERR.] PJM 1075 pmalter Invalid attribute of job: *jobid* (*attr*).**Meaning**

Jobs with the attribute *attr* are not allowed to execute on the resource unit or resource group to which they are modified.

Action

Cannot be changed to resource units or resource groups that have the following settings:.

- For jobs with the attribute io-exclusive
The job ACL function definition item define allocation-io-mode is set to no-io-exclusive and execute io-exclusive is set to disable.
- For jobs with the attribute no-io-exclusive
The job ACL function definition items define allocation-io-mode is set to io-exclusive and execute no-io-exclusive is set to disable.

Use the `pjacl` command to check the attributes of jobs that can execute on the resource units or resource groups to which they are modified.

**[ERR.] PJM 1076 pmalter *rscname=value* that have been specified in the job *jobid* cannot be specified. The value must be specified as follows:
strings,...****Meaning**

The value *value* for the resource name *rscname* specified in the `-L|--rsc-list` option for a job with the job ID *jobid* is not among the types *strings* that can be specified for the custom resources defined by the job ACL function. *strings*,... is a list of types of specifiable custom resources.

Action

Specify a resource unit or resource group that has a defined custom resource.

[ERR.] PJM 1078 pmalter *rscname=value* that have been specified in the job *jobid* does not exist.**Meaning**

The value *value* for the resource name *rscname* specified in the `-L|--rsc-list` option is for a job with the job ID *jobid* does not have a defined custom resource.

Action

Specify a resource unit name or resource group name that has a defined custom resource.

[ERR.] PJM 1079 pmalter Arbitrary character string.**Meaning**

An error occurred at the job attribute change exit that was set by the administrator. The contents of this message were set by the administrator.

Action

Use the message to resolve why the job attribute change exit failed to change the parameter.

[ERR.] PJM 1089 pmalter Not supported: *message*.

Meaning

The current version of Technical Computing Suite does not support the parameter change of the specified job.

Action

If you want to change the specified parameter, delete the target job and submit it again specifying with suitable parameter.

[ERR.] PJM 1090 pmalter PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

The system might not have been able to accept the request temporarily. Retry the command execution. If the situation does not change in about ten minutes, check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 1091 pmalter Internal error: *details*.

Meaning

An internal error occurred.

details: Detailed information for maintenance

Action

The system might not have been able to accept the request temporarily. Retry the command execution. If the situation does not change in about ten minutes, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PJM 1093 pmalter The operating command cannot be executed except system management node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a system management node or compute cluster management node.

Action

Execute the command on a system management node or compute cluster management node.

[ERR.] PJM 1095 pmalter No cluster specified on system management node.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the pjstat command, or set a cluster name in the environment variable PXMYCLST.

[ERR.] PJM 1096 pmalter Job *jobid* was rejected *opt. message*

Meaning

The parameter change for the job (job ID: *jobid*) specified in the option *opt* was rejected.

opt: -L elapse: Specified elapsed time limit value

message shows the reason why the parameter change was rejected.

Message	Meaning
Deadline schedule exists.	The attribute change for the job (job ID: <i>jobid</i>) specified in the option <i>opt</i> was rejected because a set deadline schedule already

Message	Meaning
	exists and this job execution time overlaps the deadline schedule.
Changing the elapse time limit of a running job is only supported for the job on FX server compute node.	The elapsed time limit values of running jobs executed on servers other than FX servers cannot be changed.
Internal error (<i>details</i>).	An internal error occurred.

Action

The appropriate action varies depending on the message displayed in *message*.

- When "Deadline schedule exists." is displayed
Reduce the specified (post-change) elapsed time limit value, or delete and then resubmit the job.
(Note)
After the elapsed time limit value is changed, the scheduled job end time and deadline schedule start time are times rounded off according to the setting of the minimum interval for resource map creation (CreateRscMap). You can postpone the scheduled job end time by extending the elapsed time limit value of the currently running job. However, that is not necessarily always the case since the time is rounded off according to the minimum interval setting for resource map creation. If exclusive use of the same node is scheduled in a set deadline schedule, the time cannot be postponed to any time immediately before the scheduled start time.
- When "Changing the elapse time limit of a running job is only supported for the job on FX server compute node." is displayed
If you want to change the elapsed time limit value of a running job, execute the job on an FX server.
- When "Internal error (*details*)." is displayed
Re-execute the command because the system may have been temporarily unable to accept the request. If the situation remains unchanged after about 10 minutes, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk with the collected data together with the output message.

Information Message

[INFO] PJM 1000 pmalter Job *jobid* is accepted *opt*.

Meaning

The change to the job (jobID: *jobid*) parameter specified by the option *opt* was accepted.

Action

No action is necessary.

[INFO] PJM 1081done.

Meaning

The request to change the job parameter waits for the acceptance. One dot (.) appears every three seconds. 'done.' appears when the acceptance waiting ends. If the acceptance waiting ends at less than three seconds, this message is not displayed.

Action

No action is necessary.

4.1.26 pmdumpjobinfo command

[NAME]

pmdumpjobinfo - Outputs the content of the job statistical information files.

[SYNOPSIS]

```
/usr/sbin/pmdumpjobinfo [-c clstname] [--delimiter dlmt] [--end date_time]
                        [{-g|--gname} gname[,...]] [{-g|--gid} gid[,...]]
```



```

[--jobid jobid[,...]] [--start date_time] [--quote]
[-s|-S] [{-u|--uname} uname[,...]] [{-u|--uid} uid[,...]]
[{--rscunit|--ru} rscuname[,...]]
[{--rscgrp|--rg} rscgname[,...]]
[jobinfo [ ...]]

/usr/sbin/pmdumpjobinfo --help

```

[DESCRIPTION]

This `pmdumpjobinfo` command outputs statistical information from the job statistical information file specified with *jobinfo*.

This command can be executed on the system management node or compute cluster management node.

In order to specify multiple *jobinfo* files, put them in order from oldest to most recent data by separating them with a space.

If arguments are not set, `/var/opt/FJSVtcs/shared_disk/pjm/jsti/jobinfo` will be used as an input file.

If the command is executed on the system management node, and no "/" is included in the specified *jobinfo* path, files under the directory `/var/opt/FJSVtcs/shared_disk/pjm/jsti/` of compute cluster management node should be copied to directory `/var/opt/FJSVtcs/shared_disk/pjm/jsti/` of the system management node.

The content of a job statistical information file is processed to a format that is separated with a comma (","),. Each record of the file is output to one line.

The numerical values are output in decimal number. The strings do not output any extra spaces after the text part. The reserve area outputs 0.

If there is an offset in the record, output the content which pointed by the offset instead of the offset itself.

At the beginning of each of the lines, the following character strings are output to identify records:

symbol flag	name of record
START_JACCT	Record to start outputting statistical information
JACCT_HEAD	Record of statistical information item
JI	Record of main job information
JA	Record of allocated node information
JN	Record of node information / Record of virtual node information
SHUT_JACCT	Record to stop outputting statistical information

RECORD TO START OUTPUTTING STATISTICAL INFORMATION

Record to start outputting statistical information outputs the following data.

Field	Description
1	Record name Outputs "START_JACCT" that indicates the statistical information output start record.
2	Statistical information output start time (YYYYMMDD.hhmmss)

RECORD OF STATISTICAL INFORMATION ITEM

Record of allocated node information outputs the following data.

Field	Description
1	Record name Outputs "JACCT_HEAD" that indicates the record of statistical information item.
2	"JI" indicating the job information body record or "JN" indicating the node information record/virtual node information record is output.
3 or later	item name to be output to the job information body record or node information record/virtual node information record.

RECORD OF MAIN JOB INFORMATION

When the job execution ends or the job is deleted, the record of main job information is output to the job statistical information file. For each sub job of a bulk job or step job, when the sub job execution ends or the sub job is deleted, the record of main job information is output to the job statistical information file. In addition, when all its sub jobs end, the record of main job information corresponding to the job ID is

output to the job statistical information file as a summary. Information in bulk job or step job corresponding to this job ID is called summary information on bulk job or step job.

For the record of main job information that is output, see `pjstatsinfo(7)`.

RECORD OF ALLOCATED NODE INFORMATION

When the job execution ends, the record of allocated node information is output to the job statistical information file following the record of main job information. For each sub job of a bulk job or step job, when the sub job execution ends, the record of allocated node information is output to the job statistical information file. In addition, when all its sub jobs end, the record of main job information corresponding to the job ID is output to the job statistical information file as a summary, but the allocated node information corresponding to the job ID is not output to the job statistical information file.

Record of allocated node information outputs the following data.

Field	Description
1	Record name "JA" representing record of allocated node information is output.
2	Number of allocated nodes [for jobs on FX server compute node, or node allocated jobs on PRIMERGY server compute node] Number of allocated virtual nodes [for virtual node allocated jobs on PRIMERGY server compute node]
3	Node ID Allocated node IDs and two or more node IDs are separated by spaces.

RECORD OF NODE/VIRTUAL NODE INFORMATION

When the job execution ends, the records of node/virtual node information are output to the job statistical information file following the record of main job information and the record of allocated node information. (only as many records as the number of nodes allocated [for jobs on FX server compute node, or node allocated jobs on PRIMERGY server compute node] or virtual nodes allocated [for virtual node allocated jobs on PRIMERGY server compute node]).

For each sub job of a bulk job or step job, when the sub job execution ends, the records of node/virtual node information are output to the job statistical information file and the record of allocated node information. In addition, when all its sub jobs end, the record of main job information corresponding to the job ID is output to the job statistical information file as a summary, but the node/virtual node information corresponding to the job ID is not output to the job statistical information file.

For the record of node used [for jobs on FX server compute node, or node allocated jobs on PRIMERGY server compute node] or virtual nodes used [for virtual node allocated jobs on PRIMERGY server compute node] information that is output, see `pjstatsinfo(7)`.

RECORD TO STOP OUTPUTTING STATISTICAL INFORMATION

Record to stop outputting statistical information outputs the following data.

Field	Description
1	Record name Outputs "SHUT_JACCT" that represents the statistical information output start record.
2	Statistical information output stop time (YYYYMMDD.hhmmss)

[OPTIONS]

`-c clstname`

Use this option to display the contents of statistics information files of jobs of *clstname* cluster.

If the environment variable `PXMYCLST` is set, this option can be omitted.

If the command is executed on compute cluster management node, this specified option will be disabled.

`--delimiter dlimt`

If `--data` option is specified, use *dlimt* to set the delimiter. *dlimt* is effective up to 16 characters and 17 characters or more are cut out.

`--end yyyymmdd[.HHMMSS]`

Specify the stop time of period of output. The job stop time of the statistics information record will be set according to the record which is earlier than the specified stop time. If this option is omitted, the output will cover the contents till the end of statistics information

record.

If hh:mm:ss is omitted, 24:00:00 will become effective.

{-g|--gname} *gname*[,...] | {-g|--gid} *gid*[,...]

Output the statistics information of specified group. If multiple groups are specified, use comma to separate them. Either group name(*gname*) or group ID(*gid*) will be OK.

Group names which just include digits will be treated as GID. If one group name does just include digits, use the --gname option instead of -g option.

--jobid *jobid*[,...]

Output the statistics information of specified job ID. If multiple job IDs are specified, use comma to separate them.

The sub job ID of a step job or a bulk job also can be specified.

If the job ID of a bulk job or step job is specified, the summary information on the bulk job or step job and information of all sub jobs are output.

--start *yyyymmdd[.HHMMSS]*

Specify the start time of period of output. The job stop time of the statistics information record will be set according to the record which is later than the specified start time (or equal to the specified start time). If this option is omitted, the head (old data) of the statistics information record will be used. If hh:mm:ss is omitted, 00:00:00 will become effective.

-s

Output the job information records (summary information on bulk job or step job is included) fetched from files of statistics record and display them one line per record.

-S

Output the job information records (summary information on bulk job or step job is included) and allocated node information record, node/virtual node information records fetched from files of statistics record. After the output of job information records, allocated node information record will be displayed from the next line. After the output of allocated node information records, the node/virtual node information records will be displayed one line per record from the next line.

{-u|--uname} *uname*[,...] | {-u|--uid} *uid*[,...]

Output the statistics information of specified user. If multiple users are specified, use comma to separate them. Either user name(*uname*) or user ID(*uid*) will be OK.

User names which just include digits will be treated as UID. If one user name does just include digits, use the --uname option instead of -u option.

--quote

The string item covers with a double quotation. A double quotation in the item character escapes by adding a double quotation.

{--rscunit|--ru} *rscuname*[,...]

Output the statistics information of specified resource unit name (*rscuname*).

When this option is omitted, job statistics information of all resource units is output.

{--rscgrp|--rg} *rscgname*[,...]

Output the statistics information of specified resource group name (*rscgname*).

When this option is omitted, job statistics information of all resource groups is output.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[REMARK]

If --uname, --gname and --jobid options are specified at the same time, all statistics information of corresponding jobs should be output.

e.g.) --uname usr1 --gname grp2 --jobid 100-200

The statistics information of jobs which are on condition of that the job ID is from 100 to 200, the user is usr1 and the group is grp2 should be output.

[END STATUS]

0: Normal end
1: Abnormal end

[SEE ALSO]

pjsub(1), pjstatsinfo(7)

[MESSAGES]

If an error occurs during execution of the `pmdumpjobinfo` command, one of the following messages is output to the standard error output.

Error Message

**[ERR.] PJM 5001 pmdumpjobinfo Unknown option *opt*.
Try ``pmdumpjobinfo --help`` for more information.**

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 5004 pmdumpjobinfo Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 5020 pmdumpjobinfo Cannot open the file (*path*): *code*.

Meaning

The specified file *path* cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 5030 pmdumpjobinfo Read error (*path*): *code*.

Meaning

Reading of the resource file *path* failed.

code: Internal code for maintenance

Action

The statistical information file may be damaged. Confirm the contents of the statistical information file.

[ERR.] PJM 5031 pmdumpjobinfo Seek error (*path*): *code*.

Meaning

Locating the specified file *path* failed.

code: Internal code for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 5091 pmdumpjobinfo Internal error: *details*.

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 5093 pmdumpjobinfo This command can be executed on SMM/CCM.

Meaning

Can only be executed on the system management node or compute cluster management node.

Action

Execute the command on the system management node or compute cluster management node.

[ERR.] PJM 5095 pmdumpjobinfo No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

[ERR.] PJM 5097 pmdumpjobinfo In the standby node, pmdumpjobinfo cannot be used.

Meaning

The pmdumpjobinfo command cannot be used on the standby node.

Action

Execute the pmdumpjobinfo command on the active system management node or compute cluster management node.

Warning Message

[WARN] PJM 5033 pmdumpjobinfo Invalid header or footer: *code*.

Meaning

The header or footer of the read record is invalid. Output record except invalid data.

code: Internal code for maintenance

Action

No action is necessary.

[WARN] PJM 5034 pmdumpjobinfo Illegal data.

Meaning

Invalid data was detected. Output record except invalid data.

Action

No action is necessary.

4.1.27 pmerls command

[NAME]

pmerls - The state of ERROR of the job is released.

[SYNOPSIS]

```
/usr/sbin/pmerls [-c clstname] jobid [jobid ...]  
/usr/sbin/pmerls --help
```

[DESCRIPTION]

pmerls command releases the job in the state of ERROR. This command can be executed on the system management node and the compute cluster management node.

The job that can specify it by this command is only a job in the state of ERROR. When jobs other than the state of ERROR are specified, the error is output.

When this command is executed, the job changes in the state of QUEUED, and is scheduled again.

When job ID is specified, the state of ERROR of all sub jobs of the job is released. At this time, sub jobs other than the state of ERROR disregard it.

The state of ERROR of the sub job that specifies that sub job ID is specified is released.

If the job can be specified when the command is executed two or more times specifying same job ID, everything is accepted.

However, the release processing in the state of ERROR of the job becomes only the first one time accepted with the job can specified.

This command returns after the state of ERROR of the job is released, and it shifts in the state of QUEUED.

[Multiple specification and range specification of job IDs]

You can specify a range of job IDs by using hyphen "-". Similarly, you can also specify a range of sub job IDs.

Example1: Specify a job ID range of 1 to 100

1-100

Example2: Specify a step number range of 1 to 10 for a step job (job ID = 100)

100_1-10

Example3: Specify a bulk number range of 1 to 10 for a bulk job (job ID = 101)

101[1-10]

If the specified range includes the following jobs, processing ignores the jobs and continues:

- Jobs that do not exist
- Jobs for which the user does not have privileges
- Jobs in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly (Processing is not aborted, and all specified jobs are processed).

[OPTIONS]

-c *clstname*

The job of the cluster *clstname* is processed.

When *clstname* is set to environment variable PXMYCLST, this option is omissible.

When this option is specified, this option is disregarded in execution by the compute cluster management node.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The state of ERROR of the job of job ID from 1 to 3 is released.

```
$ pmerls 1 2 3
[INFO] PJM 0900 pmerls Job 1 released.
[INFO] PJM 0900 pmerls Job 2 released.
[ERR.] PJM 0912 pmerls Job 3 does not exist.
$
```

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end (When two or more jobs are specified, this command returns abnormal termination because of the error of one job.)

[MESSAGES]

If an error occurs during the execution of the pmerls command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 0901 pmerls Unknown option or no option argument: *opt*.
Try `pmerls --help` for more information.**

Meaning

The specified *opt* is unrecognizable. Or, no argument is specified in the option.

Action

Specify the correct option. Or, specify the argument in the option.

**[ERR.] PJM 0903 pmerls Unknown option argument: *opt arg*.
Try `pmerls --help` for more information.**

Meaning

The specified *arg* is an unrecognizable argument of the option.

Action

Specify the correct argument in the option.

[ERR.] PJM 0904 pmerls Invalid option argument: *opt arg*.

Meaning

The specified argument *arg* in the option is incorrect.

Action

Specify the correct argument in the option.

**[ERR.] PJM 0910 pmerls No jobid(s) specified.
Try `pmerls --help` for more information.**

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 0911 pmerls Invalid jobid: *jobid*.**Meaning**

The format of the job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 0912 pmerls Job *jobid* does not exist.**Meaning**

The specified job does not exist. Alternatively, the user is not the owner of the specified job. Otherwise, the user does not have the privileges to cancel the error state of the specified job.

Action

Confirm that the job exists. Confirm that the user is the owner of the job. Confirm that the user has the privileges to be able to cancel the job error state.

[ERR.] PJM 0913 pmerls Job *jobid*status is not error.**Meaning**

The specified job (job ID: *jobid*) is not in the ERROR state.

Action

Check the state of the specified job.

[ERR.] PJM 0971 pmerls No permission: *jobid*.**Meaning**

The user is not permitted to operate the target job (job ID: *jobid*).

The error state cancellation operation is not permitted for the user or group executing the target job.

Action

Display the job ACL settings with the `pjacl` command. Then, confirm the permission (permit `pmerls`) for the operated objects of the `pmerls` command for the resource unit and resource group to which the target job is submitted.

[ERR.] PJM 0990 pmerls PJM daemon is not present.**Meaning**

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 0991 pmerls Internal error: *details*.**Meaning**

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 0993 pmerls The operating command cannot be executed except system management node and compute cluster management node.

Meaning

The command cannot be executed on a node other than a system management node or compute cluster management node.

Action

Execute the command on a system management node or compute cluster management node.

[ERR.] PJM 0995 pmerls No cluster specified on system management node.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the pjstat command, or set a cluster name in the environment variable PXMYCLST.

Information Message

[INFO] PJM 0900 pmerls Job *jobid* released.

Meaning

Cancellation of the error state of the specified job was accepted.

Action

No action is necessary.

4.1.28 pmexe command

[NAME]

pmexe - Runs a batch process to execute commands

[SYNOPSIS]

```
/usr/sbin/pmexe [ -c clstname ]
                  [ scope options ]
                  [ --nodetype nodetype[,...] ]
                  [ --excludetype nodetype[,...] ]
                  [ --model model[,...] ]
                  [ -P num:group[,num:group,...] ]
                  [ --active ]
                  [ --ssh-o ssh-option ]
                  [ --force-all ]
                  [ { --stdout | [ -d dirname ][ --filename name ] } ]
                  [ --direct ]
                  [ -t | -T ]
                  [ --compare ]
                  [ --chroot ]
                  { command | --ssh-keygen }

/usr/sbin/pmexe --help
```

```
scope options(exclusive):
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    --rscunit rscuname[,...]
    -n nodeid[,...] [ --giogrp | --siogrp ]
    -f filename
```

[DESCRIPTION]

The pmexe is a command which runs a batch process to execute *command* for a specified node or group.

This command outputs "Standard output" and "Standard error output" of *command* which succeeded in execution to an execution result file.

If *command* is not able to be performed, an error file will be created in order to describe the nodes or the cause of the failure.

This command is executed only on the nodes in the "on-running" or "on" power state. For nodes not included in the execution targets, the output is to the execution exclusion file.

The execution result file, error file and execution exclusion file of the command executed by the pmexe command are recorded in the respective individual files. By default, files are created in the home directory of the user who executed the pmexe command.

Name for execution result file: pmexe_result_ *command execution date and time.process ID*

Name for error file: pmexe_failed_ *command execution date and time.process ID*

Name for execution exclusion file: pmexe_excluded_ *command execution date and time.process ID*

This command executes processing for the FX server and compute nodes of PRIMERGY server through the compute cluster sub management node. In the node groups where the compute cluster sub management node has not started, the processing does not go through the compute cluster sub management node.

This command can be executed on the system management node.

[ARGUMENT]

command

Specifies an executed command.

[OPTIONS]

-c *clstname*

This option uses *clstname* to specify the cluster name of an executed object.

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to set only a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS).

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--excludetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to do not set a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS). Even if the nodes is specified for the execution

target by the `--nodetype`, the exclusion specification of this option is given to priority.

The node type that can be specified is similar to the `--nodetype`.

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--model model[,...]`

If you want to narrow down the execution targets to a specific model in the cluster, specify the model in *model*.

You can specify multiple models by delimiting them with a comma (e.g., `FT,PG`).

You can specify the following models and arbitrary models.

Abbreviation	Model
PG	PRIMERGY server
FT	FX server

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--nodegrp nodegid[,...]`

This option uses *nodegid* to specify a node group ID to set only a specific node group in a cluster as an executed object.

nodegid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x01,0x02,0x03`).

You can also specify a range with a hyphen (e.g., `0x01-0x03`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--bootgrp bootgid[,...]`

This option uses *bootgid* to specify a boot group ID to set only a specific boot group in a cluster as an executed object.

bootgid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x0101,0x0102,0x0103`).

You can also specify a range with a hyphen (e.g., `0x0101-0x0103`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`-n nodeid[,...]`

This option uses *nodeid* to specify a node ID to set only a specific node in a cluster as an executed object.

nodeid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x01010010,0x01010012,0x0101002A`).

You can also specify a range with a hyphen (e.g., `0x01010010-0x0101002A`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--giogrp`

If you want to narrow down the execution targets to only a specific GIO group in the cluster, specify this option as a sub option of the `-n` option.

The unit of execution for FX server is a GIO group, and nodes in the same GIO group as the node specified in the `-n` option are gathered in the execution targets.

For nodes that are not FX server, the `--giogrp` option specification is ignored.

`--siogrp`

If you want to narrow down the execution targets to only a specific SIO group in the cluster, specify this option as a sub option of the `-n` option.

The unit of execution for FX server is an SIO group, and nodes in the same SIO group as the node specified in the `-n` option are gathered in the execution targets.

For nodes that are not FX server, the `--siogrp` option specification is ignored.

`-f filename`

This option executes a process for a node ID written in the file that is specified with *filename*.

Each of the node IDs is written on one line of the node list by the hexadecimal number.

Lines with the comment symbol (#) written at the beginning are ignored.
 Strings other than node ID and node ID that does not exist are ignored.
 [Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit *rscuname*[,...]

This option uses *rscuname* to specify a resource unit name to set only a specific resource unit in a cluster as an executed object.
 You can specify multiple resource unit names by separating them with a comma (e.g., unitA,unitB).
 When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-P *num:group*[,*num:group*,...] | **--concurrency** *num:group*[,*num:group*,...]

The number of nodes processed at one time is specified in *num:group*. You can specify multiple *num:group* entries by delimiting them with a comma (.).
 If this option is not specified, all the nodes that are the processing target are processed in parallel.
 For *num*, specify the parallel processing quantity. For *group*, specify the parallel processing quantity you want to change in the unit of a group.
 You can specify any of the following values.

Unit	Parallel processing quantity
nodegrp	The minimum value is 1. The maximum value is the number of node groups that exist in the specified cluster.
bootgrp	The minimum value is 1. The maximum value is the maximum number of boot groups that exist in the specified cluster.
node	The minimum value is 1. The maximum value is the maximum number of nodes that exist in the specified cluster.

--active

If the redundant node is specified by you want to the execution targets to the active node, specify this option.
 When the execution target node which is the redundant configuration is specified by the unit of the node (-n or -f option) or it is not an active node, it does not make it the execution target.
 When this option is specified independently, all active nodes in a cluster are made applicable to execution.

--ssh-o *ssh-option*

Specify *ssh-option* in this option when you want an option in the same format as in the ssh command configuration file.
 Specify this option in the same format as the -o option of the ssh command.
 The following options can be specified.

```
ConnectTimeout
ConnectionAttempts
```

--force-all

Specify this option when you want to include nodes that are not in the "os-running" or "on" node state.
 Regardless of the state, however, multiuse nodes with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pxsystem.conf are always excluded from the execution targets. Their results are output to the execution exclusion file.

--stdout

This option executes standard output of an execution result.

-d *dirname*

This option uses *dirname* to store an execution result file in a specified directory.

If this option is not set, the execution result file will be output to the execution user's home directory.

--filename *name*

The name of the execution result file changes to the name specified in the argument *name*, and the "command execution date.process ID" part is omitted.

The following shows the output files when this option is specified.

Execution result file: name_result

Error file: name_failed

Execution exclusion file: name_excluded

--direct

Processing is executed without going through the compute cluster sub management node.

-t

Force pseudo-tty allocation to the execution target nodes execute *command*.

If this option is specified, command performed on the execution target nodes is also united and ended at the time of the end of pmexe command.

-T

Disable pseudo-tty allocation to the execution target nodes execute command.

If this option is specified, command performed on the execution target nodes continues and performs after the end of pmexe command.

--compare

The execution result file and the error file of command are molded in the same result style.

When the --stdout option is specified, the molded result is output to a standard output.

--chroot

When executed for the FX server, the command is executed with the root directory changed by the chroot command on the boot I/O node. Specify this option for package operations and other cases where the command must be executed after the root directory has been changed from the boot I/O node.

--ssh-keygen

This option sets SSH password-less login for the batch operation function.

When specifying this option, you cannot specify command to be executed on a node.

If this option is specified, a password must be entered to log in to each node. This is different from the normal execution of the pmexe command.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can specify only one cluster name in the environment variable PXMYCLST.

An example is shown below.

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

Exit status	Meaning
0	Normal end pmexe succeeds in execution of <i>command</i> on all target nodes. (Normal end even on execution excluded nodes)
1	Abnormal end
2	Abnormal end on all target nodes Unable to execute <i>command</i> on all target nodes. Unable to connect via ssh, or unable to execute <i>command</i> itself on target node
3	Abnormal end on some nodes and normal end on others (including execution-excluded nodes) Mixture of <i>command</i> execution successes and failures

[NOTE]

- The command execution result is once buffered and then output.
- Commands which require terminal operation such as (vim or passwd) cannot be executed.
- End status of the pmexe command is not related to End status of *command*.
- You cannot specify the --nodegrp, --bootgrp, -n, -f, and --rscunit options at the same time.
- The nodes that are execution targets are narrowed down in processing according to the combination of the options. If no node is an execution target, the processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option is invalid, processing ends abnormally with end status 1.
If the model specified in the --model option is invalid, processing ends abnormally with end status 1.
- If the ID specified in the -n, --nodegrp, or --bootgrp option is invalid, processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option does not exist in the cluster, processing ignores the node type. If the model specified in the --model option does not exist in the cluster, processing ignores that model. If the node ID specified in the -n option does not exist in the cluster, processing ignores that node.
- You cannot specify the -t and -T options at the same time.
- You can specify only a file name in the --filename option. To specify a directory as the output destination, specify the directory in the -d option.
- You cannot specify the --filename and --stdout options at the same time.
- You cannot specify the -d and --stdout options at the same time.
- SSH password-less login is not set on a system built without the install function. For this reason, the batch operation function cannot be used on that system. To use the batch operation function, you need to set SSH password-less login with the --ssh-keygen option of the pmexe command.
- To use any special character (e.g., > or &) in a command, such as in an argument, append "\" before the specified special character when specifying the command.
- In the execution target nodes, any multiuse node with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pxsystem.conf is automatically excluded from the execution targets. So the command is not executed on the node. The execution results on the node are output to the execution exclusion file.
- If the --ssh-keygen option is specified to generate an ssh key, this execute command cannot be specified.
- If many messages are output per node for commands executed by the pmexe command, the pmexe command execution does not run in parallel and takes a long time.
When you execute a command that might output more than 900 KB of messages, suppress message output by redirecting stdout and stderr to a file on each node.

[MESSAGES]

If an error occurs during the execution of the pmexe, pmscatter, and pmgather command, one of the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

The details of command errors are common to the pmexe, pmscatter, and pmgather commands.

The following is an example of an error message output by the pmexe, pmscatter, or pmgather command.

```
[ERR.] PSM 5001 COMMAND Failed to create file:filename.(details)
```

Any of pmexe, pmscatter, and pmgather can be the command name entered in *COMMAND*.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

Error Message

[ERR.] PSM 5001 COMMAND Failed to create file:filename.(details)

Meaning

Creation of the file filename failed.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 5002 COMMAND Failed to delete file:filename.(details)

Meaning

Deletion of the file filename failed.

details: Cause of the error

Action

Delete the file filename manually.

[ERR.] PSM 5003 COMMAND Invalid command.(details)

Meaning

The specified command line is invalid.

details: Specified command line

Action

Review the command line, and execute the command again.

[ERR.] PSM 5004 COMMAND Cannot be executed because it cannot check active state.

Meaning

Processing for the active node (system management node, compute cluster management node, compute cluster sub management node, or storage cluster management node) cannot be executed with the --active option specified.

Action

The reason is that the active node cannot be identified because the system monitoring function is not working on the system management node.

Restart the system management node (both the active and standby system management nodes in the case of a redundant configuration) that executed the command. Then, execute the command again for the node that is the operated object.

[ERR.] PSM 5005 COMMAND There is no target node.**Meaning**

There is no execution target node as a result of combining options.

Action

Review the options, and execute the command again.

[ERR.] PSM 5006 COMMAND Connection failed : *Node identification information (details)***Meaning**

Communication with the target node failed.

Node identification information: Information identifying the destination when communication failed (IP address or node ID)
details: Cause of the error

Action

Confirm that communication from the command execution node to the output node in *Node identification information* is normal. If the communication is normal, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 5007 COMMAND Failed to get power status of the target nodes.**Meaning**

The node power status cannot be obtained.

Action

The reason for the inability to obtain the node power status is because the system monitoring function is not working on the system management node.
Restart the system management node (both the active and standby nodes in the case of a redundant configuration) that executed the command. Then, execute the command again for the node that is the operation target.

[ERR.] PSM 5008 COMMAND Cannot executed in this system configuration.(details)**Meaning**

The batch operation command cannot be executed due to the system settings.

details: Cause of the error

Action

Review the upper limit on the number of file descriptors that can be used by the command, and execute the command again.

Warning Message

[WARN] PSM 5051 COMMAND The execution failed file was output.(filename)**Meaning**

The execution error file filename was output.

Action

Processing may not have been executed correctly. Check the execution error file contents.

Information Message

[INFO] PSM 5081 COMMAND All nodes are targets. Cluster name is *clstname*.**Meaning**

The targets are all the nodes of the target cluster *clstname*.

Action

No action is necessary.

[INFO] PSM 5082 COMMAND The execution result file was output.(*filename*)

Meaning

The execution result file *filename* was output.

Action

No action is necessary.

[INFO] PSM 5083 COMMAND The execution exclude file was output.(*filename*)

Meaning

The execution exclusion file *filename* was output.

Action

No action is necessary.

4.1.29 pmgathering command

[NAME]

pmgathering - Runs a batch process to collect files

[SYNOPSIS]

```
/usr/sbin/pmgathering [ -c clstname ]
                        [ scope options ]
                        [ --nodetype nodetype[,...] ]
                        [ --excludetype nodetype[,...] ]
                        [ --model model[,...] ]
                        [ -P num:group[,num:group,...] ]
                        [ -p ]
                        [ -r ]
                        [ --active ]
                        [ --ssh-o ssh-option ]
                        [ --force-all ]
                        [ -d dirname ] [ --filename name ]
                        srcfile[,...] dstdir

/usr/sbin/pmgathering --help

scope options (exclusive):
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    --rscunit rscuname[,...]
    -n nodeid[,...] [ --giogrp | --siogrp ]
    -f filename
```

[DESCRIPTION]

pmgathering is a command which runs a batch process to collect the srcfile for a specified node or group.

If a command fails to collect files for a node, an error file will be created in the execution user's home directory in order to describe the node or the cause of the failure.

This command is executed on the nodes in the "os-running" or "on" power state. For nodes not included in the execution targets, the output is to the execution exclusion file.

For the targets that are the FX server, the command refers to the power state of the compute node serving as the boot I/O node in the same BoB, not of the target nodes.

Name for error file: *pmgather_failed_command execution date and time.process ID*

Name for execution exclusion file: *pmgather_excluded_command execution date and time.process ID*

This command can be executed on the system management node.

ARGUMENTS

srcfile[,...]

This argument specifies the name of a collected source file or directory with full path.

Plurals can be specified by delimiting it by the comma.

dstdir

This argument specifies the name of a collected destination directory.

The collected files are placed in the following directory:

dstdir/cluster name/node group ID/boot group IS/node ID/.

When the file *dstfile* exists, it is overwritten by *srcfile*.

OPTIONS

-c *clstname*

This option uses *clstname* to specify the cluster name of a delivered object.

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodetype *nodetype*[,...]

This option uses *clstname* to specify the cluster name of an executed object.

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to set only a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS).

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

When specifying this option, be sure to specify a cluster name in *clstname* with the **-c** option, or specify a specific cluster name in the environment variable PXMYCLST.

--excludetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to do not set a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS). Even if the nodes is specified for the execution target by the **--nodetype**, the exclusion specification of this option is given to priority.

The node type that can be specified is similar to the `--nodetype`.

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--model model[,...]`

If you want to narrow down the execution targets to a specific model in the cluster, specify the model in *model*.

You can specify multiple models by delimiting them with a comma (e.g., `FT,PG`).

You can specify the following models and arbitrary models.

Abbreviation	Model
PG	PRIMERGY server
FT	FX server

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--nodegrp nodegid[,...]`

This option uses *nodegid* to specify a node group ID to set only a specific node group in a cluster as an executed object.

nodegid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x01,0x02,0x03`).

You can also specify a range with a hyphen (e.g., `0x01-0x03`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--bootgrp bootgid[,...]`

This option uses *bootgid* to specify a boot group ID to set only a specific boot group in a cluster as an executed object.

bootgid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x0101,0x0102,0x0103`).

You can also specify a range with a hyphen (e.g., `0x0101-0x0103`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`-n nodeid[,...]`

This option uses *nodeid* to specify a node ID to set only a specific node in a cluster as an executed object.

nodeid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., `0x01010010,0x01010012,0x0101002A`).

You can also specify a range with a hyphen (e.g., `0x01010010-0x0101002A`).

When specifying this option, be sure to specify a cluster name in *clstname* with the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

`--giogrp`

If you want to narrow down the execution targets to only a specific GIO group in the cluster, specify this option as a sub option of the `-n` option.

The unit of execution for FX server is a GIO group, and nodes in the same GIO group as the node specified in the `-n` option are gathered in the execution targets.

For nodes that are not FX server, the `--giogrp` option specification is ignored.

`--siogrp`

If you want to narrow down the execution targets to only a specific SIO group in the cluster, specify this option as a sub option of the `-n` option.

The unit of execution for FX server is an SIO group, and nodes in the same SIO group as the node specified in the `-n` option are gathered in the execution targets.

For nodes that are not FX server, the `--siogrp` option specification is ignored.

`-f filename`

This option executes a process for a node ID written in the file that is specified with *filename*.

Each of the node IDs is written on one line of the node list by the hexadecimal number.

Lines with the comment symbol (`#`) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

0x01010010
0x01010011
0x0101002A

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit *rscuname*[,...]

This option uses *rscuname* to specify a resource unit name to set only a specific resource unit in a cluster as an executed object.

You can specify multiple resource unit name by separating them with a comma (e.g., unitA,unitB).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-P *num:group*[,*num:group*,...] | --concurrency *num:group*[,*num:group*,...]

The number of nodes processed at one time is specified in *num:group*. You can specify multiple *num:group* entries by delimiting them with a comma (.).

If this option is not specified, all the nodes that are the processing target are processed in parallel.

For *num*, specify the parallel processing quantity. For *group*, specify the parallel processing quantity you want to change in the unit of a group.

You can specify any of the following values.

Unit	Parallel processing quantity
node	The minimum value is 1. The maximum value is the maximum number of nodes that exist in the specified cluster.

-p

This option maintains the final revise time, the final access time, and the permission of a copied source file.

-r

This option recursively copies an entire directory.

--active

If the redundant node is specified by you want to the execution targets to the active node, specify this option.

When the execution target node which is the redundant configuration is specified by the unit of the node (-n or -f option) or it is not an active node, it does not make it the execution target.

When this option is specified independently, all active nodes and no redundant nodes in a cluster are made applicable to execution.

--ssh-o *ssh-option*

Specify *ssh-option* in this option when you want an option in the same format as in the ssh command configuration file.

Specify this option in the same format as the -o option of the ssh command.

The following options can be specified.

ConnectTimeout
ConnectionAttempts

--force-all

Specify this option when you want to include nodes that are not in the "os-running" or "on" node state.

Regardless of the state, however, multiuse nodes with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pxsystem.conf are always excluded from the execution targets. Their results are output to the execution exclusion file.

-d *dirname*

This option uses *dirname* to store an execution result file in a specified directory.

If this option is not set, the execution result file will be output to the execution user's home directory.

--filename *name*

The name of the execution result file changes to the name specified in the argument *name*, and the "command execution date.process ID" part is omitted.

The following shows the output files when this option is specified.

Error file: name_failed

Execution exclusion file: name_excluded

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option.

Specification with the -c option has priority over that with the environment variable PXMYCLST.

You can specify only one cluster name in the environment variable PXMYCLST.

An example is shown below.

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The exit status is returned as follows:

Exit status	Meaning
0	Normal end Successful <i>srcfile</i> collection on all target nodes (normal end even on execution-excluded nodes)
1	Abnormal end
2	Abnormal end on all target nodes Unable to collect <i>srcfile</i> from all target nodes. Unable to connect via ssh, or collection source/destination files not found
3	Abnormal end on some nodes and normal end on others (including execution-excluded nodes) Mixture of collection successes and failures

[NOTE]

- You cannot specify the --nodegrp, --bootgrp, -n, -f, and --rscunit options at the same time.
- The nodes that are collection targets are narrowed down in processing according to the combination of the options. If no node is a collection target, the processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option is invalid, processing ends abnormally with end status 1.
If the model specified in --model option is invalid, processing ends abnormally with end status 1.
- If the ID specified in the -n, --nodegrp, or --bootgrp option is invalid, processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option does not exist in the cluster, processing ignores the node type. If the model specified in the --model option does not exist in the cluster, processing ignores that model. If the node ID specified in the -n option does not exist in the cluster, processing ignores that node.
- You can specify only a file name in the --filename option. To specify a directory as the output destination, specify the directory in the -d option.
- SSH password-less login is not set on a system built without the install function. For this reason, the batch operation function cannot be used on that system. To use the batch operation function, you need to set SSH password-less login with the --ssh-keygen option of the pmexe command.

- In the execution target nodes, any multiuse node with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pssystem.conf is automatically excluded from the execution targets. So the command is not executed on the node. The execution results on the node are output to the execution exclusion file.

[MESSAGES]

The pmexe, pmscatter, and pmgather commands share common messages. See "[MESSAGES]" in "4.1.28 pmexe command."

Also, not only the messages of the pmexe, pmscatter, and pmgather commands but common command messages are output. For details on the common command messages, see "4.2.1 Common command messages of the system management function."

4.1.30 pmjacladm command

[NAME]

pmjacladm - Configures or updates the items of the job ACL function and displays its configuration.

[SYNOPSIS]

```
/usr/sbin/pmjacladm --set [-c clstname] [--clear] -f filename
/usr/sbin/pmjacladm --set [-c clstname] control
/usr/sbin/pmjacladm --del [-c clstname] control
/usr/sbin/pmjacladm --show [--apply] [-c clstname] [--data] [--delimiter delimiter]
                        [--noname] [-o filename] [control]
/usr/sbin/pmjacladm --show --items [-c clstname]
/usr/sbin/pmjacladm --help
```

[DESCRIPTION]

pmjacladm is a command for job operation administrators to configure or update items of the job ACL function, and display its configuration.

Job ACL function has a set item in the cluster that exists in the range of operation the job manager, and the cluster administrator sets and can update all the definition items in the cluster.

The job operation administrator can set, update, and display each definition item in the cluster within the range of the object resource unit. The range of the resource unit that the job operation administrator makes a management object is a range defined by the operation management access control function of system management.

Setting/update replaces only the specified data with the content of a present job ACL function.

You can also output the content of the defined items configured in the job ACL function.

This command can be executed on the system management node and the compute cluster management node.

Job ACL function has the following as a unit of the definition.

- User unit definition
The unit of the user is defined in the job.
When the same definition as the user in the group is applied, and the same definition as all users who submit the job is applied.
- Group unit definition
The unit of the group is defined in the job.
The definition intended for the total of all users in the group and the group priority are defined.
- Whole unit definition
Definition that controls the entire job.

Job ACL function can consider the cluster, the resource unit, and the resource group to be a hierarchy, and define the above-mentioned unit definition of each hierarchy.

Only the definition "Definition item only for the cluster" only of the cluster and the cluster and the resource unit have the definition "Definition item only for the resource unit" according to the definition item. These definitions are described by the text form, and set.

[OPTIONS]

-c *clstname*

Job ACL function of cluster *clstname* is set, and displayed.

When *clstname* is set to environment variable PXMYCLST, this option is omissible.

When this option is specified, this option is disregarded in execution by the compute cluster management node.

--set

This option sets defined items.

--clear

This option initializes all the defined items that are not contained in input data. After all data is initialized, input data is overwritten.

This option is effective only for the input data of the -f option.

--del

A part of the definition item specified by the control description is deleted.

When the item in the definition section is deleted, an item concerned becomes undefined.

--show

The definition item of job ACL function is displayed.

-f *filename*

The file specified by *filename* is read as an input of the definition.

The file format conforms to the format output by the --show option.

The update process is executed in the described order.

--apply

This option shows defined values that are applied to the user or group of a displayed object of the job ACL function after completion of this command. This option enables you to check whether any application rules such as default value, etc., are correctly applied.

When two or more same definition items are set according to an undefined item and the hierarchy, the application result according to the application rule is displayed.

These items display not only values but also the application source.

It is necessary to edit it to actually use the display result as a definition.

The following specification is indispensable to control to fix the application object as the control description.

- Display intended user and the execution group (display of unit of user).

- Group for display (display of each group).

In the display of the unit of the user, the execution group is omissible. At this time, display intended user's default group is made an execution group (when "--noname" is specified, the error is returned).

It is necessary to specify the resource unit name and the resource group name for the display of all definition item. When only the resource unit name is specified, the definition item only for the resource unit is displayed.

When each specification does not exist, either only the definition item only for the cluster is displayed.

The belonging relation between specified user and group is not checked.

-o *filename*

The display of job ACL function is output to the file *filename* instead of a standard output.

When *filename* is an existing file, it is overwritten.

--items

An item list that can be defined in job ACL function and an undefined default application value are displayed.

The control specification is disregarded.

--data

The field of a set content is displayed by the form delimited by delimiter (" , ").

--delimiter "*delimiter*"

The delimiter of the display form when the --data option is specified is changed to the character specified by *delimiter*.

--noname

The name of the user name and the group name is not converted.

It processes it without confirming these though name service and /etc/passwd, etc. are confirmed from UID and GID acquired from job ACL function and the name is displayed in the display by optional --show usually.

The following items are output with id (numerical value).

- user name and group name
- Value defined in definition item as user name and group name

It makes an error of the user name and the group name specification because the name resolution is not done if this option is specified. Please specify UID,GID.

control

The control description to do the setting and the display of information of a specific definition item in job ACL function from the command line is specified. The control description is recommended to enclose the entire control with single quotation (""") to avoid the influence on the shell.

When the specification of control is omitted, nothing is displayed.

Up to 1024 characters can be set for a control description. If more than 1024 characters are set, the excessive characters will be discarded. control is written at the end of each option.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[DEFINITION]

The definition is described by the text form.

The section of each unit of the definition is separately described, and the following character strings are specified for a section name.

1. User unit definition USER:
2. Group unit definition GROUP:
3. Whole unit definition ALL:

The coverage specification for the hierarchy adds coverage to the section name.

The user unit definition is described as follows.

USER:CL[,RU=rscuname[,RG=rscgname]]

CL means the cluster, RU means the resource unit, and RG means the resource group.

The description form of each unit definition section is shown as follows.

1. User unit definition section (USER definition)

```
USER: CL[,RU=rscuname[,RG=rscgname]] {
    User definition object {
        Definition item name value
        :
    }
    User definition object {
        Definition item name value
        :
    }
}
```

2. Group unit definition section (GROUP definition)

```
GROUP: CL[,RU=rscuname[,RG=rscgname]] {
    Group definition object {
        Definition item name value
        :
    }
    Group definition object {
```



```

        Definition item name value
        :
    }
}

```

3. Whole unit definition section (ALL definition)

```

ALL: CL[,RU=rscuname[,RG=rscgname]] {
    Definition item name value
    :
}

```

Description specification

- The tab can be specified for the delimitation of each item besides space.
- The character of the line until the end is treated as a comment since "#".
- The null line is disregarded(The line only of blank/tab and the line only of the comment are disregarded).

User definition object

Either of the following definition objects are specified for the USER definition for a subsection name. However, there is an item not to be able to specify concerning the group by the definition item.

`user=<def>`

The user's default definition is specified. All the definitions of this section are applied to the user who does not have a definition in the subsection of "user=".

`user=<def>:gname`

The object group specifies the default definition of the user of *gname*.

`user=uname`

The definition of the user *uname* is specified.

`user=uname:gname`

When the user *uname* is only the group *gname*, an effective definition is specified.

Group definition object

Either of the following definition objects are specified for the GROUP definition for a subsection name.

`group=<def>`

The group's default definition is specified.

`group=gname`

The definition of the group *gname* is specified.

Definition item name

The definition item name is composed of the definition type and the item name. The definition type is shown as follows.

limit

"limit" is a definition that limits the resources made available like submitting the number of/number etc. of execution of jobs.

In the USER definition, the total limit of the number of use for each users is defined.

In the GROUP definition, the total limit of the number of use of all users in each group is defined.

In the ALL definition, the total limit of the number of use of all users is defined.

define

"define" is a general definition used so that the job manager may control the job.

"define" is a definition to treat a general set value while "limit" specializes in the limiting value.

The USER definition has a set value of each user.

The GROUP definition has a set value of the entire user of each object group (for instance, total value). The group priority is defined by "define".

The ALL definition has the total setting value of all users.

joblimit

Specifies a upper bound, a lower bound, and a default value for a job resource requirements. Those bounds limit the amount of resources that can be specified in the pjsub command. The default value will be used if resource value does not specified in the pjsub command.

Only the USER definition can be specified.

There are items can not be set the lower limit value or default value.

The values of these items are output as "-" by the option --show --items of pmjacladm command.

For these items, specify "-" as the lower limit or default value.

execute

"execute" is a definition that specifies whether the command is executable.

Only the USER definition can be specified.

"execute" might set whether the option can be specified like the pjsub command.

The submitting authority of the job is a range permitted by execute pjsub.

permit

The permission of the operational object of the job is set.

Only the USER definition can be specified.

The definition is in each command, and the operational object can be defined.

select

Specifies an allowed set of custom resource value and the default value for it. The user can only specify the allowed value in the pjsub command. If custom resource value does not specified, the default value will be used.

Definition item select can be written only in USER definition object.

The definition type and the definition object that can be specified at each definition of USER, GROUP and ALL is shown as follows

definition type	USER	GROUP	ALL
limit(limiting value)	JOB(*1)	JOB	JOB
define(set value)	JOB(*2)	JOB	JOB
joblimit(Resource limitation lower bound value/upper bound value/default value)	JOB	x	x
execute (The command execution is possible/impossible)	COM	x	x
permit (Permission for job operation)	COM	x	x
select (Type of custom resource)	JOB	x	x

- JOB

This definition targets the execution of the job.

As for user=*uname:gname* specifiable with the user definition object, the definition object of *uname* is "Job execution user", and the definition object of *gname* is "Job execution group that specified it with pjsub -g".

(*1) *gname* cannot be specified.

(*2) *gname* cannot be specified in a part of definition item.

As for group=*gname* specifiable with the group definition object, the definition object of *gname* is "Job execution group that specified it with pjsub -g".

- COM

This definition targets the execution of the command.

As for user=*uname:gname* specifiable with the user definition object, the definition object of *uname* is "Command execution user", and the definition object of *gname* is "Command execution user's current group".

Refer to "Definition item list" for the definition item name that contains the item name.

The definition item name that can be specified can be displayed by "-- show --items" option specification.

Value

The kind of the value that can be specified is shown as follows. The value that can be specified is different according to the definition item name.

Remarks

The definition item value cannot be longer than 511 characters.

Numerical value

Specification from "-9223372036854775808" to "9223372036854775807" and (integral value) is possible. However, the set value is processed according to each function of the job manager.

If a number outside the valid range is specified, it will be rounded to a number within the valid range.

The following descriptions are enabled as a unit of the numerical value.

The time designation is described in the form of HH:MM:SS. The numerical value is treated at the second.

There is no limitation in the digit number that can be specified for each of "Hour", "Minute", and "Second".

[example] 1234:58:21 (1234 hours 58 minutes 21 seconds)

12345 (12345 seconds)

The following can be described in the specification of the amount of the memory and the size of the file as a prefix of the byte.

Ki, Mi, Gi, Ti, Pi

K, M, G, T, P

[example] 1Ki (1024 byte), 1M(1000000 byte)

Character

The character is described. The front of the blank or the tab is assumed to be a value.

To contain the blank and the tab, it encloses it with a double quotation.

The specification of the multibyte character is not guaranteed.

To describe the character that starts by double quotation and "<", it escapes in the backslash.

<undef>

It means "Undefined".

It is effective for the character the type of the value. When not the setting of the value but specification and set what are assumed to be undefined, "<undef>" is specified.

All the values that the heading character starts by "<" are assumed to be "Undefined".

If the definition type is joblimit, specifying "<undef>" for either a lower limit value, an upper limit value, or a default value will result in those three values being "undefined".

A lower limit value, an upper limit value, and a default value must be specified even if one of them is specified as "<undef>".

unlimited

It means "Unlimited".

If unlimited is specified for the default value of the following definition items, it is converted to the maximum value.

joblimit node

joblimit vnode

joblimit vnode-core

joblimit proc-openfd

joblimit interact-node

joblimit interact-vnode

joblimit interact-vnode-core

joblimit interact-proc-openfd

Executable/impracticability (for execute)

The definition type can be described and either the following be described in the value of execute.

disable (impracticability)

enable (Executable)

Permission list (for permit)

The definition type can set allow (permission) and deny (refusal) to the value of permit, and it describes it as follows.

Allow object,... or a object,...

deny object,... or d object,...

The belonging relation is not confirmed though the user and the group can be specified for "object".

Two or more allow/deny can be specified. In this case, it describes by two or more lines, the same definition item name is described

since the second line, and "+" (plus) is added at the head of the value.

When two or more objects are specified in same allow/deny, the object is narrowed in specified order.

For instance, allow u(u1) and u(u2) are marks that become narrowing named u1 and u2 and are improper though permit u1 and u2 by allow u(u1, u2).

When two or more allow/deny is specified, the definition is applied in specified order.

The permission list is assumed to be the one that deny all is always attached to the head.

The object is shown as follows. Two or more objects can be specified by delimiting it by the comma.

own

Own job is targeted.

When the definition item is only permit pjsub, own belonging group is targeted.

all

All user's jobs are targeted.

g(g1, g2, ...) group(g1, g2, ...) gname(g1, g2, ...)

The execution group targets the job of *g1* and *g2*...

The name only of the figure is treated as GID. "gname" can be specified for a means to specify the group name only of the figure.

It is not confirmed whether concerned ID is registered in the system at the GID specification.

u(u1, u2, ...) user(u1, u2, ...) uname(u1, u2, ...)

The user targets the job of *u1* and *u2*...

The name only of the figure is treated as UID. "uname" can be specified for a means to specify the user name only of the figure.

It is not confirmed whether concerned ID is registered in the system at the UID specification.

<undef>

It means "Undefined".

[Description form of control]

Control of --set option

The definition change in the command line is possible according to the control specification.

The number of characters that can be specified is up to 1024, so a large-scale definition is defined by the file description.

The control description that can be specified by setting the definition item of "--set option" is shown below. Two or more definition items and the values can be described in the form by ";"(semicolon) almost thoroughly to the definition description form.

USER: CL [,RU=rscuname [,RG=rscgname]] { User-definition-object { Definition-item-name value; } }

GROUP: CL [,RU=rscuname [,RG=rscgname]] { Group-definition-object { Definition-item-name value; } }

ALL: CL [,RU=rscuname [,RG=rscgname]] { Definition-item-name; }

Control of --show option

The control description that can be specified by setting the definition item of "--show option" is shown as follows.

It is a form that omits "Value" from the description that can be specified for "--set option"(The specification of "value" is disregarded.).

USER: CL [,RU=rscuname [,RG=rscgname]] { User-definition-object { Definition-item-name; } }

GROUP: CL [,RU=rscuname [,RG=rscgname]] { Group-definition-object { Definition-item-name; } }

ALL: CL [,RU=rscuname [,RG=rscgname]] { Definition-item-name; }

All the USER definitions defined to specify, omission or "user=*:*" "User definition object" within the range of a specified section are displayed. It is correct and because the group should been fixing about specifying --apply option as the user, is necessary to specify "user=uname:gname" for "User definition object". When the group designation is omitted, it is assumed the one that user's default group was specified.

It is correct and because the group should been fixing about specifying --apply option as the group, is necessary to specify "group=gname" for "Group definition object".

The following descriptions are possible in the control description.

Description form and meaning

Description form	meaning
'*'	All data

Description form	meaning
'USER: *'	All data of definition section USER
'USER: CL'	Only CL data of definition section USER
'USER: CL, RU=*'	USER definition section all RU and data of all RG
'USER: CL, RU=rscunit1'	Definition section USER "RU" is data of rscunit1 and all RG data in rscunit1.

Within the range of an unauthorized resource unit, the display is off the subject.

Control of --del option

The form of the control description that can be specified by setting the definition item of "--del option" is equal to "--show option".

[About the application rule]

The value is decided in job ACL according to the application rule when there are two or more definitions that apply the same definition items of the user's default definition and the default definition of the user in the group, etc.

Application rule in user unit definition section

"user=" specification is applied and the definition with high priority is applied as follows in case of plurals.

Priority-high

```
user=username:gname      # Definition when user is specific group specification.
user=username            # User definition
user=<def>:gname          # Default of each user's group
user=<def>                # User's default
```

Priority-low

The item not to be able to specify concerning the group flies the specification.

Application rule in group unit definition section

"group=" specification is applied and the definition with high priority is applied as follows in case of plurals.

Priority-high

```
group=gname              # Group definition
group=<def>               # Group's default
```

Priority-low

Application rule in whole unit definition section

There is no priority object, and the definition item is always applied.

Between hierarchies

When CL (cluster), RU (resource unit), and RG (resource group) are defined, the application rule is different depending on the definition type.

The application origin of each definition item is decided based on the following rules.

limit

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority).

define

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority).

joblimit

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority).

execute

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority). If the definition item execute is not defined for the root user, the default value is 'enable', regardless of what the pmjacladm or pjacl command displays.

permit

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority). If the definition item permit is not defined for the root user, the default value is 'allow all', regardless of what the pmjacladm or pjacl command displays.

select

If the same definition specified in many sections, the highest prioritized definition will be applied. The priority of the sections is RG>RU>CL>default (the definition in RG section has the highest priority).

Special definition item and common definition value of each hierarchy

The value that can confront the definition value of the lowest layer the definition item for which the application rule can be used between hierarchies and specify it in the rank layer is called a common definition value.

For instance, the definition value of execute pjsub within the range of the resource unit becomes a common definition in each of not the value with the meaning only with the resource unit but subordinate's each resource groups though each resource group can define execute pjsub (right or wrong of the pjsub command execution).

This is called a common definition value in the resource unit. Similarly, it is a common definition to the subordinate's resource unit as for the definition of the range of the cluster, and these are called a common definition value in the cluster.

On the other hand, the one with the meaning only in the hierarchy is called a special definition item. It never has the common definition value in the hierarchy of the subordinate position though the special definition item has the common definition value in a high-ranking hierarchy.

[Definition item list]

The definition item name list of each unit of the definition that can be specified is shown.

The explanatory notes of the hierarchy that can be specified is the following.

CL = range of cluster, RU = range of resource unit, RG = range of resource group

o = Possible (special definition item), c = Possible (common definition value), x = Impossibility

User unit definition section

item	Content	Hierarchy		
		CL	RU	RG
limit ru-accept	Number of job simultaneous acceptance limitations in resource unit (*1)	c	o	x
limit ru-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource unit (*1) Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	o	x
limit ru-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource unit (*1)	c	o	x
limit ru-accept-stepsubjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource unit (*1)	c	o	x
limit ru-run-job	Number of job simultaneous execution limitations in resource unit (*1)	c	o	x
limit ru-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource unit (*1)(*2)	c	o	x
limit ru-use-node	Simultaneous number of use node limitations in resource unit (*1)(*5)	c	o	x
limit ru-use-core	Simultaneous number of CPU use core limitations in resource unit (*1)(*6)	c	o	x

item	Content	Hierarchy		
		CL	RU	RG
limit ru-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit	c	o	x
limit ru-interact-accept	Number of job simultaneous acceptance limitations in resource unit(Interactive job) (*1)	c	o	x
limit ru-interact-run-job	Number of job simultaneous execution limitations in resource unit(Interactive job) (*1)	c	o	x
limit ru-interact-use-node	Simultaneous number of use node limitations in resource unit(Interactive job) (*1)(*5)	c	o	x
limit ru-interact-use-core	Simultaneous number of CPU use core limitations in resource unit(Interactive job) (*1)(*6)	c	o	x
limit ru-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit(Interactive job)	c	o	x
limit rg-accept	Number of job simultaneous acceptance limitations in resource group (*1)	c	c	o
limit rg-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource group (*1) Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	c	o
limit rg-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource group (*1)	c	c	o
limit rg-accept-stepsubjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource group (*1)	c	c	o
limit rg-run-job	Number of job simultaneous execution limitations in resource group (*1)	c	c	o
limit rg-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource group (*1)(*2)	c	c	o
limit rg-use-node	Simultaneous number of use node limitations in resource group (*1)(*5)	c	c	o
limit rg-use-core	Simultaneous number of CPU use core limitations in resource group (*1)(*6)	c	c	o
limit rg-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource group	c	c	o
limit rg-interact-accept	Number of job simultaneous acceptance limitations in resource group(Interactive job) (*1)	c	c	o
limit rg-interact-run-job	Number of job simultaneous execution limitations in resource group(Interactive job) (*1)	c	c	o
limit rg-interact-use-node	Simultaneous number of use node limitations in resource group(Interactive job) (*1)(*5)	c	c	o
limit rg-interact-use-core	Simultaneous number of CPU use core limitations in resource group(Interactive job) (*1)(*6)	c	c	o
limit rg-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource group(Interactive job)	c	c	o
define rscunit	Default submitting resource unit name	o	x	x
define rscgroup	Default submitting resource group name	c	o	x
define interact-rscunit	Default submitting resource unit name (Interactive job)	o	x	x

item	Content	Hierarchy		
		CL	RU	RG
define interact-rscgroup	Default submitting resource group name (Interactive job)	c	o	x
define pri	User priority in resource unit (*1) When "user_prio" of the job selection policy is set, it is effective.	c	c	o
define ingroup-pri	User priority in the same group in resource unit When "usr_in_grp_prio" of the job selection policy is set, it is effective.	c	c	o
define fshare-init	User fair share initial value in resource unit (*1) When "user_fairshare" of the job selection policy is set, it is effective. (*4)	c	o	x
define fshare-recovery	User fair share recovery rate in resource unit (*1) When "user_fairshare" of the job selection policy is set, it is effective. (*4)	c	o	x
define ingroup-fshare-init	User fair share initial value in the same group in resource unit When "usr_in_grp_fairshare" of the job selection policy is set, it is effective. (*4)	c	o	x
define ingroup-fshare-recovery	User fair share recovery rate in the same group in resource unit When "usr_in_grp_fairshare" of the job selection policy is set, it is effective. (*4)	c	o	x
define allocation-mode	Default node allocation method [FX] (*11)	c	c	o
define numa-policy	NUMA allocation policy[FX][PG] (*7)	c	c	o
define load-policy	Method of selecting nodes [PG] (*8)	c	c	o
define vn-policy	Virtual node placement policy [PG] (*9)	c	c	o
define exec-policy	Execution mode policy [PG] (*10)	c	c	o
define node-priority	Priority control of allocated nodes [PG]	c	c	o
define alloc-granularity	Default node allocation granularity (*12)	c	c	o
define assign-logical-cpu	Range of logical CPUs that job processes will execute[PG] (*14)	c	c	o
define pjstat-display-mode	pjstat command job information display mode (*15)	c	c	o
define elapsed-time-mode	Specification method of default elapsed time limit value (*16)	c	c	o
define pjstat-sdt-format	Display format for the planned execution start time of a job by the pjstat command (*17)	c	c	o
define pjstat-sdt-mark	Mark of the planned execution start time of a job displayed by the pjstat command (*18)	c	c	o
define strict-mode	Default of pjsb option strict, strict-io or not specified (*19)	c	c	o
define net-route	Change of the communication path when a Tofu interconnect link goes down [FX] (*20)	c	c	o
define mpiexec-stdouterr-unit	The unit of standard output and standard error output outputted by mpiexec command [FX] (*21)	c	c	o
define mpiexec-stdout	The default standard output of mpiexec command [FX] (*22)	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
define mpiexec-stderr	The default standard error output of mpiexec command [FX] (*22)	c	c	o
define interact-mpiexec-stdout	The default standard output of mpiexec command (Interactive job) [FX] (*22)	c	c	o
define interact-mpiexec-stderr	The default standard error output of mpiexec command (Interactive job) [FX] (*22)	c	c	o
define mpiexec-std-emptyfile	Whether to create empty file by mpiexec command [FX] (*23)	c	c	o
define llio-auto-readahead	Operation of whether or not to automatically read ahead (*24)	c	c	o
define llio-async-close	Operation of whether or not to asynchronously close files on first-layer storage and second-layer storage (*24)	c	c	o
define llio-cn-read-cache	Operation of whether or not to cache a file in compute node cache when reading the file from first-layer storage or second-layer storage (*24)	c	c	o
define llio-sio-read-cache	Operation of whether or not to cache a file in first-layer storage when reading the file from second-layer storage to a compute node (*24)	c	c	o
define llio-uncompleted-fileinfo-path	Output destination for information on files that have not completed writing from cache to second-layer storage at job end (*25)	c	c	o
define llio-perf	Whether to output LLIO performance information to a file (*24)	c	c	o
define llio-perf-path	Path to the file to output LLIO performance information (*25)	c	c	o
define allocation-io-mode	Default of pjsb option no-io-exclusive, or io-exclusive [FX] (*26)	c	c	o
joblimit subjobnum	Bulk-job unit subjob number limitation (default value none)	c	c	o
joblimit priv-pri	Job unit user priority in resource unit When "job_prio" of the job selection policy is set, it is effective.	c	c	o
joblimit elapse	Limitation at elapsed time of each job	c	c	o
joblimit adaptive-elapsed-time-min	Minimum elapsed time limit for each job	c	c	o
joblimit adaptive-elapsed-time-max	Maximum elapsed time limit for each job	c	c	o
joblimit node	Limitation of number of nodes of each job	c	c	o
joblimit node-mem	Limitation of use memory of each node	c	c	o
joblimit proc-core	Corefile size limitation in each process	c	c	o
joblimit proc-cpu	Limitation at time of CPU of each process	c	c	o
joblimit proc-crproc	Number of generation of process of user limitation in each process	c	c	o
joblimit proc-data	Data segment size limitation in each process	c	c	o
joblimit proc-lockm	Locked memory size limitation in each process	c	c	o
joblimit proc-msgq	Message queue size limitation in each process	c	c	o
joblimit proc-openfd	File descriptor number limitation in each process	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
joblimit proc-psig	Limitation of number of pending signals in each process	c	c	o
joblimit proc-filesz	Limitation of size of file in each process	c	c	o
joblimit proc-stack	Stack size limitation in each process	c	c	o
joblimit proc-vmem	Virtual memory limitation in each process	c	c	o
joblimit vnode-core	Number of CPU cores of each virtual node	c	c	o
joblimit vnode-mem	Limitation of use memory of each virtual node Or limitation of "number of CPU cores per virtual node" x use memory per CPU cores"	c	c	o
joblimit vnode	Limitation of number of vnodes of each job	c	c	o
joblimit total-cores	Limitation of the total number of CPU use cores of each batch job (*13)	c	c	o
joblimit node-elapsed	Limitation of the node time product of each batch job ("Number of required nodes" x "elapsed time limit" for each batch job)	c	c	o
joblimit adaptive-node-elapsed-min	Limitation of the minimum node time product of each batch job ("Number of required nodes" x "minimum elapsed time limit" for each batch job)	c	c	o
joblimit adaptive-node-elapsed-max	Limitation of the maximum node time product of each batch job ("Number of required nodes" x "maximum elapsed time limit" for each batch job)	c	c	o
joblimit total-cores-elapsed	Limitation of the total number of CPU use cores time product of each batch job (*13) ("Number of CPU use cores" x "elapsed time limit" for each batch job)	c	c	o
joblimit custom-CustomResourceName	Limitation of the custom resource allocation	c	c	o
joblimit custom-total-CustomResourceName	Limitation of the total of the custom resource allocation of each batch job	c	c	o
joblimit custom-node-CustomResourceName	Limitation of the custom resource allocation of each node	c	c	o
joblimit custom-vnode- CustomResourceName	Limitation of the custom resource allocation of each virtual node	c	c	o
joblimit interact-elapsed	Limitation at elapsed time of each job (Interactive job)	c	c	o
joblimit interact-adaptive-elapsed-time-min	Minimum elapsed time limit for each job (Interactive job)	c	c	o
joblimit interact-adaptive-elapsed-time-max	Maximum elapsed time limit for each job (Interactive job)	c	c	o
joblimit interact-node	Limitation of number of nodes of each job (Interactive job)	c	c	o
joblimit interact-node-mem	Limitation of use memory of each node (Interactive job)	c	c	o
joblimit interact-proc-core	Corefile size limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-cpu	Limitation at time of CPU of each process (Interactive job)	c	c	o
joblimit interact-proc-crproc	Number of generation of process of user limitation in each process(Interactive job)	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
joblimit interact-proc-data	Data segment size limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-lockm	Locked memory size limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-msgq	Message queue size limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-openfd	File descriptor number limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-psig	Limitation of number of pending signals in each process (Interactive job)	c	c	o
joblimit interact-proc-filesz	Limitation of size of file in each process (Interactive job)	c	c	o
joblimit interact-proc-stack	Stack size limitation in each process (Interactive job)	c	c	o
joblimit interact-proc-vmem	Virtual memory limitation in each process (Interactive job)	c	c	o
joblimit interact-vnode-core	Number of CPU cores of each virtual node (Interactive job)	c	c	o
joblimit interact-vnode-mem	Limitation of use memory of each virtual node (Interactive job)	c	c	o
joblimit interact-vnode	Limitation of number of vnodes of each job (Interactive job)	c	c	o
joblimit interact-total-cores	Limitation of the total number of CPU use cores of each job (Interactive job) (*13)	c	c	o
joblimit interact-node-elapse	Limitation of the node time product (interactive job) ("Number of required nodes" x "elapse time limit" for each interactive job)	c	c	o
joblimit interact-adaptive-node-elapse-min	Limitation of the minimum node time product (interactive job) ("Number of required nodes" x "minimum elapse time limit" for each interactive job)	c	c	o
joblimit interact-adaptive-node-elapse-max	Limitation of the maximum node time product (interactive job) ("Number of required nodes" x "maximum elapse time limit" for each interactive job)	c	c	o
joblimit interact-total-cores-elapse	Limitation of the total number of CPU use cores time product of each job (interactive job) ("Number of CPU use cores" x "elapse time limit" for each interactive job) (*13)	c	c	o
joblimit interact-custom-CustomResourceName	Limitation of the custom resource allocation (interactive job)	c	c	o
joblimit interact-custom-total-CustomResourceName	Limitation of the total of the custom resource allocation of each batch job (interactive job)	c	c	o
joblimit interact-custom-node-CustomResourceName	Limitation of the custom resource allocation of each node (interactive job)	c	c	o
joblimit interact-custom-vnode-CustomResourceName	Limitation of the custom resource allocation of each virtual node (interactive job)	c	c	o
joblimit llio-localtmp-size	Size of the node temporary area on first-layer storage	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
joblimit llio-sharedtmp-size	Size of the shared temporary area on first-layer storage Multiply this value by the number of allocated nodes to calculate the size of the shared temporary area that can be used by 1 job.	c	c	o
joblimit llio-cn-cached-write-size	Threshold value of whether or not to cache data when writing the data to first-layer storage	c	c	o
joblimit llio-cn-cache-size	Compute node cache size on first-layer storage From the memory allocated to the job, memory of the specified size is reserved as the upper limit.	c	c	o
joblimit llio-stripe-count	Number of stripes per file when files are distributed across first-layer storage	c	c	o
joblimit llio-stripe-size	Stripe size when files are distributed across first-layer storage	c	c	o
execute pjsub	pjsub command execution authority	c	c	o
execute pjsub-interact	Interactive job submitting authority	c	c	o
execute pjsub-step	Step job submitting authority	c	c	o
execute pjsub-bulk	Bulk job submitting authority	c	c	o
execute pjsub-mswk	Master-worker job submitting authority	c	c	o
execute pjsub-fixed-elapsed-time	elapsed time limit specification submission authority	c	c	o
execute pjsub-adaptive-elapsed-time	minimum elapsed time limit specification submission authority	c	c	o
execute pjstat	pjstat command execution authority	o	x	x
execute pjdel	pjdel command execution authority	o	x	x
execute pjhold	pjhold command execution authority	o	x	x
execute pjrls	pjrls command execution authority	o	x	x
execute pjwait	pjwait command execution authority	o	x	x
execute pjalter	pjalter command execution authority	o	x	x
execute pjsig	pjsig command execution authority	o	x	x
execute pjacl	pjacl command execution authority	o	x	x
execute mpiexec-std	-of --of -std --std options execution authority of mpiexec command [FX]	c	c	o
execute mpiexec-stdout	-ofout --ofout -stdout --stdout options execution authority of mpiexec command [FX]	c	c	o
execute mpiexec-stderr	-oferr --oferr -stderr --stderr options execution authority of mpiexec command [FX]	c	c	o
execute interact-mpiexec-std	-of --of -std --std option execution authority of mpiexec command(Interactive job) [FX]	c	c	o
execute interact-mpiexec-stdout	-ofout --ofout -stdout --stdout options execution authority of mpiexec command(Interactive job) [FX]	c	c	o
execute interact-mpiexec-stderr	-oferr --oferr -stderr --stderr options execution authority of mpiexec command(Interactive job) [FX]	c	c	o
execute command-api	Authority for using command-api	o	x	x

item	Content	Hierarchy		
		CL	RU	RG
execute pjsub-P-vn-policy	-P vn-policy option execution authority of pjsub command [PG]	c	c	o
execute pjsub-P-exec-policy	-P exec-policy option execution authority of pjsub command [PG]	c	c	o
execute pjsub-torus	Torus mode job submission authority [FX]	c	c	o
execute pjsub-mesh	Mesh mode job submission authority [FX]	c	c	o
execute pjsub-noncont	Non-contiguous mode job submission authority [FX]	c	c	o
execute pjdel-enforce	--enforce option specification authority of pjdel command	c	c	o
execute pjdel-no-history	--no-history option specification authority of pjdel command	c	c	o
execute pjhold-enforce	--enforce option specification authority of pjhold command	c	c	o
execute pjsub-batch	Batch job submission authority	c	c	o
execute pjsub-normal	Normal job submission authority	c	c	o
execute pjsub-node	Node allocated job submission authority	c	c	o
execute pjsub-vnode	Virtual node allocated job submission authority	c	c	o
execute pjsub-nostrict	Authority for submitting Job that does not specify strict nor strict-io	c	c	o
execute pjsub-strict	Authority for submitting Job that specifies strict	c	c	o
execute pjsub-strict-io	Authority for submitting Job that specifies strict-io	c	c	o
execute pjsub-at	Specified execution start time job submission authority	c	c	o
execute pjsub-net-route	The authority to specify pjsub --net-route option when job submitting [FX]	c	c	o
execute pjsub-no-io-exclusive	Authority for submitting Job that specifies no-io-exclusive [FX]	c	c	o
execute pjsub-io-exclusive	Authority for submitting Job that specifies io-exclusive [FX]	c	c	o
permit pjsub	Group permission that can be specified by "pjsub -g" (Executable group) The permission of a current group when "-g option" is unspecified is contained.	c	c	o
permit pjstat	Permission of object job that can display pjstat(*3)	c	c	o
permit pjdel	Permission of executable object job of pjdel(*3)	c	c	o
permit pjhold	Permission of executable object job of pjhold(*3)	c	c	o
permit pjrls	Permission of executable object job of pjrls(*3)	c	c	o
permit pjwait	Permission of executable object job of pjwait(*3)	c	c	o
permit pjalter	Permission of executable object job of pjalter(*3)	c	c	o
permit pjsig	Permission of executable object job of pjsig(*3)	c	c	o
permit pjacl	pjacl permission of user/group that can be displayed	c	c	o
permit pmerls	Permission of executable object job of pmerls(*3)	c	c	o
permit pmalter	Permission of executable object job of pmalter(*3)	c	c	o
permit pjshowrsc	Permission of executable object job of pjshowrsc(*3)	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
select custom-CustomResourceName	Type of custom resource	c	c	o

(*1)

The group cannot specify it for the definition object in this item.(user=<def>:gname, user=uname:gname)

(*2)

The number of bulk-subjob simultaneous execution limitations is the total values of subjob in bulk-job being executing all of the object.

(*3)

The object job is permitted by the specification of execution user/execution group of a job concerned.

(*4)

The definition item related to the fair share is effective even if the job selection policy does not concern fair share values.

Multiple fair share set name can be specified by separating them with commas (.). A different value (fair share value or recovery *magnification*) can be set for each fair share set by specifying @*fairshareset* following the value.

When @*fairshareset* is omitted, @def_fs is applied.

When * (asterisc) is specified for *fairshareset*, the value is applied to all fair share sets.

When two or more settings are specified for the same fair share set, only the last value is effective.

The description example is shown as follows.

[example 1] Specifies different initial user fair share values for fair share set fs1 and fs2.

```
define fshare-init 120@fs1,90@fs2
```

Initial user fair share value of fair share set fs1: 120

Initial user fair share value of fair share set fs2: 90

[example 2] Specifies same initial user fair share value for def_fs, fs1 and fs2 by using "*".

```
define fshare-init 120@*
```

Initial user fair share value of fair share set def_fs: 120

Initial user fair share value of fair share set fs1: 120

Initial user fair share value of fair share set fs2: 120

[example 3] Specifies same initial user fair share value for fair share sets other than fs1.

```
define fshare-init 120@*,90@fs1
```

Initial user fair share value of fair share set def_fs: 120

Initial user fair share value of fair share set fs1: 90

Initial user fair share value of fair share set fs2: 120

(*5)

Virtual node allocated jobs are excluded.

When the item setting UseCoreLimit value on papjm.conf file is "all", the simultaneous number of use node limitations can not be specified.

(*6)

You can specify whether the simultaneous number of CPU core limitations are applied to the jobs that allocate virtual nodes only or all jobs. See the item UseCoreLimit value on papjm.conf file.

(*7)

Specify the value "pack" or "unpack".

(*8)

Specify the value "balancing" (balanced allocation) or "concentration" (concentrated allocation).

(*9)

Specify the value "abs_pack", "pack", "abs_unpack[=n]" or "unpack[=n]".

(*10)

Specify the value "simplex" (node monopolization), or "share" (node share).

(*11)

Specify the value "torus" (Torus mode), "mesh" (Mesh mode), or "noncont" (Non-contiguous mode).

(*12)

Specify the value "node" or "vnode".

When you submit a job by pjsub command without specifying the "node" option and "vnode" option, this setting determines whether it is the node allocated job or the virtual node allocated job.

Value	Operations when submitting a job by pjsub command without the "node" option and "vnode" option
node	Node allocated job. For batch jobs, the default value in "joblimit node" is adopted for the number of required nodes. For interactive jobs, the default value in "joblimit interact-node" is adopted for the number of required nodes.
vnode	Virtual node allocated job. For batch jobs, the default value in "joblimit vnode" is adopted for the number of required nodes. For interactive jobs, the default value in "joblimit interact-vnode" is adopted for the number of required nodes.

(*13)

The total number of CPU use core and the total number of CPU use core time product are limited for all jobs. However, it is possible to apply these limits for virtual node allocated jobs only.

See the item TotalCoresUpper value on papjm.conf file.

(*14)

This setting is effective for processing systems other than Development Studio.

value	Range of logical CPUs that job processes will execute
job	Logical CPUs for job processes that is set by CPU usage division function
all	All logical CPUs

(*15)

Specify the job information display mode of the job that is not permitted in permit pjstat.

value	Display format of job information
anonymous	Display job information that is masked.
nothing	Do not display job information.
summary	Do not display job information but the number of the jobs of the summary counts.

(*16)

Specify the value "fixed" or "adaptive".

When you submit a job by pjsub command without specifying the "elapsed" option, this setting determines whether it is the elapsedtimelimit job or elapsedtimemin job.

Value	Operations when submitting a job by pjsub command without the "elapsed" option
fixed	elapsed time limit value specified. For batch jobs, the default value in "joblimit elapsed" is adopted as the elapsed time limit value. For interactive jobs, the default value in "joblimit interact-elapsed" is adopted as the elapsed time limit value.
adaptive	elapsed time limit value specified. For batch jobs, The default value of "joblimit adaptive-elapsed-time-min" is adopted as the minimum value of elapsed time, and the default value of "joblimit adaptive-elapsed-time-max" is adopted as the maximum value of elapsed time. For interactive jobs, The default value of "joblimit interact-adaptive-elapsed-time-min" is adopted as the minimum value of elapsed time, and the default value of "joblimit interact-adaptive-elapsed-time-max" is adopted as the maximum value of elapsed time.

(*17)

Specify the display format for the planned execution start time of a job by the pjstat command.

Display format	Description
fine	Display the planned execution start time of job in minutes.
custom= <i>time_zone,time_interval</i>	Depending on the planned execution start time, you can specify the accuracy of the displayed time. Multiple display formats can be specified by delimiting them with a semicolon(;). If it is specified multiple, it must begin from the closest time zone apart from now. Example: custom=01:00:00,00:10;08:00:00,01:00;*,24:00

Specify *time_zone* and *time_interval* as follows.

- *time_zone*:

Specify the time zone targeted by this setting as the time from the current time. You can specify the range of *time_zone* from 00:01:00 to 9999:59:59.

If "*" is specified, it is assumed that all time zones are thereafter.

If "over" is specified, it is assumed that the time zone exceeds the scheduling period.

- *time_interval*:

Specify the display accuracy when the planned execution start time corresponds to *time_zone*. The accuracy is specified in the following format.

Format	Description
hh:mm	The planned execution start time of a job is rounded up at intervals of hh: Hours, mm: Minutes. hh values: 00 - 9999 mm values: 00 - 59 In this case, the display with the pjstat command is in the MM/DD hh:mm format.
DD	The planned execution start time of a job is rounded up at the interval of DD days and displayed. DD values: 0 - 9999 In this case, the display with the pjstat command is in the YYYY/MM/DD format.
"string"	The string specified by "string" is displayed as the planned execution start time of a job. It is used when you do not want to display the planned execution start time of a job. The maximum length of the string is 7 characters. The available characters are single-byte alphanumeric characters, blank, and the characters shown below. ! " # \$ % & () * + , - . / : < = > ? @ [] ^ _ ' { ~ If the string contains blank, enclose the string in double quotation escaped with backslash and enclose the display format in double quotes. Example: "custom=*,\"A B\""

(*18)

Specify a comma-delimited mark to be added to the scheduled job execution start time displayed by the pjstat command.

Specifiable mark	Meaning
@	Job for which the start time is specified with --at option of the pjsb command.
<	Backfill applied job.
#	The planned execution start time of a job exceeds the scheduling period.

If you do not add or add all marks, specify the following string.

If invalid character string is specified, it becomes all.

- all:

All marks are added.

- nothing:

All marks are not added.

(*19)

nostrict, strict or strict-io is specified.

(*20)

Specify the value "dynamic" (Dynamically change the communication path) or "static" (Do not change the communication path).

(*21)

Specify the value "mpiexec"(Outputs for each mpiexec command) or "proc"(Outputs for each rank).

(*22)

The unit of standard output and standard error output is defined by the value of "define mpiexec-stdouterr-unit". Can use following metacharacter:

Metacharacter	Description
%j	Expand the job ID.
%J	Expand the sub job ID.
%b	Expand the bulk number.
%s	Expand the step number.
%n	Expand the job name.
%o	Expand the standard output file path. Expand "%n.%J.out" in an interactive job.
%e	Expand the standard error output file path. Expand "%n.%J.err" in an interactive job.
%m	Expand the executions number of mpiexec command in a job.
%r	Expand the rank number (include spawn number). static processes: "rank number" dynamic processes: "rank number"@ "spawn number" Expand empty string if mpiexec command outputs for each mpiexec command.
%R	Expand the rank number. static processes: "rank number" dynamic processes: "rank number" Expand empty string if mpiexec command outputs for each mpiexec command.
%S	Expand the spawn number. static processes: "0" dynamic processes: "spawn number" Expand empty string if mpiexec command outputs for each mpiexec command.

In %r, %R and %S, can use following format:

Format	Description
%0"number"r,%0"number"R,%0"number"S	The value should be zero padded.
%/"number"r,%/"number"R,%/"number"S	The value should be truncated.
%0"number1"/"number2"r,%0"number1"/"number2"R,%0"number1"/"number2"S	The value should be zero padded and truncated.

(*23)

Specify on (create), off (not create), force-on (create, ignore environment variable PLE_MPI_STD_EMPTYFILE) and force-off (not create, ignore environment variable PLE_MPI_STD_EMPTYFILE).

(*24)

Specify off or on.

(*25)

Specifies an absolute path or a path relative to the job submission directory.

(*26) no-io-exclusive or io-exclusive is specified.

[FX] Only FX server is effective.

[PG] Only PRIMERGY server is effective.

Group unit definition section

item	Content	Hierarchy		
		CL	RU	RG
limit ru-accept	Number of job simultaneous acceptance limitations in resource unit	c	o	x
limit ru-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource unit Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	o	x
limit ru-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource unit	c	o	x
limit ru-accept-stepsubjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource unit	c	o	x
limit ru-run-job	Number of job simultaneous execution limitations in resource unit	c	o	x
limit ru-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource unit	c	o	x
limit ru-use-node	Simultaneous number of use node limitations in resource unit (*2)	c	o	x
limit ru-use-core	Simultaneous number of CPU use core limitations in resource unit (*3)	c	o	x
limit ru-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit	c	o	x
limit ru-interact-accept	Number of job simultaneous acceptance limitations in resource unit(Interactive job)	c	o	x
limit ru-interact-run-job	Number of job simultaneous execution limitations in resource unit(Interactive job)	c	o	x
limit ru-interact-use-node	Simultaneous number of use node limitations in resource unit(Interactive job) (*2)	c	o	x
limit ru-interact-use-core	Simultaneous number of CPU use core limitations in resource unit(Interactive job) (*3)	c	o	x
limit ru-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit(Interactive job)	c	o	x
limit rg-accept	Number of job simultaneous acceptance limitations in resource group	c	c	o
limit rg-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource group Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	c	o
limit rg-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource group	c	c	o

item	Content	Hierarchy		
		CL	RU	RG
limit rg-accept-stepsjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource group	c	c	o
limit rg-run-job	Number of job simultaneous execution limitations in resource group	c	c	o
limit rg-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource group	c	c	o
limit rg-use-node	Simultaneous number of use node limitations in resource group (*2)	c	c	o
limit rg-use-core	Simultaneous number of CPU use core limitations in resource group (*3)	c	c	o
limit rg-custom-CustomResourceName	Number of custom resource simultaneous allocation in resource in resource group	c	c	o
limit rg-interact-accept	Number of job simultaneous acceptance limitations in resource group(Interactive job)	c	c	o
limit rg-interact-run-job	Number of job simultaneous execution limitations in resource group(Interactive job)	c	c	o
limit rg-interact-use-node	Simultaneous number of use node limitations in resource group(Interactive job) (*2)	c	c	o
limit rg-interact-use-core	Simultaneous number of CPU use core limitations in resource group(Interactive job) (*3)	c	c	o
limit rg-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource group(Interactive job)	c	c	o
define pri-g	Group priority in resource unit When "group_prio" of the job selection policy is set, it is effective. (*1)	c	c	o
define fshare-init-g	Group fair share initial value in resource unit When "group_fairshare" of the job selection policy is set, it is effective.	c	o	x
define fshare-recovery-g	Group fair share recovery rate in resource unit When "group_fairshare" of the job selection policy is set, it is effective.	c	o	x

(*1)

The definition item related to the fair share is effective even if the job selection policy does not concern fair share values.

Multiple fair share set name can be specified by separating them with commas (.). A different value (fair share value or recovery magnification) can be set for each fair share set by specifying *@fairshareset* following the value.

When *@fairshareset* is omitted, *@def_fs* is applied.

When * (asterisc) is specified for *fairshareset*, the value is applied to all fair share sets.

When two or more settings are specified for the same fair share set, only the last value is effective.

The description example is shown as follows.

[example 1] Specifies different initial group fair share values for fair share set fs1 and fs2.

```
define fshare-init-g 120@fs1,90@fs2
Initial group fair share value of fair share set fs1: 120
Initial group fair share value of fair share set fs2: 90
```

[example 2] Specifies same initial group fair share value for def_fs, fs1 and fs2 by using "*".

```
define fshare-init-g 120@*
Initial group fair share value of fair share set def_fs: 120
```

Initial group fair share value of fair share set fs1: 120
Initial group fair share value of fair share set fs2: 120

[example 3] Specifies same initial group fair share value for fair share sets other than fs1.

```
define fshare-init-g 120@*,90@fs1
Initial group fair share value of fair share set def_fs: 120
Initial group fair share value of fair share set fs1: 90
Initial group fair share value of fair share set fs2: 120
```

(*2)

Virtual node allocated jobs are excluded.

When the item setting UseCoreLimit value on papjm.conf file is "all", the simultaneous number of use node limitations can not be specified.

(*3)

You can specify whether the simultaneous number of CPU core limitations are applied to the jobs that allocate virtual nodes only or all jobs. See the item UseCoreLimit value on papjm.conf file.

Whole unit definition section

item	Content	Hierarchy		
		CL	RU	RG
limit ru-accept	Number of job simultaneous acceptance limitations in resource unit	c	o	x
limit ru-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource unit Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	o	x
limit ru-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource unit	c	o	x
limit ru-accept-stepsubjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource unit	c	o	x
limit ru-run-job	Number of job simultaneous execution limitations in resource unit	c	o	x
limit ru-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource unit	c	o	x
limit ru-use-node	Simultaneous number of use node limitations in resource unit (*1)	c	o	x
limit ru-use-core	Simultaneous number of CPU use core limitations in resource unit (*2)	c	o	x
limit ru-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit	c	o	x
limit ru-interact-accept	Number of job simultaneous acceptance limitations in resource unit(Interactive job)	c	o	x
limit ru-interact-run-job	Number of job simultaneous execution limitations in resource unit(Interactive job)	c	o	x
limit ru-interact-use-node	Simultaneous number of use node limitations in resource unit(Interactive job) (*1)	c	o	x
limit ru-interact-use-core	Simultaneous number of CPU use core limitations in resource unit(Interactive job) (*2)	c	o	x
limit ru-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource unit(Interactive job)	c	o	x

item	Content	Hierarchy		
		CL	RU	RG
limit rg-accept	Number of job simultaneous acceptance limitations in resource group	c	c	o
limit rg-accept-allsubjob	Number of simultaneous acceptance limitations for all sub jobs in resource group Normal jobs and Master-Worker jobs in a batch job count as one subjob	c	c	o
limit rg-accept-bulksubjob	Number of simultaneous acceptance limitations for sub jobs of bulk jobs in resource group	c	c	o
limit rg-accept-stepsubjob	Number of simultaneous acceptance limitations for sub jobs of step jobs in resource group	c	c	o
limit rg-run-job	Number of job simultaneous execution limitations in resource group	c	c	o
limit rg-run-bulksubjob	Number of bulk-subjob simultaneous execution limitations in resource group	c	c	o
limit rg-use-node	Simultaneous number of use node limitations in resource group (*1)	c	c	o
limit rg-use-core	Simultaneous number of CPU use core limitations in resource group (*2)	c	c	o
limit rg-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource group	c	c	o
limit rg-interact-accept	Number of job simultaneous acceptance limitations in resource group(Interactive job)	c	c	o
limit rg-interact-run-job	Number of job simultaneous execution limitations in resource group(Interactive job)	c	c	o
limit rg-interact-use-node	Simultaneous number of use node limitations in resource group(Interactive job) (*1)	c	c	o
limit rg-interact-use-core	Simultaneous number of CPU use core limitations in resource group(Interactive job) (*2)	c	c	o
limit rg-interact-custom-CustomResourceName	Number of custom resource simultaneous allocation limitations in resource group(Interactive job)	c	c	o

(*1)

Virtual node allocated jobs are excluded.

When the item setting UseCoreLimit value on papjm.conf file is "all", the simultaneous number of use node limitations can not be specified.

(*2)

You can specify whether the simultaneous number of CPU core limitations are applied to the jobs that allocate virtual nodes only or all jobs. See the item UseCoreLimit value on papjm.conf file.

[EXAMPLE]

- Example of displaying all set of job ACL function item content.

```
<cmgr>$ pmjacladm -c clst1 --show '*'
#
# JOBACL definition - clst1
# [ * ]
#
USER: CL {
```

```

    user=<def> { # last update 2011-03-01 19:26:33
        define rscunit      rscunit1
        define rscgroup     rscgroup1
    }
}

USER: CL, RU=rscunit1 {
    user=<def> { # last update 2011-03-01 19:26:33
        limit ru-accept      unlimited
        limit ru-run-job     unlimited
        joblimit elapse      1          24:00:00  24:00:00
        joblimit node        1          80000     1
        joblimit node-mem    1M         10G       2G
    }

    user=<def>:group2 { # last update 2011-03-01 19:26:33
        define rscgroup      rscgroup2
    }

    user=user1 { # last update 2011-03-01 19:26:33
        permit pjstat        allow own
        permit pjstat        + allow g(group1)
        permit pjdel         allow own
        permit pjdel         + allow g(group1)
    }
}

USER: CL, RU=rscunit1, RG=rscgroup2 {
    user=user2 { # last update 2011-03-01 19:26:33
        execute pjsub        disable
    }

    user=user3 { # last update 2011-03-01 19:26:33
        permit pjhold        allow own
        permit pjrls         allow own
    }

    user=user3:group2 { # last update 2011-03-01 19:26:33
        define ingroup-pri   128
    }
}

GROUP: CL, RU=rscunit1 {
    group=group2 { # last update 2011-03-01 19:26:33
        define pri-g         160
    }
}

```

The last updated date and time of each definition subsection is displayed by the comment.

- Example of setting job ACL function

The job operation administrator edits the USER definition of the resource unit "Rscunit1" in which the authority exists in the cluster, and the definition is reflected in job ACL function (merge).

```

<cmgr>$ pmjacladm -c clst1 --show -o jobacl.txt 'USER: CL, RU=rscunit1'
<cmgr>$ vi jobacl.txt (A necessary definition item is edited.)

```

```

<cmgr>$ pmjacladm -c clst1 --set -f jobacl.txt # The edit data is reflected(merge).

```

The entry of "NewuserB" is newly added to the USER definition of the resource unit "Rscunit1".

As for the definition item, a set value of job ACL function is all the same before the definition is added because of undefined.

```

$ pmjacladm -c clst1 --set 'USER: CL, RU=rscunit1 { user=newuserB }'

```

If the definition item setting is done to an undefined user, the user definition is newly registered.

```
$ pmjacladm -c clst1 --set 'USER: CL, RU=rscunit1 { user=newuserC {joblimit elapse 1 3600 3600 } }'
```

The limitation value upper bound of the elapsed time of userD of the USER definition is changed at 3600 seconds, the default value is changed at 1800 seconds, and the number of job simultaneous execution limitations is changed undefined.

```
$ pmjacladm -c clst1 --set 'USER: CL, RU=rscunit1 { user=userD {joblimit elapse 1 3600 1800; limit ru-run-job <undef> } }'
```

- Example of confirming application value by --show --apply option

The resource unit, the resource group, the user, and the group designation are necessary for "--apply option".

When the specification of the group is omitted, user's default group is applied. (It returns because of the error when "--noname option" is specified.)

When only the resource unit is specified, only the definition item only for the resource unit is displayed.

The following points are different from an actual definition display. The edit of the addition of an appropriate section etc. when using it as a definition is needed.

- The resource unit and the resource group are displayed in coverage, and the definition item only for the cluster and the item only for the resource unit are output.
- The item application display (following description) is added between the definition item and the value.
user=*uname:gname*" is displayed in the user definition object, and the item not to be able to specify concerning the group(user=<def>:*gname*, user=*uname:gname*) is displayed.

```
<cmgr>$ pmjacladm --show --apply 'USER: CL, RU=rscunit1, RG=rscgroup1 {user=user1:group1 }'
#
# JOBACL applied result - clst1
# [ USER: CL, RU=rscunit1, RG=rscgroup1 { user=user1:group1 } ]
#
USER: CL, RU=rscunit1, RG=rscgroup1 {
    user=user1:group1 {
        limit ru-accept                <RU->  unlimited
        limit ru-accept-allsubjob     <RU->  unlimited
        limit ru-accept-bulksubjob     <RU->  unlimited
        limit ru-accept-stepsubjob     <RU->  unlimited
        limit ru-run-job               <RU->  unlimited
        limit ru-run-bulksubjob        <--->  unlimited
        limit ru-use-node              <--->  unlimited
        limit ru-interact-accept       <--->  unlimited
        limit ru-interact-run-job      <--->  unlimited
        limit ru-interact-use-node     <--->  unlimited
    (...snip...)
        define rscunit                 <CL->  rscunit1
        define rscgroup                <CL->  rscgroup1
        define interact-rscunit        <CL->  rscunit1
        define interact-rscgroup       <CL->  rscgroup1
        define pri                     <RUu>  128
        define ingroup-pri             <RGug> 100
    (...snip...)
        joblimit interact-adaptive-elapsed-time-min <---> 00:00:01 23:00:00 01:00:00
    (...snip...)
        permit pjsub                  <--->  allow own
        permit pjstat                 <RUu>  allow own
        permit pjstat                 <RUu>  + allow g(group1)
        permit pjdel                  <RU-g>  allow own
        permit pjdel                  <RU-g>  + allow g(group1)
        permit pjhold                 <--->  deny all
        permit pjrls                  <--->  deny all
        permit pjwait                 <--->  allow own
        permit pjalter                 <--->  allow own
        permit pjsig                  <--->  allow own
        permit pjacl                  <--->  allow own
        permit pmerls                 <--->  deny all
    }
```

```

    permit pmalter          <---> deny all
    permit pjshowrsc        <---> allow all
}

```

The item application display is added between the definition item name and the value. The content of the display is as follows.

Item application display of user unit definition section (USER:)

Display content	meaning
<--->	The default value of the system is applied.
<--->	The application display of the cluster is not seen excluding the cluster administrator (The value is displayed).
<CL->	Application from definition section "user=<def>" of cluster
<CL-g>	Application from definition section "user=<def>:gname" of cluster
<CLu>	Application from definition section "user=uname" of cluster
<CLug>	Application from definition section "user=uname:gname" of cluster
<RU->	Application from definition section "user=<def>" of resource unit
<RU-g>	Application from definition section "user=<def>" of resource unit
<RUu>	Application from definition section "user=uname" of resource unit
<RUug>	Application from definition section "user=uname:gname" of resource unit
<RG->	Application from definition section "user=<def>" of resource group
<RG-g>	Application from definition section "user=<def>:gname" of resource group
<RGu>	Application from definition section "user=uname" of resource group
<RGug>	Application from definition section "user=uname:gname" of resource group

Item application display of group unit definition section (GROUP:)

Display content	meaning
<--->	The default value of the system is applied.
<--->	The application display of the cluster is not seen excluding the cluster administrator (The value is displayed).
<CL->	Application from definition section "group=<def>" of cluster
<CLg>	Application from definition section "group=gname" of cluster
<RU->	Application from definition section "group=<def>" of resource unit
<RUg>	Application from definition section "group=gname" of resource unit
<RG->	Application from definition section "group=<def>" of resource group
<RGg>	Application from definition section "group=gname" of resource group

Item application display of whole unit definition section (ALL:)

Display content	meaning
<--->	The default value of the system is applied.
<--->	The application display of the cluster is not seen excluding the cluster administrator (The value is displayed).
<CL->	Application from definition section of cluster.
<RU->	Application from definition section of resource unit.
<RG->	Application from definition section of resource group.

- Display by form delimited by delimiter (";") by - show --data option .

```
$ pmjacladm -c clst1 --show --data --delimiter ";" '*'
USER:;CL
define;rscunit;rscunit1
define;rscgroup;rscgroup1
USER:;CL;RU=rscunit1
user=<def>
limit;ru-accept;unlimited
limit;ru-run-job;unlimited
joblimit;elapse;1;24:00:00;24:00:00
joblimit;node;1;80000;1
joblimit;node-mem;1M;10G;2G
(...snip...)
user=user1
permit;pjstat;allow own
permit;pjstat;+ allow g(group1)
permit;pjdel;allow own
permit;pjdel;+ allow g(group1)
(...snip...)
```

The definition item name delimits the definition type and the item name and is displayed.

The lower limit value, the upper limit value and the default value of the definition type "joblimit" is delimited and displayed.

The list of the value of the definition type "Permit" is displayed as it is by one item.

- Confirmation of set item name by -- show --items option

The definition item list that can be set is displayed by the section form.

```
$ pmjacladm -c clst1 --show --items
#
# JOBACL items list - clst1
#
#USER: CL {                                <- Definition item only for cluster
#   user=<def> {
#   user=<def>:group {
#   user=user {
#   user=user:group {
#       define rscunit                    rscunit1      # CL
#       define interact-rscunit          rscunit1      # CL
#       execute pjstat                   enable         # CL
#       execute pjdel                    enable         # CL
(...snip...)
#   }
#}
#USER: CL, RU=runame {                    <- Definition item only for resource unit
#   user=<def> {
#   user=<def>:group {
#   user=user {
#   user=user:group {
#       limit ru-accept                  unlimited      # CLRU nogrp
#       limit ru-accept-allsubjob       unlimited      # CLRU nogrp
#       limit ru-accept-bulksubjob       unlimited      # CLRU nogrp
#       limit ru-accept-stepsubjob       unlimited      # CLRU nogrp
#       limit ru-run-job                 unlimited      # CLRU nogrp
#       limit ru-run-bulksubjob          unlimited      # CLRU nogrp
#       limit ru-use-node                unlimited      # CLRU nogrp
(...snip...)
#   }
#}
#USER: CL, RU=runame, RG=rgrname {        <- Definition item only for resource group
#   user=<def> {
#   user=<def>:group {
#   user=user {
#   user=user:group {
```

```

#      limit rg-accept                unlimited    # CLRURG nogrp
#      limit rg-accept-allsubjob      unlimited    # CLRURG nogrp
#      limit rg-accept-bulksubjob       unlimited    # CLRURG nogrp
#      limit rg-accept-stepsubjob       unlimited    # CLRURG nogrp
#      limit rg-run-job                 unlimited    # CLRURG nogrp
(...snip...)
#      define pri                      128          # CLRURG nogrp
#      define ingroup-pri              128          # CLRURG
#      define allocation-mode          torus         # CLRURG
#      define alloc-granularity        node         # CLRURG
#      joblimit subjobnum -             unlimited -   # CLRURG
#      joblimit priv-pri -             255          128 # CLRURG
(...snip...)
#    }
#}
#GROUP: CL, RU=runame {
#  group=<def> {
#    group=group {
#      limit ru-accept                unlimited    # CLRU
#      limit ru-accept-allsubjob      unlimited    # CLRU
#      limit ru-accept-bulksubjob       unlimited    # CLRU
#      limit ru-accept-stepsubjob       unlimited    # CLRU
#      limit ru-run-job                 unlimited    # CLRU
(...snip...)
#    }
#}
#GROUP: CL, RU=runame, RG=rgname {
#  group=<def> {
#    group=group {
#      limit rg-accept                unlimited    # CLRURG
#      limit rg-accept-allsubjob      unlimited    # CLRURG
#      limit rg-accept-bulksubjob       unlimited    # CLRURG
#      limit rg-accept-stepsubjob       unlimited    # CLRURG
#      limit rg-run-job                 unlimited    # CLRURG
(...snip...)
#    }
#}
#ALL: CL, RU=runame {
#  limit ru-accept                unlimited    # CLRU
#  limit ru-accept-allsubjob      unlimited    # CLRU
#  limit ru-accept-bulksubjob       unlimited    # CLRU
#  limit ru-accept-stepsubjob       unlimited    # CLRU
#  limit ru-run-job                 unlimited    # CLRU
(...snip...)
#}
#ALL: CL, RU=runame, RG=rgname {
#  limit rg-accept                unlimited    # CLRURG
#  limit rg-accept-allsubjob      unlimited    # CLRURG
#  limit rg-accept-bulksubjob       unlimited    # CLRURG
#  limit rg-accept-stepsubjob       unlimited    # CLRURG
#  limit rg-run-job                 unlimited    # CLRURG
(...snip...)
#}

```

The hierarchy that can be specified is marked as a comment by the end of each definition item.

CL: Definition item only for cluster

CLRU: Definition item only for resource unit (It is possible to specify it for a common definition value within the range of the cluster).

CLRURG: Resource group definition item (It is possible to specify it for a common definition value within the range of the hierarchy of the high rank).

nogrp: User unit definition item for which definition concerning group cannot be specified (user=<def>:group, user=user:group.

- Example of deleting set item by --del option
The job operation administrator becomes an object only the item with the authority.

Example of deleting all data

```
$ pmjacladm -c clst1 --del '*' # Same as "pmjacladm --clear -f /dev/null"
```

Example of deleting all data of USER section

```
$ pmjacladm -c clst1 --del 'USER: *'
```

Example of deleting data of USER section all resource unit (The data of the resource group is deleted).

```
$ pmjacladm -c clst1 --del 'USER: CL, RU=*'
```

Example of deleting only default section of user in USER section cluster data.

```
$ pmjacladm -c clst1 --del 'USER: CL { user=<def> }'
```

Example of deleting value of definition item "joblimit node" of resource group "Rscgroup" in USER section resource unit "Rscunit" (initialization).

```
$ pmjacladm -c clst1 --del 'USER: CL, RU=rscunit, RG=rscgroup { user=user1 {joblimit node } }'
```

- Save and restoration of job ACL function
The cluster administrator should execute save/restoration. When the job operation manager executes save/restoration, only the part with the authority is saved.

```
$ pmjacladm -c clst1 --show -o jobaclbackup.data '*'
```

```
$ pmjacladm -c clst1 --set --clear -f jobaclbackup.data
```

[END STATUS]

The end status is returned as follows:

- 0: Normal end
- 1: Abnormal end

[REMARKS]

- If the job ACL function may be updated on a large scale, it is recommended that the user save the current content of the job ACL function in advance. Pay sufficient attention when reflecting updates, etc., to the job ACL function.
- It does not recommend it because there is a possibility that the update and the reflection not intended with both parties are generated though two or more job operation managers can do the update work from plurals at the same time when there is an authority in the definition in the same resource unit.
- When the mistake is found even by one in the defining statement of the updated job ACL function, all the data specified when this command is executed is not reflected.
- The *rscuname* specified for RU and the *rscgname* specified for RG must be a string of 1 to 63 characters consisting of alphanumeric characters, hyphens "-" and underscores "_".

[SEE ALSO]

pjacl(1), papjm.conf(5)

[MESSAGES]

If an error occurs during the execution of the pmjacladm command, the following messages is output to the standard error output.

Error Message

**[ERR.] PJM 3101 pmjacladm Unknown option *opt*.
Try `pmjacladm --help` for more information.**

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

**[ERR.] PJM 3102 pmjacladm Combination of option is illegal.
Try `pmjacladm --help` for more information.**

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

[ERR.] PJM 3104 pmjacladm Argument format error: *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 3120 pmjacladm Cannot open the file (*path*): *code*.

Meaning

The specified file *path* cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 3132 pmjacladm Write error (*path*): *code*.

Meaning

Writing to the specified file *path* failed.

code: Internal code for maintenance

Action

Confirm that the user has the privileges to write to the specified file and directory. Also confirm the contents of the specified file.

[ERR.] PJM 3151 pmjacladm Control error: *details*.

Meaning

The specified argument (*control*) is incorrect.

details: The following error causes are output.

Output contents	Cause
invalid section <i>SEC</i> .	The specified section or subsection is invalid. <i>SEC</i> shows the character string indicating the invalid section.

Output contents	Cause
user(<i>USER</i>) not found	The user was not found. <i>USER</i> shows the user name.
group(<i>GROUP</i>) not found	The group was not found. <i>GROUP</i> shows the group name.
invalid uid(<i>NUM</i>)	The specified user ID is invalid. <i>NUM</i> shows the invalid numerical value.
invalid gid(<i>NUM</i>)	The specified group ID is invalid. <i>NUM</i> shows the invalid numerical value.
invalid item <i>ITEM</i>	The defined item name is incorrect. <i>ITEM</i> shows the defined item name (including the defined type).
<i>ITEM</i> is cluster(<i>CL</i>) item only.	An exclusive defined item is specified on a layer outside the specifiable range. <i>ITEM</i> shows the defined item name (including the defined type).
<i>ITEM</i> is cluster(<i>CL</i>) or rscunit(<i>RU</i>) item only.	
the default value is greater than the upper value for <i>ITEM</i> .	The defined type is joblimit, and the default value is greater than the upper limit value. <i>ITEM</i> shows the defined item name (including the defined type).
the default value is less than the lower value for <i>ITEM</i> .	The defined type is joblimit, and the default value is less than the lower limit value. <i>ITEM</i> shows the defined item name (including the defined type).
the lower value is greater than the upper value for <i>ITEM</i> .	The defined type is joblimit, and the lower limit value is greater than the upper limit value. <i>ITEM</i> shows the defined item name (including the defined type).
invalid value for <i>ITEM</i> (<i>TYPE</i>).	The specified value is incorrect. A numerical value is specified in an incorrect format or without a unit. Alternatively, the specified numerical value is greater than the maximum value. <i>ITEM</i> shows the defined item name (including the defined type). <i>TYPE</i> shows any of lower, upper, and default only if the defined type is joblimit. <i>TYPE</i> shows nothing if the defined type is other than joblimit.
a value cannot be specified for <i>ITEM</i> (<i>TYPE</i>).	The specified value for an item is one that cannot be specified. <i>ITEM</i> shows the defined item name (including the defined type). <i>TYPE</i> shows any of lower, upper, and default only if the defined type is joblimit. <i>TYPE</i> shows nothing if the defined type is other than joblimit.
format error in <i>ITEM</i>	An error was found in the specification format of a value. If the defined type is upper, an upper limit value and default value are required. If the defined type is joblimit, a lower limit value, upper limit value, and default value are required. <i>ITEM</i> shows the defined item name (including the defined type).
format error.	A specification error other than the above was found.

Action

Check the contents of the specification according to the cause indicated in the details, and specify the correct argument (*control*).

[ERR.] PJM 3152 pmjacladm File format error. line:*lineno*: *details*.

Meaning

The format of coding in the input file is incorrect.

lineno: Number of the line containing an error

details: The following error causes are output

Output contents	Cause
invalid section <i>SEC</i> .	The specified section or subsection is invalid. <i>SEC</i> shows the character string indicating the invalid section.
user(<i>USER</i>) not found	The user was not found. <i>USER</i> shows the user name.
group(<i>GROUP</i>) not found	The group was not found. <i>GROUP</i> shows the group name.
invalid uid(<i>NUM</i>)	The specified user ID is invalid. <i>NUM</i> shows the invalid numerical value.
invalid gid(<i>NUM</i>)	The specified group ID is invalid. <i>NUM</i> shows the invalid numerical value.
invalid item <i>ITEM</i>	The defined item name is incorrect. <i>ITEM</i> shows the defined item name (including the defined type).
<i>ITEM</i> is cluster(<i>CL</i>) item only.	An exclusive defined item is specified on a layer outside the specifiable range. <i>ITEM</i> shows the defined item name (including the defined type).
<i>ITEM</i> is cluster(<i>CL</i>) or rscunit(<i>RU</i>) item only.	
the default value is greater than the upper value for <i>ITEM</i> .	The defined type is joblimit, and the default value is greater than the upper limit value. <i>ITEM</i> shows the defined item name (including the defined type).
the default value is less than the lower value for <i>ITEM</i> .	The defined type is joblimit, and the default value is less than the lower limit value. <i>ITEM</i> shows the defined item name (including the defined type).
the lower value is greater than the upper value for <i>ITEM</i> .	The defined type is joblimit, and the lower limit value is greater than the upper limit value. <i>ITEM</i> shows the defined item name (including the defined type).
invalid value for <i>ITEM</i> (<i>TYPE</i>).	The specified value is incorrect. A numerical value is specified in an incorrect format or without a unit. Alternatively, the specified numerical value is greater than the maximum value. <i>ITEM</i> shows the defined item name (including the defined type). <i>TYPE</i> shows any of lower, upper, and default only if the defined type is joblimit. <i>TYPE</i> shows nothing if the defined type is other than joblimit.
a value cannot be specified for <i>ITEM</i> (<i>TYPE</i>).	The specified value for an item is one that cannot be specified. <i>ITEM</i> shows the defined item name (including the defined type). <i>TYPE</i> shows any of lower, upper, and default only if the defined type is joblimit. <i>TYPE</i> shows nothing if the defined type is other than joblimit.
format error in <i>ITEM</i>	An error was found in the specification format of a value. If the defined type is upper, an upper limit value and default value are required. If the defined type is joblimit, a lower limit value, upper limit value, and default value are required. <i>ITEM</i> shows the defined item name (including the defined type).
format error.	A specification error other than the above was found.

Action

Check the contents according to the cause indicated in the details, and write the file in the correct format.

[ERR.] PJM 3191 pmjacladm Internal error: *details*.

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 3193 pmjacladm This command can be executed on SMM/CCM.

Meaning

Can only be executed on the system management node or compute cluster management node.

Action

Execute the command on the system management node or compute cluster management node.

[ERR.] PJM 3194 pmjacladm Communication failed to system management function.

Meaning

Information retrieval from the system management function failed.

Action

Check the setting configuration of the system management function is correct on the command execution node or the compute cluster management node. If the cause is unknown, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 3195 pmjacladm No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

4.1.31 pmpjmadm command

[NAME]

pmpjmadm - Shows the setting and content of a configuration file for job operations of each resource unit.

[SYNOPSIS]

```
/usr/sbin/pmpjmadm --set [-c clstname] [--rscunit|--ru] rscuname [--force]
/usr/sbin/pmpjmadm --show [-c clstname] [--rscunit|--ru] rscuname
/usr/sbin/pmpjmadm --check [-c clstname] [--file filename | [--rscunit|--ru] rscuname}
/usr/sbin/pmpjmadm --help
```

[DESCRIPTION]

pmpjmadm is a command for job operation administrators to configure or update a configuration file for job operations of an entire cluster, and display its configuration.

This command can be executed only on the system management node.

[OPTIONS]

-c *clstname*

Configure, update or show a configuration file which is for job operations for cluster *clstname*.

If *clstname* is set in environment variable PXMYCLST, this option can be omitted.

`{--rscunit|--ru} rscuname`

The resource unit name *rscuname* which is the executing target of this command is set.

--set

The content of a current configuration file is set and shows the result of the configuration (either success or failure).

The following content which is each resource unit configuration in the configuration file is set.

Resource unit setting `/etc/opt/FJSVtcs/Rscunit.d/resource unit name/pmpjm.conf`

If this command is executed without the `--force` option set, a message always appears with a yes/no prompt asking whether to execute it.

Resource unit which is setting target is decided by `--rscunit` option.

--force

This option forcibly executes commands without displaying a yes/no prompt asking whether to execute them.

--show

This option shows currently set contents in the form of a configuration file. For a default setting that is not explicitly set, "# default" must be added to the end of the line.

--check

This option verifies the format of a current configuration file, compares the file to the actual configuration, and displays the result of these checks.

If the `--file` option is set, this option will verify the format of a specified file, compare the file to the actual configuration, and display the result of these checks.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

An example of `--set` option is as follows:

```
$ pmpjmadm --set --rscunit unit1
[WARNING]
pmpjmadm is going to set config
Do you really want to continue (y/n)? y
```

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[REMARKS]

If you add, modify, or delete custom resources, the information of jobs in the EXIT/CANCEL/REJECT state is deleted.

[SEE ALSO]

`pmpjm.conf(5)`, `papjmadm(8)`

[MESSAGES]

If an error occurs during the execution of the `pmpjmadm` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 6001 pmpjmadm Unknown option *opt*.
Try ``pmpjmadm --help`` for more information.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 6002 pmpjmadm Combination of option is illegal.
Try `pmpjmadm --help` for more information.

Meaning

The combination of specified options is invalid.

Action

Specify the correct combination of options.

[ERR.] PJM 6003 pmpjmadm Unknown option argument *arg*.
Try `pmpjmadm --help` for more information.

Meaning

For the option, the specified argument *arg* is unknown, or no argument is specified.

Action

Specify the correct argument.

[ERR.] PJM 6004 pmpjmadm Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 6020 pmpjmadm Cannot open the file (*path*): *code*.

Meaning

The specified file *path* cannot be opened.

code: Internal code for maintenance

Action

Confirm that the file name and file path are correct.

[ERR.] PJM 6021 pmpjmadm File pathname too long.

Meaning

The file name is too long.

Action

Specify a file name within 4096 characters.

[ERR.] PJM 6022 pmpjmadm File format error: *details*.

Meaning

The format of coding in the configuration file is incorrect.

details:

invalid section detected

The section name is incorrect.

invalid item data detected

The item name is incorrect, or there are duplicate lines.

no section detected or no section

The file does not contain any section.

no ResourceUnitName

ResourceUnitName is not defined.

itemcnt

In ResourceGroup, JobSelectPolicy, PrologueEpilogue or the JobEvaluation section, two or more values are defined in the same item in the same section.

ERROR:line = *line-number submessage*

In the line shown with the line number, there is an error.

Action

Write the file in the correct format.

When "ERROR:line = *line-number submessage*" is output in *details*, remove the cause of the error referring to information on the submessage. After that, write the file in the correct format.

[ERR.] PJM 6029 pmpjmadm Cannot get the file: *code*.

Meaning

The configuration file for the specified index does not exist.

code: Internal code for maintenance

Action

Review the argument for the option, and execute the command again.

[ERR.] PJM 6050 pmpjmadm Operation failed: *code*.

Meaning

Command processing failed.

code: Internal code for maintenance

Action

If a pmscatter or pmexe command message is output together with this message, see the corrective action of the respective message. If command execution fails even after taking the above corrective action, or if no pmscatter or pmexe command message is output, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6075 pmpjmadm Too many custom resources defined.

Meaning

There are too many custom resource names.

Action

Reduce the number of custom resource names to 64 or less.

[ERR.] PJM 6090 pmpjmadm PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 6091 pmpjmadm Internal error: *details*.

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6093 pmpjmadm This command can be executed on SMM.

Meaning

Can only be executed on the system management node.

Action

Execute the command on the system management node.

[ERR.] PJM 6095 pmpjmadm No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

Information Message

[INFO] PJM 6000 pmpjmadm Operation completed.

Meaning

Command processing ended normally.

Action

No action is necessary.

4.1.32 pmpjmopt command

[NAME]

pmpjmopt - Job manager optional control command for job operation administrators

[SYNOPSIS]

Operation of resource unit and resource group

```
/usr/sbin/pmpjmopt [-c clstname] --set-rsc-ug [--rscunit|--ru] rscuname  
                  [--all-rsc-groups|{--rscgrp|--rg} rscgname[,...]]  
                  {{enable|disable}[,{start|stop}]}|[enable|disable],[start|stop}}  
/usr/sbin/pmpjmopt [-c clstname] --show-rsc-ug [--rscunit|--ru] rscuname  
                  [--all-rsc-groups|{--rscgrp|--rg} rscgname[,...]]
```

Operation of fair share value

```
/usr/sbin/pmpjmopt [-c clstname] --set-fairshare [--rscunit|--ru] rscuname
```

```

[--all-fairshares|--fairshare fairshareset]
{--value newvalue|--reset-value}
{{-g group[,...]|-u user[,...]|--gu group:user[,...]}|
{--all-groups|--all-users|--all-gu}}
/usr/sbin/pmpjmopt [-c clstname] --show-fairshare [--rscunit|--ru] rscuname
[--all-fairshares|--fairshare fairshareset]
{{-g group[,...]|-u user[,...]|--gu group:user[,...]}|
{--all-groups|--all-users|--all-gu}}
/usr/sbin/pmpjmopt --help

```

[DESCRIPTION]

pmpjmopt is a command for job operation administrators to set and display whether jobs can be submitted or can be executed, and to set and display the fair share value. This command can be executed only on the system management node.

- Setting and display of enable,start/disable,stop of resource unit and resource group
The execution permission "start" or "stop" that cannot be executed is set specifying the resource unit and the resource group and the setting and the submit permission "enable" or no submit permission "disable" is set.
Moreover, a set content is displayed.
Even if "stop" is set, the job that has already executed it continues processing.
Even if "disable" is set, the job that has already submitted it continues processing.
When optional --all-rsc-groups is specified, all the resource groups in the specified resource unit are effective. Processing does not interrupt even if the error occurs during processing when there are two or more objects of processing (The error message is output).
- Combination of submit (enable/disable) and execute(stop/start) of job.

submit	execute	Processing of job in resource group
-	start	Execution: o.
-	stop	Execution: x (The execution job continues. The execution waiting job maintains the existing state of things.)
enable	-	Submit: o
enable	start	Submit: o, Execution: o
enable	stop	Submit: o, Execution: x (The execution job continues. The execution waiting job maintains the existing state of things.)
disable	-	Submit: x
disable	start	Submit: x, Execution of job of submit: o
disable	stop	Submit: x, Execution: x (The execution job continues. The execution waiting job maintains the existing state of things.)

The set content can be referred to by the --show-rsc-ug option specification.

- Operation of fair share value
The change of the fair share value and the set fair share value are displayed.

OPTIONS

-c *clstname*

The resource unit, the resource group, and the fair share value of the cluster *clstname* are set, and displayed.
When *clstname* is set to environment variable PXMYCLST, this option is omissible.

--set-rsc-ug

The specified value is set to the resource unit and the resource group.
The specified argument is shown as follows.
Argument that can be specified.

argument	description
start	Execution:o
stop	Execution:x
enable	Submit:o
disable	Submit:x

Start and stop cannot be specified simultaneously.

enable and disable cannot be specified simultaneously.

It is possible to combine "start,stop" and "enable,disable" and to specify it simultaneously.

--show-rsc-ug

A set value of the resource unit and the resource group that specifies it is displayed.

{--rscunit|--ru} *rscuname*

The resource unit *rscuname* is specified.

--all-rsc-groups

Specification of all permitted resource groups.

When this option is specified, the specification of the resource group is disregarded.

{--rscgrp|--rg} *rscgname*

The resource group *rscgname* is specified.

--set-fairshare

The fair share value of the specified resource unit is set.

The fair share set to be update can be specified with --all-fairshares or --fairshare *fairshareset* option. When both --all-fairshares option and --fairshare *fairshareset* option are not given, fair share value of the fair share set def_fs is updated.

--show-fairshare

The fair share value is displayed.

The fair share set to be displayed can be specified with --all-fairshares or --fairshare *fairshareset* option. When both --all-fairshares option and --fairshare *fairshareset* option are not given, fair share value of all fair share sets in the resource unit is displayed.

--value *newvalue*

The fair share value is changed to *newvalue*.

--reset-value

The fair share value is reset in an initial value.

--all-groups

The fair share value of all groups is targeted.

--all-users

The fair share value of all users is targeted.

--all-gu

All the fair share values of the user in the group are targeted.

-g *group*[,...]

The fair share value of the group name *group* is targeted.

When the group name starts by the figure, it treats as GID.

When the specified group does not exist, the error is returned.

-u *user*[,...]

The fair share value of the user name *user* is targeted.

When the user name starts by the figure, it treats as UID.

When the specified user does not exist, the error is returned.

`--gu group:user[,...]`

The fair share value of the groupname:username *group:user* is targeted.
When the user name starts by the figure, it treats as GID and UID.
When the specified group or user does not exist, the error is returned.

`--all-fairshares`

Specifies the target of `--set-fairshare` or `--show-fairshare` as all fair share sets in the resource unit.

`--fairshare fairshareset`

Specifies the target of `--set-fairshare` or `--show-fairshare` as fair share set *fairshareset* in the resource unit.

`--help`

This option displays the usage of this command.
If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

- The job execution by the resource unit "unit1" is set in impossibility, and job submit is set possible.

```
$ pmpjmopt -c clst1 --set-rsc-ug --rscunit unit1 enable,stop
[INFO] PJM 6100 pmpjmopt Operation completed.
$
```

- The state of the resource unit "unit1" is confirmed.

```
$ pmpjmopt -c clst1 --show-rsc-ug --rscunit unit1
ResourceUnitName = unit1
Apply Value :ENABLE,STOP
$
```

- The state of the resource group "group1" is confirmed.

```
$ pmpjmopt -c clst1 --show-rsc-ug --rscunit unit1 --rscgrp group1
ResourceGroupName = group1
Apply Value :ENABLE,STOP
$
```

- Changes the user fair share value of "user1" and "user2" of fair share set fs_1 to 1500.

```
$ /usr/sbin/pmpjmopt -c clst1 --set-fairshare --value 1500 --rscunit unit1 -u user1,user2 \
--fairshare fs_1
[INFO] PJM 6100 pmpjmopt Operation completed.
$
```

- Changes the user fair share value of "user1" and "user2" of all fair share sets to 1500.

```
$ /usr/sbin/pmpjmopt -c clst1 --set-fairshare --value 1500 --rscunit unit1 -u user1,user2 \
--all-fairshares
[INFO] PJM 6100 pmpjmopt Operation completed.
$
```

- Initialize the user fair share value of "user1" and "user2" of fair share set fs_1.

```
$ /usr/sbin/pmpjmopt -c clst1 --set-fairshare --rscunit unit1 --reset-value --all-users \
--fairshare fs_1
[INFO] PJM 6100 pmpjmopt Operation completed.
$
```

- Initialize the user-in-group fair share value of "user9" and "user11" that belong to "group1" of all fair share sets.

```
$ /usr/sbin/pmpjmopt -c clst1 --set-fairshare --rscunit unit1 --reset-value \
--gu group1:user9,group1:user11 --all-fairshares
```

```
[INFO] PJM 6100 Operation completed.  
$
```

[END STATUS]

The end status is returned as follows:

0: Normal end

1: Abnormal end

[MESSAGES]

If an error occurs during the execution of the `pmpjmopt` command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

**[ERR.] PJM 6101 pmpjmopt Unknown option *opt*.
Try ``pmpjmopt --help`` for more information.**

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 6104 pmpjmopt Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 6190 pmpjmopt PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 6191 pmpjmopt Internal error: *details*.

Meaning

An internal error was detected.

details: detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 6193 pmpjmopt This command can be executed on SMM.

Meaning

Can only be executed on the system management node.

Action

Execute the command on the system management node.

[ERR.] PJM 6195 pmpjmopt No cluster specified on SMM.

Meaning

No cluster name is specified. Execution on a system management node requires a cluster name.

Action

Specify a cluster name in the -c option of the command, or set a cluster name in the environment variable PXMYCLST.

Information Message

[INFO] PJM 6100 pmpjmopt Operation completed.

Meaning

Command processing ended normally.

Action

No action is necessary.

4.1.33 pmresume command

[NAME]

pmresume - Resume a suspended job.

[SYNOPSIS]

```
/usr/sbin/pmresume [-c clstname] jobid [jobid ...]  
/usr/sbin/pmresume --help
```

[DESCRIPTION]

The pmresume command resumes specified jobs or sub jobs.

This command accepts only jobs or sub jobs in the SUSPENDED state. If some of the specified jobs or sub jobs that are not in the SUSPENDED state, this command outputs an error message and exits abnormally.

When this command is executed, the job manager daemon starts resuming operation, then the states of the specified jobs or sub jobs become RESUME. When the operation has completed, the states become RUNNING, and suspended processes start running again.

This command returns immediately when the resume operation has started. If all operation object jobs or sub jobs are in RUNNING state, the operation has completed. The pjstat command can be used to confirm the states.

For a step job or bulk job, the job ID or the sub job ID can be specified for *jobid*.

This command can be executed only on the system management node or the compute cluster management node.

[Specifying multiple or a range of job IDs]

You can specify a job ID range by using a hyphen (-).

Example: To specify a range from job ID 1 to 100

1-100

When specifying a range, even if you include jobs such as those listed below within the range, no error occurs. Instead, those jobs are ignored, and processing continues.

- Job that does not exist
- Job in a state that cannot be specified

For cases like the above when multiple job IDs are specified, an error message is output since the job IDs are specified explicitly. (Processing is not aborted, and all specified jobs are processed.)

[OPTIONS]

`-c clstname`

This option specifies the target cluster name *clstname*. When the command is executed on the system management node and the environment parameter PXMYCLST has been set, then this option can be omitted.

When the command is executed on the node with other types, this option will be ignored.

`--help`

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following is a use example of the pmresume command:

```
<jmgr>$ pmresume 1
[INFO] PJM 1400 pmresume Accepted job 1.
```

```
<jmgr>$ pmresume 2-3 4
[INFO] PJM 1400 pmresume Accepted job 2-3.
[INFO] PJM 1400 pmresume Accepted job 4.
```

```
<jmgr>$ pjstat 1-4
JOB_ID ... JOB_NAME MD ST USER ...
1 ... jobname1 NM RUN user1 ...
2 ... jobname2 NM RUN user2 ...
3 ... jobname4 NM RUN user3 ...
4 ... jobname4 NM RUN user4 ...
```

[END STATUS]

0: Normal end

1: Abnormal end (If multiple job IDs are specified, an abnormal end is returned when even a single job ends with an error.)

[SEE ALSO]

pmsuspend(8)

[MESSAGES]

If an error occurs during the execution of the pmresume command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 1401 pmresume Unknown option *opt*.
Try `pmresume --help` for more information.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 1410 pmresume No jobid(s) specified.
Try `pmresume --help` for more information.

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 1411 pmresume Invalid jobid syntax *jobid*.

Meaning

The specified job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 1412 pmresume Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job is correctly specified.

[ERR.] PJM 1413 pmresume Job *jobid* state error.

Meaning

The command cannot be executed because the state of the specified job (job ID: *jobid*) is not SUSPENDED.

Action

Confirm whether the state of the job is SUSPENDED.

[ERR.] PJM 1490 pmresume PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 1491 pmresume Internal error: *details*.

Meaning

An internal error was detected.

details: detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PJM 1493 pmresume This command can be executed on SMM/CCM.

Meaning

Can only be executed on the system management node or compute cluster management node.

Action

Execute the pmresume command on the system management node or compute cluster management node.

[ERR.] PJM 1495 pmresume No cluster specified on SMM.

Meaning

No cluster name is specified.

Action

When executing the command on the system management node, specify a cluster name.

Information Message

[INFO] PJM 1400 pmresume Accepted job *jobid*.

Meaning

The request to resume the specified job was accepted.

Action

No action is necessary.

4.1.34 pmrscadm command

[NAME]

pmrscadm - Command that operates configuration file for the job operation manager.

[SYNOPSIS]

```
/usr/sbin/pmrscadm --set [--force] [--rscunit|--ru] rscuname  
/usr/sbin/pmrscadm --set -c clstname --rscunit rscuname -f nodelist [--force]  
/usr/sbin/pmrscadm --show [--rscunit|--ru] rscuname  
/usr/sbin/pmrscadm --check [--file filename] [--rscunit|--ru] rscuname  
/usr/sbin/pmrscadm --help
```

[DESCRIPTION]

The pmrscadm command is a command to operate the configuration file for the job operation manager by the option. This command is for system administrator and job operation administrator. It can be used only in the system management node.

[OPTIONS]

--set

The content of the configuration file of the resource unit specified in *rscuname* is reflected in all nodes in the resource unit.

When executed without the --force option specified, the command inquires whether to continue execution (y/n).

If reflection of the configuration file fails, the node ID and the reason for the failure are output to an error file.

This command is executed only on the nodes in the "os-running" or "on" power state. For nodes not included in the execution targets, the node ID and the reason for the exclusion are output to the error file.

The error file is output to \$HOME/pmrscadm_<cluster name>.<command execution date and time>.<process ID>_failed.

--force

The command is compulsorily executed without hearing right or wrong of execution.

--show

The content of a present configuration file of the resource unit corresponding to *rscuname* is displayed.

When the default value exists in the item not described in the configuration file, an item concerned is displayed.

When a past configuration file is displayed, number (*idx*) of the history is specified.

--check

The comparison of the confirmation of a confirmation of the format and an actual composition is checked against the content of a configuration file now at the time of being in the place of default, and the result is displayed.

When you check the comparison of a past configuration file, number (*idx*) of the history and the resource unit name (*rscuname*) is specified.

--file *filename*

The file name to be checked is specified.

{--rscunit|--ru} *rscuname*

The resource unit name of the object is specified.

-c *clstname*

This option is specified with the -f option, and specify the cluster name that reflects the settings.

-f *odelist*

To describe the node ID for which you want to reflect the settings on specified *odelist* file.

Ignore and continue processing that nonexistent node ID line in which a character string other than node ID is described.

EXAMPLE

```
0x01010010
0x01010011
0x01010012
```

--help

The format on this command are displayed.

When this option is specified, the argument and other options are all disregarded.

[EXAMPLE]

The content of the configuration file is reflected in all nodes under the control of the resource unit.

```
# pmrscadm --set --rscunit runit1
[WARNING]
Do you really want to continue (y/n)? y
[INFO] PRM 2202 pmrscadm The processing of the configuration file was completed
```

[END STATUS]

The end status is returned as follows:

0: Normal end

Other than 0: Abnormal end

[SEE ALSO]

pmrsc.conf(5), parscaadm(8)

[MESSAGES]

If an error occurs during the execution of the pmrscadm command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PRM 2401 pmrscadm Duplicated option: *opt*

Meaning

The option *opt* is specified more than once.

opt: Option name

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 2402 pmrscadm Option error

Meaning

The specified option is invalid.

Action

Review the option, and execute the command again.

[ERR] PRM 2403 pmrscadm Invalid cluster name: *clstname*

Meaning

Incorrect cluster name *clstname* specified.

clstname: Cluster name

Action

Specify the correct cluster name, and execute the command again.

[ERR.] PRM 2404 pmrscadm Not a file: *fname*

Meaning

The file specified by *fname* is not a file.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 2408 pmrscadm Not found file: *fname*

Meaning

The file specified by *fname* does not exist.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 2409 pmrscadm Cannot be used on this node

Meaning

The pmrscadm command cannot be used on this node.

Action

Execute the command on the system management node.

[ERR.] PRM 2412 pmrscadm The cluster other than compute cluster is specified: *clstname*

Meaning

The specified cluster name *clstname* is not for a compute cluster.

Action

Specify the cluster name of a compute cluster, and execute the command again.

[ERR.] PRM 2413 pmrscadm There are no contents of the file: *fname*

Meaning

The file contains nothing.

fname: File name

Action

Specify the correct file name, and execute the command again.

[ERR.] PRM 2414 pmrscadm Read conffile failed: *fname*(*reason*)

Meaning

The configuration file specified by *fname* contains a syntax error.

fname: File name

reason: Error reason

Action

Take the corrective action described in *reason*, and execute the command again.

[ERR.] PRM 2415 pmrscadm item name *itemname* is invalid

Meaning

The item name or section name specified by *itemname* is incorrect.

itemname: Item name or section name

Action

Specify the correct item name or section name, and execute the command again.

[ERR.] PRM 2416 pmrscadm *itemname* is duplicate

Meaning

The item name or section name specified by *itemname* has been specified repeatedly.

itemname: Item name or section name

Action

Delete either of the repeated items or sections, and execute the command again.

[ERR.] PRM 2417 pmrscadm *itemname* is not found

Meaning

The *itemname* item must be specified but has not been specified.

itemname: Item name

Action

Add the *itemname* item, and execute the command again.

[ERR.] PRM 2418 pmrscadm item value *itemname* is invalid

Meaning

The value specified by *itemname* cannot be set.

itemname: Item name

Action

Correct the item value specified by *itemname*, and execute the command again.

[ERR.] PRM 2419 pmrscadm No adjustment of the value between items(*itemname1* *itemname2*)

Meaning

The values specified by *itemname1* and *itemname2* are not consistent with each other.

itemname1: Compared item name
itemname2: Compared item name

Action

Correct the item values specified by *itemname1* and *itemname2*, and execute the command again.

[ERR.] PRM 2425 pmrscadm Connection error: *clstname*

Meaning

Operation of the configuration file failed because communication was not possible with the active compute cluster management node in the cluster *clstname*.

Action

Check the state of the compute cluster management node in the cluster *clstname* by using the pashowclst command. If the error occurs even though the compute cluster management node is in the normal state, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PRM 2426 pmrscadm Failed in the distribution of the script

Meaning

Distribution of the script failed.

Action

After starting the node described in the error file, try again to reflect the configuration file.

[ERR.] PRM 2427 pmrscadm Setting of job memory failed: *filename*

Meaning

The memory for jobs cannot be set for the BoB described in the file.

filename: File name

Action

Failed to set memory for jobs for target boot group. Execute the pmrscadm command again with the --set option. If the problem persists, contact a custom engineer (CE) to verify that the BMC configuration is correct.

After the problem is resolved, set the *odelist* specified with the -f option of the pmrscadm command to the node IDs of all nodes in BoB that failed to be set and execute the command again.

[ERR.] PRM 2430 pmrscadm Invalid rscunit name: *rscuname*

Meaning

The specified resource unit name *rscuname* is incorrect.

rscuname: Resource unit name

Action

Specify the correct resource unit name, and execute the command again.

[ERR.] PRM 2433 pmrscadm User ID, group ID or permission of the script *scriptname* is wrong

Meaning

The user ID, group ID, or access permission for the script *scriptname* is invalid.

Action

Confirm that the owner of the script *scriptname* is a user with job operation administrator privileges or higher and that the access permission is 0660.

[ERR.] PRM 2434 pmrscadm script file does not exist: *path***Meaning**

The script does not exist in *path*.

Action

Place the script in *path* and execute the command again.

[ERR.] PRM 2435 pmrscadm Too many ExitFunc defined.**Meaning**

The number of the ExitFunc item of the job resource manager exit function exceeds 100.

Action

Reduce the number of the ExitFunc item of the job resource manager exit function to 100 or less.

[ERR.] PRM 2436 pmrscadm It failed in ExitFunc file access: *path***Meaning**

Access under *path* failed.

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PRM 2440 pmrscadm Because the job operation has not stopped, the configuration file cannot be reflected.**Meaning**

The configuration file cannot be reflected because job operations have not stopped.

Action

Stop job operations and make sure that there are no running jobs in the target cluster. Then, execute the command again. For details on how to reflect the configuration file, see "Job Operation Software Administrator's Guide for Job Management."

[ERR.] PRM 2450 pmrscadm Internal error (*details*)**Meaning**

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 2460 pmrscadm No memory (*details*)**Meaning**

Memory acquisition failed.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 2461 pmrscadm Not found ExitFuncScriptDir directory: *path***Meaning**

The specified directory *path* does not exist in the ExitFuncScriptDir item of the job resource manager exit function in the pmrsc.conf file.

Action

In the ExitFuncScriptDir item, specify the directory containing the job resource manager exit script.

[ERR.] PRM 2462 pmrscadm Not a directory path: *path***Meaning**

The *path* path specified in the ExitFuncScriptDir item of the job resource manager exit function in the pmrsc.conf file is not a directory.

Action

Specify a directory in the ExitFuncScriptDir item.

[ERR.] PRM 2463 pmrscadm Resource Group *rscgname* does not exist**Meaning**

The *rscgname* name specified in the ResourceGroupName item is not defined in the pmpjm.conf file.

rscgname : Resource group name

Action

Specify a resource group name defined in the pmpjm.conf file.

[ERR.] PRM 2464 pmrscadm Too long ExitFuncScriptDir directory path: *path***Meaning**

The path name under the directory specified in the ExitFuncScriptDir item of the job resource manager exit function in the pmpjm.conf file is too long. (The path name combines the directory name and the exit script name.)

Action

With up to 2,048 characters, specify a path name combining the directory name and exit script name.

[ERR.] PRM 2799 pmrscadm Internal error (*details*)**Meaning**

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

Warning Message

[WARN] PRM 2301 pmrscadm item name *itemname* is invalid**Meaning**

The item name or section name specified by *itemname* is incorrect.

itemname: Item name or section name

Action

Specify the correct item name or section name, and execute the command again.

[WARN] PRM 2302 pmrscadm *itemname* is duplicate**Meaning**

The item name or section name specified by *itemname* has been specified repeatedly.

itemname: Item name or section name

Action

Delete either of the repeated items or sections, and execute the command again.

[WARN] PRM 2303 pmrscadm *itemname* is not found**Meaning**

The *itemname* item must be specified but has not been specified.

itemname: Item name

Action

Add the *itemname* item, and execute the command again.

[WARN] PRM 2304 pmrscadm item value *itemname* is invalid**Meaning**

The value specified by *itemname* cannot be set.

itemname: Item name

Action

Correct the item value specified by *itemname*, and execute the command again.

[WARN] PRM 2305 pmrscadm No adjustment of the value between items(*itemname1* *itemname2*)**Meaning**

The values specified by *itemname1* and *itemname2* are not consistent with each other.

itemname1: Compared item name

itemname2: Compared item name

Action

Correct the item values specified by *itemname1* and *itemname2*, and execute the command again.

[WARN] PRM 2306 pmrscadm User ID, group ID or permission of the script *scriptname* is wrong**Meaning**

The user ID, group ID, or access permission for the script *scriptname* is invalid.

Action

Confirm that the owner of the script *scriptname* is a user with job operation administrator privileges or higher and that the access permission is 0660.

[WARN] PRM 2307 pmrscadm script file does not exist: *path***Meaning**

The script does not exist in *path*.

Action

Place the script in *path* and execute the command again.

[WARN] PRM 2308 pmrscadm Not a file: *fname***Meaning**

The file specified by *fname* is not a file.

Action

Specify the correct file name, and execute the command again.

[WARN] PRM 2309 pmrscadm There are failed nodes in reflection of configuration file: *fname***Meaning**

The configuration file cannot be reflected on the node described in *fname*.

fname: File name

Action

After starting the node described in *fname*, try again to reflect the configuration file.

[WARN] PRM 2311 pmrscadm There are nodes that failed in reflection of configuration file(Internal error): *fname***Meaning**

The configuration file cannot be reflected to the node described in *fname* due to an internal error.

fname: File name

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

Information Message

[INFO] PRM 2201 pmrscadm You need to reboot nodes to reflect the setting. Execute pashowclst with soft-maintenance mode for rebooting nodes to which the setting should be reflected**Meaning**

Restart the node to reflect the configuration file.

Action

Restart the compute cluster management node, compute cluster sub management node, and all compute nodes of the resource unit in the cluster to which the resource unit belongs. For details on how to restart, see "Shifting to the Maintenance Mode and Its Recovery" in "Details of the System Management Function" in "Job Operation Software Administrator's Guide System Administration."

[INFO] PRM 2202 pmrscadm The processing of the configuration file was completed**Meaning**

pmrscadm command ended normally.

Action

No action is necessary.

[INFO] PRM 2203 pmrscadm Configuration file processing not executed**Meaning**

The parscadm command was not executed.

Action

No action is necessary.

4.1.35 pmscatter command

[NAME]

pmscatter - Runs a batch process to deliver files

[SYNOPSIS]

```
/usr/sbin/pmscatter [ -c clstname ]
                    [ scope options ]
                    [ --nodetype nodetype[,...] ]
                    [ --excludetype nodetype[,...] ]
                    [ --model model[,...] ]
                    [ -P num:group[,num:group,...] ]
                    [ -p ]
                    [ -r ]
                    [ --active ]
                    [ --ssh-o ssh-option ]
                    [ --force-all ]
                    [ -d dirname ] [ --filename name ]
                    [ --direct ]
                    srcfile[,...] dstfile

/usr/sbin/pmscatter --help

scope options (exclusive):
    --nodegrp nodegid[,...]
    --bootgrp bootgid[,...]
    --rscunit rscuname[,...]
    -n nodeid[,...] [ --giogrp | --siogrp ]
    -f filename
```

[DESCRIPTION]

pmscatter is a command which runs a batch process to deliver *srcfile* to a specified node or group.

If a command fails to deliver files to a node, an error file will be created in the execution user's home directory in order to describe the node and the cause of the failure.

This command is executed on the nodes in the "os-running" or "on" power state. For nodes not included in the execution targets, the output is to the execution exclusion file.

For the targets that are the FX server, the command refers to the power state of the compute node serving as the boot I/O node in the same BoB, not of the target nodes.

Name for error file: pmscatter_failed_command execution date and time.process ID

Name for execution exclusion file: pmscatter_excluded_command execution date and time.process ID

This command executes processing for the FX server and compute nodes of PRIMERGY server through the compute cluster sub management node. In the node groups where the compute cluster sub management node has not started, the processing does not go through the compute cluster sub management node.

This command can be executed on the system management node.

[ARGUMENTS]

srcfile[,...]

This argument specifies the name of a delivered source file or directory with full path.

Plurals can be specified by delimiting it by the comma. When delivered file name or two or more directory names are specified for *srcfile*, it delivers it in order specifying it.

The file is not processed for the file name that *srcfile* is illegal.

dstfile

This argument specifies the name of a delivered destination file, path or directory in full path.

When the file *dstfile* exists, it is overwritten by *srcfile*.

[OPTIONS]

-c *clstname*

This option uses *clstname* to specify the cluster name of a delivered object.

If this option is not specified, the cluster name set by the environment variable PXMYCLST is used.

--nodetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to set only a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS).

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--excludetype *nodetype*[,...]

This option uses *nodetype* to specify a node type to do not set a specific node type in a cluster as an executed target.

You can specify multiple node types by separating them with a comma (e.g., CCM,CCS). Even if the nodes is specified for the execution target by the --nodetype, the exclusion specification of this option is given to priority.

The node type that can be specified is similar to the --nodetype.

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--model *model*[,...]

If you want to narrow down the execution targets to a specific model in the cluster, specify the model in *model*.

You can specify multiple models by delimiting them with a comma (e.g., FT,PG).

You can specify the following models and arbitrary models.

Abbreviation	Model
PG	PRIMERGY server
FT	FX server

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--nodegrp *nodegid*[,...]

This option uses *nodegid* to specify a node group ID to set only a specific node group in a cluster as an executed object.

nodegid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x01,0x02,0x03).

You can also specify a range with a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--bootgrp *bootgid*[,...]

This option uses *bootgid* to specify a boot group ID to set only a specific boot group in a cluster as an executed object.

bootgid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x0101,0x0102,0x0103).

You can also specify a range with a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-n *nodeid*[,...]

This option uses *nodeid* to specify a node ID to set only a specific node in a cluster as an executed object.

nodeid must be the hexadecimal number.

You can specify multiple IDs by separating them with a comma (e.g., 0x01010010,0x01010012,0x0101002A).

You can also specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--giogrp

If you want to narrow down the execution targets to only a specific GIO group in the cluster, specify this option as a sub option of the -n option.

The unit of execution for FX server is a GIO group, and nodes in the same GIO group as the node specified in the -n option are gathered in the execution targets.

For nodes that are not FX server, the --giogrp option specification is ignored.

--siogrp

If you want to narrow down the execution targets to only a specific SIO group in the cluster, specify this option as a sub option of the -n option.

The unit of execution for FX server is an SIO group, and nodes in the same SIO group as the node specified in the -n option are gathered in the execution targets.

For nodes that are not FX server, the --siogrp option specification is ignored.

-f *filename*

This option executes a process for a node ID written in the file that is specified with *filename*.

Each of the node IDs is written on one line of the node list by the hexadecimal number.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

[Example]

```
0x01010010
0x01010011
0x0101002A
```

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

--rscunit *rscuname*[,...]

This option uses *rscuname* to specify a resource unit name to set only a specific resource unit in a cluster as an executed object.

You can specify multiple resource unit name by separating them with a comma (e.g., unitA,unitB).

When specifying this option, be sure to specify a cluster name in *clstname* with the -c option, or specify a specific cluster name in the environment variable PXMYCLST.

-P *num:group*[,*num:group*,...] | --concurrency *num:group*[,*num:group*,...]

The number of nodes processed at one time is specified in *num:group*. You can specify multiple *num:group* entries by delimiting them with a comma (,).

If this option is not specified, all the nodes that are the processing target are processed in parallel.

For *num*, specify the parallel processing quantity. For *group*, specify the parallel processing quantity you want to change in the unit of

a group.

You can specify any of the following values.

Unit	Parallel processing quantity
nodegrp	The minimum value is 1. The maximum value is the number of node groups that exist in the specified cluster.
node	The minimum value is 1. The maximum value is the maximum number of nodes that exist in the specified cluster.

-p

This option maintains the final revise time, the final access time, and the permission of a copied source file.

-r

This option recursively copies an entire directory.

--active

If the redundant node is specified by you want to the execution targets to the active node, specify this option.

When the execution target node which is the redundant configuration is specified by the unit of the node (-n or -f option) or it is not an active node, it does not make it the execution target.

When this option is specified independently, all active nodes in a cluster are made applicable to execution.

--ssh-o *ssh-option*

Specify *ssh-option* in this option when you want an option in the same format as in the ssh command configuration file.

Specify this option in the same format as the -o option of the ssh command.

The following options can be specified.

ConnectTimeout

ConnectionAttempts

--force-all

Specify this option when you want to include nodes that are not in the "os-running" or "on" node state.

Regardless of the state, however, multiuse nodes with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pxsystem.conf are always excluded from the execution targets. Their results are output to the execution exclusion file.

-d *dirname*

This option uses *dirname* to store an execution result file in a specified directory.

If this option is not set, the execution result file will be output to the execution user's home directory.

--filename *name*

The name of the execution result file changes to the name specified in the argument *name*, and the "command execution date.process ID" part is omitted.

The following shows the output files when this option is specified.

Error file: name_failed

Execution exclusion file: name_excluded

--direct

Processing is executed without going through the compute cluster sub management node.

--help

This option displays information on how to use this command.

If this option is specified, all the arguments and other options are ignored.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option. Specification with the -c option has priority over that with the environment variable PXMYCLST. You can specify only one cluster name in the environment variable PXMYCLST. An example is shown below.

[Example]

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The end status is returned as follows:

Exit status	Meaning
0	Normal end Successful <i>srcfile</i> delivery on all target nodes (normal end even on execution-excluded nodes)
1	Abnormal end
2	Abnormal end on all target nodes Unable to deliver <i>srcfile</i> on all target nodes. Unable to connect via ssh, or delivery source/destination files not found.
3	Abnormal end on some nodes and normal end on others (including execution-excluded nodes) Mixture of delivery successes and failures

[NOTE]

- You cannot specify the --nodegrp, --bootgrp, -n, -f, and --rscunit options at the same time.
- The nodes that are delivery targets are narrowed down in processing according to the combination of the options. If no node is a delivery target, the processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option is invalid, processing ends abnormally with end status 1. If the model specified in the --model option is invalid, processing ends abnormally with end status 1.
- If the ID specified in the -n, --nodegrp, or --bootgrp option is invalid, processing ends abnormally with end status 1.
- If the node type specified in the --nodetype or --excludetype option does not exist in the cluster, processing ignores the node type. If the model specified in the --model option does not exist in the cluster, processing ignores that model. If the node ID specified in the -n option does not exist in the cluster, processing ignores that node.
- You can specify only a file name in the --filename option. To specify a directory as the output destination, specify the directory in the -d option
- SSH password-less login is not set on a system built without the install function. For this reason, the batch operation function cannot be used on that system. To use the batch operation function, you need to set SSH password-less login with the --ssh-keygen option of the pmexe command.
- In the execution target nodes, any multiuse node with the yes (enabled) setting for the PowerControlOnly setting item in the cluster configuration definition file pxsystem.conf is automatically excluded from the execution targets. So the command is not executed on the node. The execution results on the node are output to the execution exclusion file.

[MESSAGES]

The pmexe, pmscatter, and pmgather commands share common messages. See "[MESSAGES]" in ["4.1.28 pmexe command."](#)

Also, not only the messages of the pmexe, pmscatter, and pmgather commands but common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

4.1.36 pmsetjobinfo command

[NAME]

pmsetjobinfo - Command that sets job information from the resource management exit function

[SYNOPSIS]

```
/usr/sbin/pmsetjobinfo --type typename --value values  
/usr/sbin/pmsetjobinfo --help
```

[DESCRIPTION]

This command sets the information specified by options for job statistical information.
You can use this command by calling it from the resource management exit function.

[OPTIONS]

--type *typename*

In *typename*, specify the item you want to set for statistical information. You can specify the following:

reason: REASON item of job statistical information

Note that the item specified in reason is set only when the end code of the exit script indicates an error (cases 1 to 4). Otherwise, the setting is ignored.

--value *values*

Specify a value to set in job statistical information.

The value that can be specified in *values* has the following relationship with the value specified in --type:

reason: The number of characters that can be specified is 63 characters or less.

--help

This option displays a synopsis of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

An example of setting "Does not exist image file" in the REASON item of job statistical information is shown below.
The following is written in the resource management exit function and executed:

```
/usr/sbin/pmsetjobinfo --type reason --value "Does not exist image file"
```

[END STATUS]

The following end codes are returned.

0: Normal end

1: Abnormal end (option error)

255: Abnormal end (internal error)

[SEE ALSO]

pmrsc.conf(5), prealloc(5), predel(5), postfree(5)

[MESSAGES]

If an error occurs during the execution of the pmsetjobinfo command, the following messages is output to the standard error output.

Error Message

[ERR.] PRM 8201 pmsetjobinfo Cannot be used by this user

Meaning

The user does not have the privileges to execute the pmsetjobinfo command.

Action

Review the user's privileges to execute the command. Then, execute it again.

[ERR.] PRM 8202 pmsetjobinfo Cannot be used on this node

Meaning

The pmsetjobinfo command cannot be used on this node.

Action

Execute the command on the compute node.

[ERR.] PRM 8203 pmsetjobinfo Duplicated option: *opt*

Meaning

The option *opt* is specified more than once.

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 8204 pmsetjobinfo Invalid option: *opt*

Meaning

The specified option *opt* is invalid.

Action

Review the option, and execute the command again.

[ERR.] PRM 8205 pmsetjobinfo Invalid argument: *opt arg*

Meaning

The specified argument *arg* for the option *opt* is invalid.

Action

Review the argument for the option *opt*, and execute the command again.

[ERR.] PRM 8206 pmsetjobinfo Invalid combination of options

Meaning

The combination of specified options is invalid.

Action

Review the option, and execute the command again.

[ERR.] PRM 8207 pmsetjobinfo No memory(*details*)

Meaning

Memory acquisition failed.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 8208 pmsetjobinfo Cannot be used by calling program

Meaning

You can not use this command by calling it from other than the resource management exit function.

Action

Call it from the resource management exit function.

[ERR.] PRM 8299 pmsetjobinfo Internal error ([details](#))

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

4.1.37 pmsetstats command

[NAME]

pmsetstats - Command that sets job information from the resource management exit function

[SYNOPSIS]

```
/usr/sbin/pmsetstats --record record_name --item item_name --type value_type --value value  
/usr/sbin/pmsetstats --help
```

[DESCRIPTION]

This command sets the information specified by options for job statistical information.
You can use this command by calling it from the resource management exit function.
This command can be executed on the job master node or job slave node.
However, the job statistical information settings are only taken from the job master node.

[OPTIONS]

--record *record_name*

In *record_name*, specify the unit you want to set for statistical information. You can specify the following:

JI: Job statistical information
JN: Node statistical information

--item *item_name*

In *item_name*, specify the identification name you want to set for statistical information. You can specify the identification name that specified for ItemName in papjstats.conf.

--type *value_type*

In *value_type*, specify the type of values you want to set for statistical information. You can specify the following:

char, int8, uint8, int16, uint16, int32, uint32, int64, uint64, double, float, timespec, time, size, string

--value *value*

Specify a value to set in job statistical information. Specify it in 4,095 characters or less.

When specifying "timespec" in the --type option, specify seconds and nanoseconds by delimiting them with a decimal point (".").

--help

This option displays a synopsis of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

An example of setting "10" in the license-used item (int32 type) of job statistical information is shown below.
The following is written in the resource management exit function and executed:

```
/usr/sbin/pmsetstats --record JI --item license-used --type int32 --value 10
```

[END STATUS]

The following end codes are returned.

- 0: Normal end
- 1: Abnormal end (option error)
- 2: Abnormal end (memory shortage)
- 255: Abnormal end (internal error)

[NOTE]

This command does not check consistency with input arguments and papjstats.conf file.

For example, when undefined item is specified in the --item option or when the item specified in the --item option and the type of values specified in the --type option are not consistent with each other, this command is not an error.

[SEE ALSO]

pmrsc.conf(5), prealloc(5), predel(5), postfree(5), papjstats.conf(5), pmsetjobinfo(8)

[MESSAGES]

If an error occurs during the execution of the pmsetstats command, the following messages is output to the standard error output.

[ERR.] PRM 8101 pmsetstats Cannot be used by this user

Meaning

The user does not have the privileges to execute the pmsetstats command.

Action

Review the user's privileges to execute the command. Then, execute it again.

[ERR.] PRM 8102 pmsetstats Cannot be used on this node

Meaning

The pmsetstats command cannot be used on this node.

Action

Execute the command on the compute node.

[ERR.] PRM 8103 pmsetstats Duplicated option: *opt*

Meaning

The option *opt* is specified more than once.

Action

Specify the option correctly, and execute the command again.

[ERR.] PRM 8104 pmsetstats Invalid option: *opt*

Meaning

The specified option *opt* is invalid.

Action

Review the option, and execute the command again.

[ERR.] PRM 8105 pmsetstats Invalid argument: *opt arg*

Meaning

The specified argument *arg* for the option *opt* is invalid.

Action

Review the argument for the option *opt*, and execute the command again.

[ERR.] PRM 8106 pmsetstats Invalid combination of options

Meaning

The combination of specified options is invalid.

Action

Review the option, and execute the command again.

[ERR.] PRM 8107 pmsetstats No memory(*details*)

Meaning

Memory acquisition failed.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

[ERR.] PRM 8108 pmsetstats Cannot be used by calling program

Meaning

You can not use this command by calling it from other than the resource management exit function.

Action

Call it from the resource management exit function.

[ERR.] PRM 8199 pmsetstats Internal error (*details*)

Meaning

An internal error was detected.

details: Detail information for maintenance

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data together with the output message.

4.1.38 pmsuspend command

[NAME]

pmsuspend - Suspend a running job.

[SYNOPSIS]

```
/usr/sbin/pmsuspend [-c clstname] [-R|--reason reasonmessage] jobid [jobid ...]  
/usr/sbin/pmsuspend --help
```

[DESCRIPTION]

pmsuspend command suspend specified jobs or sub jobs.

Interactive jobs cannot be suspended.

If the specified job is a job that can not be suspended, an error message is output.

This command accepts only jobs or sub jobs in the RUNNING state. If some of the specified jobs or sub jobs that are not in the RUNNING state, this command outputs an error message and exits abnormally.

When this command is executed, the job manager daemon starts suspending operation, then the states of the specified jobs or sub jobs become SUSPEND. When the operation has completed, the states become SUSPENDED, and the process suspends execution.

This command returns immediately when the suspend operation has starts. If all operation object jobs or sub jobs are in SUSPENDED state, the operation has completed. The pjstat command can be used to confirm the states.

Only jobs in the RUNNING state can be specified in this command. If the specified job is a job in another state, an error message is output.

Resources of jobs suspended by this command are not released and can not be assigned to other jobs. Jobs or sub jobs suspended by this command will never transition until the job is resumed using the pmresume command.

For a step job or bulk job, the job ID or the sub job ID can be specified for *jobid*.

This command can be executed on the system management node and the compute cluster management node.

[Specifying multiple or a range of job IDs]

You can specify a job ID range by using a hyphen (-).

Example: To specify a range from job ID 1 to 100

1-100

When specifying a range, even if you include jobs such as those listed below within the range, no error occurs. Instead, those jobs are ignored, and processing continues.

- Job that does not exist
- Job in a state that cannot be specified

For cases like the above when multiple jobs are specified, an error message is output since the IDs are specified explicitly. (Processing is not aborted, and all specified jobs are processed.)

[OPTIONS]

-c *clstname*

This option specifies the target cluster name *clstname*. When the command is executed on the system management node and the environment parameter PXMYCLST has been set, then this option can be omitted.

When the command is executed on the node with other types, this option will be ignored.

-R|--reason *reasonmessage*

It outputs the suspend reason of the jobs or sub jobs to the REASON column of the pjstat command.

When the job is suspended normally, the specified string by this option is output in the REASON column of the pjstat command in the form "command execution user name: reasonmessage". The string must be shorter than 64 characters. The string can only contain single-byte alphabet and numeric character.

--help

This option displays the usage of this command.

If this option is specified, all the arguments and other options are ignored.

[EXAMPLE]

The following is a use example of the pmsuspend command:

```
<jmgr>$ pmsuspend -R XXXX 1  
[INFO] PJM 1300 pmsuspend Accepted job 1.
```

```
<jmgr>$ pmsuspend 2-3 4
[INFO] PJM 1300 pmsuspend Accepted job 2-3.
[INFO] PJM 1300 pmsuspend Accepted job 4.
```

```
<jmgr>$ pjstat -v 1-4
JOB_ID ... JOB_NAME MD ST USER ... REASON
1 ... jobname1 NM SPD user1 ... root:XXXX
2 ... jobname2 NM SPD user2 ... root:
3 ... jobname4 NM SPD user3 ... root:
4 ... jobname4 NM SPD user4 ... root:
```

[END STATUS]

0: Normal end

1: Abnormal end (If multiple job IDs are specified, an abnormal end is returned when even a single job ends with an error.)

[SEE ALSO]

pmresume(8)

[MESSAGES]

If an error occurs during the execution of the pmsuspend command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.

Error Message

[ERR.] PJM 1301 pmsuspend Unknown option *opt*.
Try `pmsuspend --help` for more information.

Meaning

The specified option *opt* is unrecognizable.

Action

Specify the correct option.

[ERR.] PJM 1303 pmsuspend Unknown option argument *arg*.
Try `pmsuspend --help` for more information.

Meaning

For the option, the specified argument *arg* is unknown, or no argument is specified.

Action

Specify the correct argument.

[ERR.] PJM 1304 pmsuspend Argument format error *arg*.

Meaning

The setting value of the argument *arg* in the option is incorrect.

Action

Specify the correct setting value.

[ERR.] PJM 1310 pmsuspend No jobid(s) specified.
Try `pmsuspend --help` for more information.

Meaning

No job ID is specified.

Action

Specify a job ID.

[ERR.] PJM 1311 pmsuspend Invalid jobid syntax *jobid*.

Meaning

The specified job ID *jobid* is incorrect.

Action

Specify the correct job ID.

[ERR.] PJM 1312 pmsuspend Job *jobid* does not exist.

Meaning

The specified job (job ID: *jobid*) does not exist.

Action

Confirm that the job is correctly specified.

[ERR.] PJM 1313 pmsuspend Job *jobid* state error.

Meaning

The command cannot be executed because the state of the specified job (job ID: *jobid*) is not RUNNING.

Action

Confirm whether the state of the job is RUNNING.

[ERR.] PJM 1314 pmsuspend Job *jobid* type error.

Meaning

The specified job is a type of job that cannot be suspended (interactive job).

Action

Batch jobs are the only type of job that can be suspended.

[ERR.] PJM 1315 pmsuspend Job *jobid* has already been accepted.

Meaning

The operation for the specified job (job ID: *jobid*) has already been accepted.

Action

The operation is ignored. No action is necessary.

[ERR.] PJM 1389 pmsuspend Not supported: *message*.

Meaning

The current version of Technical Computing Suite cannot suspend the job with the job ID specified in the option.

Action

When you want to release the resource that is used by the target job, end the job.

[ERR.] PJM 1390 pmsuspend PJM daemon is not present.

Meaning

The job manager function is not working, or communication with the job manager function is disabled.

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 1391 pmsuspend Internal error: *details*.

Meaning

An internal error was detected.

details: detail information for maintenance

Action

Check the operating state of the job manager function and the system state according to the "Job Operation Software Administrator's Guide for System Management."

[ERR.] PJM 1393 pmsuspend This command can be executed on SMM/CCM.

Meaning

Can only be executed on the system management node or compute cluster management nodes.

Action

Execute the pmsuspend command on the system management node or compute cluster management node.

[ERR.] PJM 1395 pmsuspend No cluster specified on SMM.

Meaning

No cluster name is specified.

Action

When executing the command on the system management node, specify a cluster name.

Information Message

[INFO] PJM 1300 pmsuspend Accepted job *jobid*.

Meaning

The request to suspend the specified job was accepted.

Action

No action is necessary.

4.1.39 pxinst command

[NAME]

pxinst - Installation function

[SYNOPSIS]

```
/usr/sbin/pxinst <sub-command>
                [--help]
```

<sub-command>

{init, repo, initrd, rootfs, svcs, kmod, node, status, sync, check}

[DESCRIPTION]

The `pxinst` command is a command for system administrators to configure settings for OS installation. This command can be used on the active system management node.

[OPTIONS]

`--help`

This option displays how to use this command.

[SUBCOMMAND]

<Initial installation function settings>

`init <action> [options]`

action:

all

This action configures all the settings at once.

dhcp

This action configures dhcp server settings.

tftp

This action configures tftp server settings.

ntp

This action configures ntp server settings.

http

This action configures http server settings.

ssh

This action creates an ssh key used by the installation function,

inst

This action configures the environment for operation as the installer node.

options:

`--help`

This option displays how to use this subcommand.

<Repository management function>

`repo <action> [options]`

action:

create -R *reponame* [-S *isoimage*] [--rootfs *rootfs*] [--enable]

This action creates a new repository.

list [-R *reponame*]

This action displays information on a created repository.

update -R *reponame*

This action updates a repository to the latest state.

pkgkit -R *reponame* --kit *kitpath*

This action registers an RPM package file with a repository. The file is included in a package kit.

delete -R *reponame* [--force]

This action deletes a repository.

options:

-R *reponame*, --repo *reponame*

Specify the name of the repository targeted for processing.

-S *isoimage*, --iso *isoimage*

Specify the absolute or relative path to the ISO image file of an OS.

--rootfs *rootfs*

Specify the absolute or relative path of the created *rootfs* to register with a repository.

Specify this option to automatically incorporate an existing FX server *rootfs* or existing *rootfs* created in a different system into the repository.

--kit *kitpath*

Specify the absolute or relative path to a package kit.

--enable

Specify this option to enable a repository for the command execution node.

--force

This option forcibly executes the command without displaying a y/n prompt asking to confirm the execution.

--help

This option displays how to use this subcommand.

<initrd creation function>

initrd <action> [options]

action:

create -R *reponame*

This action creates an initrd for the OS version corresponding to the specified repository.

options:

-R *reponame*, --repo *reponame*

Specify the name of the repository targeted for processing.

--help

This option displays how to use this subcommand.

<Root file system creation function>

rootfs <action> [options]

action:

create -R *reponame* [--use-image *rootfs*]

This action creates a new rootfs for the OS version corresponding to the specified repository.

addpkg -R *reponame* -P *pkgname* [,...]

This action adds arbitral packages to the rootfs for the OS version corresponding to the specified repository.

updpgk -R *reponame*

This action updates the package installed in the rootfs for the OS version corresponding to the specified repository.
It is equivalent to executing "yum upgrade" for the rootfs.

unpack -R *reponame* -d *dstdir*

This action unpacks the rootfs for the OS version corresponding to the specified repository, in the directory specified by the -d option.

If you want to implement arbitrary settings for the rootfs, use this function.

pack -R *reponame* -d *srcdir*

This action creates a rootfs for the OS version corresponding to the repository specified by the --repo option. "/" designates the directory specified by the -d option.

options:

-R *reponame*, --repo *reponame*

Specify the name of the repository targeted for processing.
A FX server repository cannot be specified.

-P *pkgname*[,...], --pkg *pkgname*[,...]

Specify the name of a package or group to add to the rootfs.
Specify the group name in the "@group name" format.
When specifying multiple names, delimit them with a comma.
Each specified package must be registered with the repository in advance.

-d *dstdir* | -d *srcdir*

Specify the directory targeted for processing.
If this option is specified together with the unpack action, an existing directory cannot be specified.
If this option is specified together with the pack action, only an existing directory can be specified.

--use-image *rootfs*

Specify the absolute or relative path to a created rootfs to register with the repository.
This option executes processing on the basis of a created rootfs, without creating a new rootfs.

The option must be specified for a repository with a different architecture than that of the installer node.
It is necessary to make or obtain rootfs specified beforehand.

--help

This option displays how to use this subcommand.

<ServerView Suite linkage function>

svs <action> [*options*]

action:

setup -R *reponame* [-S *isoimage*]

This action incorporates the kernel module and bundled software that are provided by ServerView.

list -R *reponame*

This action displays a list of the applicable ServerView bundled software.

options:

-R *reponame*, --repo *reponame*

Specify the name of the repository targeted for processing.

-s isoimage, --iso isoimage

Specify the absolute or relative path to the ISO image file of ServerView Suite DVD1.

--help

This option displays how to use this subcommand.

<Kernel module incorporation function>

kmod <action> [options]

action:

add -R reponame --rpm rpmfile[,...]

This action incorporates the kernel module into the initrd.

del -R reponame --mod modname[,...]

This action deletes the kernel module from the initrd.

options:

-R reponame, --repo reponame

Specify the name of the repository targeted for processing.

--rpm rpmfile[,...]

Specify the absolute or relative path to the RPM package file.

When specifying multiple files, delimit their paths with a comma.

If the RPM package file does not exist, the command ends abnormally.

--mod modname[,...]

Specify the name of the kernel module to be deleted.

If you want to delete "mod1.ko", specify "mod1" with ".ko" excluded.

When specifying multiple names, delimit them with a comma.

--help

This option displays how to use this subcommand.

<Node information registration function>

node <action> [options]

action:

add -R reponame [-T {img | pkg | disk}]
[-N nodedef [-P partdef]]
[--svs no[,...]]
[--bootopt param]
[--no-svs]
[--bios]

This action registers PRIMERGY server node information. If the same host name is already registered, the settings are failed.

add -R reponame -N nodedef --ft
[--pool startaddr/mask-endaddr/mask]

This action registers FX server node information. If the same host name is already registered, the settings are failed.

del {-m hostname[,...] | -f hostfile}

This action deletes registered node information.

If the specified node is not registered, a warning message is output and processing continues.

update -T {img | pkg | disk}

{-m hostname[,...] | -f hostfile}

This action changes the operation at the next node startup time when reconfiguring a node where the OS is already installed. The PXE configuration file corresponding to the node is updated.

You cannot specify pkg for FX server node.

In the case of host names or IP addresses with consecutive numbers, it is possible to specify the range.

(Example: If you want to specify the nodes from host001 to host100, specify host[001-100])

If the specified node is not registered, a warning message is output and processing continues.

options:

-R reponame, --repo reponame

Specify the name of the repository targeted for processing.

-T {img | pkg | disk}, --type {img | pkg | disk}

Specify a provisioning method. You can specify the following values.

If this option is omitted, img is assumed specified.

Specifiable value	Description
pkg	Package-based installation
img	Image-based installation
disk	Disk boot (*)

(*) If disk is specified, the disk boot is done without running installation processing at node startup.

Specify this value to control installation processing to temporarily stop running.

In this situation, the provisioning method needs to be changed by the update action before installation.

-N nodedef, --nodedef nodedef

Specify the absolute or relative path to the node information definition file or the FX server node information definition file.

If img or pkg is specified, this option is required when "node add" is specified.

Since the FX server node information definition file is divided into multiple files, you can also specify a directory.

-P partdef, --partdef partdef

Specify the absolute or relative path to the file that defines disk partition information.

If img or pkg is specified, this option is required when "node add" is specified.

This option is available only with the PRIMERGY server.

--svs no[,...]

Specify the software numbers, delimited with a comma, of the ServerView bundled software to apply to the node.

If this option is omitted, the bundled software is not applied.

This option is available only with the PRIMERGY server.

--bootopt param

Specify an OS boot option.

Specify this option to, for example, set "--bootopt console=ttyS0" to enable remote console connection.

--ft

Specify this option to register a FX server node.

-m hostname[,...], --hostname hostname[,...]

Specify the host name or IP address for provisioning to the node targeted for processing.

When specifying multiple host names or IP addresses, delimit them with a comma.

You can specify host names or IP addresses with sequential numbers as a range.

(Example: To process host001 to host100, specify host[001-100].)

If a specified host name is not registered, a warning message is output and processing continues.

-f *hostfile*

Specify the file that contains the host name or IP address of the node targeted for processing.
Write one node per line.
(Example)

```
host001
host002
host003
```

--pool *startaddr/mask-endaddr/mask*

Specify an IP address pool (range of assigned IP addresses) to use dynamic IP addresses for FX server installation.
Use the "start IP address/mask bits - end IP address/mask bits" format.
The start mask bits must match the end mask bits.

(Example)

To dynamically assign 192.0.2.10/24 to 192.0.2.225/24
--pool 192.0.2.10/24-192.0.2.225/24

--no-svs

Specify this option to use an OS version or hardware that is not supported by ServerView Suite.
If this option is specified, the kernel module and bundled software that are provided by ServerView Suite are not applied.

--bios

Specify this option for startup in BIOS mode.
If BIOS mode is set for hardware, specify this option.
If the option is omitted, startup in UEFI mode is assumed.

--help

This option displays how to use this subcommand.

<Installation completion confirmation function>

status <action> [*options*]

action:

show

This action displays the installation completion status of nodes.
The following example shows the display.

```
HOSTNAME OPERATION STATUS STATE DETAIL
-----
host01 ImgProvision Completed Succeeded -
host02 ImgProvision InProgress Installing -
host03 ImgProvision Completed Failed Detailed information
```

Item name	Description
HOSTNAME	Host name of the target node for installation
OPERATION	Type of operation PkgProvision: Package-based OS installation ImgProvision: Image-based OS installation
STATUS	Operation status NotStarted: Operation not yet started InProgress: Operation in progress Completed: Operation completed
STATE	Processing status Waiting: Waiting for node startup to beginning of installation processing

Item name	Description
	Preparing: Preparing for installation processing Partitioning: Creating disk partition Installing: Installing OS Posting: Post-installation processing in progress Rebooting: Restarting Succeeded: Normal end Failed: Abnormal end PostFailed: post processing failed
DETAIL	Detailed information. If STATE is neither Failed nor PostFailed, a hyphen (-) is displayed. The displayed information on the cause of an error consists of up to 255 characters.

clear [-s status] | [-m hostname[,...]] | -f hostfile]

This action clears the installation completion status of a node so that the status is no longer displayed.
 By default, the only information that can be cleared is information about nodes where installation has not begun or has completed successfully.
 Information about nodes where installation is in progress or has failed is not cleared.
 To clear information about nodes where installation is in progress or has failed, specify the --hostname option or -f option.

options:

-s status, --status status

Specify the STATUS targeted for clearing.
 All node information of the specified STATUS is no longer displayed.
 You can specify NotStarted or Completed for the STATUS.
 If Completed is specified, only the node information in the succeeded STATE is cleared.

-m hostname[,...], --hostname hostname[,...]

Specify the host name of the node targeted for processing.
 When specifying multiple host names, delimit them with a comma.
 You can specify host names or IP addresses with sequential numbers as a range.
 (Example: To process host001 to host100, specify host[001-100].)
 If this option is specified, information is cleared regardless of the STATUS or STATE.
 If status information is not registered for the specified host name, a warning message is output and processing continues.

-f hostfile

Specify the file that contains the host names of the nodes targeted for processing.
 Write one node per line.
 If this option is specified, information is cleared regardless of the STATUS or STATE.
 If status information is not registered for the specified host name, a warning message is output and processing continues.
 (Example)

```
host001
host002
host003
```

--help

This option displays how to use this subcommand.

<Replication function>

sync <action> [options]

action:

all [-m hostname[,...]] | -f hostfile]

This action synchronizes all of the installation function settings.
 It is equivalent to executing both repo and srv.

repo [-R *reponame*] [-m *hostname*[,...] | -f *hostfile*]

This action synchronizes a repository.

srv [-m *hostname*[,...] | -f *hostfile*]

This action synchronizes the server settings required by the installation function.

options:

-R *reponame*, --repo *reponame*

Specify the name of the repository targeted for processing.

-m *hostname*[,...], --hostname *hostname*[,...]

Specify the host name or IP address for provisioning to the node targeted for processing.

When specifying multiple host names or IP addresses, delimit them with a comma.

You can specify host names or IP addresses with sequential numbers as a range.

(Example: To process host001 to host100, specify host[001-100].)

If the specified host names include an unregistered node, the command ends abnormally.

-f *hostfile*

Specify the file that contains the host names or IP addresses for provisioning to the nodes targeted for processing.

Write one node per line.

If the specified host names include an unregistered node, the command ends abnormally.

(Example)

host001

host002

host003

--help

This option displays how to use this subcommand.

<Service status check function>

check [*options*]

action:

check [-m *hostname*[,...] | -f *hostfile*]

This action checks the status of services for the install function to operate.

options:

-m *hostname*[,...], --hostname *hostname*[,...]

Specify the host name or IP address for provisioning to the node targeted for processing.

When specifying multiple host names or IP addresses, delimit them with a comma.

You can specify host names or IP addresses with sequential numbers as a range.

(Example: To process host001 to host100, specify host[001-100].)

If the specified host names include an unregistered node, the command ends abnormally.

-f *hostfile*

Specify the file that contains the host names or IP addresses for provisioning to the nodes targeted for processing.

Write one node per line.

If the specified host names include an unregistered node, the command ends abnormally.

(Example)

host001

host002
host003

--help

This option displays how to use this subcommand.

[EXIT STATUS]

The following end codes are returned.

0: Normal end
1: Abnormal end

[SEE ALSO]

pxinst.conf(5)

[MESSAGES]

If an error occurs during the execution of the pxinst command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.2 Common command messages of the install function."](#)

Error Message

[ERR.] INST 1000 pxinst Cannot find configuration file.

Meaning

The install function does not have an environment configuration file.

Action

Place the environment configuration file of the install function at `"/etc/opt/FJSVtcs/inst/pxinst.conf,"` and execute the command again.

[ERR.] INST 1001 pxinst File format is not valid.(*details*)

Meaning

The format of the environment configuration file of the install function is incorrect.

details: Cause of the error

Action

Based on the yaml syntax error information displayed in *details*, review the environment configuration file of the install function, and execute the command again.

[ERR.] INST 1002 pxinst Invalid parameter.(key=*KEY*,value=*VALUE*)

Meaning

The *VALUE* value for the *KEY* key in the environment configuration file of the install function is incorrect.

Action

Review the *VALUE* value, and execute the command again.

[ERR.] INST 1003 pxinst *KEY* is required.

Meaning

KEY must be specified.

Action

Specify *KEY* in the environment configuration file of the install function, and execute the command again.

[ERR.] INST 1004 pxinst *KEY* is duplicated.

Meaning

KEY is specified multiple times.

Action

Specify *KEY* in the environment configuration file of the install function, and execute the command again.

[ERR.] INST 1005 pxinst The specified repository is already exists.

Meaning

The specified repository already exists.

Action

If the specified repository name is incorrect, review the repository name, and execute the command again. If the specified repository name is correct, delete the repository, and then execute the command again.

[ERR.] INST 1006 pxinst Invalid repository name.

Meaning

The specified repository name is incorrect.

Action

Review the specification of the repository, and execute the command again.

[ERR.] INST 1007 pxinst An installation media is not found.

Meaning

The OS installation media was not found.

Action

Confirm that the OS installation media is inserted in the active system management node, and execute the command again.

[ERR.] INST 1008 pxinst The specified iso image is not found.

Meaning

The specified ISO image was not found.

Action

Review the specification of the ISO image, and execute the command again.

[ERR.] INST 1009 pxinst The specified iso image is not supported OS.

Meaning

The specified ISO image is an ISO image of an unsupported OS. Alternatively, the specified image is not an ISO image of an OS.

Action

Specify an ISO image of a supported OS, and execute the command again.

[ERR.] INST 1010 pxinst Failed to create repository.(*details*)

Meaning

Creation of the repository failed.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 1011 pxinst The specified repository is not found.

Meaning

The specified repository was not found.

Action

Review the specification of the repository, and execute the command again.

[ERR.] INST 1012 pxinst Failed to update repository.(details)

Meaning

Repository update failed.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 1013 pxinst The specified file is not package-kit.

Meaning

The file specified in a command option is not a package kit.

Action

Review the specification of the command option, and execute the command again.

[ERR.] INST 1014 pxinst Failed to add package-kit.(details)

Meaning

Registration of the package kit failed.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 1015 pxinst Failed to delete repository.(details)

Meaning

Deletion of the repository failed.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 1016 pxinst A ServerView Suite media is not found.

Meaning

The ServerView Suite media was not found.

Action

Confirm that the ServerView Suite media is inserted in the active system management node, and execute the command again. The media is supplied with the hardware.

[ERR.] INST 1017 pxinst The specified iso image is not found.

Meaning

The specified ISO image was not found.

Action

Review the specification of the ISO image, and execute the command again.

[ERR.] INST 1018 pxinst The specified iso image or disk is not valid.

Meaning

The specified media or ISO image is not the media or ISO image of ServerView Suite.

Action

If using the media, check the contents of the media inserted in the command execution node, and execute the command again. If using the ISO image, review the specification of the ISO image, and execute the command again.

[ERR.] INST 1019 pxinst Invalid ServerView software number.

Meaning

The specified bundled ServerView software number is incorrect.

Action

Check the bundled ServerView software number specified with the pxinst svcs list command, and execute the command again.

[ERR.] INST 1020 pxinst The file *filename* is not a valid RPM package.

Meaning

filename is not a valid RPM package file.

Action

Review the contents of the specified option, and execute the command again.

[ERR.] INST 1022 pxinst The Node definition file is not found.

Meaning

The node information definition file was not found.

Action

Review the contents of the specified --nodedef option, execute the command again.

[ERR.] INST 1023 pxinst The disk partition definition file is not found.

Meaning

The disk partition information definition file was not found.

Action

Review the contents of the specified --partdef option, execute the command again.

[ERR.] INST 1024 pxinst The rootfs is not found.

Meaning

rootfs (root file system) was not found.

Action

Using the function for creating the root file system, create a root file system corresponding to the repository used. Then, execute the install process again.

[ERR.] INST 1025 pxinst Connection error occurred.(*IP address*)

Meaning

Processing stopped because a connection to the node of displayed IP address could not be established.

Action

Check the node operation status displayed with the IP address, remove the cause, and then execute the installation processing again.

[ERR.] INST 1026 pxinst Specified repository does not support this command.

Meaning

The specified repository does not support this command.

Action

Review the repository name. If necessary, execute the command again.

[ERR.] INST 1027 pxinst The "updates.img" is not found.

Meaning

The updates.img corresponding to the specified repository was not found.

Action

Execute the pxinst svcs setup command, and create updates.img. Then, execute the command again.

[ERR.] INST 1028 pxinst The RPM package file does not found.

Meaning

The specified RPM package file was not found.

Action

Review the contents of the specified option, and execute the command again.

[ERR.] INST 1029 pxinst The specified host name is not registered.(*hostname*)

Meaning

The specified host name *hostname* is not registered.

Action

Review the specified host name, and execute the command again.

[ERR.] INST 1030 pxinst hostname must be either installer node, relay node, or "BIO".

Meaning

The host name *hostname* must be for any of the following: installer node, relay node, and boot I/O (BIO) node.

Action

Review the specified host name. If necessary, execute the command again.

[ERR.] INST 1031 pxinst The IP address range specified with the pool option has already been registered. (range)

Meaning

range specified in the --pool option includes an already registered IP address. Multiple pools that include the same IP address cannot be registered.

Action

Review the settings in dhcpd.conf and the specification of the --pool option. If necessary, execute the command again.

[ERR.] INST 1032 pxinst service *service* restart failed.

Meaning

Deletion of the service *service* failed.

Action

Remove the cause, referring to the detailed message output immediately after this message and also to the service status display command and service log file of the OS used. Then, start the service. If the cause is unknown, collect investigation data from the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or the Fujitsu Support Desk with the collected data together with the output message.

[ERR.] INST 1033 pxinst *hostname* must be either installer node or relay node.

Meaning

The host name *hostname* must be for either the installer node or relay node.

Action

Review the specified host name. If necessary, execute the command again.

[ERR.] INST 1034 pxinst initrd has not been created.

Meaning

Processing stopped because initrd had not been created.

Action

Use the initrd creation function to create initrd. Then, execute the command again.

[ERR.] INST 1035 Specified rootfs is not found.

Meaning

The specified rootfs (root file system) was not found.

Action

Review the contents of the specified option, and execute the command again.

[ERR.] INST 1036 The specified hostname is already registered.

Meaning

The specified hostname is already registered.

Action

Review the hostname and execute the command again.

[ERR.] INST 1037 The specified Tofu is already registered.

Meaning

The specified Tofu is already registered.

Action

Review the compute network IP address and execute the command again.

[ERR.] INST 1038 The specified MNG_NET is already registered.

Meaning

The specified MNG_NET is already registered.

Action

Review the management network IP address and execute the command again.

Warning Message

[WARN] INST 1500 pxinst Replication failed node exists.(*filename*)

Meaning

Synchronization processing failed on an existing node.

Action

Remove the cause, referring to *filename*, in which the host name and cause for the synchronization failure has been output. Then, execute synchronization again.

[WARN] INST 1501 pxinst Abnormal service status node exists.

Meaning

An existing node has a problem in the service operation status.

Action

From the contents displayed by the command, identify the node with the problem, and check the node service status. If necessary, start the service.

[WARN] INST 1502 pxinst The specified hostname is not registered.(*hostname*)

Meaning

The specified host name *hostname* is not registered.

Action

Ignore *hostname*, and continue processing.

[WARN] INST 1503 pxinst The following packages is not found.

pkg pkg ...

Meaning

The repository does not have the RPM package file *pkg*.

Action

Skip the RPM package file that does not exist, and continue processing. If necessary, execute the command again.

[WARN] INST 1504 pxinst The following packages is already installed.

pkg pkg ...

Meaning

The RPM package file *pkg* is already installed.

Action

Skip the already installed RPM package file, and continue processing. If necessary, execute the command again.

Information Message

[INFO] INST 1801 pxinst Initialization completed successfully.(*details*)

Meaning

The install function was successfully initialized. The initialized functions appear in *type*.

Action

No action is necessary.

[INFO] INST 1802 pxinst Please run the 'pxinst sync' command to update the repository changes to slave servers.

Meaning

Execute synchronization so that the standby installer node and relay node reflect the repository changes.

Action

If necessary, use the pxinst sync command to execute synchronization.

[INFO] INST 1803 pxinst Adding nodes completed.(*nodenum=number of nodes*)

Meaning

Addition of nodes has completed. The number of nodes is the number of added nodes.

Action

No action is necessary.

[INFO] INST 1804 pxinst Deleting nodes completed.(*nodenum=number of nodes*)

Meaning

Deletion of nodes has completed. The number of nodes is the number of deleted nodes.

Action

No action is necessary.

[INFO] INST 1805 pxinst File replication completed.

Meaning

File replication processing has completed.

Action

No action is necessary.

[INFO] INST 1806 pxinst Service state check completed.

Meaning

Service state check processing has completed.

Action

No action is necessary.

4.1.40 pxsystemadm command

[NAME]

pxsystemadm - Cluster configuration management command

[SYNOPSIS]

```
/usr/sbin/pxsystemadm --convert node_file
/usr/sbin/pxsystemadm --create conf_file [--ft ft_dirname] [--force]
/usr/sbin/pxsystemadm --set-hardware [-c clstname ] [scope options ] [--force]
/usr/sbin/pxsystemadm --check-hardware [-c clstname] [scope options]
/usr/sbin/pxsystemadm --set [-c clstname] [scope options] [--force]
/usr/sbin/pxsystemadm --help
```

scope options (exclusive):

```
--nodegrp nodegid[,...]
--bootgrp bootgid[,...]
-n nodeid[,...]
-f filename
```

[DESCRIPTION]

The `pxsystemadm` command is a command for system administrators to set a cluster configuration for the Job Operation Software. Administrators can register a new cluster, and add and delete nodes in a cluster. They can also delete a cluster that is no longer necessary. The configuration of an added or deleted cluster is written in the cluster configuration definition file. In addition, the `pxsystemadm` command hardware settings of the FX server. After defining or changing a cluster configuration by using the `--create` option, use the `--set` option to reflect the cluster configuration to all nodes. This command can be used on the active system management node.

[OPTIONS]

`--convert node_file`

This option creates a template for the cluster configuration definition file from the node configuration information files output by the installation design sheet.

In the argument, specify any path to the place for the node configuration information files to create a template in the current directory at the command execution time.

`--create conf_file`

This option sets the contents of the cluster configuration definition file created by the system administrator, to the system management node.

To add a cluster or to add nodes that includes the FX server, specify the `--ft` option, and specify the absolute path to the storage directory of the FX server node configuration information files in the argument.

`--set-hardware`

This option configures control settings of the FX server.

Use this option when communication with the hardware defined in the cluster configuration is possible.

If this setting fails with any node, an error file is output to the current directory.

`--check-hardware`

This option checks whether control setting has configured for every FX server in the specified range.

If it has not been configured for any FX server, an error file is output to the current directory.

`--set`

This option reflects the defined node configuration to each node.

If this setting fails with any node, an error file is output to the current directory.

If the `--force` option is not specified, a yes/no prompt asks whether to continue execution.

`--force`

This option forcibly executes an instruction without a yes/no prompt asking whether to continue the execution.

-c *clstname*

Specify the execution target cluster in *clstname*.

If the `--set-hardware`, `--check-hardware`, or `--set` option is specified and this option is not specified, the set cluster name in the environment variable `PXMYCLST` is used.

--nodegrp *nodegid*[,...]

Specify the ID of a node group in *nodegid* if you want only that specific node group as the execution target in the cluster.

Write *nodegid* in hexadecimal.

You can specify multiple node group IDs by delimiting them with a comma (e.g., 0x01,0x02,0x03).

You can specify a range of IDs by using a hyphen (e.g., 0x01-0x03).

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

--bootgrp *bootgid*[,...]

Specify the ID of a boot group in *bootgid* if you only that specific boot group as the execution target in the cluster.

Write *bootgid* in hexadecimal.

You can specify multiple boot group IDs by delimiting them with a comma (e.g., 0x0101,0x0102,0x0103).

You can specify a range of IDs by using a hyphen (e.g., 0x0101-0x0103).

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

-n *nodeid*[,...]

Specify the ID of a node in *nodeid* if you want only that specific node as the execution target in the cluster.

Write *nodeid* in hexadecimal.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010012).

You can specify a range of IDs by using a hyphen (e.g., 0x01010010-0x0101002A).

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

-f *filename*

This option executes processing on the node IDs written in the file specified in *filename*.

Write one node ID in hexadecimal per line in the file *filename*.

Lines with the comment symbol (#) written at the beginning are ignored.

Strings other than node ID and node ID that does not exist are ignored.

(Example)

0x01010010

0x01010011

0x0101002A

When specifying this option, be sure to specify a cluster name in *clstname* in the `-c` option, or specify a specific cluster name in the environment variable `PXMYCLST`.

--help

This option displays how to use this command.

If this option is specified, all the arguments and other options are ignored.

[NOTE]

- The `--convert`, `--create`, `--set-hardware`, `--check-hardware`, and `--set` options cannot be specified at the same time.
- The `--nodegrp`, `--bootgrp`, `-n`, and `-f` options cannot be specified at the same time.
- If the `--set-hardware` or `--set` option is executed but fails to set any node, you need to eliminate the cause of the failure and re-execute the command on the failed node.

[ENVIRONMENT]

PXMYCLST

You can specify a cluster name in the environment variable PXMYCLST instead of the -c option. Specification with the -c option has priority over that with the environment variable PXMYCLST. You can specify only one cluster name in the environment variable PXMYCLST. An example is shown below.
(Example)

```
export PXMYCLST=cluster1
```

[EXIT STATUS]

The following end codes are returned.

- 0: Normal end
- 1: Abnormal end
- 2: Setting failure on node

[SEE ALSO]

pxsystem.conf(5)

[MESSAGES]

If an error occurs during the execution of the pxsystemadm command, the following messages is output to the standard error output. Also, the information message that reports process completion, etc. is output to standard output.



See

Not only the following messages but also common command messages are output. For details on the common command messages, see ["4.2.1 Common command messages of the system management function."](#)

Error Message

[ERR.] PSM 1001 pxsystemadm There are nodes which has not setting of the hardware.(*filepath*)

Meaning

Some nodes do not have hardware control settings.

filepath: Error file name

Action

Referring to the error file output to the current directory, configure hardware control on the nodes that do not have the hardware control settings.

[ERR.] PSM 1002 pxsystemadm Processing failed.(*details*)

Meaning

Command processing failed.

details: Cause of the error

Action

Remove the cause of error indicated in the *details*, and execute the command again.

Warning Message

[WARN] PSM 1051 pxsystemadm Reboot the cluster to reflect the settings.

Meaning

The cluster must be restarted to reflect the setting changes.

Action

Restart the target cluster.

[WARN] PSM 1052 pxsystemadm Reboot the node to reflect the settings.

Meaning

The node must be restarted to reflect the setting changes.

Action

Restart the target node.

Information Message

[INFO] PSM 1081 pxsystemadm The file was output.(*filename*)

Meaning

The *filename* model of the cluster configuration definition file was output to the current directory.

Action

No action is necessary.

[INFO] PSM 1082 pxsystemadm The setting was completed.

Meaning

The setting has completed.

Action

No action is necessary.

[INFO] PSM 1083 pxsystemadm The setting check succeeded.

Meaning

All the specified nodes are already set.

Action

No action is necessary.

[INFO] PSM 1084 pxsystemadm Deletion of the last cluster include SMM.

Meaning

Cluster configuration information was deleted from the system management node because all the clusters were deleted.

Action

No action is necessary.

4.2 Common command messages

The system management function and the install function have common command messages.

4.2.1 Common command messages of the system management function

The following commands of the system management function output not only messages specific to each command but also common command messages:

- pachecker
- paclstmgr
- paconsole
- padumpmgr
- paeventadm
- paledctl
- pamonitor_notice
- pamonitoradm
- pamopuginadm
- papwradm
- papwrctl
- parender
- paserviceadm
- pashowclst
- pasnap
- pastart
- pastop
- pmexe
- pmgather
- pmscatter
- pxsystemadm

If an error occurs during the execution of any of these commands, one of the following messages is output to the standard error output. Also, an information message that reports processing completion, etc. is output to the standard output.

The descriptions of command errors are common to the commands.

The following is example of an output common command message.

```
[ERR.] PSM 0001 COMMAND Internal error occurred.(details)
```

The name of the executed command is input in *COMMAND*.



See

Each command reference describes messages specific to one of the above-mentioned commands. See the respective command reference.

Error Message

[ERR.] PSM 0001 *COMMAND* Internal error occurred.(*details*)

Meaning

An internal error occurred.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

[ERR.] PSM 0002 COMMAND Cannot be used on this node.**Meaning**

The specified command is not available on this node.

Action

Review the types of node for executing the command, and execute the command again.

[ERR.] PSM 0003 COMMAND Invalid option.(*details*)**Meaning**

The specified value of the option is invalid.

details: Invalid option

Action

Confirm the correct option, and execute the command again.

[ERR.] PSM 0004 COMMAND Invalid combination of option.**Meaning**

The combination of specified options is invalid.

Action

Confirm the correct option, and execute the command again.

[ERR.] PSM 0005 COMMAND Invalid argument.(*details*)**Meaning**

The specified argument is invalid.

details: Invalid character string

Action

Review the argument, and execute the command again.

[ERR.] PSM 0006 COMMAND Invalid cluster name.(*clstname*)**Meaning**

The specified cluster name *clstname* is invalid.

Action

Review the cluster name, and execute the command again.

[ERR.] PSM 0007 COMMAND No such file.(*filename*)**Meaning**

The specified file *filename* does not exist.

Action

Check whether the specified file exists, and execute the command again.

[ERR.] PSM 0008 COMMAND No such directory.(*dirname*)**Meaning**

The specified directory *dirname* does not exist.

Action

Check whether the specified directory exists, and execute the command again.

[ERR.] PSM 0009 COMMAND No data available.(details)

Meaning

No applicable information, such as execution target nodes, exists.

details: Specified information

Action

Review the option argument, and execute the command again.

[ERR.] PSM 0010 COMMAND Node configuration error occurred.(details)

Meaning

The node configuration definition is incorrect.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 0011 COMMAND Node configuration is not set.

Meaning

The node configuration is not set.

Action

Set the node configuration.

[ERR.] PSM 0012 COMMAND Configuration file is not found.

Meaning

There is no configuration file.

Action

Prepare a configuration file.

[ERR.] PSM 0013 COMMAND Set command is not executed.

Meaning

The configuration file contents were not reflected.

Action

Reflect the configuration file, and execute the command again.

[ERR.] PSM 0014 COMMAND Invalid file format.(details)

Meaning

The format of the configuration file is incorrect.

details: Information on the incorrect format in the file

Action

Review the format of the configuration file, and execute the command again.

[ERR.] PSM 0015 *COMMAND* Already running.(*details*)**Meaning**

The *COMMAND* command is running.

detail: Detailed error information

Action

Confirm that the *COMMAND* command is not running. Then, execute the command again.

[ERR.] PSM 0016 *COMMAND* Memory allocation failed.(*details*)**Meaning**

Memory acquisition failed.

detail: Detailed error information

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 0017 *COMMAND* Failed to read configuration file.(*details*)**Meaning**

Reading of the configuration file failed.

details: Detailed error information

Action

Check for an error in the configuration file.

[ERR.] PSM 0018 *COMMAND* Permission denied.(*filename*)**Meaning**

The privileges of the configuration file are not appropriate.

filename: Invalid file name

Action

Check the file permissions.

[ERR.] PSM 0019 *COMMAND* Node configuration is being changed.**Meaning**

The command cannot be executed because the configuration is being changed.

Action

After waiting a moment, execute the command again.

[ERR.] PSM 0020 *COMMAND* There are nodes that failed in the setting.(*filename*)**Meaning**

Some nodes experienced a setting failure.

filename: Output error file path

Action

Remove the cause, referring to the error file *filename*, and execute the command again.

[ERR.] PSM 0021 COMMAND I/O error occurred file=*filename*.(*details*)**Meaning**

An I/O error occurred in the operation for *filename*.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again. If the cause is unknown, collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) or Fujitsu Support Desk with the collected data together with the output message.

[ERR.] PSM 0022 COMMAND Unfailover setting is incomplete.**Meaning**

The failover suppression setting is incomplete.

Action

Use the --set-failover option of the pacstmgr command to set failover suppression again.

[ERR.] PSM 0023 COMMAND Invalid file name.(*details*)**Meaning**

The file name does not meet naming conventions.

Action

In accordance with naming conventions, define a file name with up to 64 characters, including single-byte alphanumeric characters, the hyphen, the underscore, and the period. However, the first character cannot be a hyphen.

Warning Message

[WARN] PSM 0051 COMMAND There are nodes that failed in the setting.(*filename*)**Meaning**

Some nodes experienced a setting failure.

filename: Output error file path

Action

Remove the cause, referring to the error file *filename*, and execute the command again.

Information Message

[INFO] PSM 0081 COMMAND Command canceled.**Meaning**

The command was canceled.

Action

No action is necessary.

[INFO] PSM 0082 COMMAND The processing of the configuration file was completed.**Meaning**

The configuration file was processed successfully.

Action

No action is necessary.

4.2.2 Common command messages of the install function

The following commands of the install function output not only messages specific to each command but also common command messages:

- pxinst
- paclone

If an error occurs during the execution of any of these commands, one of the following messages is output to the standard error output. Also, an information message that reports processing completion, etc. is output to the standard output.

The descriptions of command errors are common to the commands.

The following is example of an output common command message.

```
[ERR.] INST 0001 COMMAND Option analysis error.(details)
```

The name of the executed command is input in *COMMAND*.



See

.....
Each command reference describes messages specific to one of the above-mentioned commands. See the respective command reference.
.....

Error Message

[ERR.] INST 0001 *COMMAND* Option analysis error.(*details*)

Meaning

This is an option analysis error.

details: Invalid contents

Action

Referring to *details*, review the options, and execute the command again.

[ERR.] INST 0002 *COMMAND* Invalid argument.(*arg*)

Meaning

The specified argument *arg* is invalid.

Action

Review the argument, and execute the command again.

[ERR.] INST 0003 *COMMAND* Invalid combination of options.

Meaning

The combination of specified options is invalid.

Action

Review the option, and execute the command again.

[ERR.] INST 0004 *COMMAND* Invalid argument *arg* for *opt*

Meaning

Invalid argument *arg* for *opt*

Action

Review the argument *arg* for the option *opt*, and execute the command again.

[ERR.] INST 0005 *COMMAND* sub-command is not specified.

Meaning

No valid sub command is specified.

Action

Specify a valid sub command, and execute the command again.

[ERR.] INST 0006 **COMMAND** This command can be executed only on active installer node.

Meaning

The command was executed on a node other than the active installer node.

Action

Execute the command on the active installer node.

[ERR.] INST 0007 **COMMAND** sub-command failed.(*operation*)

Meaning

An error occurred in the processing of *COMMAND* sub-command.

operation: Operation description

Action

Remove the cause, referring to the information output immediately after the error, and execute the command again.

[ERR.] INST 0008 **COMMAND** Multiple sub-command are specified.

Meaning

Multiple sub commands are specified.

Action

Review the executed command line, and execute the command again.

[ERR.] INST 0009 **COMMAND** System call error occurred.(*details*)

Meaning

A system call error occurred.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again. If the cause is unknown, collect investigation data about the system management node according to the "Job Operation Software_Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] INST 0010 **COMMAND** I/O error occurred. file=*filename*.(*details*)

Meaning

An I/O error occurred in the operation for *filename*.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again. If the cause is unknown, collect investigation data about the system management node according to the "Job Operation Software_Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] INST 0011 COMMAND Failed to parse the configuration file.(filename, details)**Meaning**

An error occurred in the analysis of the configuration file *filename*.

details: Cause of the error

Action

Remove the cause, referring to *details*, and execute the command again.

[ERR.] INST 0012 COMMAND Not found configuration file.(filename)**Meaning**

The configuration file *filename* does not exist.

Action

Place the configuration file *filename*, and execute the command again.

[ERR.] INST 0013 COMMAND action is not specified.**Meaning**

No valid action is specified.

Action

Specify a valid action, and execute the command again.

[ERR.] INST 0014 COMMAND Specified directory is not found.(dir for opt)**Meaning**

The *dir* directory specified in the *opt* option was not found.

Action

Review the specification of the directory, specify an existing directory, and execute the command again.

[ERR.] INST 0015 COMMAND Specified directory is already exists.(dir for opt)**Meaning**

The *dir* directory specified in the *opt* option already exists.

Action

Review the specification of the directory, specify a directory that does not exist, and execute the command again.

**[ERR.] INST 0016 COMMAND Command execute error.(cmd)
detail=details****Meaning**

The install function executed the *cmd* command internally, but execution failed.

Action

Remove the cause, referring to *details*, and execute the command again. If the cause is unknown, collect investigation data about the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] INST 0017 COMMAND Node configuration error occurred.(details)

Meaning

The node configuration definition is incorrect.

details: Cause of the error

Action

Collect investigation data about the system management node according to the "Job Operation Software Administrator's Guide for Maintenance." Then, contact a Fujitsu systems engineer (SE) or FUJITSU Support Desk with the collected data on the system management node together with the output message.

[ERR.] INST 0018 **COMMAND** Structure configuration file is not found.

Meaning

The hierarchical structure definition file does not exist.

Action

Create the hierarchical structure definition file /var/opt/FJSVtcs/inst/configure/structure.yaml, and execute the command again.

[ERR.] INST 0499 **COMMAND** Internal error occurred.(*details*)

Meaning

An internal error occurred.

details: Cause of the error

Action

Collect investigation data according to the "Job Operation Software Administrator's Guide for Maintenance," and then contact a Fujitsu systems engineer (SE) with the collected data together with the output message.

Information Message

[INFO] INST 0800 **COMMAND** sub-command completed successfully.(*operation*)

Meaning

The operation of *COMMAND* sub-command completed successfully.

operation: Operation description

Action

No action is necessary.

[INFO] INST 0801 **COMMAND** Created a configuration backup.(*filepath*)

Meaning

A backup of the configuration file was created.

filepath: Path for placing the created backup file. The path may be a file or directory, depending on the executed processing.

Action

No action is necessary. If you need to restore the configuration file to the state before command execution, overwrite the configuration file with the *filepath* contents.

4.3 Configuration Files

This section describes how to write and reference configuration files for administrators.

4.3.1 How to Write in the Configuration Files

The configuration files of the Job Operation Software use the following format (with a few exceptions) for the written contents.

```
section name {  
    item1 = value  
    item2 = value  
    subsection name {  
        item3 = value  
        item4 = value  
    }  
}
```

The section name is a keyword showing what kind of structural definition is enclosed in curly brackets ("{" "}"). Determine this name according to the definition contents. Depending on the definition item, a section will contain further subsections. For details on definition items, see the configuration file reference.

Notes on writing in a configuration file are shown below.

- Be sure to specify a section name.
The curly bracket "{" represents the beginning of a section, and it must be written on the same line as the name. The section name must be marked as separate from section names and subsection names by one or more spaces or tab characters. The curly bracket "}" represents the end of a section, and it must be written alone on a single line.
- Write one single-byte space between an item and an equal sign, and between an equal sign and a setting value.
- A line beginning with a number sign("#") or the part after a number sign("#") on a line is regarded as a comment.
- Spaces written at the beginning are ignored.
- Specify a name and ID for the scope of each section. The scope (setting item used for identification) reflects the settings in the section. Write the name and ID at the beginning of the section.
For details on setting items used for identification, see the configuration file reference in "[4.3 Configuration Files](#)."
- Write dates in the "year-month-day hour:minute:second" format.
Details on how to write the date are as follows.
 - Write the year with four digits.
 - Write the month, day, hour, minute, and second with two digits each. Even if the first digit is 0, write 0.
 - Write the hour in 24-hour format.
 - Connect the year, month, and day with a hyphen("-"). Connect the hour, minute, and second with a colon(":"). Insert a space between the day and hour.
- Write the disk size in powers of 10 and the memory size in powers of 2.
For details on units, see "[Representation of units](#)" in [Preface](#).
- After editing a configuration file, unless specified otherwise, set the file owner to user root and group root, and set access permission to 0600 to prevent malicious users from tampering with the configuration file.

The following is a reference to the configuration file.

Configuration files are listed in alphabetical order.

4.3.2 paevent.conf file

[NAME]

paevent.conf - Log monitoring condition configuration file

[SYNOPSIS]

/etc/opt/FJSVtcs/paevent.conf

[DESCRIPTION]

paevent.conf is a file for setting the conditions to monitor events (specific message patterns) output to a log.
If no configuration is defined, all nodes are excluded from monitoring target for logs.

Cluster

This section defines a cluster. This definition is required.

You can specify the following sections and items:

ClusterName

NodeType section

Node section

Either the NodeType section or the Node section must be specified.

You can specify multiple NodeType or Node definitions in a single Cluster section.

If you specify multiple definitions for the same node, all the definitions for the node are set.

ClusterName

This option defines the name of the cluster. This definition is required.

NodeType

This section defines the node types that are monitoring targets for logs. You can omit this definition.

You can specify multiple node types by delimiting them with a comma. It sets the following items and sections.

NodeTypeName

Event section

NodeTypeName

This section defines the node types that are monitoring targets for logs. This definition is required.

You can specify multiple node types by delimiting them with a comma.

The following node types can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
CN	Compute node
LN	Login node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
SCM	Storage cluster management node
OSS	OSS node
MDS	MDS node
MGS	MGS node

Node

This section defines the nodes that are monitoring targets for logs. You can omit this definition.

It sets the following items and sections.

NodeID

Event section

NodeID

This section defines the node IDs that are monitoring targets for logs. This definition is required.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010011,0x0101002A).

You also can specify a range with a hyphen (e.g., 0x01010010-0x0101002A).

Event

This section defines the log monitoring conditions.

Items in NodeType and other sections must always be defined. This section can have multiple instances of a definition.

It sets the following items.

EventName

FileName

ChkString

IgnoreString

Timer

Counter

ExecCmd

EventName

This item defines label names of messages to be monitored. This definition is required.

It is used for output to a command execution log when a message pattern is detected.

Use a different label name for each monitoring condition.

The Event section can contain only one definition of this item.

The label name must not begin with a '-' (hyphen), it must consist of up to 63 alphanumeric characters and '-'

FileName

This item defines the absolute path of a log file that is a monitoring target. You can omit this definition.

Specify a Job Operation Software or user-specific log file.

If this definition is omitted, /var/log/messages is the target.

The Event section can contain only one definition of this item.

ChkString

This section defines the messages to be monitored. This definition is required.

You can use POSIX extended regular expressions. (*)

The substring that corresponds to the entire regular expression is stored in PAEVENT_STR0.

The substrings of string that correspond to the parenthesized subexpressions of pattern (i.e."(...)") is stored from PAEVENT_STR1 to PAEVENT_STR10.

The maximum number of substring that corresponds to the parenthesized subexpressions of pattern is 10.

You can reference the stored environment variables on the ExecCmd command line.

(*) When using regular expressions, please describe the escape character (\) as necessary.

IgnoreString

This item describes a specific pattern included in messages that match the monitoring condition defined by ChkString but should be ignored. You can omit this definition.

You can use POSIX extended regular expressions. (*)

You can specify up to 32,767 definitions in the Event section.

(*) When using regular expressions, please describe the escape character (\) as necessary.

Timer

This item specifies an interval in seconds between message output to be monitored and command execution. The command is executed when the message is output Counter times. You can omit this definition.

Timer starts each time a message occurs and counts messages that occurred within the time.

If you want to execute the commands specified by ExecCmd when a monitoring condition log is output even just once, specify 0.

You can specify multiple intervals by delimiting them with a comma. In this case, be sure to write as many values as the counter values specified in Counter.

The same time cannot be specified multiple times. In addition, this item must be specified in ascending order.

You can specify a value ranging from 0 seconds, minimum, to 31536000, maximum.

If this setting is omitted, the default value of 0 is used.

The Event section can contain only one definition of this item.

If any of the specified conditions is satisfied, the message counter is reset and resumes counting from the next message.

Counter

This item specifies how many occurrences of the messages that are monitoring targets will trigger command execution. The messages must occur within the number of seconds specified by Timer. You can omit this definition.

You can specify multiple intervals by delimiting them with a comma. In this case, be sure to write as many values as the timer values specified in Timer.

You can not specify the same number of times. In addition, it must be specified in ascending order.

You can specify a value ranging from 1 seconds, minimum, to 32767, maximum.

If this setting is omitted, the default value of 1 is used.

The Event section can contain only one definition of this item.

ExecCmd

This item describes the command line in double quotes to run on the detected node, when a message meeting monitoring conditions is detected. This definition is required.

The command line described in this item is executed by sh.

Commands are executed with /var/opt/FJSVtcs/psm/tmp as the current directory.

The Event section can contain only one definition of this item.

You can reference the following environment variables on the command line at execution.

Environment variable name	Description
PAEVENT_NODEID	Node ID of the found string.
PAEVENT_IPADDRESS	IP address for the nodeid which is detected message. (BIO,GIO,SIO,CN:IP address for tofu network.The other nodes:IP address for Management network)
PAEVENT_LINE	Line of the pattern matching (It does not include line breaks).
PAEVENT_STR0	Matching string of the Regular expressions.
PAEVENT_STR1	Matching substring of the first group.
PAEVENT_STR2	Matching substring of the 2nd group.
....
PAEVENT_STR10	Matching substring of the 10th group.

[EXAMPLE]

The following example shows how to define the paevent.conf file.

```
Cluster {
  ClusterName = cluster1
  NodeType {
    NodeTypeName = CCM,CCS
    Event {
      EventName = Down_Check
      ChkString = "(eth[0-9]?): NIC Link is Down"
      ExecCmd = "echo $PAEVENT_STR1 is Down > node_check.txt"
    }
  }
}
Node {
  NodeID = 0x01010010
  Event {
    EventName = clock
    ChkString = "kernel:.*clock"
    ExecCmd = "touch /var/log/aaa.txt"
  }
}
```

[SEE ALSO]

paeventadm(8)

4.3.3 pamonitor.conf file

[NAME]

pamonitor.conf - System monitoring configuration file

[SYNOPSIS]

`/etc/opt/FJSVtcs/pamonitor.conf`

[DESCRIPTION]

pamonitor.conf is a file that defines the actions of the system monitoring function.

Cluster

This section defines each cluster. This item must be specified. It sets the following items and sections.

ClusterName

This item specifies a cluster name. This item must be specified.

Monitor

This section defines the monitoring settings. This item must be specified.

NodeInterval

This item specifies the interval in seconds for alive monitoring of the OS.

After the specified period has elapsed since the completion of the previous alive check, the next alive check is performed. Specify an integer ranging from 60 to 86400. If this item is omitted, 60 is the value for this action.

NodeTimeout

This item specifies the timeout period in seconds for alive monitoring of the OS.

If the OS alive check process does not complete within the specified period, the alive check of the OS fails. Specify an integer ranging from 30 to 86400. If this item is omitted, 300 is the value for this action.

NodeRetry

This item specifies the retry count for alive monitoring of the OS.

If the OS alive check process does not complete, the alive check of the OS fails. Then, the check is retried the number of times specified by NodeRetry.

The maximum period for the alive check of the OS is (NodeTimeout / NodeRetry) (seconds).

Specify an integer ranging from 0 to NodeTimeout. If 0 is specified, do not retry. If this item is omitted, 10 is the value for this action.

ServiceInterval

This item specifies the interval in seconds for alive monitoring of the services in except compute nodes.

After the specified period has elapsed since the completion of the previous service check, the next service check is performed. Specify an integer ranging from 60 to 86400. If this item is omitted, 60 is the value for this action.

ServiceInterval_CN

This item specifies the interval in seconds for alive monitoring of the services on compute nodes.

After the specified period has elapsed since the completion of the previous service check, the next service check is performed. If the compute node is also configured as an IO node, the ServiceInterval value is used.

Specify an integer ranging from 60 to 86400. If this item is omitted, 60 is the value for this action.

ServiceTimeout

This item specifies the number of seconds in the period for confirming each service.

If the service confirmation process does not complete within the specified period, an error is considered to have occurred. Specify an integer ranging from 30 to 86400. If this item is omitted, 60 is the value for this action.

TofuDynamicReroute

This item specifies whether to allow dynamic changes to the communication paths of the jobs. Specify either of the following values:

0: Do not allow dynamic changes to the communication paths of the jobs.

1: Allow dynamic changes to the communication paths of the jobs.

If this specification is omitted, 0 is assumed specified.

MasterLogLevel

This item specifies one of the following values as the monitoring master level:

Log level	Usage
1:	Output only information messages.
2:	Output debug information but minimize output as much as possible considering the system load.
3:	Output the most detailed debug information without considering the system load. Normally, specify 1. If this item is omitted, 1 is the value for this action.

SubLogLevel

This item specifies the monitoring sub log level.

Log level	Usage
1:	Output only information messages.
2:	Output debug information but minimize output as much as possible considering the system load.
3:	Output the most detailed debug information without considering the system load. Normally, specify 1. If this item is omitted, 1 is the value for this action.

SlaveLogLevel

This item specifies the monitoring slave log level.

Log level	Usage
1:	Output only information messages.
2:	Output debug information but minimize output as much as possible considering the system load.
3:	Output the most detailed debug information without considering the system load. Normally, specify 1. If this item is omitted, 1 is the value for this action.

IC

This section defines the condition to judge IC service is abnormal. Change the monitoring conditions only if you need to.

DeviceName

This item specifies the name of the device. Specify "mlx" for EDR.

PortNumber

This item specifies the port number. Specify "1,2" for EDR.

PeakRate

This item specifies the performance value in which the state of IB is considered to be normal. The state of IB is treated as a degeneracy when not agreeing to the specified performance value.

Specify either of the followings.

1. When 100 Gb/sec is normalized by EDR(Service state display: IC(o)), "100 Gb/sec" is specified

Service state	Performance value
Normal IC(o)	100 Gb/sec (4X EDR)
Degeneration IC(!)	56 Gb/sec (4X FDR) 40 Gb/sec (4X FDR10) 40 Gb/sec (4X QDR) 20 Gb/sec (4X DDR) 14 Gb/sec (1X FDR) 10 Gb/sec (1X FDR10) 10 Gb/sec (1X QDR) 10 Gb/sec (4X)

2. When 56 Gb/sec is normalized by EDR or FDR(IC(o)), "100 Gb/sec, 56 Gb/sec " is specified.

[EDR HCA]

Service state	Performance value
Normal IC(o)	100 Gb/sec (4X EDR) 56 Gb/sec (4X FDR)
Degeneration IC(!)	40 Gb/sec (4X FDR10) 40 Gb/sec (4X QDR) 20 Gb/sec (4X DDR) 14 Gb/sec (1X FDR) 10 Gb/sec (1X FDR10) 10 Gb/sec (1X QDR) 10 Gb/sec (4X)

[FDR HCA]

Service state	Performance value
Normal IC(o)	56 Gb/sec (4X FDR)
Degeneration IC(!)	40 Gb/sec (4X FDR10) 40 Gb/sec (4X QDR) 20 Gb/sec (4X DDR) 14 Gb/sec (1X FDR) 10 Gb/sec (1X FDR10) 10 Gb/sec (1X QDR) 10 Gb/sec (4X)

[EXAMPLE]

The following example shows how to define the pamonitor.conf file.

```
Cluster {
  ClusterName = cluster1
  Monitor {
    NodeInterval = 60
    NodeTimeout = 300
    NodeRetry = 10
    ServiceInterval = 60
    ServiceInterval_CN = 60
  }
}
```

```

        ServiceTimeout = 60
        TofuDynamicReroute = 0
        MasterLogLevel = 1
        SubLogLevel = 1
        SlaveLogLevel = 1
    }
    IC {
        DeviceName = "mlx"
        PortNumber = "1,2"
        PeakRate = "100 Gb/sec"
    }
}

```

[SEE ALSO]

pamonitoradm(8)

4.3.4 pamoplugin.conf file

[NAME]

pamoplugin.conf - Monitor plugin configuration file

[SYNOPSIS]

/etc/opt/FJSVtcs/pamoplugin.conf

[DESCRIPTION]

pamoplugin.conf is the file that defines the actions of monitor plugin.

Cluster

This section defines each cluster. This item must be specified. It sets the following items and sections.

ClusterName

This item specifies a cluster name. This item must be specified.

Concurrency

This item specifies the maximum number of concurrent calls of the monitor plugin command. If the number of calls exceeds the specified value, the calling instance waits for a running instance of the command to complete.

Specify an integer ranging from 1 to 99999999. If this item is omitted, the value is 1024.

AutoNodeOff

This item specifies whether to enable the automatic node off function.

If the "on" value is specified, the transition of the Boot I/O node of FX server which is including compute to down stops the under nodes(compute nodes, Global I/O nodes which is including compute and Storage I/O nodes which is including compute).

on: Enable

off: Disable

If this item is omitted, the value is on.

PluginCmd

This section defines error exit. The specification of PluginCmd or Dump is required.

ServiceName

This item obtains the target service name. The pashowclst command displays the service name. If this item is omitted, all the services that are monitoring targets are targets. You can specify multiple service names by delimiting them with a comma.

NodeType

This item specifies the node type that is a target. If this item is omitted, all node types are targets. You can specify multiple node types by delimiting them with a comma.

ExcludeType

This item specifies the node type that is a except target. You can specify multiple node types by delimiting them with a comma. Specify for one or more node types that become objects together with the NodeType specification value to exist. You can omit this definition.

Status

This item specifies a service or OS state.

The monitor plugin command runs when a service or the OS enters the specified state. Specify one of the following values. You can specify multiple service or OS states by delimiting them with a comma.

Setting value (Service or OS state)	Description
o	Service running
x	Service stopped or error occurred
f	Failover in progress
-	Service is not monitoring target or OS stopped
!	Degeneracy occurred (reduced function)
w	Startup has completed on the standby node. The system can be switched back
*	Operating on one side
a	Layered file system I/O abnormal state

If this item is omitted, the value is "Service stopped (unavailable state): x."

Cmd

This item specifies the full path name of the command to be executed.

Enclose it with double quart (") when you specify an optional command etc.

If the specified file does not exist or there is no authority for execution, this item is ignored.

The following environment variable is set up to the execution process of Cmd.

Environment variable name	Description
PACLSTNAME	The following environment variable is set up to the execution process of Cmd.
PANODENUM	Number of nodes whose state has changed. The state change of each node is set in the environment variable PANODEn.
PANODEn	Change to the state of a node. As many specification values as the number of nodes whose state has changed are set in order from PANODE1 to PANODEn. Maximum value of n: PANODENUM

The setting is a colon (:) separated value.

node ID:node type:service state[:service state...]

If the one node holds some the node types, the state of each of the node types is set as a comma (,) separated value.(example: BIO,CN)

The state of service is set by "Service name=state". If the states of multiple services have changed, the state of each of the services is set as a colon (:) separated value. (example: PJM=o:PLE=o)

Dump

This section defines the automatic dump collection settings. The specification of PluginCmd or Dump is required.

NodeType

This item specifies a node type. If this item is omitted, all node types are targets.

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

ExcludeType

This item specifies the node type that is a except target. You can specify multiple node types by delimiting them with a comma. Specify for one or more node types that become objects together with the NodeType specification value to exist.

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

AutoDump

This item specifies whether to enable the automatic dump collection command.

on: Enable

off: Disable

If this item is omitted, the value is on. Regardless of this setting, automatic dump collection is disabled for nodes that have entered the software maintenance mode by specifying `--soft-mainte` option of the `paclstmgr` command, or transitioned to hardware maintenance mode by specifying `--hard-mainte` option of the `paclstmgr` command.

AutoForceOff

This item uses the following values to specify whether to forcibly power off in the event of no transition to the dump collection state within the period specified by `WaitTime` after automatic dump collection starts.

on: Enable
off: Disable

If this item is omitted, the value is off.

WaitTime

This item specifies the interval in seconds until force power-off, which is specified by `AutoForceOff`. Specify an integer ranging from 60 to 99999999. If this item is omitted, the value is 600.

[EXAMPLE]

The following example shows how to define the `pamoplugin.conf` file.

```
Cluster {
  ClusterName = cluster1
  AutoNodeOff = on
  PluginCmd {
    NodeType = CN
    ExcludeType = BIO,GIO,SIO
    ServiceName = OS
    Status = x
    Cmd = "/work/test_CN.sh > /dev/null"
  }
  PluginCmd {
    NodeType = BIO,GIO,SIO
    ServiceName = OS
    Status = x
    Cmd = "/work/test_IO.sh > /dev/null"
  }
  Dump {
    NodeType = CN
    AutoDump = on
    AutoForceOff = off
  }
}
```

[SEE ALSO]

`pamopluginadm(8)`

4.3.5 papjm.conf file

[NAME]

`papjm.conf` - Defines the operation environment for job operation in the entire cluster.

[SYNOPSIS]

`/etc/opt/FJSVtcs/papjm.conf`

[DESCRIPTION]

The `papjm.conf` file configures on the system management node, defines configuration to execute job operation in the entire cluster. This configuration is reflected by executing the `papjmadm` command on the system management node.

Cluster managers configure the definition of this file.
The lines in this file must be shorter than 639 characters.

Cluster

This section describes the setting values of job operation per cluster.

ClusterName

This specifies the cluster name.

LogLevel

This is a unique log output level for the job manager.

level	application
1	It is a usual log level. An important message that there is an influence in operation is output.
2	Detailed information for debugging in addition to level 1 is output. Detailed information is output to an internal log for the investigation.
3	Information for debugging is output more in detail compared with level 2.

You can set from 1 to 3 for the log output level.
If this level is not set, the value will be 1.

JstiOutData

This specifies whether the statistical information record is output.
When the statistical information record is output, yes is specified. When the statistical information record is not output, no is specified.
If this is not set, the value will be yes.

JstiRejectData

This specify the save setting of the information of the job for which job submission was rejected.
Specify yes to save information on jobs for which job submission was rejected, or no for not saving jobs.
If this is not set, the value will be no.

LLIOPerfData

This specifies whether data is output to the LLIO performance information record file.
When data is output to the LLIO performance information record file, yes is specified. When it is not output to the file, no is specified.
If the specification of whether data is output to the LLIO performance information record file is omitted, the value will be no.

KeepJobData

The storage period of job information(EXIT/CANCEL/REJECT) takes the day as the unit.
It can be set from 0 to 365. If the storage period of job information is not set, it will be set to 10.
However, the following operations delete the information of jobs in the EXIT/CANCEL/REJECT state:

- Adding, modifying, and deleting custom resource settings with the pmpjmadm command
- Modifying statistics settings with the papjmstatsadm command

RunJobAlterElapse

The valid/invalid of the function to change the elapsed time limit value for jobs in the RUNNING state. On is valid, off is invalid.
The default value is off.

PjdelNoHistory

For the deleted job that were in QUEUED state, this item sets whether to suppress job information output to the job history information which is output by the pjstat -H option.
Specify yes to suppress output or no to not suppress output. However, even if no is specified, the --no-history option of the pjdel command suppresses the output. The default value is no.

PjdelNoStats

For the deleted job that were in QUEUED state, this item sets whether to suppress output of the job statistical information file (.stats file).

Specify yes to suppress output or no to not suppress output. However, even if no is specified, the --no-stats option of the pjdel command suppresses the output. The default value is no.

SaveScript

The valid/invalid of saving ended job script files is specified. On is valid, off is invalid.

The default value is off.

TotalCoresUpper

This item specifies the target jobs on the total number of CPU use cores of each job or the total number of CPU use cores time product of each job. This setting is applied to the following job ACL limitations.

total-cores

total-cores-elapsed

interact-total-cores

interact-total-cores-elapsed

Value	Meaning
vnode	Only the virtual node allocated jobs are targets.
all	All jobs are targets. Number of the CPU use cores for the node allocated job is (number of CPU cores on a node) * (number of requested nodes).

The default value is "all".

ResourceUnit

This section describes setting values of job operation per resource unit.

LogLevel

This is a unique log output level for the job scheduler.

level	application
1	It is a usual log level. An important message that there is an influence in operation is output.
2	Detailed information for debugging in addition to level 1 is output. Detailed information is output to an internal log for the investigation.
3	Information for debugging is output more in detail compared with level 2.

You can set from 1 to 3 for the log output level. If this level is not set, the value will be 1.

Backfill

This specifies whether the backfill function (outstripping of the job) can be set. When outstripping of the job by the backfill function is enabled, yes is specified. When outstripping of the job by the backfill function is not enabled, no is specified.

If this is not specified, the value will be yes.

BackfillTarget

The job that the backfill function (outstripping of the job) targets is specified. When you judge whether to outstrip the job by the backfill function, the outstripping range of the job is specified.

Value	Meaning
rscunit	All jobs in the resource unit are outstripped.
rscgrp	All jobs in the resource group to which the job belongs are outstripped.

The default value is rscgrp.

DecidedGap

The time (buffer) added at the expected completion time of the job is specified.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:01:00

Grace

Specifies the grace time to start re-scheduling when the job ends.

When the value for which expected completion time of the job (Limit at job beginning time + elapsed time + time changed by DecidedGap) and an actual difference at the finish time are specified with Grace is exceeded, the job scheduler schedules the execution waiting job again.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:02:00

AdaptiveElapsedTimeJobTerminateGrace

Specify the forced termination delay time of the job that specified the minimum value of elapsed time limit when submitting the job.

If jobs specified as *elapsedtimemin*- or *elapsedtimemin-elapsedtimemax* at job submission exceeded the minimum elapsed time limit, they are forcibly terminated by starting another job or deadline using the same resource.

In this case, SIGTERM will be sent before forced termination, SIGKILL will be sent after *AdaptiveElapsedTimeJobTerminateGrace* has passed.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:00:10

CreateRscMap

The section set resource map making interval.

The format is "*time_zone,time_interval*". The format of *time_zone* and *time_interval* is "hour:minute:second".

time_zone is specified to the time which is from now time to ending time of counting time zone, *time_interval* is specified to resource map interval which is made in that time zone.

Resource map making interval can be specified in several numbers. If it is specified multiple, it must begin from the closest time zone apart from now.

Settable values:

time_zone

00:01:00 to 9999:59:59 or "*" (* will be set in the last line)

time_interval

00:01:00 to 9999:59:59

The default value is: (It is set as the following 4 lines)

01:00:00,00:10:00

24:00:00,00:30:00

240:00:00,01:00:00

*,24:00:00

SchedulePeriod

Scheduling period is set.

The time in the future which is later than current time added with this period will not be assigned.

Settable values: 00:01:00 to 9999:59:59. The format is "hour:minute:second".

The default value is 240:00:00.

DynamicSchedulePeriod

Dynamic scheduling period is set.

The format is "ratio,minimum scheduling period".

Determines the scheduling period by job maximum elapsed time limit instead of SchedulePeriod values.
Ratio is settable values: 1 to 2147484347.
Minimum scheduling period is settable values: 00:01:00 to 9999:59:59. The unit is time.
The default value is "2,24:00:00".

RestartNormal

Automatically re-executes a normal job by default.
If yes is specified, a normal job is automatically re-executed by default. If no is specified, a normal job is not automatically re-executed by default.
The default value is yes.

RestartStep

Automatically re-executes a step job by default.
If yes is specified, a step job is automatically re-executed by default. If no is specified, a step job is not automatically re-executed by default.
The default value is yes.

RestartBulk

Automatically re-executes a bulk job by default.
If yes is specified, a bulk job is automatically re-executed by default. If no is specified, a bulk job is not automatically re-executed by default.
The default value is no.

RestartMswk

Automatically re-executes a master-worker job by default.
If yes is specified, a master-worker job is automatically re-executed by default. If no is specified, a master-worker job is not automatically re-executed by default.
The default value is yes.

MailSend

Whether all related mails can be sent to jobs in the current resource unit is set.
When mails are sent, yes is specified. When mails are not sent, no is specified.
The default value is yes.

Fairshare

Specifies whether the fair share value of the all resource units in this cluster is reduced or not.
When on is specified, the fair share value will be reduced. When off is specified, the fair share value will not be reduced.
The default value is off.

FshareRecoveryValue

Fair share recovery value is set.
Job using resource (node) which is treated as fair share * elapse time is recovery quantity in the period. In the operation, after assuming scale job has been executed, the period which lasts until execution priority having completely recovered will be took as standard.
Settable values: 0 to 18446744073709551615.
The default value is 236.(*)

(*) The recovery quantity which JOB using 165888 nodes by 24 hours recovers in one week.
$$(165888 * 24 \text{ hours}) / (24 \text{ hours} * 7 \text{ days}) / \text{recovery ratio} 100 = 236$$

For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the fair share value is calculated using the maximum elapse limit.

UseCoreLimit

This item specifies the target jobs on the simultaneous number of the CPU use cores. This setting is applied to the following job ACL limitations.

ru-use-core
rg-use-core

ru-interact-use-core
rg-interact-use-core

Value	Meaning
vnode	Virtual node allocated jobs are targets.
all	All jobs are targets. Simultaneous number of the CPU use cores for node allocated jobs is (number of CPU cores on a node) * (number of requested nodes).

The default value is "vnode".

StartTimeGuarantee

The valid/invalid of the function to guarantee of planned execution start time is specified. On is valid, off is invalid.
The default value is on.

SuspendPlannedPeriod

Specifies the planned suspend period.

The planned suspend period is used to schedule suspended jobs. When a job is suspended, the job will be scheduled as if it will be resumed after the planned suspend period. The time that the job is assumed to be going to be resumed is called as "the planned resume time". If the job will have not been resumed by the planned resume time, the time will be postponed for the planned suspend period.

Settable values: 00:01:00 to 9999:59:59. The format is "hour:minute:second".

Default value: 01:00:00.

SuspendFailForceError

Specifies whether or not the job that failed to be suspended or to be resumed goes into ERROR state.

Value	Meaning
no	If the operation failed due to a recoverable error, the job will remain in its current state (RUNNING state or SUSPENDED state). If the operation failed due to an unrecoverable error, the job will be aborted. The aborted job will be re-executed if it is automatically re-executable.
yes	The job will be aborted and will go into ERROR state.

Default value: no

HoldAcceptDate

Specifies the interpretation of the job submission time in scheduling when the job is released from the HOLD state.

Value	Meaning
release	Time when the job is released from the HOLD state.
accept	Time when the job is submitted.

The default value is release.

ShelfPackingNodeAllocation

Only FX server is effective.

This specifies whether the free nodes on shelves allocated other I/O-sharing jobs are preferentially allocated for I/O-sharing jobs.

Value	Meaning
yes	When allocating 48 or fewer nodes to an I/O-sharing job, allocate free nodes preferentially on shelves where other I/O-sharing jobs are already allocated.
no	Allocate free nodes to the I/O-sharing job without considering the above.

Default is no.

JobSchedulingTargetLimit

This section specifies the maximum number of scheduled jobs, excluding interactive jobs. Jobs that are not to be scheduled this time will be rescheduled at the next scheduling time. Interactive jobs are always scheduled.

You can specify a value ranging from 0 to 2147483647 for the maximum number of jobs. If you specify 0, all jobs are scheduled. The default value is 10000.

JobSchedulingTargetMode

This section specifies the method of selecting the jobs for scheduling with JobSchedulingTargetLimit.

You can specify either `aprio` or `jobselectpolicy`.

The default value is `jobselectpolicy`.

- **aprio**

Determines the subjects of scheduling by order of job priority in the resource unit and the job submission time.

- **jobselectpolicy**

Determines the subjects of scheduling according to the specification in the JobSelectPolicy section. The following policies are supported. Other policies are ignored.

- `fcfs`
 - `node`
 - `group_prio`
 - `user_prio`
 - `job_aprio`
 - `rscgrp_prio`
 - `elapsed_limit`
 - `node_times_elapsed`
 - `at`
 - `interact`
 - `job_epoint`

JobSelectPolicy

This section set job select policy.

- **name**

The policy name is specified. The character number can be set less than 15.(single-byte alphabet and numeric character, hyphen and underscore)

If it is not set, it will use the default definition.

- **fcfs = { *order* [, *asc* | *desc*] | *off* }**

Whether need to do the job selection evaluation according to job submitted time. If this is not set, it needs to evaluate according to job submitted time.

- order*: the evaluating order of item(1~256)
 - asc*: sort in ascending order(default)
 - desc*: sort in descending order
 - off*: not evaluate

- **node = { *order* [, *asc* | *desc*] | *off* }**

Only FX server is effective.

Whether need to do the job selection evaluation according to required node number.

- order*: the evaluating order of item(1~256)
 - asc*: sort in ascending order(default)
 - desc*: sort in descending order
 - off*: not evaluate

- **group_prio = { *order* [, *asc* | *desc*] | *off* }**

Whether need to do the job selection evaluation according to group priority.

- order*: the evaluating order of item(1~256)
 - asc*: sort in ascending order
 - desc*: sort in descending order(default)
 - off*: not evaluate

- **user_prio = { *order* [, *asc* | *desc*] | *off* : }**

Whether need to do the job selection evaluation according to user priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

- **usr_in_grp_prio = { *order* [, *asc* | *desc*] | *off* : }**

Whether need to do the job selection evaluation according to user priority in the group.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set group_prio as the policy to be evaluated before the evaluation order set in usr_in_grp_prio. For example, if the evaluation order of usr_in_grp_prio is 5, set the evaluation order of group_prio to 4. When setting usr_in_grp_prio and usr_in_grp_faishare, be sure to evaluate usr_in_grp_prio and usr_in_grp_faishare consecutively, and be sure to set group_prio as the policy to be evaluated before the previous one. For example, set the evaluation order of usr_in_grp_prio and usr_in_grp_faishare to 5 and 6, and set the evaluation order of group_prio to 4. For all groups that run the job, set a different value for the group priority within the resource unit (define pri-g) of the job ACL function.

- **job_prio = { *order* [, *asc* | *desc*] | *off* : }**

Whether need to do the job selection evaluation according to job priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set user_prio as the policy to be evaluated before the evaluation order set in job_prio. For example, if the evaluation order of job_prio is 5, set the evaluation order of user_prio to 4. For all users that run the job, set a different value for the user priority within the resource unit (define pri) of the job ACL function.

- **job_aprio = { *order* [, *asc* | *desc*] | *off* }**

Whether need to do the job selection evaluation according to job priority in resource unit.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

- **rscgrp_prio = { *order* [, *asc* | *desc*] | *off* : }**

Whether need to do the job selection evaluation according to resource group priority.

The resource group priority is set in the ResourceGroup section.

order: the evaluating order of item(1~256)
asc: sort in ascending order
decs: sort in descending order(default)
off: not evaluate

- **group_fairshare = { *order* [, *asc* | *desc*] | *off* : }**

Whether need to do the job selection evaluation according to group fair share.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

- **usr_in_grp_fairshare = { *order* [, *asc* | *desc*] *off* : }**

Whether need to do the job selection evaluation according to user fair share in the group.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set group_prio as the policy to be evaluated before the evaluation order set in usr_in_grp_fairshare. For example, if the evaluation order of usr_in_grp_fairshare is 5, set the evaluation order of group_prio to 4. When setting usr_in_grp_prio and usr_in_grp_faishare, be sure to evaluate usr_in_grp_prio and usr_in_grp_faishare consecutively, and be sure to set group_prio as the policy to be evaluated before the previous one. For example, set the evaluation order of usr_in_grp_prio and usr_in_grp_faishare to 5 and 6, and set the evaluation order of group_prio to 4. For all groups that run the job, set a different value for the group priority within the resource unit (define pri-g) of the job ACL function.

- **user_fairshare = { *order* [, *asc* | *desc*] *off* : }**

Whether need to do the job selection evaluation according to user fair share.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

- **elapse_limit = { *order* [, *asc* | *desc*] *off* : }**

Whether need to do the job selection evaluation according to elapse limit.

For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the maximum elapse limit is used.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)
desc: sort in descending order
off: not evaluate

- **node_times_elapse = { *order* [, *asc* | *desc*] *off* : }**

Only FX server is effective.

Whether need to do the job selection evaluation according to node times accumulation.

For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the node times accumulation is calculated with the maximum elapse limit.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)
desc: sort in descending order
off: not evaluate

- **at = { *order* [, *true* | *false*] *off* : }**

Whether need to do the job selection evaluation according to execution start time.

order: the evaluating order of item(1~256)
true: has a preference for execution start time(default)
false: has no preference for execution start time
off: not evaluate

- **interact = { *order* [, *true* | *false*] *off* : }**

Whether need to do the job selection evaluation according to submitting interact jobs.

order: the evaluating order of item(1~256)
true: has a preference for interact jobs(default)
false: has a preference for non-interact jobs
off: not evaluate

- **job_epoint = { *order* [, *asc* | *desc*] *off* }**

This specifies whether there is a comprehensive evaluation using multiple items.

order: the evaluating order of item(1~256)
asc: sort in ascending order

desc: sort in descending order(default)
off: not evaluate

JobEvaluation

This section defines how to calculate job evaluation points (job evaluation definitions) regarding the job selection policy item job_epoint.

The job evaluation definitions can be specified up to 32.

- **name = job evaluation definition name**

Name for a job evaluation definition.

The job evaluation definition name must consist of 1 to 127 characters. Also, each job evaluation definition name must be unique within its cluster.

A job evaluation definition is defined as a cluster setting in the papjm.conf file. You can use that definition in the pmpjm.conf file, which has settings for each resource unit, by specifying the job evaluation definition name.

If the setting of a job evaluation definition name is omitted, the common definition within the cluster is used.

- **item name = weight[, value 1@point 1: value 2@point 2: ... : value N@point N]**

This defines the points corresponding to the values of the items to evaluate.

The definition is written in the "value@point" format. When writing multiple definitions, delimit them with a colon (:).

You can specify up to 64 definitions.

The points for items that have numerical values are determined as follows.

Value	Point
-----	-----
Minimum value for item value <= value 1	point 1
value 2 < value <= value 3	point 2
...	...
value N-1 < value <= value N	point N

Define each point such that $value\ n-1 < value\ n$ is true.

The points for items that do not have a numerical value are indicated by the definitions with their corresponding values.

Also, the values of multiple definitions must differ from one another.

The specifiable item names and values are as follows.

Item	Meaning	Range of specifiable values
node elapse_limit waittime	Number of nodes requested by a job Elapsed time limit value for job execution (seconds) Time from job acceptance to the start of job scheduling or rescheduling (seconds)	1 to INT32_MAX
group_prio user_prio usr_in_grp_prio job_prio job_aprio rscgrp_prio	Priority of the group to which the user (job submitter) belongs Priority of the user (job submitter) Priority of the user (job submitter) within the group Job priority Job priority within the resource unit Resource group priority	0 to 255
node_times_elapse	Product of the number of nodes requested by a job (node) and the elapsed time limit value for job execution (elapse_limit)	1 to INT64_MAX
at interact	Whether the execution start time is specified Whether the job is an interactive job	yes or no

If "*" is the specified value, it means that the value is greater than $value\ n-1$ (for a numerical value) or other than $value\ n-1$ (for a non-numerical value).

[Example] 10@1:20@2:*:3

If the value is greater than 20, the point is point 3.

[Example] no@0,*@1

If the value is other than no, the point is point 1.

You can specify a value ranging from 1 to 1024 for the point.

The job evaluation points are the totals from the multiplication of *weight* for the point of each item. If the point scales differ according to the item or if you want to emphasize a specific item, use *weight* to adjust the balance between different points. You can specify a value ranging from -2147483648 to 2147483647 for *weight*.

If the total of an individual point is below the lower point limit (-9223372036854775807) or over the upper point limit (9223372036854775807), the job evaluation point is the lower limit or upper limit, respectively.

[EXAMPLE]

A definition example of the papjm.conf file is as follows:

```
$ cat /etc/opt/FJSVtcs/papjm.conf
Cluster {
    ClusterName = clusterA
    LogLevel = 1
    JstiOutData = yes
    JstiRejectData = no
    KeepJobData = 10
    SaveScript = off
    RunJobAlterElapse = off
    PjdelNoHistory = no
    PjdelNoStats = no
    ResourceUnit {
        LogLevel = 1
        Backfill = yes
        BackfillTarget = rscgrp
        DecidedGap = 00:01:00
        Grace = 00:02:00
        CreateRscMap = "01:00:00, 00:10:00"
        CreateRscMap = "24:00:00, 01:00:00"
        CreateRscMap = "*", 24:00:00"
        SchedulePeriod = 25:00:00
        DynamicSchedulePeriod = 2,24:00:00
        RestartNormal = yes
        RestartStep = yes
        RestartBulk = no
        RestartMswk = yes
        MailSend = yes
        Fairshare = off
        FshareRecoveryValue = 236
        StartTimeGuarantee = on
        HoldAcceptDate = release
        ShelfPackingNodeAllocation = no
        JobSchedulingTargetLimit = 10000
        JobSchedulingTargetMode = jobselectpolicy
    }
}
# default Job Select Policy
JobSelectPolicy {
    job_aprio = 2,desc
#    group_fairshare = 3,desc
#    user_fairshare = 4,desc
#    group_prio = 5,desc
#    usr_in_grp_prio = 6,desc
#    usr_in_grp_fairshare = 7,desc
#    interact = 8,true
#    node_times_elapse = 9,asc
#    node = 10,asc
#    elapse_limit = 11,asc
#    user_prio = 12,desc
#    job_prio = 13,desc
```

```

#         at =                      14,true
#         fcfs =                    256,asc
#     }
# Job Select Policy 1
#     JobSelectPolicy {
#         name = policy1
#         group_fairshare =          3,desc
#         fcfs =                    256,asc
#     }
# Job Select Policy 2
#     JobSelectPolicy {
#         name = policy2
#         fcfs =                    256,asc
#         job_aprio =                2,desc
#     }
# Job Select Policy 3
#     JobSelectPolicy {
#         name = policy3
#         group_fairshare =          3,desc
#         group_prio =              4,desc
#         usr_in_grp_fairshare =    5,desc
#         fcfs =                    256,asc
#     }
# Job Select Policy 4
#     JobSelectPolicy {
#         name = policy4
#         job_epoint =              1,desc
#         fcfs =                    255,asc
#     }
#     JobEvaluation {
#         name = jobevall
#         waittime = 1,1800@1:*@2
#     }
# }

```

[SEE ALSO]

papjmadm(8)

4.3.6 papjmstats.conf file

[NAME]

papjmstats.conf - File that defines statistical information settings in a cluster

[SYNOPSIS]

/etc/opt/FJSVtcs/papjmstats.conf

[DESCRIPTION]

The papjmstats.conf file, which is placed on the system management node, defines the statistical information settings of an entire cluster. The settings are reflected by the papjmstatsadm command executed on the system management node. The cluster administrator sets the definitions in this file. The number of characters on each line in this file must be 511 or less.

Cluster

This section writes values for statistical information settings in a cluster.

Record

This section writes the settings to save in a statistical information record.

PATH

This item specifies the path to the file that is the save destination for statistical information records.

* If duplicate PATHs are defined within the same Record section, an error occurs at the setting time.

Similarly, if different Record sections specify the same path, an error occurs at the setting time.

Command

This section writes the display setting of `pjstat -s/-S`.

* If this section is defined in duplicate, an error occurs at the setting time.

File

This section writes the settings to save in a statistical information file (.stats file).

* If this section is defined in duplicate, an error occurs at the setting time.

You can write the following sections in the above Record, Command, and File sections.

JI

This section writes the settings to save for job statistical information.

* If this section is defined in duplicate within the same section, an error occurs at the setting time.

JN

This section writes the settings to save for node statistical information.

* If this section is defined in duplicate within the same section, an error occurs at the setting time.

You can write either of the following items in the above JI and JN sections.

- ITEM

This item specifies the statistical information items to save.

- NOT_SAVE_ITEM

This item specifies the statistical information items not to save.

* If unable to write ITEM or NOT_SAVE_ITEM on a single line, you can split the contents to write them in multiple ITEM or NOT_SAVE_ITEM items.

Item

This section writes information on an item newly defined by the job operation administrator.

One of the following items is set.

* If any item is defined in duplicate within the same section, an error occurs at the setting time.

Definition item name	Description	Value	Default value
ItemName	Identification name used as follows to identify an item defined by the job operation administrator: - Setting of a value for an item defined by the job operation administrator - Setting to save for statistical information	Character string with up to 79 characters. It can use single-byte alphanumeric characters, a hyphen, and an underscore. The string cannot begin with a hyphen. Since names that begin with "CR-" or "pjm-" are reserved words, they cannot be used to setting of a value for an item defined by the job operation administrator.	Cannot be omitted
ItemNameDisp	Item name displayed by <code>pjstat -s/-S</code> and a statistical information file (.stats file). [Example] Displayed by <code>pjstat -s PRIORITY: 127</code>	Character string with up to 79 characters. It can use single-byte alphanumeric characters, a hyphen, and an underscore.	Identification name defined in ItemName

Definition item name	Description	Value	Default value
	The above example is the part showing PRIORITY.		
RecordNameList	Specifies the records for setting the values of the relevant items.	Specify records by delimiting them with a comma (.). You can specify the following 2 types of records: JI: Job statistical information JN: Node statistical information	JI,JN
DataType	Specifies a value type for the relevant item.	See (*1).	PJMX_DATATYPE_UINT64
DispFormat	Specifies a display format for pjstat -s/-S and the statistical information file (.stats file) of the relevant item.	See (*2). If the selected display format is not compatible with DataType, this item is disabled.	Default display format of each DataType
DispFormatOption	Can specify an actual system of units when dec, an SI prefix, or a binary prefix is specified in DispFormat. [Example] B written: GB, MB, KB. W written: KW, MW, KW.	Character string with up to 15 characters. If the selected DispFormat is other than dec, prefix_kilo, prefix_kibi, prefix_mega, prefix_mebi, prefix_giga, prefix_gibi, prefix_tera, or prefix_tebi, this item is disabled.	System of units not added

(*1) You can specify the following values in DataType.

Specified value	Meaning	With DispFormat not specified and display by pmdumpjobinfo
PJMX_DATATYPE_CHAR	The item has a char type value.	Displays specified characters. DispFormat is disabled.
PJMX_DATATYPE_INT8	The item has an int8 type value.	Shown as decimal
PJMX_DATATYPE_UINT8	The item has a uint8 type value.	Shown as decimal
PJMX_DATATYPE_INT16	The item has an int16 type value.	Shown as decimal
PJMX_DATATYPE_UINT16	The item has a uint16 type value.	Shown as decimal
PJMX_DATATYPE_INT32	The item has an int32 type value.	Shown as decimal
PJMX_DATATYPE_UINT32	The item has a uint32 type value.	Shown as decimal
PJMX_DATATYPE_INT64	The item has an int64 type value.	Shown as decimal
PJMX_DATATYPE_UINT64	The item has a uint64 type value.	Shown as decimal
PJMX_DATATYPE_FLOAT	The item has a float type value.	6 digits after decimal point
PJMX_DATATYPE_DOUBLE	The item has a double type value.	6 digits after decimal point
PJMX_DATATYPE_TIMESPEC	The item has a struct timespec type value.	In units of nsec
PJMX_DATATYPE_TIME	The item has a time_t type value.	Shown as decimal
PJMX_DATATYPE_SIZE	The item has a size_t type value.	Shown as decimal

Specified value	Meaning	With DispFormat not specified and display by pmdumpjobinfo
PJMX_DATATYPE_STRING	The item has a char * type value.	Displays specified characters. DispFormat is disabled.

(*2) You can specify the following values in DispFormat.

Specified value	Meaning
dec	<p>Display in decimal notation.</p> <p>This is valid when DataType is one of the following:</p> <p>PJMX_DATATYPE_INT8 PJMX_DATATYPE_UINT8 PJMX_DATATYPE_INT16 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_INT32 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_INT64 PJMX_DATATYPE_UINT64 PJMX_DATATYPE_FLOAT PJMX_DATATYPE_DOUBLE PJMX_DATATYPE_TIME PJMX_DATATYPE_SIZE</p>
bin	<p>Display in binary notation.</p> <p>Numbers are not padded with 0's.</p> <p>[Example] Displayed for a value of 255 11111111</p> <p>This is valid when DataType is one of the following:</p> <p>PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64</p>
bin_padding	<p>Display in binary notation.</p> <p>Numbers are padded with 0's according to the number of bits of DataType.</p> <p>[Example] Displayed for a value of 255 8 bits: 11111111 16 bits: 0000000011111111 32 bits: 00000000000000000000000011111111 64 bits: 0011111111</p> <p>This is valid when DataType is one of the following:</p> <p>PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64</p>
oct	<p>Display in octal notation.</p> <p>Numbers are not padded with 0's.</p> <p>[Example] Displayed for a value of 255 0377</p> <p>This is valid when DataType is one of the following:</p> <p>PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64</p>
oct_padding	<p>Display in octal notation.</p> <p>Numbers are padded with 0's according to the number of bits of DataType.</p> <p>[Example] Displayed for a value of 255</p>

Specified value	Meaning
	8 bits: 0377 16 bits: 0000377 32 bits: 000000000377 64 bits: 0000000000000377 This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8 PJM_X_DATATYPE_UINT16 PJM_X_DATATYPE_UINT32 PJM_X_DATATYPE_UINT64
hex	Display in hexadecimal notation. Numbers are not padded with 0's. [Example] Displayed for a value of 255 0xFF This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8 PJM_X_DATATYPE_UINT16 PJM_X_DATATYPE_UINT32 PJM_X_DATATYPE_UINT64
hex_padding	Display in hexadecimal notation. Numbers are padded with 0's according to the number of bits of DataType. [Example] Displayed for a value of 255 8 bits: 0xFF 16 bits: 0x00FF 32 bits: 0x000000FF 64 bits: 0x00000000000000FF This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8 PJM_X_DATATYPE_UINT16 PJM_X_DATATYPE_UINT32 PJM_X_DATATYPE_UINT64
prefix_kilo	Display in SI prefix kilo notation. Values before conversion are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1024 1.1 K (1024) This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8 PJM_X_DATATYPE_UINT16 PJM_X_DATATYPE_UINT32 PJM_X_DATATYPE_UINT64
prefix_kibi	Display in binary prefix kibi notation. Values before conversion are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1024 1.0 Ki (1024) This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8 PJM_X_DATATYPE_UINT16 PJM_X_DATATYPE_UINT32 PJM_X_DATATYPE_UINT64
prefix_mega	Display in SI prefix mega notation. Values before conversion are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1048576 1.1 M (1048576) This is valid when DataType is one of the following: PJM_X_DATATYPE_UINT8

Specified value	Meaning
	PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
prefix_mebi	Display in binary prefix mebi notation. Values before conversion are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1024 1.0 Mi (1024) This is valid when DataType is one of the following: PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
prefix_giga	Display in SI prefix giga notation. Values in units of mega are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1073741824 1.1 G (1074) This is valid when DataType is one of the following: PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
prefix_gibi	Display in binary prefix gibi notation. Values in units of mebi are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 1073741824 1.0 Gi (1024) This is valid when DataType is one of the following: PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
prefix_tera	Display in SI prefix tera notation. Values in units of giga are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 109951162776 1.1 T (1100) This is valid when DataType is one of the following: PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
prefix_tebi	Display in binary prefix tebi notation. Values in units of gibi are shown in parentheses. The displayed values have been rounded up to the nearest tenth. [Example] Displayed for a value of 109951162776 1.0 Ti (1024) This is valid when DataType is one of the following: PJMX_DATATYPE_UINT8 PJMX_DATATYPE_UINT16 PJMX_DATATYPE_UINT32 PJMX_DATATYPE_UINT64
date	Date display. The display format is YYYY/MM/DD hh:mm:ss. [Example] Displayed for a value of 0 1970/01/01/ 00:00:00 This is valid when DataType is the following: PJMX_DATATYPE_TIMESPEC PJMX_DATATYPE_TIME


```

    }
    File {
        JI {
            ITEM = jid,jnam,elpl,nnuma
        }
    }
    Item {
        ItemName = ITEM_A
        RecordNameList = JI,JN
    }
    Item {
        ItemName = ITEM_B
        RecordNameList = JI
    }
    Item {
        ItemName = ITEM_C
        RecordNameList = JN
    }
}

```

[SEE ALSO]

papjstatsadm(8)

4.3.7 papwr.conf file

[NAME]

papwr.conf - File that defines the system control settings

[SYNOPSIS]

/etc/opt/FJSVtcs/papwr.conf

[DESCRIPTION]

papwr.conf is a file that defines the settings for power control. The default values are used for undefined clusters.

Cluster

This section defines a cluster.

ClusterName

Define a cluster name. This definition is required.

If the specified cluster does not exist, the entire Cluster section is ignored as an invalid section.

Model

This section defines a model.

Model

Specify a model name. This definition is required.

You can define the following models by the following model names:

PRIMERGY server: PG

The operation of a model not defined with Model uses the default values for IPMI user name, IPMI password for console connection.

User

Write an IPMI user name. You can omit this definition.

If this definition is omitted, the name is the default value of admin.

This definition becomes effective when PG is specified as Model.

Password

Write an IPMI password.

User and Password must match the corresponding settings of the main unit of the nodes subject to power control.

This definition becomes effective when PG is specified as Model.

OptionLimit

This section defines the default values and limits for the setting values of options.

Concurrency

This section defines the default values and limits for the number of concurrent executions.

Control

Describe the target power operation. This definition is required.

Specify any of the following: on, off, force-off, cold-reboot, reset, cycle, init, and warm-reboot. You can specify multiple values by delimiting them with a comma. "on" also applies to the pastart command. "off" also applies to the pastop command without the --emergency option specified. "force-off" also applies to the pastop command with the --emergency option specified.

Default

The default value for the number of concurrent executions specified in the -P option of the papwrctl command, pastart command, and pastop command changes to this specified value.

You can omit this definition. When you omit this definition, it is judged that 24 is set for PRIMERGY server and nodes of arbitrary model. In the case of an FX server, it is assumed to be set to 30.

When the number of nodes in the cluster is less than these values, the number of nodes in the cluster is set.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

Default = *num* [: *model*]

UpperLimit

Specify the upper limit on the number of concurrent executions specified in the -P option of the papwrctl command, pastart command, and pastop command.

If specified with a value larger than the upper limit, the papwrctl command terminates abnormally. You can omit this definition.

When you omit this definition, it is judged that number of nodes in the cluster is set, and operates.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

UpperLimit = *num* [: *model*]

LowerLimit

Specify the lower limit on the number of concurrent executions specified in the -P option of the papwrctl command, pastart command, and pastop command.

If specified with a value smaller than the lower limit, the papwrctl command terminates abnormally. You can omit this definition.

When you omit this definition, it is judged that 1 is set, and operates.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

LowerLimit = *num* [: *model*]

Interval

This section defines the default values and limits for an interval.

Control

Describe the target power operation. This definition is required.

Specify any of the following: on, off, force-off, cold-reboot, reset, cycle, init, and warm-reboot. You can specify multiple values by delimiting them with a comma. "on" also applies to the pastart command. "off" also applies to the pastop command without the --emergency option specified. "force-off" also applies to the pastop command with the --emergency option specified.

Default

The default value with millisecond for the interval specified in the --interval option of the papwrctl command, pastart command, and pastop command changes to this specified value. You can omit this definition. When you omit this definition, it is judged that 0 is set, and operates.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

Default = *num* [: *model*]

UpperLimit

Specify the upper limit on the interval specified in the --interval option of the papwrctl command, pastart command, and pastop command with millisecond.

If specified with a value larger than the upper limit, the papwrctl command terminates abnormally. You can omit this definition. When you omit this definition, it is judged that 86400000 is set, and operates.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

UpperLimit = *num* [: *model*]

LowerLimit

Specify the lower limit on the interval specified in the --interval option of the papwrctl command, pastart command, and pastop command with millisecond.

If specified with a value smaller than the lower limit, the papwrctl command terminates abnormally. You can omit this definition. When you omit this definition, it is judged that 0 is set, and operates.

You can vary the specified value by model. If no model is specified, all models are targets.

SYNOPSIS

LowerLimit = *num* [: *model*]

Plugin

This section defines the power control plugin.

Model

Specify a model name. This definition is required.

You can specify arbitrary model name that exists in the node configuration information.

Path

Write the full path of the power control plugin command. This definition is required.

Power control on nodes of arbitrary model by the papwrctl command uses the written power control plugin command information.

[EXAMPLE]

The following example shows how to define the papwr.conf file.

```
Cluster {
  ClusterName = cluster1
  Model {
    Model = PG
    User = admin
    Password = admin
  }
  OptionLimit {
    Concurrency {
      Control = on
      Default = 48:FT
      UpperLimit = 96:FT
      LowerLimit = 1:FT
    }
    Concurrency {
      Control = off,force-off
    }
  }
}
```

```

        Default = 480:FT
        UpperLimit = 960:FT
        LowerLimit = 1:FT
    }
    Interval {
        Control = on
        Default = 100:FT
        UpperLimit = 10000:FT
        LowerLimit = 100:FT
    }
    Interval {
        Control = off,force-off
        Default = 0:FT
        UpperLimit = 10000:FT
        LowerLimit = 0:FT
    }
}
Plugin {
    Model = VE
    Path = /root/pwrctl_virtual
}
}

```

[SEE ALSO]

papwradm(8)

4.3.8 papwrm.conf file

[NAME]

papwrm.conf - File that defines the operation of the power management function

[SYNOPSIS]

/etc/opt/FJSVtcs/papwrm.conf

[DESCRIPTION]

The papwrm.conf file defines setting values of the power management function.

Format

Use the following synopsis to write papwrm.conf:

```

Section_name {
    Item1 = "Setting_value1"
    Item2 = "Setting_value2"
    Subsection_name {
        Item3 = "Setting_value3"
        Item4 = "Setting_value4"
    }
}

```

A section name is a keyword indicating what structure is defined by the part enclosed in curly brackets { }. Section names have already been determined according to what is defined. A section may contain subsections, depending on the defined items. Writing two or more same items in a section or subsection causes an error. Quote Setting value using double quote("").

SystemPower

This section specifies the setting values of the system power data collection function. If two or more sections are set, an error occurs when the settings are attempted to be applied, and they are not applied. If this section is omitted, the function stops collecting system power data. If you are configuring multiple clusters, configure the Cluster subsection separately for each cluster. If you are configuring multiple resource units, configure the ResourceUnit subsection separately for each resource unit as well.

In this section, set the following items.

StartTime

Write the number of seconds (0-59) for the start of processing to collect power consumption information and calculate power consumption.

The processing to collect power consumption information and calculate power consumption begins at the specified time (in seconds) every minute. For example, when the value is 40, the processing to collect starts at 00min 40sec, 01min 40sec, and so on.

Default value is 30.

LogLevel

Write a log level.

Log level	Usage
1	Output of information messages only
2	Debug-level messages Debug information condensed as much as possible in output considering system load
3	Debug-level messages The most detailed debug information that is output without considering system load

Default value is 1.

AcceptableRange

Write, in units of seconds (0-3600), an acceptable range of variation in the measurement time of power consumption used to calculate a total value.

The basic guidelines are that the range be at least 60 seconds, which is the operation interval of system power data collection, and the minimum period among the power measurement intervals for external devices.

Default value is 600.

CommandLine

Write a command to collect external device power consumption data (command full path and arguments).

The power data collection function that operates on the active system management node calls and executes this command.

The command full path is a character string that does not contain commas (,) and is a valid file path name. If an invalid value or nothing is specified, no information is collected about the power consumption of external devices.

Command full path is valid character string with up to 255 characters. The command arguments are character strings with up to 1,023 characters.

When the system management node is redundantly configured, place the same file on the standby system management node.

You can omit this item. However, to register an external device name in an external device file (/etc/opt/FJSVtcs/pwrn/extdev), the item is required.

PowerGroup

This subsection specifies power group setting values.

You can write multiple instances of this subsection. You can set one power group in each PowerGroup section. Up to 10,000 power groups can be registered.

If this subsection is omitted, power groups are not registered.

In this subsection, set the following items.

PowerGroupName

Write a power group name.

A power group name is a character string consisting of 1 to 63 characters. The name can use alphanumeric characters (uppercase/lowercase), a hyphen (-), and an underscore (_). However, a hyphen cannot be specified at the beginning of the string.

The power group name cannot be the same as one specified in another PowerGroup subsection. A duplicate name causes an error.

This item cannot be omitted.

ClusterName

Write the compute cluster name of the compute nodes to be registered in the power group. The nodes belong to that compute cluster.

You can omit this item. However, to register compute nodes in the power group, the item is required.

NodeList

Write the name of the file containing the node IDs of the compute nodes to be registered in the power group.

Specify a file name by removing the directory name (/etc/opt/FJSVtcs/pwrn). For the file name, write a character string that consists of up to 255 characters and does not contain the directory name, a comma (,), or a slash (/).

Node IDs of multiple compute nodes can be written in the file by separating them with line breaks. A hyphen (-) can be used to specify a range (e.g., 0x01010001-0x0101000F).

If this item is specified, the compute nodes which are specified are registered in the power group. If this item is omitted but Cluster Name is written, all the compute nodes belonging to the compute cluster specified in ClusterName are registered.

This item requires to set the item ClusterName.

ExternalDeviceList

Write the name of the file containing the external device names of the external devices to be registered in the power group.

Specify a file name by removing the directory name. For the file name, specify a character string that consists of up to 255 characters and does not contain the directory name, a comma (,), or a slash (/).

Only the external equipment names registered in /etc/opt/FJSVtcs/pwrn/extdev can be written in the file.

This item can be omitted.

PowerSave

This section is for writing the setting values of the auto power control function for compute nodes.

If two or more sections are set, an error occurs when the settings are attempted to be applied, and they are not applied. If this section is omitted, the auto power control function is disabled. If the auto power control function is set to multiple clusters, Cluster subsection must be defined for each cluster. As the same, if the auto power control function is set to multiple resource units, ResourceUnit subsection must be defined for each resource unit.

In this section, set the following items.

Cluster

In this subsection, write about the compute cluster targeted by the auto power control function.

This subsection is optional. The omitted calculation cluster calculation node automatic power control feature is disabled.

In this subsection, set the following items.

ClusterName

Write the name of the compute cluster targeted by the auto power control function.

If a non-existent cluster name is specified, an error occurs. This item can not be omitted.

LogLevel

Write a log level.

Log level	Usage
1	Output of information messages only
2	Debug-level messages Debug information condensed as much as possible in output considering system load
3	Debug-level messages The most detailed debug information that is output without considering system load

The default value is 1.

ResourceUnit

In this subsection, write about the resource unit targeted by the auto power control function.

This subsection is optional. The skipped resource unit calculation node auto power control function is disabled.

In this subsection, set the following items.

Item name	Description
ResourceUnitName	Write the name of the resource unit targeted by the auto power control function. If a non-existent resource unit name is specified, an error occurs. This item can not be omitted.

Item name	Description
NodeList	<p>Specify the file containing the target compute node IDs for the auto power control function for compute nodes.</p> <p>Specify a file name by removing the directory name (/etc/opt/FJSVtcs/pwrm/). The file name is a character string that consists of up to 255 characters and does not contain a comma (,) or slash (/).</p> <p>The file is located directly under /etc/opt/FJSVtcs/pwrm/ on the system management node.</p> <p>You can specify multiple compute node IDs in the file by using a line break. You can specify a range of IDs by using a hyphen (-). (Example: 0x01FF0001-0x01FF0010)</p> <p>An invalid or duplicate compute node ID causes an error. However, it does not cause an error if it is in a specified range of compute node.</p> <p>Multiple files cannot be specified.</p> <p>If this item is omitted, all the compute nodes in the resource unit are targets of the auto power control function for compute nodes.</p>

JobPowerEstimation

This section is for writing the setting values of the job power estimate function and power limit scheduling function (both are referred to below as the job power estimate function). If two or more sections are set, an error occurs when the settings are attempted to be applied, and they are not applied. If this section is omitted, the job power estimate function and power limit scheduling function are disabled. If you are configuring multiple clusters, configure the Cluster subsection separately for each cluster. If you want to configure more than one resource unit, you can also use the Separate settings for each source unit.

In this section, set the following items.

Cluster

In this subsection, write about the compute cluster targeted by the job power estimate function.

You can omit this subsection. The job power estimate function is disabled for omitted compute clusters.

In this subsection, set the following items.

ClusterName

Write the name of the compute cluster targeted by the job power estimate function.

If a non-existent compute cluster name is specified, the Cluster subsection is ignored as invalid subsection.

If a non-existing compute cluster name is specified, an error occurs.

This item can not be omitted.

LogLevel

Write a log level.

Log level	Usage
1	Output of information messages only
2	Debut-level messages Debug information condensed as much as possible in output considering system load
3	Debug-level messages The most detailed debug information that is output without considering system load

The default value is 1.

DbAuth

Write the user name and password used when connecting to the job power database.

Set the user name and password connected by a colon (:). (Example: user:password)

This item can nor be omitted.

ResourceUnit

In this subsection, write about the resource unit targeted by the job power estimate function. You can specify multiple resource units as targets of the job power estimate function by specifying multiple instances of this subsection.

For a resource unit that is not specified to this subsection, the job power estimate function is disabled.
In this subsection, set the following items.

Item name	Description
ResourceUnitName	Write a resource unit name. If a non-existing resource unit is specified, an error occurs. This item can not be omitted.
DefaultNodePower	Write job power (0-100000, Unit:W) per compute node. If the job power estimate function failed to estimate job power, the DefaultNodePower value is used as the power estimate value. The value is also used for interactive jobs and virtual node allocated jobs. This item can not be omitted.
ErrorNoticeMode	Write the return value for indicating the abnormal end of a function of the job power estimate library. 0: An abnormal end is reported by a return value. 1: An abnormal end is not reported by a return value. The job can be executed. The default value is 0.

PowerKnob

This section is for writing the setting values for power control with the power knob at job start and end and also for the operable range of Power API (power knob operations) by users. This section can be set for FX servers. They are referred below to as the power knob operation function.

If two or more sections are set, an error occurs when the settings are attempted to be applied, and they are not applied. If this section is omitted, power is not controlled with the power knob at job start and end, and the operable range of Power API by users is not set.

If you are configuring multiple clusters, configure the Cluster subsection separately for each cluster. If you are configuring multiple resource units, configure the ResourceUnit subsection separately for each resource unit as well.

In this section, set the following items. For details of each items, refer to the manual "Job Operation Software API user's Guide for Power API."

Cluster

In this subsection, write about the compute cluster targeted by the power knob operation function. For the cluster that is not specified to this subsection, power is not controlled with the power knob. Likewise, the operable range of power knob operations by users is not set.

In this subsection, set the following items.

ClusterName

Write the name of the compute cluster targeted by the power knob operation function.

If a non-existent cluster name is specified, an error occurs.

This item can not be omitted.

LogLevel

Write a log level.

Log level	Usage
1	Output of information messages only
2	Debug-level messages Debug information condensed as much as possible in output considering system load
3	Debug-level messages The most detailed debug information that is output without considering system load

The default value is 1.

ResourceUnit

In this subsection, write about the resource unit targeted by the power knob operation function. You can specify multiple resource units as targets of the power knob operation function by specifying multiple instances of this section.

For the resource unit that is not specified this subsection, the power knob operation function is disabled.
In this subsection, set the following items.

- **ResourceUnitName**

Write the name of the resource unit targeted by power knob operation function. If a non-existent resource unit name is specified, an error occurs.

This item can not be omitted.

- **AllowSharedKnob**

Specify whether to allow users to perform power knob operations in cases with a node-sharing job.

0: Do not allow users of node-sharing jobs to perform power knob operations.

1: Allow even users of node-sharing jobs to perform power knob operations.

If this setting is 1, users in the same node can operate the power knob as they like.

The default value is 0.

- **ComputeNode**

In this subsection, write power knob values for compute nodes.

Item name	Description
IdleState	<p>This section has the power knob settings for the node idle state.</p> <p>freq Sets a CPU frequency (MHz). This item can not be omitted. The value from 0 to 10000 is can be specified.</p> <p>throttling_state Restricts HBM access. The value from 0 to 9 is can be specified. The value 0 means no restriction (maximum performance). The value 9 means the maximum restriction (10% of the number of requests, minimum performance). This item can not be omitted.</p> <p>issue_state Restricts instructions issued for compute cores. The value 0 or 1 is can be specified. 0: 4 instructions (maximum performance) 1: 2 instructions (minimum performance) This item can not be omitted.</p> <p>ex_pipe_state Restricts the EX number for compute cores. The value 0 or 1 is can be specified. 0: Use AB (maximum EX number, maximum performance). 1: Use only A.(minimum performance). This item can not be omitted.</p> <p>eco_state Sets an eco mode state for compute cores. The value from 0 to 2 is can be specified. 0: Turn off eco mode and use FLAB (maximum performance). 1: Turn off eco mode and use only FLA. 2: Turn on eco mode and use only FLA (minimum performance). This item can not be omitted.</p> <p>retention_state Sets a Retention state for compute cores. The value 0 or 1 is can be specified. 0: Non-Retention mode 1: Retention mode This item can not be omitted.</p> <p>retention_state_acores Sets a Retention state for assistant cores excluding Core 0. The value 0 or 1 is can be specified. 0: Non-Retention mode 1: Retention mode This item can not be omitted.</p>

Item name	Description
RunningState	<p>This subsection has power knob specifications for job execution on compute nodes. The same items as the IdleState subsection can be set.</p> <p>The following items restrict the power knob values that can also be specified by users in power knob operations.</p> <p>freq_min Sets the lower limit for the CPU frequency (MHz). If this item is omitted, the value of the item freq is used. The value from 0 to 10000 is can be specified.</p> <p>freq_max Sets the upper limit for the CPU frequency (MHz). If this item is omitted, the value of the item freq is used. The value from 0 to 10000 is can be specified.</p> <p>throttling_state_min Sets the lower limit for restricting HBM access. If this item is omitted, the value of the item throttling_state is used. The value from 0 to 9 is can be specified.</p> <p>throttling_state_max Sets the upper limit for restricting HBM access. If this item is omitted, the value of the item throttling_state is used. The value from 0 to 9 is can be specified.</p> <p>issue_state_min Sets the lower limit for restricting instructions issued for compute cores. If this item is omitted, the value of the item issue_state is used. The value 0 or 1 is can be specified.</p> <p>issue_state_max Sets the upper limit for restricting instructions issued for compute cores. If this item is omitted, the value of the item issue_state is used. The value 0 or 1 is can be specified.</p> <p>ex_pipe_state_min Sets the lower limit for restricting the EX number for compute cores. If this item is omitted, the value of the item ex_pipe_state is used. The value 0 or 1 is can be specified.</p> <p>ex_pipe_state_max Sets the upper limit for restricting the EX number for compute cores. If this item is omitted, the value of the item ex_pipe_state is used. The value 0 or 1 is can be specified.</p> <p>eco_state_min Sets the lower limit for the eco mode state of compute cores. If this item is omitted, the value of the item eco_state is used. The value from 0 to 2 is can be specified.</p> <p>eco_state_max Sets the upper limit for the eco mode state of compute cores. If this item is omitted, the value of the item eco_state is used. The value from 0 to 2 is can be specified.</p> <p>retention_state_min Sets the lower limit for the Retention state of compute cores. If this item is omitted, the value of the item retention_state is used. The value 0 or 1 is can be specified.</p> <p>retention_state_max Sets the upper limit for the Retention state of compute cores. If this item is omitted, the value of the item retention_state is used. The value 0 or 1 is can be specified.</p>

IONode

In this subsection, write power knob values for compute nodes that also serve as I/O nodes. The same items as the ComputeNode subsection can be set.

This subsection is described to set a value different from the compute node about a certain item. If this subsection or its item is omitted, the values specified in ComputeNode are used.

Item name	Description
IdleState	This subsection has the power knob settings for the node idle state.
RunningState	This subsection has power knob specifications for job execution on compute nodes that also serve as I/O nodes.

[EXAMPLE]

An example of the papwrn.conf file is shown below.

```
SystemPower {
    StartTime = "50"
    LogLevel = "1"
    CommandLine = "/usr/sbin/cmd arg"
    AcceptableRange = "600"
    PowerGroup {
        PowerGroupName = "pwrgrp1"
        ClusterName = "cluster"
        NodeList = "sysnodelist1.txt"
        ExternalDeviceList = "extlist1.txt"
    }
}
PowerSave {
    Cluster {
        ClusterName = "cluster"
        LogLevel = "1"
        ResourceUnit {
            ResourceUnitName = "unit1"
            NodeList = "savenodelist1.txt"
        }
    }
}
JobPowerEstimation {
    Cluster {
        ClusterName = "cluster"
        LogLevel = "1"
        DbAuth = "pwr:password"
        ResourceUnit {
            ResourceUnitName = "unit1"
            DefaultNodePower = "160"
            ErrorNoticeMode = "0"
        }
    }
}
PowerKnob {
    Cluster {
        ClusterName = "cluster"
        LogLevel = "1"
        ResourceUnit {
            ResourceUnitName = "unit1"
            AllowSharedKnob = "0"
            ComputeNode {
                IdleState {
                    freq = "1600"
                    throttling_state = "9"
                    issue_state = "1"
                    ex_pipe_state = "1"
                    eco_state = "2"
                    retention_state = "1"
                    retention_state_acores = "1"
                }
                RunningState {
```


The papwrmgradm command cannot configure the operations corresponding to any items that cannot be omitted but are not written in this file.
Only single-byte alphanumeric characters and symbols can be written in the configuration file.

[FILE]

/etc/opt/FJSVtcs/papwrm.conf

[SEE ALSO]

papwrmgradm(8), pasyspwr(8)

4.3.9 parsc.conf file

[NAME]

parsc.conf - File that defines operation of job resource management function in cluster.

[SYNOPSIS]

/etc/opt/FJSVtcs/parsc.conf

[DESCRIPTION]

parsc.conf In the file, a set value of the job resource management function in the cluster is defined.
The cluster manager does the definition setting of this file.
Set the owner as root/root, and the permission as 0600.Please arrange this file in the system management node.
Please reflect the configuration file by using the parscadm command in the system management node.

Cluster

It is a section that describes a set value of the job resource management function in the cluster.

ClusterName

This specifies a cluster name to reflect setting values of the resource management function.

JobMem

This specifies the amount of memory assigned to jobs using a percentage value.
You can assign certain percentages of the real memory amount to jobs. The settable range of this value is 1-100 (%). It is not mandatory to set this value. If the setting value is not set, the value will be 90 (%).
After executing the parscadm command, reboot the compute cluster management node and all compute nodes of the cluster whose configuration has been changed.
When the execution did -f option of parscadm command, must be reboot of the nodelist written by node ID.

LogLevel

This sets the log output level of the resource management function.
The settable values are as follows:

Log level	Usage
1	Operation mode
2	Debug mode
3	Detailed debug mode

If the setting value is not set, the operation mode will be used.

RscWatchInterval

This specifies the monitoring interval of the resource usage state in minutes.
The resource management function collects the usage state of the compute resource for each of the compute nodes at assigned

intervals. The settable range of this value is 0-1440 (minutes). It is not mandatory to set this value. If the setting value is not set, the value will be 10 (minutes). If the setting value is 0, there is no monitoring.

MemFailJobDel

It is specified whether the job is forcibly terminated at the time of an excess of a quota memory or the time of OOM-Killer execution. on is specified when forcing a job to terminate. off is specified when not forcing a job to terminate. If the setting value is not set, the value will be off.

In the execution does parscadm command before change setting cluster, please the job itself turning on and the execution control does not exist.

The job turning on and the execution control must use the pmpjmopt command.

JobRscUsage

This specifies setting of whether to periodically collect job statistical information. on is specified collecting job statistical information. off is specified not collecting job statistical information. If the setting value is not set, the value will be off. The collected information is output to /var/opt/FJSVtcs/prm/jobrscusage on the compute cluster management node.

The output format of the information is the CSV format shown below.

Item	Description	
Time	Acquisition time of information	
JOBID	Job ID	
Bulk Number	Bulk number For jobs other than bulk jobs, this item is blank.	
Step Number	Step number For jobs other than step jobs, this item is blank.	
Retry count	Job retry count	
Job submission time	Submission time of the job	
Information tag	JIS: Indicates a snapshot of information by job. JNS: Indicates a snapshot of information by node used by 1 job.	
Information body	Information is output in the order shown in the following table. For details, see pjstatsinfo(7). If the information tag is JIS, fields of 1-2 and 20-29 are blank.	
	Number	Corresponding information name displayed by pjstat -s/-S
	1	NODE ID
	2	VNODE ID
	3	CPU NUM (USE)
	4	USER CPU TIME (USE)
	5	SYSTEM CPU TIME (USE)
	6	MAX MEMORY SIZE (USE)
	7	FJ PROFILER
	8	SECTOR CACHE
	9	INTRA NODE BARRIER
	10	POWER CONSUMPTION STATE
	11	AVG POWER CONSUMPTION OF NODE (IDEAL)
	12	MAX POWER CONSUMPTION OF NODE (IDEAL)
	13	MIN POWER CONSUMPTION OF NODE (IDEAL)
	14	ENERGY CONSUMPTION OF NODE (IDEAL)
15	AVG POWER CONSUMPTION OF NODE (MEASURED)	

Item	Description	
	16	MAX POWER CONSUMPTION OF NODE (MEASURED)
	17	MIN POWER CONSUMPTION OF NODE (MEASURED)
	18	ENERGY CONSUMPTION OF NODE (MEASURED)
	19	UTILIZATION INFO OF POWER API
	20	PROC NUM
	21	PERF COUNT 1
	22	PERF COUNT 2
	23	PERF COUNT 3
	24	PERF COUNT 4
	25	PERF COUNT 5
	26	PERF COUNT 6
	27	PERF COUNT 7
	28	PERF COUNT 8
	29	PERF COUNT 9

[EXAMPLE]

```
Cluster {
  ClusterName = clusterA
  JobMem = 90
  LogLevel = 1
  RscWatchInterval = 10
  MemFailJobDel = off
  JobRscUsage = off
}
```

[SEE ALSO]

pjstatsinfo(7), parscadm(8), pmpjmopt(8)

4.3.10 paservice.conf file

[NAME]

paservice.conf - User definition service configuration file

[SYNOPSIS]

/etc/opt/FJSVtcs/paservice.conf

[DESCRIPTION]

paservice.conf is the file that defines the setting of user definition service.

Cluster

This section defines each cluster. This item must be specified. It sets the following items and sections.

ClusterName

Service section

ClusterName

This item specifies a cluster name. This item must be specified.

Service

This section defines user definition service. It is necessary to define in Cluster sections one or more, and plurals can be defined. The following items can be specified.

Name

NodeID

NodeType

ExcludeType

ChkPath

Name

It sets service name of the user definition. This item must be specified.

The service name is defined by an alphabetic capital letter and the figure within five characters two characters or more.

The set service name is displayed by the pashowclst command. An existing service name cannot be set.

NodeID

This item specifies the node ID that is a user definition service target. This item can be omitted.

You can specify multiple node IDs by delimiting them with a comma (e.g., 0x01010010,0x01010011,0x0101002A).

You also can specify a range with a hyphen(e.g., 0x01010010-0x0101002A).

NodeType

This item specifies the node type that is a user definition service target. You can specify multiple node types by delimiting them with a comma. This item can be omitted.

The following node types can be specified. Also, the node type of a multiuse node can be specified.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
MGS	MGS node
MDS	MDS node
OSS	OSS node

ExcludeType

This item specifies the node type that is a except target. You can specify multiple node types by delimiting them with a comma. Specify for one or more node types that become objects together with the NodeType specification value to exist.

You can omit this definition.

ChkPath

The command name for the system monitoring function to check service is specified in the full path.

[EXAMPLE]

The paservice.conf file example is the following.

The NFS client service is defined in the login node of the compute cluster. The service of LDAP is defined in the node of a multiuse cluster. (ID:0xFFFF0001,0xFFFF0002)

```

Cluster {
    ClusterName = cluster1
    Service {
        Name = NFS
        NodeType = LN
        ChkPath= /usr/local/bin/nfsclient_check.sh
    }
}
Cluster {
    ClusterName = otherclst
    Service {
        Name = LDAP
        NodeID = 0xFFFF0001,0xFFFF0002
        ChkPath= /usr/local/bin/ldapserver_check.sh
    }
}

```

[SEE ALSO]

paserviceadm(8)

4.3.11 patrouble.conf file

[NAME]

patrouble.conf - Trouble case configuration file

[SYNOPSIS]

/etc/opt/FJSVtcs/patrouble.conf

[DESCRIPTION]

patrouble.conf is a configuration file used by the pasnap command to define trouble cases.

When trouble occurs, investigation materials must be collected from nodes to investigate the trouble. Here, these nodes are referred to as related nodes.

(*) Don't operate deletion or name change to patrouble.conf when you correct the component files by specifying trouble case.

Trouble

This section defines trouble cases. These definitions are required.

You can specify the following section and items:

TroubleName

Define a name of the trouble case. This definition is required.

This name is used as an argument in the --case option of the pasnap command.

Using alphanumeric characters and the hyphen (-), the name of the trouble name must consist of up to 63 characters and not begin with a hyphen.

NodeType

Define the node type of the node where the trouble occurs. This definition is required.

You can specify the following node types. You can also specify the node types of a multiuse node.

Abbreviation	Node type
SMM	System management node
CCM	Compute cluster management node
CCS	Compute cluster sub management node
BIO	Boot I/O node
SIO	Storage I/O node

Abbreviation	Node type
GIO	Global I/O node
CN	Compute node
LN	Login node
SCM	Storage cluster management node
OSS	OSS node
MDS	MDS node
MGS	MGS node

RelatedNode

This section defines conditions for related nodes. This section is required.

Set the following items:

NodeTypename

Define the node type name of related nodes. This definition is required.

You can specify the same node type as in NodeTypename described above.

Condition

Define conditions for related nodes. This definition is required.

You can specify the following conditions.

Condition name	Description
target_node	Node where trouble occurs
same_bootgrp	Nodes that belong to same boot group as node where trouble occurs
same_nodegrp	Nodes that belong to same node group as node where trouble occurs
same_cluster	Nodes that belong to same cluster as node where trouble occurs

Component

Define a component name for collection. This definition is required.

You can specify multiple component names by delimiting them with a comma (e.g., INST,PSM).

You can specify the following components.

Abbreviation	Component
INST	Installation function
PSM	System management function
PWRM	Power management function
PLE	Parallel execution environment
PRM	Resource management function
PJM	Job manager function
FEFS	Fujitsu Exabyte File System
OS	OS
ALL	Covers all above components

[EXAMPLE]

An example of the patrouble.conf file is shown below.

```

Trouble {
    TroubleName = CN-NODE-DOWN
    NodeTypeName = CN
    RelatedNode {
        NodeTypeName = CN
        Condition = target_node
        Component = ALL
    }
    RelatedNode {
        NodeTypeName = BIO
        Condition = same_bootgrp
        Component = ALL
    }
    RelatedNode {
        NodeTypeName = CCS
        Condition = same_nodegrp
        Component = ALL
    }
    RelatedNode {
        NodeTypeName = SMM
        Condition = same_cluster
        Component = ALL
    }
}

```

4.3.12 pmpjcmd.conf file

[NAME]

pmpjcmd.conf - File that defines the API operating environment

[SYNOPSIS]

/etc/opt/FJSVtcs/pjm/pmpjcmd.conf

[DESCRIPTION]

The pmpjcmd.conf file defines settings for command APIs. The file is placed on the system management, compute cluster management, and login nodes.

The settings are read when a command API is used on each node.

The job operation administrator sets the definitions in this file.

The number of characters on each line in this file must be 511 or less.

Create this file with root/root as the owner/group and 0644 as the file mode.

OwnNode

This section writes local node setting values.

EXECUTE_INTERVAL

Specify the minimum time interval (milliseconds) before a command API sends a request to the job manager function.

Administrators set a value to restrict the request interval so that numerous operation requests are not sent to the job manager function in a short period.

The command API function pjcmd_xxxx_execute() sends requests to the job manager function. When the function is called, if the time specified by this parameter has not elapsed after the previous request, the current request sleeps in the function until that time has elapsed.

Call intervals other than that of the pjcmd_xxxx_execute() function are not restricted.

The upper limit is 1000 (milliseconds). If 0 is specified, requests are processed immediately without waiting.

If this parameter is not set on the node that calls the command API function (including cases where the set value is inappropriate or the configuration file does not exist), the setting for the compute cluster management node applies. If the parameter is also not set for the compute cluster management node, the default value of 100 milliseconds applies.

MAX_JOB_RESULT_NUM

Specify the upper limit on the number of jobs that are included in job operation results. Job operation APIs (submit, delete, lock, unlock, send signal, wait for completion, and change job parameters) return the results of individual jobs.

If a range of job IDs is specified as the target of the operation, the large volume of resulting information, which also includes non-existent jobs, which may increase traffic to and from the job manager function. Administrators set this parameter to restrict the data volume of operation results.

If the quantity of the jobs that are the target of the operation exceeds this value, the operation is performed, but no results are returned for the part of the operation exceeding the upper limit. The operation results of up to MAX_JOB_RESULT_NUM jobs are returned in order from the smallest job ID.

If this parameter is not set on the node that calls the command API function (including cases where the value is inappropriate or the configuration file does not exist), the setting for the compute cluster management node applies. If the parameter is also not set for the compute cluster management node, the default value of 100,000 jobs applies.

JOBINFO_PRINT_SUSPEND_STATUS

Specify whether to display the SUSPND column (number of jobs during a suspension period) in summary information when the command API function `pjcmd_jobinfo_print_resp()` displays job information.

Use this parameter when you want to display the SUSPND column, on only a system that uses a suspend-resume function.

Specified value	Meaning
yes	Display
no	Do not display (Default)

This parameter also applies to summary information displayed by the `pjstat` command.

JOBINFO_PRINT_VNODE_ITEMS

Specify whether to display virtual node allocation-related items (*) when the command API function `pjcmd_jobinfo_print_resp()` displays job information. Use this parameter when you do not want to display these unnecessary items on a system that does not use virtual node allocation.

(*) Items: VNODE, CORE, V_MEM, V_POL, E_POL, and RANK

Specified value	Meaning
yes	Display (Default)
no	Do not display

This parameter also applies to summary information displayed by the `pjstat` command.

[EXAMPLE]

The following example shows definitions in the `pmpjcmd.conf` file.

```
$ cat /etc/opt/FJSVtcs/pjm/pmpjcmd.conf
OwnNode {
    EXECUTE_INTERVAL = 100
    MAX_JOB_RESULT_NUM = 100000
    JOBINFO_PRINT_SUSPEND_STATUS = no
    JOBINFO_PRINT_VNODE_ITEMS = yes
}
```

4.3.13 pmpjconf file

[NAME]

`pmpjconf` - Defines the operation environment for job operation per resource unit.

[SYNOPSIS]

/etc/opt/FJSVtcs/Rscunit.d/rscuname/pmpjm.conf

rscuname: The resource unit name which the operation of job operation management function is set in.

[DESCRIPTION]

The pmpjm.conf file configures on the system management node, defines the configuration for executing job operation per resource unit. This configuration is reflected by executing the pmpjmadm command on the system management node.

The job operation managers configure the definition of this file.

The lines in this file must be shorter than 639 characters.

ResourceUnit

This section describes setting values of job operation per resource unit.

ResourceUnitName

Resource unit name that sets the job operation is specified.

It is necessary to specify resource unit name.

LogLevel

This is a unique log output level for the job scheduler.

level	application
1	It is a usual log level. An important message that there is an influence in operation is output.
2	Detailed information for debugging in addition to level 1 is output. Detailed information is output to an internal log for the investigation.
3	Information for debugging is output more in detail compared with level 2.

You can set from 1 to 3 for the log output level. If this level is not set, the value will be 1.

Backfill

This specifies whether the backfill function (outstripping of the job) can be set for the resource unit. When outstripping of the job by the backfill function is enabled, yes is specified. When outstripping of the job by the backfill function is not enabled, no is specified.

When this setting is omitted, the setting value of Backfill in the pmpjm.conf that sets the entire cluster is applied.

BackfillTarget

The job that the backfill function (outstripping of the job) targets is specified for the resource unit. When you judge whether to outstrip the job by the backfill function, the outstripping range of the job is specified.

Value	Meaning
rscunit	All jobs in the resource unit are outstripped.
rscgrp	All jobs in the resource group to which the job belongs are outstripped.

When this setting is omitted, the setting value of BackfillTarget in the pmpjm.conf that sets the entire cluster is applied.

DecidedGap

The time (buffer) added at the expected completion time of the job is specified.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:01:00

Grace

Specifies the grace time to start re-scheduling when the job ends.

When the value for which expected completion time of the job (Limit at job beginning time + elapsed time + time changed by DecidedGap) and an actual difference at the finish time are specified with Grace is exceeded, the job scheduler schedules the

execution waiting job again.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:02:00

AdaptiveElapsedTimeJobTerminateGrace

Specify the forced termination delay time of the job that specified the minimum value of elapsed time limit when submitting the job. If jobs specified as *elapsedtimemin-* or *elapsedtimemin-elapsedtimemax* at job submission exceeded the minimum elapsed time limit, they are forcibly terminated by starting another job or deadline using the same resource.

In this case, SIGTERM will be sent before forced termination, SIGKILL will be sent after *AdaptiveElapsedTimeJobTerminateGrace* has passed.

This format is "hour:minute:second" as shown below:

Settable values: 00:00:00 - 99:59:59

Default value: 00:00:10

CreateRscMap

The section set resource map making interval.

The format is "*time_zone, time_interval*". The format of *time_zone* and *time_interval* is "hour:minute:second".

time_zone is specified to the time which is from now time to ending time of counting time zone, *time_interval* is specified to resource map interval which is made in that time zone.

Resource map making interval can be specified in several numbers. If it is specified multiple, it must begin from the closest time zone apart from now.

Settable values:

time_zone

00:01:00 to 9999:59:59 or "*" (* will be set in the last line)

time_interval

00:01:00 to 9999:59:59

If this is not set, the default value is: (It is set as the following 3 lines)

01:00:00,00:10:00

24:00:00,01:00:00

*,24:00:00

SchedulePeriod

Scheduling period is set.

The time in the future which is later than current time added with this period will not be assigned.

Settable values: 00:01:00 to 9999:59:59. The unit is time.

The default value is 25:00:00.

DynamicSchedulePeriod

Dynamic scheduling period is set.

The format is "ratio,minimum scheduling period".

Determines the scheduling period by job maximum elapsed time limit instead of SchedulePeriod values.

Ratio is settable values: 1 to 2147484347.

Minimum scheduling period is settable values: 00:01:00 to 9999:59:59. The unit is time.

The default value is "2,24:00:00".

RestartNormal

Automatically re-executes a normal job by default.

If yes is specified, a normal job is automatically re-executed by default. If no is specified, a normal job is not automatically re-executed by default.

The default value is yes.

RestartStep

Automatically re-executes a step job by default.

If yes is specified, a step job is automatically re-executed by default. If no is specified, a step job is not automatically re-executed

by default.
The default value is yes.

RestartBulk

Automatically re-executes a bulk job by default.
If yes is specified, a bulk job is automatically re-executed by default. If no is specified, a bulk job is not automatically re-executed by default.
The default value is no.

RestartMswk

Automatically re-executes a master-worker job by default.
If yes is specified, a master-worker job is automatically re-executed by default. If no is specified, a master-worker job is not automatically re-executed by default.
The default value is yes.

MailSend

Whether all related mails can be sent to jobs in the current resource unit is set.
When mails are sent, yes is specified. When mails are not sent, no is specified.
The default value is yes.

Fairshare

Specifies the fair share set whose fair share value will be reduced when a job submitted to this resource unit will start running.

- on[*@fairshareset*]

Fair share value of the fair share set *fairshareset* will be reduced when a job submitted to this resource unit will start running.
Only one fair share set name can be specified.
fairshareset must follow the following rules:

- 1-15 characters
- alphanumeric characters, hyphens (-) or underscores (_)

If *fairshareset* is not specified, fair share value of fair share set *def_fs* will be reduced.
Fair share sets with the same name each other are treated as a different fair share set only if they belong to different resource units.
If the same fair share set name is specified for Fairshare item (in the ResourceUnit section) and ResourceGroupFairshare item (in the ResourceGroup section), the name specified for ResourceGroupFairshare item is ignored, so that the fair share value will be reduced only once when a job submitted to the resource group will start.
When the resource unit is different even if fair share set name is the same, it is treated as another fair share value respectively.

- off

Fair share value of the resource unit will not be reduced when a job submitted to this resource unit will start running.

When this setting is omitted, the value of Fairshare item in *papjm.conf* is applied.

FshareRecoveryValue

Fair share recovery value is set.
Job using resource (node) which is treated as fair share * elapse time is recovery quantity in the period. In the operation, after assuming scale job has been executed, the period which lasts until execution priority having completely recovered will be taken as standard.
Settable values: 0 to 18446744073709551615.
For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the fair share value is calculated using the maximum elapse limit.
This item affects all fair share sets in the resource unit.
When this setting is omitted, the value of FshareRecoveryValue item in *papjm.conf* is applied.

AllocType

The type of the resource unit is specified.
When the compute node is FX server, node is specified. When the compute node is PRIMERGY server, vnode is specified.
The default value is node.

StartTimeGuarantee

The valid/invalid of the function to guarantee of planned execution start time is specified for the resource unit. On is valid, off is invalid.

When this setting is omitted, the setting value of StartTimeGuarantee in the papjm.conf that sets the entire cluster is applied.

JobStderrMsgLevel

The error when the job is executed is output to a standard error of the job.

level	application
0	Nothing is output.
1	An ERR message is output.
2	A WARN message in addition to level 1 is output.
3	An INFO message in addition to level 2 is output.

If this level is not set, the value will be 0.

SuspendPlannedPeriod

Specifies the planned suspend period.

The planned suspend period is used to schedule suspended jobs. When a job is suspended, the job will be scheduled as if it will be resumed after the time that the job is assumed to be going to be resumed is called as "the planned resume time".

If the job will have not been resumed by the planned resume time, the time will be postponed for the planned suspend period.

Settable values: 00:01:00 to 9999:59:59. The format is "hour:minute:second".

Default value: 01:00:00.

SuspendFailForceError

Specifies whether or not the job that failed to be suspended or to be resumed goes into ERROR state.

Value	Meaning
no	If the operation failed due to a recoverable error, the job will remain in its current state (RUNNING state or SUSPENDED state). If the operation failed due to an unrecoverable error, the job will be aborted. The aborted job will be re-executed if it is automatically re-executable.
yes	The job will be aborted and will go into ERROR state.

Default value: no

HoldAcceptDate

Specifies the interpretation of the job submission time in scheduling when the job is released from the HOLD state.

Value	Meaning
release	Time when the job is released from the HOLD state.
accept	Time when the job is submitted.

When this setting is omitted, the HoldAcceptDate setting applied in the papjm.conf file that sets the entire cluster is applied.

SchedulerPluginLoadPath

Specifies the directory paths where the plugin libraries are located. Job scheduler will load plugin libraries in the specified directories.

To specify multiple directories, separate the paths with ":".

The default value is "/etc/opt/FJSVtcs/plugin/pjm/pjsd/normal_mode".

ShelfPackingNodeAllocation

Only FX server is effective.

This specifies whether the free nodes on shelves allocated other I/O-sharing jobs are preferentially allocated for I/O-sharing jobs.

Value	Meaning
yes	When allocating 48 or fewer nodes to an I/O-sharing job, allocate free nodes preferentially on shelves where other I/O-sharing jobs are already allocated.
no	Allocate free nodes to the I/O-sharing job without considering the above.

When this setting is omitted, the ShelfPackingNodeAllocation setting applied in the papjm.conf file that sets the entire cluster is applied.

JobSchedulingTargetLimit

This section specifies the maximum number of scheduled jobs, excluding interactive jobs. Jobs that are not to be scheduled this time will be rescheduled at the next scheduling time. Interactive jobs are always scheduled.

You can specify a value ranging from 0 to 2147483647 for the maximum number of jobs. If you specify 0, all jobs are scheduled. The default value is 10000.

JobSchedulingTargetMode

This section specifies the method of selecting the jobs for scheduling with JobSchedulingTargetLimit.

You can specify either aprio or jobselectpolicy.

The default value is jobselectpolicy.

aprio

Determines the subjects of scheduling by order of job priority in the resource unit and the job submission time.

jobselectpolicy

Determines the subjects of scheduling according to the specification in the JobSelectPolicy section. The following policies are supported. Other policies are ignored.

fcfs
node
group_prio
user_prio
job_aprio
rscgrp_prio
elapse_limit
node_times_elapse
at
interact
job_epoint

JobSelectPolicy

This section set job select policy.

You can define only one job select policy in the ResourceUnit section, and it is either of the following.

- Specify the name of the job select policy to apply to a resource unit, among the job select policy that are defined in the papjm.conf file. The file has settings for the entire cluster.
- Set only the items to be changed or added in the common definitions of the cluster defined in the papjm.conf file. In this case, a new policy name cannot be associated with the set contents.

name

The policy name is defined by papjm.conf. The job select policy is becoming effective according to the content set by that policy name.

When job select policy is set in various resource units, the following definition is defined.

fcfs = {order[,asc |,desc]|off}

Whether need to do the job selection evaluation according to job submitted time. If this is not set, it needs to evaluate according to job submitted time.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)

desc: sort in descending order
off: not evaluate

node = {*order*[,asc |,desc][off]}

Only FX server is effective.

Whether need to do the job selection evaluation according to required node number.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)
desc: sort in descending order
off: not evaluate

group_prio = {*order*[,asc |,desc][off:]}

Whether need to do the job selection evaluation according to group priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

user_prio = {*order*[,asc |,desc][off:]}

Whether need to do the job selection evaluation according to user priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

usr_in_grp_prio = {*order*[,asc |,desc][off:]}

Whether need to do the job selection evaluation according to user priority in the group.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set group_prio as the policy to be evaluated before the evaluation order set in usr_in_grp_prio. For example, if the evaluation order of usr_in_grp_prio is 5, set the evaluation order of group_prio to 4. When setting usr_in_grp_prio and usr_in_grp_faishare, be sure to evaluate usr_in_grp_prio and usr_in_grp_faishare consecutively, and be sure to set group_prio as the policy to be evaluated before the previous one. For example, set the evaluation order of usr_in_grp_prio and usr_in_grp_faishare to 5 and 6, and set the evaluation order of group_prio to 4. For all groups that run the job, set a different value for the group priority within the resource unit (define pri-g) of the job ACL function.

job_prio = {*order*[,asc |,desc][off:]}

Whether need to do the job selection evaluation according to job priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set user_prio as the policy to be evaluated before the evaluation order set in job_prio. For example, if the evaluation order of job_prio is 5, set the evaluation order of user_prio to 4. For all users that run the job, set a different value for the user priority within the resource unit (define pri) of the job ACL function.

job_aprio = {*order*[,asc |,desc][off]}

Whether need to do the job selection evaluation according to job priority in resource unit.

order: the evaluating order of item(1~256)
asc: sort in ascending order

desc: sort in descending order(default)
off: not evaluate

rsgrp_prio = {order[,asc |,desc]|off:}

Whether need to do the job selection evaluation according to resource group priority.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

group_fairshare[@fairshareset] = {order[,asc |,desc]|off}

Specifies whether the job selection based on the group fair share value is enabled.

The group fair share value of the specified fair share set is used in job selection to evaluate a job priority. This setting is effective even when Fairshare item (in the ResourceUnit section) and ResourceGroupFairshare item (in the ResourceGroup section) are off. Two or more settings with different fair share set name can be specified if they have different evaluation order. The form of fair share set name is same with Fairshare item in the ResourceUnit section.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

usr_in_grp_fairshare[@fairshareset] = {order[,asc |,desc]|off}

Specifies whether the job selection based on the user-in-group fair share value is enabled.

The user-in-group fair share value of the specified fair share set is used in job selection to evaluate a job priority.

This setting is effective even when Fairshare item (in the ResourceUnit section) and ResourceGroupFairshare item (in the ResourceGroup section) are off. Two or more settings with different fair share set name can be specified if they have different evaluation order. The form of fair share set name is same with Fairshare item in the ResourceUnit section.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

Note

Be sure to set group_prio as the policy to be evaluated before the evaluation order set in usr_in_grp_fairshare. For example, if the evaluation order of usr_in_grp_fairshare is 5, set the evaluation order of group_prio to 4. When setting usr_in_grp_prio and usr_in_grp_faishare, be sure to evaluate usr_in_grp_prio and usr_in_grp_faishare consecutively, and be sure to set group_prio as the policy to be evaluated before the previous one. For example, set the evaluation order of usr_in_grp_prio and usr_in_grp_faishare to 5 and 6, and set the evaluation order of group_prio to 4. For all groups that run the job, set a different value for the group priority within the resource unit (define pri-g) of the job ACL function.

user_fairshare[@fairshareset] = {order[,asc |,desc]|off}

Specifies whether the job selection based on the user fair share value is enabled.

The user fair share value of the specified fair share set is used in job selection to evaluate a job priority. This setting is effective even when Fairshare item (in the ResourceUnit section) and ResourceGroupFairshare item (in the ResourceGroup section) are off. Two or more settings with different fair share set name can be specified if they have different evaluation order. The form of fair share set name is same with Fairshare item in the ResourceUnit section.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

elapse_limit = {order[,asc |,desc]|off:}

Whether need to do the job selection evaluation according to elapse limit.

For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the maximum elapse limit is used.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)

desc: sort in descending order
off: not evaluate

node_times_elapse = {*order*[,asc |,desc][off:]}

Only FX server is effective.

Whether need to do the job selection evaluation according to node times accumulation.

For jobs with the elapse limit specified in the range (For example, "10-" or "10-100"), the node times accumulation is calculated with the maximum elapse limit.

order: the evaluating order of item(1~256)
asc: sort in ascending order(default)
desc: sort in descending order
off: not evaluate

at = {*order*[,true |,false][off:]}

Whether need to do the job selection evaluation according to execution start time.

order: the evaluating order of item(1~256)
true:has a preference for execution start time(default)
false:has no preference for execution start time
off: not evaluate

interact = {*order*[,true |,false][off:]}

Whether need to do the job selection evaluation according to submitting interact jobs.

order: the evaluating order of item(1~256)
true:has a preference for interact jobs(default)
false:has a preference for non-interact jobs
off: not evaluate

job_epoint = {*order*[,asc |,desc][off]}

This specifies whether there is a comprehensive evaluation using multiple items.

order: the evaluating order of item(1~256)
asc: sort in ascending order
desc: sort in descending order(default)
off: not evaluate

JobEvaluation

This section defines how to calculate job evaluation points (job evaluation definitions) regarding the job selection policy item *job_epoint*.

You can define only one job evaluation definition in the ResourceUnit section, and it is either of the following.

- Specify the name of the job evaluation definition to apply to a resource unit, among the job evaluation definitions that are defined in the papjm.conf file. The file has settings for the entire cluster.
- Set only the items to be changed or added in the common definitions of the cluster defined in the papjm.conf file. In this case, a new job evaluation definition name cannot be associated with the set contents.

name = *job evaluation definition name*

This specifies the name of the job evaluation definition to apply to a resource unit, among the job evaluation definitions that are defined in the papjm.conf file. The file has settings for the entire cluster.

The omission of this item means that the contents of the JobEvaluation section are changes or additional settings for the common definitions of the cluster.

item name = *weight*[,*value 1@point 1: value 2@point 2: ... : value N@point N*]

This defines the points corresponding to the values of the items to evaluate.

The definition is written in the "*value@point*" format. When writing multiple definitions, delimit them with a colon (:). You can specify up to 64 definitions.

The points for items that have numerical values are determined as follows.

Value	Point
-----	-----
Minimum value for item value <= value <= value 1	point 1
value 2 < value <= value 3	point 2
...	...
value N-1 < value <= value N	point N

Define each point such that $value\ n-1 < value\ n$ is true.

The points for items that do not have a numerical value are indicated by the definitions with their corresponding values.

Also, the values of multiple definitions must differ from one another.

The specifiable item names and values are as follows.

Item	Meaning	Range of specifiable values
node elapse_limit waittime	Number of nodes requested by a job Elapsed time limit value for job execution (seconds) Time from job acceptance to the start of job scheduling or rescheduling (seconds)	1 to INT32_MAX
group_prio user_prio usr_in_grp_prio job_prio job_aprio rscgrp_prio	Priority of the group to which the user (job submitter) belongs Priority of the user (job submitter) Priority of the user (job submitter) within the group Job priority Job priority within the resource unit Resource group priority	0 to 255
node_times_elapse	Product of the number of nodes requested by a job (node) and the elapsed time limit value for job execution (elapse_limit)	1 to INT64_MAX
at interact	Whether the execution start time is specified Whether the job is an interactive job	yes or no

If "*" is the specified value, it means that the value is greater than $value\ n-1$ (for a numerical value) or other than $value\ n-1$ (for a non-numerical value).

[Example] 10@1:20@2:*:3

If the value is greater than 20, the point is point 3.

[Example] no@0,*@1

If the value is other than no, the point is point 1.

You can specify a value ranging from 1 to 1024 for the point.

The job evaluation points are the totals from the multiplication of *weight* for the point of each item. If the point scales differ according to the item or if you want to emphasize a specific item, use *weight* to adjust the balance between different points. You can specify a value ranging from -2147483648 to 2147483647 for *weight*.

If the total of an individual point is below the lower point limit (-9223372036854775807) or over the upper point limit (9223372036854775807), the job evaluation point is the lower limit or upper limit, respectively.

PrologueEpilogue

This section describes setting values of prologue and epilogue function.

ShellName

The execution shell of the prologue epilogue script is specified.

The default value is /bin/sh.

ExecUser

The execution user of the prologue epilogue script is specified.

ROOT: root user

JOBUSER: job execution user
The default value is ROOT.

PrologueName

Path of the prologue script is specified.
There is no default value.

EpilogueName

Path of the epilogue script is specified.
There is no default value.

PrologueTime

The rough estimate of execution time(seconds) of the prologue script is specified.
The value can be specified 0-2147483648.
The default value is 0.

EpilogueTime

The rough estimate of execution time(seconds) of the epilogue script is specified.
The value can be specified 0-2147483648.
The default value is 0.

ContainElapse

It is specified whether to include the prologue epilogue at the elapsed time.
yes: The execution elapsed time contains the elapsed time of the prologue epilogue.
no: The execution elapsed time does not contains the elapsed time of the prologue epilogue.
The default value is no.

ResourceGroup

This section describes setting values of job operation per resource group.

ResourceGroupName

Resource group name is set. This can not be omitted.

ResourceGroupNode

Node Resource of resource group is set. When this setting is omitted, all nodes in the resource unit are applied.
The specification method is different in FX server and PRIMERGY server.

[FX server]

Three setting types are shown as following: the node number, ratio and shape.

When the node number is set, value is specified. (Exp: 10)

When the ratio is set, "%" is specified. (Exp: 100%)

When the shape is set, the minimum coordinate-maximum coordinate of the tofu unit is specified. (Exp: 0,0,0-2,3,2)

Minimum coordinate is X1,Y1,Z1 and maximum coordinate is X2,Y2,Z2 describe "X1,Y1,Z1-X2,Y2,Z2".

The number of nodes of the resource group is rounded up to the Tofu unit (12 nodes) when specifying it at the number of nodes or ratio.

[PRIMERGY server]

Node ID is specified.

Specified node ID is delimited by ",".

You can also specify node ID ranges. The range is formed as *start-nodeid-end-nodeid*. When specifying a range, only nodes of the same node group can be specified. When specifying nodes in different node groups, specify a range for each node group.

ResourceGroupNode can be described more than once in the same section. If you specify more than one, all specifications are valid.

ResourceGroupPrio

The resource group priority is specified (0-255).

When the value is large, priority is high.

The following selection evaluation (for instance, fcfs) between the resource groups as which priority is the same is adopted.

The default value is 127.

ResourceGroupTsha

Only FX server is effective.

The allocation mode whether multiple jobs can share one Tofu unit is specified.

Value	Meaning
1	Single node jobs can only share one Tofu unit.
2	Jobs that require 1 to 11 nodes share one Tofu unit.

This setting is enable for any job in torus mode and not io-exclusive mode.

The default value is 1.

ResourceGroupExecPolicy

Only FX server is effective.

Execution mode policy of the resource group is specified.

Value	Meaning
share	Nodes are shared with resource groups of the specified number of nodes or node rate.
simplex	Nodes are not shared with resource groups of the specified number of nodes or node rate.

The default value is share.

This setting is enable for shape-specified resource groups.

Nodes are shared with other resource groups for any resource group of the specified number of nodes or node rate, regardless of the setting.

Backfill

This specifies whether the backfill function (outstripping of the job) can be set for the resource group. When outstripping of the job by the backfill function is enabled, yes is specified. When outstripping of the job by the backfill function is not enabled, no is specified.

When this setting is omitted, the setting of Backfill applied to the resource unit is applied.

StartTimeGuarantee

The valid/invalid of the function to guarantee of planned execution start time is specified for the resource group. On is valid, off is invalid.

When this setting is omitted, the setting of StartTimeGuarantee applied to the resource unit is applied.

JobStderrMsgLevel

The error when the job is executed is output to a standard error of the job.

level	application
0	Nothing is output.
1	An ERR message is output.
2	A WARN message in addition to level 1 is output.
3	An INFO message in addition to level 2 is output.

When this setting is omitted, the setting of JobStderrMsgLevel applied to the resource unit is applied.

ResourceGroupFairshare

Specifies the fair share set whose fair share value will be reduced when a job submitted to this resource group will start running. The form of this item is same with Fairshare in the ResourceUnit section.

If *fairshareset* is same as Fairshare item (in the ResourceUnit section), the fair share value of *fairshareset* will be reduced once when a job submitted to this resource group will start running. If *fairshareset* is different from Fairshare item (in the ResourceUnit section), the fair share value of *fairshareset* in both this setting and Fairshare item (in the ResourceUnit section) will be reduced when a job submitted to this resource group will start running.

When this setting is omitted, the value of Fairshare item (in the ResourceUnit section) is applied.

Note

If the Fairshare setting item for the resource unit is on, the fair share value is subtracted when job execution starts, even when the ResourceGroupFairshare item is off.

HoldAcceptDate

Specifies the interpretation of the job submission time in scheduling when the job is released from the HOLD state.

Value	Meaning
release	Time when the job is released from the HOLD state.
accept	Time when the job is submitted.

When this setting is omitted, the HoldAcceptDate setting applied in the resource unit is applied.

ShelfPackingNodeAllocation

Only FX server is effective.

This specifies whether the free nodes on shelves allocated other I/O-sharing jobs are preferentially allocated for I/O-sharing jobs. The format is the same format as the ShelfPackingNodeAllocation setting in the ResourceUnit section.

When this setting is omitted, the ShelfPackingNodeAllocation setting applied in the resource unit is applied.

JobSelectPolicy

This section set job select policy.

You can define only one job select policy in the ResourceGroup section.

The format is the same format as the JobSelectPolicy section in the ResourceUnit section.

JobEvaluation

This section defines how to calculate job evaluation points (job evaluation definitions) regarding the job selection policy item job_epoint.

You can define only one job evaluation definition in the ResourceGroup section.

The format is the same format as the JobEvaluation section in the ResourceUnit section.

CustomResource

This section defines a custom resource.

A custom resource can be defined in both ResourceUnit section and ResourceGroup section.

You can define up to 64 CustomResource sections in one pmpjm.conf.

Name

Specifies a name of a custom resource. This can not be omitted.

NodeID

Specifies the nodes associated with the custom resource.

If the custom resource is associated with some nodes, the custom resource is available only on those nodes. If no nodes are associated, the resource is available on the entire nodes belonging to the resource unit (or the resource group).

You can specify multiple nodes delimited by commas (,).

You can also specify node ID ranges.

The range is formed as *start-nodeid-end-nodeid*.

When specifying a range, the following rules must be followed.

[FX server]

Only nodes of the same boot group can be specified.

When specifying nodes in different boot groups, specify a range for each boot group.

[PRIMERGY server]

Only nodes of the same node group can be specified.

When specifying nodes in different node groups, specify a range for each node group.

NodeID can be described more than once in the same section.

If you specify more than one, all specifications are valid.

ValueType

Specifies the type of custom resource value. This can not be omitted.

Value	Meaning
numeric	Specifies a custom resource requirement as a numerical value (total amount or amount per node).
string	Specifies a custom resource requirement as a string value (one of the administrator-defined strings).

Value

Specifies an amount or strings of the custom resource. This can not be omitted.

Specifiable value depends on ValueType.

[ValueType=numeric]

Specifies an amount of the custom resource as a number.

A value ranging from 1 to 99999999999 or "unlimited" can be specified.

When specifying the NodeID, specify the number of resources for each node.

Otherwise, specify the number of resources per resource unit or resource group.

[ValueType=string]

Defines a set of valid values for the custom resource value.

You can specify multiple valid values delimited by commas (.). The value must consist of alphanumeric characters, hyphens, or underscores.

Each string must not exceed 63 characters, and the entire string must not exceed 511 characters (including commas and spaces).

"-" Can not specify only one character. Spaces before and after commas are ignored.

ExitFunc

This section describes the settings of the job manager exit function or the job scheduler exit function.

The ResourceUnit and ResourceGroup sections may describe the values in this section.

ExitFuncLib

Specifies the name of the library of the job manager exit function or job scheduler exit function. This cannot be omitted.

Place the library in /var/opt/FJSVtcs/shared_disk/pjm/hook on the compute cluster management node.

ExitFuncPri

Specifies the execution priority of the job manager exit function or job scheduler exit function.

You can specify a value ranging from 0 to 255 for the priority. The lowest priority is 0, and the highest priority is 255.

The exit function is executed in descending order of priority.

The default value is 127.

ExitFuncType

Specifies the type of exit function.

Value	Meaning
pjm	Job manager exit function
pjs	Job scheduler exit function

Scheduler

This is a section for specifying the plugin libraries to be loaded into the job scheduler.

This section must be defined in the ResourceUnit section.

Name

Specifies the scheduler name of the plugin libraries. This item can not be omitted.

The specified name must follow the following conditions:

At least 1 character and not more than 63 characters

All characters are alphanumeric characters or hyphens "-" and underscores "_".

Plugins

Specifies the file names of plugin libraries to be loaded into job scheduler.

To specify multiple plugin libraries, separate the file names with ",".

The default value is "" (no libraries will be loaded).

[EXAMPLE]

A definition example of the pmpjm.conf file is as follows:

```
$ cat /etc/opt/FJSVtcs/Rscunit.d/unit1/pmpjm.conf
ResourceUnit {
    ResourceUnitName = unit1
    LogLevel = 1
    Backfill = yes
    BackfillTarget = rscgrp
    DecidedGap = 00:01:00
    Grace = 00:02:00
    CreateRscMap = "01:00:00, 00:10:00"
    CreateRscMap = "24:00:00, 01:00:00"
    CreateRscMap = "*", 24:00:00"
    SchedulePeriod = 25:00:00
    DynamicSchedulePeriod = 2,24:00:00
    RestartNormal = yes
    RestartStep = yes
    RestartBulk = no
    RestartMswk = yes
    MailSend = yes
    Fairshare = off
    FshareRecoveryValue = 118
    StartTimeGuarantee = on
    JobStderrMsgLevel = 0
    HoldAcceptDate = release
    SchedulerPluginLoadPath = "/etc/opt/FJSVtcs/plugin/pjm/pjsd/normal_mode"
    ShelfPackingNodeAllocation = no
    JobSchedulingTargetLimit = 10000
    JobSchedulingTargetMode = jobselectpolicy
    JobSelectPolicy {
        name = policy2
    }
}
# ResourceGroup {
#     ResourceGroupName =
#     ResourceGroupNode =
#     Backfill =
#     StartTimeGuarantee =
#     ResourceGroupFairshare =
#     ShelfPackingNodeAllocation=
#     JobSelectPolicy {
#         name =
#     }
#     JobEvaluation {
#         name =
#     }
#     CustomResource {
#         Name =
#         ValueType =
#         Value =
#     }
#     ExitFunc {
#         ExitFuncLib =
#         ExitFuncPri =
```

```
#         ExitFuncType =
#     }
# }
# Scheduler {
#     Name =
#     Plugins =
# }
# }
```

[SEE ALSO]

pmpjmadm(8)

4.3.14 pmrsc.conf file

[NAME]

pmrsc.conf - File that defines operation of job resource management function in resource unit.

[SYNOPSIS]

```
/etc/opt/FJSVtcs/runitname/pmrsc.conf
runitname:Resource unit name in which operation of job resource management function is defined.
```

[DESCRIPTION]

pmrsc.conf The pmrsc.conf file, a set value of the job resource management function in the resource unit is defined. The job operation manager does the definition setting of this file.

Set the owner as root/root, and the permission as 0600. Please arrange this file in the system management node.

Please reflect the configuration file by using the pmrscadm command in the system management node.

Cluster

It is a section that describes a set value of the job resource management function in the cluster.

ClusterName

The cluster name in which a set value of the job resource management function is reflected is specified.

ResourceUnit

It is a section that describes a set value of the job resource management function in the resource unit.

ResourceUnitName

The resource unit name in which a set value of the job resource management function is reflected is specified.

JobMem

This specifies the amount of memory assigned to jobs using a percentage value.

You can assign certain percentages of the real memory amount to jobs. The settable range of this value is 1-100 (%). It is not mandatory to set this value. If the setting value is not set, the value will be 90 (%).

if to be reflect of setting item of the JobMem, after executing the pmrscadm command, please resource unit for which the setting has been changed , Compute cluster management node to which the resource unit belongs and all compute nodes, there does reboot.

When the execution did -f option of parcadm command, must be reboot of the *nodelist* written by node ID.

RscWatchInterval

This specifies the monitoring interval of the resource usage state in minutes.

The resource management function collects the usage state of the compute resource for each of the compute nodes at assigned intervals. The settable range of this value is 0-1440 (minutes). It is not mandatory to set this value. If the setting value is not set, the value will be 10 (minutes). If the setting value is 0, there is no monitoring.

RespWaitTime

This specifies the maximum response wait time in minutes for node-to-node communication of resource management.

The settable range of this value is between 120 and 1440 (minutes). It is not mandatory to set this value. If no setting value is set, the value will be 120 (minutes).

ExitFunc

This section describes the definition of the resource management exit function.

A total of up to 100 definitions for this section can be registered in the resource unit section and the resource group section.

If the placement directory has multiple descriptions of the same section, the function is called as many times as the number of descriptions.

ExitFuncScriptDir

This specifies the placement directory for the executed exit function.

Place the prealloc, postfree, and predel files immediately under this directory. Only the exit functions placed here are enabled. For details on each exit function, see prealloc(5), predel(5), and postfree(5).

If there is no exit function in the specified directory, an error occurs.

Also, if the full path of a file or directory in the specified directory exceeds 2,048 characters, an error occurs.

ExitFuncPri

This specifies the priority for the executed exit function.

You can specify a value ranging from 0 (minimum) to 255 (maximum) for the priority. The execution order within each exit function type (prealloc, postfree, predel) is determined according to the specified priority.

The description of the setting value can be omitted. If it is omitted, 127 is applied.

If the same priority is defined for multiple exit functions, the execution order is determined according to the following rules.

- If their descriptions are in different sections, the description in the resource unit section is executed with priority over the description in the resource group section.
- If their descriptions are in the same section, they are executed in the order described. Note that if one of the following events occurs during execution of an exit function, the exit functions with lower priority are not executed.
- The execution time exceeds the timeout value.
- The exit function end code is any value from 1 to 4.

ExitFuncTimer

This specifies the timeout value for the execution time of an executed exit function.

You can set a value ranging from 1 to 1800 (seconds). The description of the setting value can be omitted. If it is omitted, 10 (seconds) is applied.

Note that the value calculated by multiplying the number of jobs that can be executed simultaneously on one node by the total value of ExitFuncTimer values may affect the execution of a job. If that value is set and it exceeds the RespWaitTime value, the job may exceed the response wait time of node-to-node communication for resource management. In turn, the compute node may disconnect from operation.

Therefore, calculate the above value and set ExitFuncTimer so that it does not exceed RespWaitTime.

ResourceGroup

This section describes the setting values for the function of job resource management in a resource group.

ResourceGroupName

This specifies the resource group name that reflects the setting values of the job resource management function.

Specify the resource group name specified in pmpjm.conf.

[EXAMPLE]

```
Cluster {
  ClusterName = clusterA
  ResourceUnit {
    ResourceUnitName = runit1
    JobMem = 90
    RscWatchInterval = 10
    RespWaitTime = 120
  }
}
```

```

ExitFunc {
    ExitFuncScriptDir = /work/exitfunc1
    ExitFuncPri = 100
    ExitFuncTimer = 10
}

ResourceGroup {
    ResourceGroupName = groupA

    ExitFunc {
        ExitFuncScriptDir = /work/exitfunc2
        ExitFuncPri = 120
        ExitFuncTimer = 30
    }
}
}
}
}

```

[SEE ALSO]

prealloc(5), predel(5), postfree(5), pmpjm.conf(5), pmrscadm(8), pmpjmopt(8)

4.3.15 pxinst.conf file

[NAME]

pxinst.conf - Installation function configuration file

[SYNOPSIS]

/etc/opt/FJSVtcs/pxinst.conf

[DESCRIPTION]

The pxinst.conf file is a configuration file that defines operating environment settings for the installation function and basic settings for the installation destination nodes.

The pxinst.conf file is a text file written in yaml format.

A keyword must be followed by one colon and one single-byte space (a tab cannot be used).

Keywords in this file are case-sensitive.

Comments can be inserted anywhere in the file (except inside quotation marks).

A comment begins with the # character and ends at the end of the line.

<System settings>

The system settings define the operating environment for the installation function.

repodir

Specify the absolute path to the directory for placing the repository that stores the RPM package file and images for installation.

The default value is /var/repos.

tftpdn

Specify the absolute path to the directory for storing the TFTP server configuration files and materials for network installation.

The default value is /tftpboot/pxinst.

provision_nic

Specify the device name assigned to the provision network used by the DHCP server.

No default value is defined, so you need to manually set a value.

ntp_servers

Specify the IP address or host name of the NTP server referenced by the installer node.

For the IP address, specify a dotted IPv4 address.

For multiple IP addresses, specify them delimited with a comma.
No default value is defined, so you need to manually set a value.

update_hosts

Specify an automatic update setting for /etc/hosts.
Specify whether to automatically edit entries in /etc/hosts when node information is registered, updated, or deleted.
To work in linkage with an external server, such as LDAP, or to manually manage /etc/hosts, specify False.
The default value is False.

<Installation parameters>

The installation parameters define basic settings for the installation destination nodes.
These settings apply to all the installation destination nodes.

lang

Specify the default language of the system.
The default value is en_US.UTF-8.
(Example)

ja_JP.UTF-8

timezone

Specify the time zone of the system.
The default value is Asia/Tokyo.
(Example)

Asia/Tokyo

keyboard

Specify a keyboard layout.
The default value is jp106.
(Example)

jp106

skipx

To not configure the X Window System, specify true.
To configure the X Window System, specify false.
The default value is true.

passalgo

Specify the hash algorithm for password encryption.
You need to specify the same algorithm as that for the installer node.
The default value is sha512.
(Example)

sha512

[EXAMPLE]

```
# Install Function Settings
#-----#
# system settings
#-----#
# specify the directory path to create RPM repositories.
# You need to specify an absolute path.
repodir: '/var/repos'

# Specify the directory for the setting of the tftp server.
# You need to specify an absolute path.
tftpdirdir: '/tftpboot/pxinst'

# Specify the network interface name for provision network.
```



```

provision_nic: 'eno1'

# specify the IP address or host name to be used the Network Time Protocol servers.
# Comma-separated in the case of multiple.
# e.g. 192.0.2.1,192.0.2.2
ntp_servers: '192.0.2.1'

#-----#
# install parameters
#-----#
# SystemLanguage
lang: 'ja_JP.UTF-8'

# SystemTimezone
timezone: 'Asia/Tokyo'

# Keyboard Layout
keyboard: 'jp106'

# X Window System
# If present true, X is not configured on the installed system.
# If present false, X is configured on the installed system.
skipx: true

# Password hash algorithm
# e.g. sha512 md5
passalgo: 'sha512'

```

[SEE ALSO]

pxinst(8)

4.3.16 pxsystem.conf file

[NAME]

pxsystem.conf - Cluster configuration definition file

[SYNOPSIS]

/etc/opt/FJSVtcs/pxsystem.conf

[DESCRIPTION]

The pxsystem.conf file defines a cluster configuration.

Operation

The operations to be performed are written in this section.

This section must be specified. It has the following item.

Type

Specify the type of operation to perform. You can specify the following values.

Value	Description
add-clst	Adds clusters.
del-clst	Deletes clusters.
add-node	Adds nodes.
del-node	Deletes nodes.
update	Rewrites configuration information.

When specifying "add-clst" in the Type item of the Operation section, write in the following sections and items.

Cluster

Cluster definitions are written in this section. It has the following items.

ClusterName

Specify the name of a cluster. This item must be specified.

Use the following naming convention for cluster names:

- Up to 63 characters
- Single-byte alphanumeric characters (case-sensitive)
- Hyphen (-) can be used (except as the first character)
- Underscore (_) cannot be used

ClusterType

Specify the cluster type. You can specify the following values.

Cluster name	Value
Compute cluster	compute
Storage cluster	storage
Multiuse cluster	multiuse

UseStorageClusterName

Specify the name of the storage cluster used by the defined cluster.

HA

This section configures a redundancy function. It has the following item.

Node

This section defines nodes in a redundant configuration. It has the following items.

Item name	Description
NodeType	Specifies the node type of the target nodes for the mount setting.
Mount	<p>This section defines mount information. It has the following items.</p> <p>Device Specifies the block special device to be mounted. Instead of explicitly specifying the device, you can specify the file system for this mounted device by using either a UUID, volume label or by-id. (Example) UUID : Device = "UUID=xxxxxxxx-0000-xxxx-0000-xxxxxxxxxxxxx" volume label : Device = "LABEL=xxxxxxxxx" by-id : Device = "/dev/disk/by-id/xxx-oooooooooooooooooooooooooooooooooooo"</p> <p>MountPoint Specifies the mount point for the file system to be mounted. (Example) /var/opt/FJSTcs/shared_disk</p> <p>FSType Specifies the type of file system. (Example) ext4</p> <p>OptList Specifies the options to specify for the mount command. You can specify multiple options by delimiting them with a comma (.).</p>

SMMGroup

The system management node is written in this section. It has the following items.

ManageNetMaster

Specify a representative IP address for the management network.

If the system management node is in a redundant configuration, this item is required.

SMM

This section defines the configuration of the system management nodes. It has the following items.

Item name	Description
ControlNetSrv	Specifies the IP address of the control network (for software communication).
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item.
ManageNet	Specifies the IP address of the management network.
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.
CombinedType	Specifies a dual use of node types in this item. "SCM" can be specified.

MasterGroup

Compute cluster management, storage cluster management, MGS, MDS, OSS, login, and compute nodes are written in this section. For the compute nodes that belong to a node group, write in the NodeGroup section.

ManageNetMaster

Specify the representative IP address of the compute cluster management node in the management network.

When the node configuration applies to the following conditions, this item is required.

- Compute cluster management node is in a redundant configuration
- The storage cluster management node does not combine system management node and it is in a redundant configuration.

SCM

The configuration of the storage cluster management node is written in this section. It has the following items.

Item name	Description
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item.
ManageNet	Specifies the IP address of the management network.
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.

CCM

The configuration of the compute cluster management node is written in this section. It has the following items.

Item name	Description
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item.
ManageNet	Specifies the IP address of the management network.
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.

MGS

The configuration of the MGS node is written in this section. It has the following item.

Item name	Description
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.

The writing method depends on whether IP addresses are individually assigned to nodes or multiple addresses are batch assigned in accordance with rules.

To assign IP addresses on an individual basis, write in the following items.

Item name	Description
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item.
ManageNet	Specifies the IP address of the management network.
CombinedType	Specifies a dual use of node types in this item. "MDS" or "MDS,OSS" can be specified.

To assign IP addresses on the basis of rules, write in the following section.

Item name	Description
ManageNetRule	<p>This section defines IP address assignment rules for the management network. It has the following items.</p> <p>NetworkAddress Specifies sequential network addresses to assign to nodes.</p> <p>NetMask Specifies the netmask value.</p> <p>StartIPAddress Specifies the first IP address in the sequence.</p> <p>MaskUnit Specifies the unit for the mask boundary. You can specify the following values: Node group : NodeGroup Boot group : BootGroup Node : Node</p> <p>If this item is omitted, no boundary is defined for the mask.</p> <p>NumberOfUnit Specifies the number of units for the mask boundary. Unless MaskUnit is specified, this item cannot be specified.</p> <p>IncrementMask Specifies the increment in the IP address mask value when the sequential numbers reach the mask boundary. If MaskUnit is specified, this item must be specified.</p>
ControlNetRule	<p>This section defines IP address assignment rules for the control network. It has the same items as ManageNetRule.</p>

MDS

The configuration of the MDS node is written in this section.

It has the same items as the MGS section.

However, CombinedType cannot be specified.

OSS

The configuration of the OSS node is written in this section.
It has the same items as the MGS section.
However, CombinedType cannot be specified.

LN

The configuration of the login node is written in this section.
It has the same items as the MGS section.
However, CombinedType cannot be specified.

CN

The configuration of a compute node is written in this section.
It has the same items as the MGS section.
However, CombinedType cannot be specified.

MULTIUSE

The configuration of a multiuse node is written in this section. It has the following items.

Item name	Description
NodeTypename	Specifies a node type name. Consisting of 3 to 5 characters, the name can use uppercase letters and numbers. The node type names defined for the Job Operation Software cannot be used.
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.
PowerControlOnly	Specifies the use of only the control function on multiuse nodes of any model if "yes" is specified for this item. If this item is omitted, "no" is specified.

To assign IP addresses on an individual basis, write in the following items.

Item name	Description
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item.
ManageNet	Specifies the IP address of the management network.

To assign IP addresses on the basis of rules, write in the following items.

Item name	Description
ManageNetRule	This section defines IP address assignment rules for the management network. It has the same items as the ManageNetRule subsection of the MGS section.
ControlNetRule	This section defines IP address assignment rules for the control network.

It has the same items as the ControlNetRule subsection of the MGS section.

BootGroup

Boot groups are written in this section. It has the following items.

Item name	Description
Model	Specifies a model name. Specify "FT".
NumberOfBootGroups	Specifies the number of boot groups to set.

NodeGroup

A node group is written in this section.

Model

Specify a model name. Specify "FT" only when defining the FX server.

NumberOfNodeGroups

If you define multiple node groups by assigning IP addresses on the basis of rules in a NodeGroup section, specify the number of node groups to set.

NumberOfBootGroups

Specify the number of boot groups to set.

ManageNetMaster

Specify the representative IP address of the compute cluster sub management node in the management network.
If the compute cluster sub management node is in a redundant configuration, this item is required.

ManageNetRule

This section is written to assign IP addresses on the basis of rules to the nodes in a node group.
It has the same items as the ManageNetRule subsection of the MGS section.

ControlNetRule

This section is written to assign IP addresses on the basis of rules to the nodes in a node group.
It has the same items as the ControlNetRule subsection of the MGS section.

CCS

The configuration of the compute cluster sub management node is written in this section. It has the following items.

Item name	Description
Model	Specifies a model name. Specify a 2-character name consisting of single-byte uppercase letters and numbers. "FT", "FH", "FX", and "MP" cannot be specified.

To assign IP addresses on an individual basis, write in the following items.

Item name	Description
ControlNet	Specifies the IP address of the control network. When using a power control plugin for a node with neither "PG" nor "FT" as its model name, you can define any character string for this item. If this item is omitted, an IP address is assigned in accordance with the IP address assignment rules defined in the NodeGroup section.
ManageNet	Specifies the IP address of the management network. If this item is omitted, an IP address is assigned in accordance with the IP address assignment rules defined in the NodeGroup section.

To assign IP addresses on the basis of rules, write in the following items.

Item name	Description
ManageNetRule	This section is written to assign an IP address for the management network on the basis of rules. It has the same items as the ManageNetRule subsection of the MGS section.
ControlNetRule	This section is written to assign an IP address for the control network on the basis of rules. It has the same items as the ControlNetRule subsection of the MGS section.
NumberOfNodes	Specifies the number of compute cluster sub management nodes defined in a node group.

CN

The configuration of a compute node is written in this section.
It has the same items as the CCS section.

When specifying "del-clst" in the Type item of the Operation section, write in the following section and item.

Cluster

Cluster definitions are written in this section. It has the following item.

ClusterName

Specify the name of the cluster to delete.

When specifying "add-node" in the Type item of the Operation section, write in the following sections and items.

Cluster

Cluster definitions are written in this section. It has the following items.

ClusterName

Specify the name of the cluster where a node will be added.

HA

This section configures a redundancy function.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

MasterGroup

Compute cluster management, storage cluster management, MGS, MDS, OSS, login, and compute nodes are written in this section. For the compute nodes that belong to a node group, write in the NodeGroup section.

ManageNetMaster

Specify the representative IP address of the compute cluster management node or storage cluster management node in the management network.

If the compute cluster management node or storage cluster management node is added, this item is required.

CCM

The configuration of the compute cluster management node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

SCM

The configuration of the storage cluster management node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

MGS

The configuration of the MDS node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

MDS

The configuration of the MDS node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

OSS

The configuration of the OSS node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

LN

The configuration of the login node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

CN

The configuration of a compute node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

MULTIUSE

The configuration of a multiuse node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

BootGroup

Boot groups are written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

NodeGroup

A node group is written in this section. In the case of adding a node group, the section has the same items as when specifying "add-clst" in the Type item of the Operation section. To add a compute node to a defined node group, write in the following items.

NodeGroupID

To add a compute node to a defined node group, write in the following items.

Model

Specify a model name. Specify "FT" only when defining the FX server.

NumberOfBootGroups

Specify the number of boot groups to add.

CN

The configuration of a compute node is written in this section.

It has the same items as when specifying "add-clst" in the Type item of the Operation section.

When specifying "del-node" in the Type item of the Operation section, write in the following sections and items.

Cluster

Cluster definitions are written in this section. It has the following items.

ClusterName

Specify the name of the cluster where a node will be deleted.

HA

This section configures a redundancy function. It has the following item.

Node

This section defines nodes in a redundant configuration. It has the following items.

Item name	Description
NodeType	Specifies the node type of the nodes that have the mount setting to be deleted.
Mount	This section defines mount information. It has the following item. Device Specifies the name of the device to delete.

Node

The node to delete is written in this section.

If you are deleting multiple nodes, write this section for each node you want to delete.

It has the following item.

NodeID

Specify the node ID to delete.

Node IDs of system management, compute cluster management, compute cluster sub management, storage cluster management, and nodes of FX server cannot be specified.

NodeGroup

The node group to delete is written in this section.

If you are deleting multiple node groups, write this section for each node group you want to delete.

It has the following item.

NodeGroupID

Specify the ID of the node group to delete.

BootGroup

The boot group to delete is written in this section.

If you are deleting multiple boot groups, write this section for each boot group you want to delete.

It has the following item.

BootGroupID

Specify the ID of the boot group to delete.

When specifying "update" in the Type item of the Operation section, write in the following sections and items.

Cluster

Cluster definitions are written in this section. It has the following items.

ClusterName

Specify the name of a cluster.

HA

This section defines nodes in a redundant configuration. It has the following item.

Node

This section defines nodes in a redundant configuration. It has the following items.

Item name	Description
NodeType	Specifies the node type of the target nodes for the mount setting.
Mount	<p>This section defines mount information. It has the following items.</p> <p>Device Specifies the name of the device to rewrite. Instead of explicitly specifying the device, you can specify the file system for this mounted device by using either a UUID, volume label or by-id.</p> <p>MountPoint Specifies rewriting, in this item, at the mount point for the file system to be mounted.</p> <p>FSType Specifies rewriting of the type of file system in this item.</p> <p>OptList Specifies rewriting, in this item, of the options specified in the mount command. You can specify multiple options by delimiting them with a comma (.).</p>

Node

The node for rewriting information is written in this section. It has the following items.

NodeID

Specify the ID of the node to change.

ManageNetMaster

Specify the representative IP address for the changed management network.

ManageNet

Specify the IP address for the changed management network.

ControlNetSrv

Specify the IP address for the changed control network (communication).

ControlNet

Specify the IP address for the changed control network.

TofuNet

Specify the IP address for the changed Tofu network.

Model

Specify the changed model name. The model name of nodes defined with the model name "FT" cannot be changed.

PowerControlOnly

Specify "yes" in this item to use only the control function on multiuse nodes. If this item is omitted, "no" is specified.

[EXAMPLE]

An example of the pxsystem.conf file is shown below.

```
Operation {
    Type = "add-clst"
}
Cluster {
    ClusterName = "cluster1"
    ClusterType = "compute"
    UseStorageClusterName = "storage"
    HA {
        Node {
            NodeType = "SMM"
            Mount {
                Device = "oooo-xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
                MountPoint = "/var/opt/FJSVtcs/shared_disk"
                FSType = "ext4"
                OptList = "rw"
            }
        }
    }
}
SMMGroup {
    ManageNetMaster = "192.0.2.1"
    SMM {
        ControlNetSrv = "198.51.100.1"
        ControlNet = "198.51.100.3"
        ManageNet = "192.0.2.2"
        Model = "PG"
    }
    SMM {
        ControlNetSrv = "198.51.100.2"
        ControlNet = "198.51.100.4"
        ManageNet = "192.0.2.3"
        Model = "PG"
    }
}
MasterGroup {
    ManageNetMaster = "192.0.2.4"
    CCM {
        ControlNet = "198.51.100.5"
        ManageNet = "192.0.2.5"
        Model = "PG"
    }
    CCM {
        ControlNet = "198.51.100.6"
        ManageNet = "192.0.2.6"
        Model = "PG"
    }
}
LN {
```

```

        ControlNet = "198.51.100.7"
        ManageNet = "192.0.2.7"
        Model = "PG"
    }
    MULTIUSE {
        NodeType = "NFS"
        ControlNet = "198.51.100.8"
        ManageNet = "192.0.2.8"
        Model = "PG"
    }
}
NodeGroup {
    Model = "FT"
    NumberOfNodeGroups = "3"
    NumberOfBootGroups = "24"
    CCS {
        Model = "PG"
        NumberOfNodes = "2"
        ControlNetRule {
            NetworkAddress = "198.51.100.0"
            NetMask = "255.255.255.0"
            StartIPAddress = "198.51.100.11"
            MaskUnit = "NodeGroup"
            NumberOfUnit = "2"
            IncrementMask = "255.255.255.0"
        }
        ManageNetRule {
            NetworkAddress = "192.0.2.0"
            NetMask = "255.255.255.0"
            StartIPAddress = "192.0.2.11"
            MaskUnit = "NodeGroup"
            NumberOfUnit = "2"
            IncrementMask = "255.255.255.0"
        }
    }
}
}
}
}

```

The following is an example of deleting nodes:

```

Operation {
    Type = "del-node"
}
Cluster {
    ClusterName = "cluster1"
    Node {
        NodeID = "0xFFFF0004"
    }
    Node {
        NodeID = "0xFFFF0005"
    }
}
}

```

[NOTE]

- To not define a control network for a node that does not have "PG" or "FT" as its model name, specify "-" in the ControlNet item.

[SEE ALSO]

pxsystemadm(8)

4.4 Exit Scripts

The job resource management exit function of the Job Operation Software provides scripts (exit script) such as shell and perl for use in exit processing.

The following is a reference to the exit script.

Exit scripts are listed in alphabetical order.

4.4.1 postfree

[NAME]

postfree - File executed after job resource release

[SYNOPSIS]

`/userdir/postfree`

`userdir`: Directory specified by `ExitFuncScriptDir` in the `pmrsc.conf` file

[DESCRIPTION]

Postfree describes the processing executed after job resource release. The content of the file is arbitrary.

The job operation administrator performs the definition setting of this file.

Specify the job operation administrator as the owner of this file and set the permission to 0660.

Also, locate this file on the system management node.

Perform the definition setting on the system management node and use the `pmrscadm` command to reflect it in the configuration file.

Postfree is executed as system processes.

Postfree is executed in the following environment.

- Execution node
It is executed on all the compute nodes within the job allocation range.
- Execution timing
It is executed after job resource release (including the time of the job suspending execution). The epilogue script is executed before postfree.
- Execution authority
It is executed under the root authority.
- Execution result
The standard output or the standard error output in postfree does not output the result.
- Execution directory
It is executed on the home directory of the root user.
- Environment variable
The following variables can be referenced in postfree.

Environment variable name	Description
PJM_ADAPTIVE_ELAPSED_TIME_MAX [FX]	Maximum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_ADAPTIVE_ELAPSED_TIME_MIN [FX]	Minimum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_APPNAME	Character string specified in pjsub --appname
PJM_ASSIGN_LOGICAL_CPU [PG]	The range of logical CPUs that job processes to work job: Job processes can use only logical CPUs for the job on allocated CPU cores. all: Job processes can use all logical CPUs on allocated CPU cores.
PJM_ASSIGN_ONLINE_NODE [FX]	1 is set if you specify in pjsub --mpi assign-online-node.

Environment variable name	Description
PJM_AT	Job execution start time specified in the --at option of the pjsub command
PJM_AUXGID	Auxiliary group ID
PJM_BULKNUM	Bulk number (set only for bulk job)
PJM_COMMENT	Character string specified in the --comment option of the pjsub command
PJM_CORE_MEM_LIMIT [PG]	The amount of memory per CPU core (byte) specified in pjsub -L core-mem
PJM_CPUINF	Information on allocated CPU core
PJM_CUSTOM_RESOURCES	Custom resource information specified in pjsub -L Custom resource name and number of custom resource or kind of custom resource are set. In the case of the custom resources set to node, "/n" is added to custom resource name.
PJM_DPREFIX	Prefix that indicates that it is an instruction line in a job script for job operation software
PJM_ELAPSED_TIME_MODE [FX]	Method of specifying the executable time for the job specified with the -L elapse option of the pjsub command adaptive : For specifications like "-L elapse=30:00-" or "-L elapse=30:00-1:00:00" fixed : For specifications like "-L elapse=30:00"
PJM_ELAPSE_LIMIT	Maximum executable time of the job (second) specified in pjsub -L elapse
PJM_ENVIRONMENT	"BATCH" for batch job and "INTERACT" for interactive job
PJM_EXEC_POLICY [PG]	Execution mode policy specified in pjsub -P exec-policy simplex: SIMPLEX (A physical node is monopolized) share: SHARE (A physical node is shared)
PJM_FSNAME	Character string specified in pjsub --fs
PJM_GID	Group ID of the job execution user
PJM_JOBDIR	Path of the current directory at the start of job script execution
PJM_JOBEXIT	Job termination code (invalid when the job suspend or an error before job starts is the trigger for execution) 1: Normal termination 2: CPU time limit violation 3: Elapsed time limit violation 4: KILL due to node service failure 5: Unable to start the job due to process generation failure 6: Forcible release requested by PJM 7: Hardware error (ICC error) 8: Termination due to OOM-Killer behavior 9: Collection failure of extended CPU statistic information 11: Voluntary memory acquisition failure 21: Epilogue execution failure 255: Other error
PJM_JOBID	Job ID
PJM_JOBNAME	Job name
PJM_JOBNUM	Number of jobs in execution on the job allocation compute node
PJM_LLIO_ASYNC_CLOSE	Operation of whether or not to asynchronously close files on first-layer storage and second-layer storage on: Asynchronous close off: Synchronous close

Environment variable name	Description
PJM_LLIO_AUTO_READAHEAD	Operation of whether or not to automatically read ahead when a job is going to try reading consecutive areas again on first-layer storage or second-layer storage multiple times on: Read ahead. off: Do not read ahead.
PJM_LLIO_CN_CACHED_WRITE_SIZE	If the output size (in bytes) of data to write to first-layer storage is smaller than this value, the data is not immediately output to the storage but instead stored temporarily in cache on the compute node.
PJM_LLIO_CN_CACHE_SIZE	Size (in bytes) of the cache on the compute node in layered storage
PJM_LLIO_CN_READ_CACHE	Specifies whether a file read from layered storage by a job is cached on the compute node on: Cached off: Not cached
PJM_LLIO_LOCALTMP_SIZE	Size (in bytes) of the node temporary area on first-layer storage
PJM_LLIO_PERF	Specifies whether to output the LLIO performance information file on: Output off: Do not output
PJM_LLIO_PERF_PATH	Path to the LLIO performance information file
PJM_LLIO_SHAREDTMP_SIZE	Size (in bytes) of the shared temporary area on first-layer storage
PJM_LLIO_SIO_READ_CACHE	Specifies whether a file read from second-layer storage to the compute node is cached in first-layer storage on: Cached off: Not cached
PJM_LLIO_STRIPE_COUNT	Number of stripes per file in distributed placement of a file in first-layer storage
PJM_LLIO_STRIPE_SIZE	Stripe size (in bytes) in distributed placement of a file in first-layer storage
PJM_LLIO_UNCOMPLETED_FILEINFO_PATH	Path to the file for an output list of the file names of any files that did not complete output to second-layer storage and remained in the cache at job end
PJM_LOCALTMP	Path of a temporary area in a node on first-layer storage
PJM_MAILOPTION	Operation specified with the -m option of the pjsub command for mail notification
PJM_MAILSENDTO	Mail destination user specified in the --mail-list option of the pjsub command
PJM_MPI_PROC	Maximum number of processes generated when an MPI program starts specified in pjsub --mpi proc
PJM_MPI_SHAPE_X [FX]	The number of nodes in the X-axis direction specified in pjsub --mpi shape
PJM_MPI_SHAPE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsub --mpi shape
PJM_MPI_SHAPE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsub --mpi shape
PJM_NET_ROUTE [FX]	Operation when a link goes down for the Tofu interconnect used as the communication path by a job dynamic: Change the communication path, and job execution continues. static: Do not change the communication path. The job ends abnormally.
PJM_NODE	Value specified in pjsub -L node
PJM_NODE_ALLOCATION_MODE [FX]	Node allocation methods specified in pjsub -L node torus: torus mode

Environment variable name	Description
	mesh: mesh mode noncont: non-contiguous mode
PJM_NODE_MEM_LIMIT [FX]	The upper limit of memory usage (byte) for each node specified in pjsub -L node-mem
PJM_NODE_TYPE	Type of compute node 0: Compute node where a job script does not start 1: Compute node where a job script starts
PJM_NODE_X [FX]	The number of nodes in the X-axis direction specified in pjsub -L node
PJM_NODE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsub -L node
PJM_NODE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsub -L node
PJM_NORESTART	1 is set if you specify in pjsub --norestart
PJM_O_HOME	Environment variable HOME of the user who executed the pjsub command
PJM_O_HOST	Name of the host where the pjsub command is executed
PJM_O_LANG	Environment variable LANG of the user who executed the pjsub command
PJM_O_LOGNAME	Environment variable LOGNAME of the user who executed the pjsub command
PJM_O_MAIL	Environment variable MAIL of the user who executed the pjsub command
PJM_O_NODEINF [PG]	Path of the allocated node list file
PJM_O_PATH	Environment variable PATH of the user who executed the pjsub command
PJM_O_SHELL	Environment variable SHELL of the user who executed the pjsub command
PJM_O_TZ	Environment variable TZ of the user who executed the pjsub command
PJM_O_WORKDIR	Current directory when the pjsub command is executed
PJM_PROC_BY_NODE	Maximum number of processes generated by an MPI program specified in pjsub --mpi max-proc-per-node
PJM_RANK_MAP_BYCHIP	The rules on assigning nodes for the ranks specified in pjsub --mpi rank-map-bychip
PJM_RANK_MAP_BYNODE	The rules on assigning nodes for the ranks specified in pjsub --mpi rank-map-bynode
PJM_RETRYNUM	Number of job execution retry times
PJM_RSCGRP	Resource group name specified in pjsub -L rscgrp or -L rg
PJM_RSCUNIT	Resource unit name specified in pjsub -L rscunit or -L ru
PJM_SHAREDTMP	Path to a shared temporary area
PJM_SHELL	Path of the shell which executes a job
PJM_SHELLEXIT	Shell termination code (invalid when the job suspend or an error before job starts is the trigger of the execution)
PJM_SPDINF	Trigger to execute the applicable exit function 1: Triggered by operations other than the suspend and resume operations 2: Triggered by the suspend and resume operations
PJM_STDERR_PATH	The path name of the job standard error output file specified in pjsub -e
PJM_STDOUT_PATH	The path name of the job standard output file specified in pjsub -o
PJM_STEPNUM	Step number (set only for step job)
PJM_SUBJOBID	Sub job ID (set only for step job and bulk job)
PJM_UID	User ID of the job execution user

Environment variable name	Description
PJM_VNID	Virtual node ID
PJM_VNODE	Value specified in pjsub -L vnode
PJM_VNODE_CORE	Value specified in pjsub -L vnode-core or -L vnode=(core=)
PJM_VNODE_MEM_LIMIT [PG]	The amount of memory (byte) for each virtual node specified in pjsub -L vnode-mem (or -L mem)
PJM_VN_POLICY [PG]	The virtual node placement policy specified in pjsub -P vn-policy abs-pack: Absolutely PACK pack: PACK abs-unpack: Absolutely UNPACK unpack: UNPACK
Arbitrary environment variable	Environment variable specified in pjsub -x or pjsub -X (Only environment variables starting with "PJM_")

Execution of a job is controllable by the exit code of postfree.

0: Do nothing

1: Changing job state to ERROR

2: Re-executing the jobs

3: Changing job state to HOLD

4: Forcibly terminated the job

Other : Do nothing

[NOTE]

In a setting where postfree operates, RUNNING or RUNNING-E state may take longer for postfree processing.

In addition, other job states may remain unchanged on the compute node where postfree is being executed, until postfree processing terminates.

[SEE ALSO]

prealloc(5), predel(5), pmrsc.conf(5), pmpjm.conf(5), pmrscadm(8)

4.4.2 prealloc

[NAME]

prealloc - File executed before job resource allocation

[SYNOPSIS]

`/userdir/prealloc`

`userdir`: Directory specified by ExitFuncScriptDir in the pmrsc.conf file

[DESCRIPTION]

Prealloc describes the processing executed before job resource allocation. The content of the file is arbitrary.

The job operation administrator performs the definition setting of this file.

Specify the job operation administrator as the owner of this file and set the permission to 0660.

Also, locate this file on the system management node.

Perform the definition setting on the system management node and use the pmrscadm command to reflect the configuration file.

Prealloc is executed as system processes.

Prealloc is executed in the following environment.

- Execution node

It is executed on all the compute nodes within the job allocation range.

- Execution timing

It is executed before job resource allocation (including the time of the job resuming execution). It is executed before the prologue script.

- Execution authority
It is executed under the root authority.
- Execution result
The standard output or the standard error output in prealloc does not output the result.
- Execution directory
It is executed on the home directory of the root user.
- Environment variable
The following variables can be referenced in prealloc.

Environment variable name	Description
PJM_ADAPTIVE_ELAPSED_TIME_MAX [FX]	Maximum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_ADAPTIVE_ELAPSED_TIME_MIN [FX]	Minimum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_APPNAME	Character string specified in pjsub --appname
PJM_ASSIGN_LOGICAL_CPU [PG]	The range of logical CPUs that job processes to work job: Job processes can use only logical CPUs for the job on allocated CPU cores. all: Job processes can use all logical CPUs on allocated CPU cores.
PJM_ASSIGN_ONLINE_NODE [FX]	1 is set if you specify in pjsub --mpi assign-online-node.
PJM_AT	Job execution start time specified in the --at option of the pjsub command
PJM_AUXGID	Auxiliary group ID
PJM_BULKNUM	Bulk number (set only for bulk job)
PJM_COMMENT	Character string specified in the --comment option of the pjsub command
PJM_CORE_MEM_LIMIT [PG]	The amount of memory per CPU core (byte) specified in pjsub -L core-mem
PJM_CPUINF	Information on allocated CPU core
PJM_CUSTOM_RESOURCES	Custom resource information specified in pjsub -L Custom resource name and number of custom resource or kind of custom resource are set. In the case of the custom resources set to node, "/n" is added to custom resource name.
PJM_DPREFIX	Prefix that indicates that it is an instruction line in a job script for job operation software
PJM_ELAPSED_TIME_MODE [FX]	Method of specifying the executable time for the job specified with the -L elapse option of the pjsub command adaptive: For specifications like "-L elapse=30:00-" or "-L elapse=30:00-1:00:00" fixed: For specifications like "-L elapse=30:00"
PJM_ELAPSE_LIMIT	Maximum executable time of the job (second) specified in pjsub -L elapse
PJM_ENVIRONMENT	"BATCH" for batch job and "INTERACT" for interactive job
PJM_EXEC_POLICY [PG]	Execution mode policy specified in pjsub -P exec-policy simplex: SIMPLEX (A physical node is monopolized) share: SHARE (A physical node is shared)
PJM_FSNAME	Character string specified in pjsub --fs
PJM_GID	Group ID of the job execution user
PJM_JOBDIR	Path of the current directory at the start of job script execution
PJM_JOBID	Job ID

Environment variable name	Description
PJM_JOBNAME	Job name
PJM_JOBNUM	Number of jobs in execution on the job allocation compute node
PJM_LLIO_ASYNC_CLOSE	Operation of whether or not to asynchronously close files on first-layer storage and second-layer storage on: Asynchronous close off: Synchronous close
PJM_LLIO_AUTO_READAHEAD	Operation of whether or not to automatically read ahead when a job is going to try reading consecutive areas again on first-layer storage or second-layer storage multiple times on: Read ahead. off: Do not read ahead.
PJM_LLIO_CN_CACHED_WRITE_SIZE	If the output size (in bytes) of data to write to first-layer storage is smaller than this value, the data is not immediately output to the storage but instead stored temporarily in cache on the compute node.
PJM_LLIO_CN_CACHE_SIZE	Size (in bytes) of the cache on the compute node in layered storage
PJM_LLIO_CN_READ_CACHE	Specifies whether a file read from layered storage by a job is cached on the compute node on: Cached off: Not cached
PJM_LLIO_LOCALTMP_SIZE	Size (in bytes) of the node temporary area on first-layer storage
PJM_LLIO_PERF	Specifies whether to output the LLIO performance information file on: Output off: Do not output
PJM_LLIO_PERF_PATH	Path to the LLIO performance information file
PJM_LLIO_SHAREDTMP_SIZE	Size (in bytes) of the shared temporary area on first-layer storage
PJM_LLIO_SIO_READ_CACHE	Specifies whether a file read from second-layer storage to the compute node is cached in first-layer storage on: Cached off: Not cached
PJM_LLIO_STRIPE_COUNT	Number of stripes per file in distributed placement of a file in first-layer storage
PJM_LLIO_STRIPE_SIZE	Stripe size (in bytes) in distributed placement of a file in first-layer storage
PJM_LLIO_UNCOMPLETED_FILEINFO_PATH	Path to the file for an output list of the file names of any files that did not complete output to second-layer storage and remained in the cache at job end
PJM_LOCALTMP	Path of a temporary area in a node on first-layer storage
PJM_MAILOPTION	Operation specified with the -m option of the pjsb command for mail notification
PJM_MAILSENDTO	Mail destination user specified in the --mail-list option of the pjsb command
PJM_MPI_PROC	Maximum number of processes generated when an MPI program starts specified in pjsb --mpi proc
PJM_MPI_SHAPE_X [FX]	The number of nodes in the X-axis direction specified in pjsb --mpi shape
PJM_MPI_SHAPE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsb --mpi shape
PJM_MPI_SHAPE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsb --mpi shape
PJM_NET_ROUTE [FX]	Operation when a link goes down for the Tofu interconnect used as the communication path by a job

Environment variable name	Description
	dynamic: Change the communication path, and job execution continues. static: Do not change the communication path. The job ends abnormally.
PJM_NODE	Value specified in pjsub -L node
PJM_NODE_ALLOCATION_MODE [FX]	Node allocation methods specified in pjsub -L node torus: torus mode mesh: mesh mode noncont: non-contiguous mode
PJM_NODE_MEM_LIMIT [FX]	The upper limit of memory usage (byte) for each node specified in pjsub -L node-mem
PJM_NODE_TYPE	Type of compute node 0: Compute node where a job script does not start 1: Compute node where a job script starts
PJM_NODE_X [FX]	The number of nodes in the X-axis direction specified in pjsub -L node
PJM_NODE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsub -L node
PJM_NODE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsub -L node
PJM_NORESTART	1 is set if you specify in pjsub --norestart
PJM_O_HOME	Environment variable HOME of the user who executed the pjsub command
PJM_O_HOST	Name of the host where the pjsub command is executed
PJM_O_LANG	Environment variable LANG of the user who executed the pjsub command
PJM_O_LOGNAME	Environment variable LOGNAME of the user who executed the pjsub command
PJM_O_MAIL	Environment variable MAIL of the user who executed the pjsub command
PJM_O_NODEINF [PG]	Path of the allocated node list file
PJM_O_PATH	Environment variable PATH of the user who executed the pjsub command
PJM_O_SHELL	Environment variable SHELL of the user who executed the pjsub command
PJM_O_TZ	Environment variable TZ of the user who executed the pjsub command
PJM_O_WORKDIR	Current directory when the pjsub command is executed
PJM_PROC_BY_NODE	Maximum number of processes generated by an MPI program specified in pjsub --mpi max-proc-per-node
PJM_RANK_MAP_BYCHIP	The rules on assigning nodes for the ranks specified in pjsub --mpi rank-map-bychip
PJM_RANK_MAP_BYNODE	The rules on assigning nodes for the ranks specified in pjsub --mpi rank-map-bynode
PJM_RETRYNUM	Number of job execution retry times
PJM_RSCGRP	Resource group name specified in pjsub -L rscgrp or -L rg
PJM_RSCUNIT	Resource unit name specified in pjsub -L rscunit or -L ru
PJM_SHAREDTMP	Path to a shared temporary area
PJM_SHELL	Path of the shell which executes a job
PJM_SPDINF	Trigger to execute the applicable exit function 1: Triggered by operations other than the suspend and resume operations 2: Triggered by the suspend and resume operations
PJM_STDERR_PATH	The path name of the job standard error output file specified in pjsub -e
PJM_STDOUT_PATH	The path name of the job standard output file specified in pjsub -o

Environment variable name	Description
PJM_STEPNUM	Step number (set only for step job)
PJM_SUBJOBID	Sub job ID (set only for step job and bulk job)
PJM_UID	User ID of the job execution user
PJM_VNID	Virtual node ID
PJM_VNODE	Value specified in pjsub -L vnode
PJM_VNODE_CORE	Value specified in pjsub -L vnode-core or -L vnode=(core=)
PJM_VNODE_MEM_LIMIT [PG]	The amount of memory (byte) for each virtual node specified in pjsub -L vnode-mem (or -L mem)
PJM_VN_POLICY [PG]	The virtual node placement policy specified in pjsub -P vn-policy abs-pack: Absolutely PACK pack: PACK abs-unpack: Absolutely UNPACK unpack: UNPACK
Arbitrary environment variable	Environment variable specified in pjsub -x or pjsub -X (Only environment variables starting with "PJM_")

Execution of a job is controllable by the exit code of prealloc.

0: Do nothing

1: Changing job state to ERROR

2: Re-executing the jobs

3: Changing job state to HOLD

4: Forcibly terminated the job

Other: Do nothing

[NOTE]

In a setting where prealloc operates, RUNNING-A may take longer for prealloc processing.

In addition, other job states may remain unchanged on the compute node where prealloc is being executed, until prealloc processing terminates.

[SEE ALSO]

predel(5), postfree(5), pmrsc.conf(5), pmrscadm(8)

4.4.3 predel

[NAME]

predel - File executed when the job is deleted

[SYNOPSIS]

```
/userdir/predel
```

userdir: Directory specified by ExitFuncScriptDir in the pmrsc.conf file

[DESCRIPTION]

Predel describes the processing executed when the job is deleted. The content of the file is arbitrary.

The job operation administrator performs the definition setting of this file.

Specify the job operation administrator as the owner of this file and set the permission to 0660.

Also, locate this file on the system management node.

Perform the definition setting on the system management node and use the pmrscadm command to reflect it in the configuration file.

Predel is executed as system processes.

Predel is executed in the following environment.

- Execution node
It is executed on all the compute nodes within the job allocation range.
- Execution timing
It is executed when:
 - the pjdcl command is executed;
 - the pjhold command is executed;
 - a job resource (elapsed time, CPU time, or memory) limit is exceeded;
 - the compute node (or its service) fails; and
 - the job is deleted by the system for a reason other than the above.
- * Environment variable PJM_DELINEF indicates the execution timing.
- Execution authority
It is executed under the root authority.
- Execution result
The standard output or the standard error output in predel does not output the result.
- Execution directory
It is executed on the home directory of the root user.
- Environment variable
The following variables can be referenced in predel.

Environment variable name	Description
PJM_ADAPTIVE_ELAPSED_TIME_MAX [FX]	Maximum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_ADAPTIVE_ELAPSED_TIME_MIN [FX]	Minimum value (in seconds) of the elapsed time limit for the job specified with the -L elapse option of the pjsub command
PJM_APPNAME	Character string specified in pjsub --appname
PJM_ASSIGN_LOGICAL_CPU [PG]	The range of logical CPUs that job processes to work job: Job processes can use only logical CPUs for the job on allocated CPU cores. all: Job processes can use all logical CPUs on allocated CPU cores.
PJM_ASSIGN_ONLINE_NODE [FX]	1 is set if you specify in pjsub --mpi assign-online-node.
PJM_AT	Job execution start time specified in the --at option of the pjsub command
PJM_AUXGID	Auxiliary group ID
PJM_BULKNUM	Bulk number (set only for bulk job)
PJM_COMMENT	Character string specified in the --comment option of the pjsub command
PJM_CORE_MEM_LIMIT [PG]	The amount of memory per CPU core (byte) specified in pjsub -L core-mem
PJM_CPUINF	Information on allocated CPU core
PJM_CUSTOM_RESOURCES	Custom resource information specified in pjsub -L Custom resource name and number of custom resource or kind of custom resource are set. In the case of the custom resources set to node, "/n" is added to custom resource name.
PJM_DELINEF	Indicates which is the trigger for executing the predel script. 1: The pjdcl command execution is the trigger for execution. 2: The pjhold command execution is the trigger for execution.

Environment variable name	Description
	<p>3: A deletion request by the job manager function or the job scheduler function is the trigger for execution.</p> <p>4: An error on the compute node is the trigger for execution.</p> <p>6: An elapsed time limit violation is the trigger for execution.</p> <p>7: Excessive memory use is the trigger for execution.</p> <p>8: OOM-Killer execution is the trigger for execution.</p> <p>9: The detection of an error in the job execution environment is the trigger for execution. [McKernel only]</p>
PJM_DPREFIX	Prefix that indicates that it is an instruction line in a job script for job operation software
PJM_ELAPSED_TIME_MODE [FX]	<p>Method of specifying the executable time for the job specified with the -L elapse option of the pjsub command</p> <p>adaptive : For specifications like "-L elapse=30:00-" or "-L elapse=30:00-1:00:00"</p> <p>fixed : For specifications like "-L elapse=30:00"</p>
PJM_ELAPSE_LIMIT	Maximum executable time of the job (second) specified in pjsub -L elapse
PJM_ENVIRONMENT	"BATCH" for batch job and "INTERACT" for interactive job
PJM_EXEC_POLICY [PG]	<p>Execution mode policy specified in pjsub -P exec-policy</p> <p>simplex: SIMPLEX (A physical node is monopolized)</p> <p>share: SHARE (A physical node is shared)</p>
PJM_FSNAME	Character string specified in pjsub --fs
PJM_GID	Group ID of the job execution user
PJM_JOBDIR	Path of the current directory at the start of job script execution
PJM_JOBID	Job ID
PJM_JOBNAME	Job name
PJM_LLIO_ASYNC_CLOSE	<p>Operation of whether or not to asynchronously close files on first-layer storage and second-layer storage</p> <p>on: Asynchronous close</p> <p>off: Synchronous close</p>
PJM_LLIO_AUTO_READAHEAD	<p>Operation of whether or not to automatically read ahead when a job is going to try reading consecutive areas again on first-layer storage or second-layer storage multiple times</p> <p>on: Read ahead.</p> <p>off: Do not read ahead.</p>
PJM_LLIO_CN_CACHED_WRITE_SIZE	If the output size (in bytes) of data to write to first-layer storage is smaller than this value, the data is not immediately output to the storage but instead stored temporarily in cache on the compute node.
PJM_LLIO_CN_CACHE_SIZE	Size (in bytes) of the cache on the compute node in layered storage
PJM_LLIO_CN_READ_CACHE	<p>Specifies whether a file read from layered storage by a job is cached on the compute node</p> <p>on: Cached</p> <p>off: Not cached</p>
PJM_LLIO_LOCALTMP_SIZE	Size (in bytes) of the node temporary area on first-layer storage
PJM_LLIO_PERF	<p>Specifies whether to output the LLIO performance information file</p> <p>on: Output</p> <p>off: Do not output</p>
PJM_LLIO_PERF_PATH	Path to the LLIO performance information file

Environment variable name	Description
PJM_LLIO_SHAREDTMP_SIZE	Size (in bytes) of the shared temporary area on first-layer storage
PJM_LLIO_SIO_READ_CACHE	Specifies whether a file read from second-layer storage to the compute node is cached in first-layer storage on: Cached off: Not cached
PJM_LLIO_STRIPE_COUNT	Number of stripes per file in distributed placement of a file in first-layer storage
PJM_LLIO_STRIPE_SIZE	Stripe size (in bytes) in distributed placement of a file in first-layer storage
PJM_LLIO_UNCOMPLETED_FILEINFO_PATH	Path to the file for an output list of the file names of any files that did not complete output to second-layer storage and remained in the cache at job end
PJM_LOCALTMP	Path of a temporary area in a node on first-layer storage
PJM_MAILOPTION	Operation specified with the -m option of the pjsub command for mail notification
PJM_MAILSENDTO	Mail destination user specified in the --mail-list option of the pjsub command
PJM_MPI_PROC	Maximum number of processes generated when an MPI program starts specified in pjsub --mpi proc
PJM_MPI_SHAPE_X [FX]	The number of nodes in the X-axis direction specified in pjsub --mpi shape
PJM_MPI_SHAPE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsub --mpi shape
PJM_MPI_SHAPE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsub --mpi shape
PJM_NET_ROUTE [FX]	Operation when a link goes down for the Tofu interconnect used as the communication path by a job dynamic: Change the communication path, and job execution continues. static: Do not change the communication path. The job ends abnormally.
PJM_NODE	Value specified in pjsub -L node
PJM_NODE_ALLOCATION_MODE [FX]	Node allocation methods specified in pjsub -L node torus: torus mode mesh: mesh mode noncont: non-contiguous mode
PJM_NODE_MEM_LIMIT [FX]	The upper limit of memory usage (byte) for each node specified in pjsub -L node-mem
PJM_NODE_TYPE	Type of compute node 0: Compute node where a job script does not start 1: Compute node where a job script starts
PJM_NODE_X [FX]	The number of nodes in the X-axis direction specified in pjsub -L node
PJM_NODE_Y [FX]	The number of nodes in the Y-axis direction specified in pjsub -L node
PJM_NODE_Z [FX]	The number of nodes in the Z-axis direction specified in pjsub -L node
PJM_NORESTART	1 is set if you specify in pjsub --norestart
PJM_O_HOME	Environment variable HOME of the user who executed the pjsub command
PJM_O_HOST	Name of the host where the pjsub command is executed

Environment variable name	Description
PJM_O_LANG	Environment variable LANG of the user who executed the pjsub command
PJM_O_LOGNAME	Environment variable LOGNAME of the user who executed the pjsub command
PJM_O_MAIL	Environment variable MAIL of the user who executed the pjsub command
PJM_O_NODEINF [PG]	Path of the allocated node list file
PJM_O_PATH	Environment variable PATH of the user who executed the pjsub command
PJM_O_SHELL	Environment variable SHELL of the user who executed the pjsub command
PJM_O_TZ	Environment variable TZ of the user who executed the pjsub command
PJM_O_WORKDIR	Current directory when the pjsub command is executed
PJM_PROC_BY_NODE	Maximum number of processes generated by an MPI program specified in pjsub --mpi max-proc-per-node
PJM_RANK_MAP_BYCHIP	The rules on assigning nodes for the ranks specified in pjsub --mpi rankmap-bychip
PJM_RANK_MAP_BYNODE	The rules on assigning nodes for the ranks specified in pjsub --mpi rank-map-bynode
PJM_RETRYNUM	Number of job execution retry times
PJM_RSCGRP	Resource group name specified in pjsub -L rscgrp or -L rg
PJM_RSCUNIT	Resource unit name specified in pjsub -L rscunit or -L ru
PJM_SHAREDTMP	Path to a shared temporary area
PJM_SHELL	Path of the shell which executes a job
PJM_SPDINF	Trigger to execute the applicable exit function 1: Triggered by operations other than the suspend and resume operations 2: Triggered by the suspend and resume operations
PJM_STDERR_PATH	The path name of the job standard error output file specified in pjsub -e
PJM_STDOUT_PATH	The path name of the job standard output file specified in pjsub -o
PJM_STEPNUM	Step number (set only for step job)
PJM_SUBJOBID	Sub job ID (set only for step job and bulk job)
PJM_UID	User ID of the job execution user
PJM_VNID	Virtual node ID
PJM_VNODE	Value specified in pjsub -L vnode
PJM_VNODE_CORE	Value specified in pjsub -L vnode-core or -L vnode=(core=)
PJM_VNODE_MEM_LIMIT [PG]	The amount of memory (byte) for each virtual node specified in pjsub -L vnode-mem (or -L mem)
PJM_VN_POLICY [PG]	The virtual node placement policy specified in pjsub -P vn-policy abs-pack: Absolutely PACK pack: PACK abs-unpack: Absolutely UNPACK unpack : UNPACK
Arbitrary environment variable	Environment variable specified in pjsub -x or pjsub -X (Only environment variables starting with "PJM_")

Execution of a job is controllable by the exit code of predel.

0: Do nothing

1: Changing job state to ERROR

2: Re-executing the jobs

3: Changing job state to HOLD

4: Forcibly terminated the job

Other: Do nothing

[NOTE]

In a setting where predel operates, RUNNING-P, RUNNING, or RUNNING-E state may take longer for predel processing.

If predel is executed on a job in RUNNING state, the elapsed time includes the predel processing.

There are multiple predel execution triggers and predel is executed for every execution trigger. Therefore, predel may be executed twice or more in one job.

[SEE ALSO]

`pjdel(1)`, `pjhold(1)`, `prealloc(5)`, `postfree(5)`, `pmrsc.conf(5)`, `pmrscadm(8)`