

Fujitsu Software Technical Computing Suite V4.0L20

Job Operation Software Glossary

J2UL-2550-01ENZ0(04)
September 2024

Preface

Purpose of This Manual

This manual describes the term of "Job Operation Software" included in Technical Computing Suite.

See the respective manuals for hardware and related software terms.

Intended Readers

This manual is intended for all users of Job Operation Software.

The manual assumes that readers have a basic knowledge of Linux.

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.

Terms

The terms are listed in alphabetical order.

Manual Name

In this manual, the "Job Operating Software" part of the manual name is abbreviated.

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.
- All other trademarks are the property of their respective owners.

Date of publication and Version

Version	Manual Code
September 2024, Version 1.4	J2UL-2550-01ENZ0(04)
March 2021, Version 1.3	J2UL-2550-01ENZ0(03)
September 2020, Version 1.2	J2UL-2550-01ENZ0(02)
June 2020, Version 1.1	J2UL-2550-01ENZ0(01)
February 2020, First version	J2UL-2550-01ENZ0(00)

Copyright

Copyright FUJITSU LIMITED 2020-2024

Update history

Changes	Location	Version
Added NVIDIA H100 as GPU.	-	1.4
Changed NVIDIA Tesla to NVIDIA V100 and A100.	-	1.3
Added the following descriptions: <ul style="list-style-type: none">- I/O-exclusive mode- I/O-exclusive job- I/O-sharing mode- I/O-sharing job	-	1.2
Added the following descriptions: <ul style="list-style-type: none">- Job master node- Job slave nodes	-	1.1

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

Contents

Glossary	1
--------------------------------	---

Glossary

This section describes terms related to Job Operation Software and Job Operation Software. Parentheses [] are the main manual for a term or related description.

Administrator

Administrators are users who operate and manage the hardware and software configuring the system. The Job Operation Software categorizes administrators into three types (system administrator, cluster administrator, job operation administrator) depending on the range of responsibility. [Overview]

AOC (Active Optical Cable)

Product that optical transceiver combines with optical fiber cable. [Administrator's Guide for Maintenance]

Arbitrary model

In the Job Operation Software, the node other than PRIMERGY server and FX server. An arbitrary model name can be defined by using the configuration management function. [Setup Guide, Administrator's Guide for System Management]

Arena

In order to improve the performance for user memory allocation / free processing, a memory area (memory pool) allocated by the large page library is used beforehand. Arena refers to the combined area of this memory pool and the management area of the pool. [Administrator's Guide for HPC Extensions]

Assistant core [FX]

This system-dedicated core is designed to prevent system daemons and the processing of interrupts generated on the system in the CPU core on which jobs run. System daemons and interrupt processing, etc. are implemented on the assistant core. This reduces system noise that affects operation performance. [End-user's Guide, Administrator's Guide for Job Management]

Auto power control function for compute nodes

The auto power control function for compute nodes saves power by automatically powering off compute nodes not scheduled to be executing any jobs. It will power them on again later. [Administrator's Guide for Power Management]

Backfill function

The backfill function can execute lower-priority jobs prior to higher-priority jobs only if there are sufficient open resources. [Administrator's Guide for Job Management]

Backup/Restore function

The backup/restore function backs up the system disk as a disk image and restores from the disk image. [Overview, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Batch job

A batch job is a job which is executed un-interactively. In contrast, a job in which the interactive operation is possible for the program is called a "interactive job." [End-user's Guide]

BMC (Baseboard Management Controller)

Controller to control the power and to monitor an abnormality in the FX server and PRIMERGY server. [Administrator's Guide for System Management]

BoB (Bunch of Blade) [FX]

Control unit used in FX sever. A BoB is configured with 16 nodes. [Overview]

Boot group [FX]

A boot group is a collective unit for managing compute nodes. It is monitored by the boot I/O node. [Overview]

Boot I/O node (BIO: Boot I/O node) [FX]

I/O node that acts as the boot server of nodes in the FX server. [Overview]

Bulk job

A bulk job consists of multiple instances of the same job submitted at the same time for execution. A bulk job is used to execute individual jobs with different input parameters. Each job is called a sub job. [End-user's Guide]

Cache coloring

Physical page management method to positively increase the cache hit ratio. [End-user's Guide for HPC Extensions]

Capping

Capping is a mechanism that controls power consumption to prevent it from exceeding a set upper limit value. [Overview, Administrator's Guide for Power Management]

Cluster

A cluster is the unit of a group of nodes sharing job operations and files. The Job Operation Software has a compute cluster for executing jobs, a storage cluster for sharing files and a multiuse cluster that can be used for any purpose. [Overview, Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Job Management]

CMG (Core Memory Group) [FX]

CMG is the name of the group of compute cores, assistant core, cache, memory controller, and memory in the FX server CPU. [Administrator's Guide for Job Management]

CMU (CPU Memory Unit) [FX]

Unit with CPUs, memory, and AOCs. [Overview]

Cold reboot

Operation to restart the hardware after it has been powered off in restarting a node. In contrast, restarting the OS without powering off the hardware is called "warm reboot." [Administrator's Guide for System Management]

Command API

The command API is an API for calling the command interface of the job operation management function from an application. [Overview, End-user's Guide, Administrator's Guide for Job Management, API user's Guide for Command API]

Compute cluster

A group of nodes on which to execute the job. [Overview]

Compute cluster management node (CCM: Compute Cluster Management node)

Node that manages information relating to job operation in a compute cluster. [Overview]

Compute cluster sub management node (CCS: Compute Cluster Sub Management node)

Node for reducing the load of service monitoring by the compute cluster management node. [Overview]

Compute network

High-speed network used by parallel programs like MPI programs to communicate between nodes. [Overview, Setup Guide]

Compute node (CN: Compute node)

Node that executes jobs. [Overview]

Computer resource

Computer resources are the required resources for executing applications in jobs. These resources include CPUs, memory, disk resources, networks, CPU time, etc. Sometimes, they are simply called resources. [Overview, End-user's Guide, Administrator's Guide for Job Management]

Configuration management function

The configuration management function manages the configuration information of the nodes and networks in the system. [Overview, Administrator's Guide for System Management]

Control network

Network used to control the hardware of nodes, such as power control, notification of abnormalities. [Overview, Setup Guide]

Custom resource

An arbitrary resource that can be allocated to execute a job. This resource is set by the administrator. [End-user's Guide, Administrator's Guide for Job Management, Administrator's Guide for Power Management]

Deadline scheduling

A deadline schedule is one of the job scheduling functions. This function guarantees that jobs end by the specified deadline. Suppose the system needs to be stopped at a certain time for maintenance. The administrator sets a deadline, and jobs are scheduled according to the deadline. [Administrator's Guide for Job Management]

Detailed error information

Detailed error information is information that holds error details for the command API. [API user's Guide for Command API]

Development Studio

Development Studio is software that supports high-performance parallel application programs, from development to execution, in Fortran, C, and C++. [Overview]

Diskless node [FX]

Node that uses an area in the boot I/O node of FX server by mounting it via NFS. [Setup Guide, Administrator's Guide for System Management]

Docker mode

Docker mode is a job execution environment. By preparing an environment, called a Linux container, isolated from the host OS, you can execute a program on a different OS (container image) than the host OS. Other job execution environments are "normal mode", "McKernel mode", and "KVM mode." [End-user's Guide, Administrator's Guide for Job Management]

Dump generation management [FX]

A function that enables the resource (disk capacity) of compute node to be used efficiently by limiting the number of maintenance materials collected. [Overview]

EDR

One of the transfer rate of IB. EDR is an abbreviation for Enhanced Data Rate. [Administrator's Guide for System Management]

End-user (user)

End-users are users who execute their user applications and obtain the results in the job execution environment provided by the Job Operation Software. [Overview]

ETERNUS

ETERNUS is a Fujitsu storage product. [Setup Guide, Software Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Execution mode policy [PG]

Function to shorten the startup time of FX server and increase the occupancy ratio. [End-user's Guide, Administrator's Guide for Job Management]

Exit function

Exit processing for the job manager function and job scheduler exit function. C language library functions created by the administrator. Exit functions are called from the job manager function or job scheduler function at specific times of job execution. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

Exit function library

An exit function library is a library of exit functions. The job manager exit function and job scheduler exit function require administrators to create and register exit function libraries. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

Exit process

Exit process is a generic term for processing created by an administrator in a hook. Depending on the type of hook, they are called exit functions, exit scripts, and so on. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

Exit script

An exit script is an exit process for the job resource management exit function. This script is written by the administrator using a shell or perl. There are 3 types of exit scripts (prealloc, predel, and postfree scripts) that are called at different times. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

External device

An external device is associated with a node other than a compute node in a system. The term refers not only to a disk unit, network switch, or other device directly required for executing a job. It also refers to a cooling device for computers or other device not directly associated with job execution. [Administrator's Guide for Power Management]

Failover

Switching the active and standby node in the node with the redundant configuration of active/standby configuration. Moreover, put the FEFS service aside by the node of the mutual-standby configuration. [Administrator's Guide for System Management]

Fair share function

The fair share function is a function for determining the priorities of users and groups according to job execution results so that users and groups each have fair use of the system. After a user or group executes jobs frequently or a large-scale job, the user's or group's subsequent jobs have a lower execution priority. [End-user's Guide, Administrator's Guide for Job Management]

Fast restart [FX]

A function to shorten the startup time of the FX server and increase the occupancy ratio. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

FDR (Fourteen Data Rate)

One of the transfer rate of IB. FDR is an abbreviation for Fourteen Data Rate. [Administrator's Guide for System Management]

FDR10 (Fourteen Data Rate 10)

One of the transfer rate of IB. FDR10 is an abbreviation for Fourteen Data Rate 10. [Administrator's Guide for System Management]

FEFS

FEFS is a parallel distributed file system developed by Fujitsu. [Overview]

File System (Technical Computing Suite)

File system products in the Technical Computing Suite. File systems realize layered storage. There are 2 file systems: LLIO (first-layer storage) and FEFS (second-layer storage). [Overview]

First-layer storage

The first-layer storage is the fast storage of FX servers that can be accessed from within a job. Used as temporary storage during job execution. [Overview, End-user's Guide]

GIO group [FX]

Compute node group that uses the same global I/O node in compute cluster. [Overview]

Global file system

The global file system is a large-scale file system shared between the nodes in the system. The FEFS configures the global file system. [End-user's Guide]

Global I/O node (GIO: Global I/O node) [FX]

Node for relaying input/output for second-layer storage in the FX server. [Overview]

GPU

Abbreviation for Graphics Processing Unit. For the Job Operation Software, "GPU" refers to the NVIDIA V100, A100 or H100. [Setup Guide, Administrator's Guide for Job Management]

Handle

A handle is information that holds the operation request content to the job operation management function from the command API. [API user's Guide for Command API]

HCA (Host Channel Adapter)

An HCA is a communication card for connecting a node to an IB switch. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Hook

A hook is a mechanism that executes an administrator-prepared process (exit process) at a specific time during processing by the job operation management function. In the Job Operation Software, mechanisms such as job manager exit function and job resource management exit function that call arbitrary processing are collectively called hooks. Hooks for job operations management include job manager exit, job scheduler exit, Prologue and epilogue function, and job resource management exit functions. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

HPC extension function [FX]

A generic term of functions that make FX server specific functions available for OS and Technical Computing Suite so that this system can be used efficiently. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

HPC tag address override function [FX]

A part of HPC extend function that controls the processor-specific HPC tag address override function installed on the FX server. [Overview, End-user's Guide for HPC Extensions]

I/O network

High-speed network for using the shared file system provided by a storage cluster. [Overview]

I/O-exclusive job [FX]

The job in I/O-exclusive mode. [End-user's Guide, Administrator's Guide for Job Management]

I/O-exclusive mode [FX]

A method that uses storage I/O nodes exclusively for processing I/O for a single job. The job is allocated nodes on a per-shelf basis on the FX server. The job in I/O-exclusive mode is called "I/O-exclusive job". In contrast, a node allocation method in which a single job does not exclusively use storage I/O nodes for I/O processing is called "I/O-sharing mode". [End-user's Guide, Administrator's Guide for Job Management]

I/O-sharing job [FX]

The job in I/O-sharing mode. [End-user's Guide, Administrator's Guide for Job Management]

I/O-sharing mode [FX]

A method of sharing storage I/O node that a job uses for I/O processing with other jobs. The job in I/O-sharing mode is called "I/O-sharing job". In contrast, a method in which storage I/O nodes are exclusively used for I/O processing for a single job is called "I/O-exclusive mode". [End-user's Guide, Administrator's Guide for Job Management]

IB

See InfiniBand.

IB-HCA

See Mellanox HCA.

ICC (Inter Connect Controller) [FX]

LSI that takes charge of connection between nodes and IO function in FX server. In the Job Operation Software, the state of ICC can be displayed by using the system management function. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

InfiniBand

InfiniBand is an interconnect that connects nodes together. It is abbreviated to IB. The Job Operation Software supports it as an IO network. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Install function

The install function supports the system installation, such as OS installation and configuration. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Installer node

A node with install function. This node acts as an installer server. An example of an installer node is the system management node. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Interactive job

An interactive job is a job in which the interactive operation is possible for the program. In contrast, jobs that execute un-interactively are called "batch job." [End-user's Guide]

Inter-core hardware barrier driver/library [FX]

The inter-core hardware barrier is a function for high-speed synchronization between cores in a CPU. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

IPMI (Intelligent Platform Management Interface)

IPMI is a common interface for server management hardware. Through this interface, users can monitor operating characteristics like temperature, voltage, and the fan status. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Power Management]

Job

A job is the unit of execution of a program created by a user. The job that execution is requested (submitted) to the Job Operation Software is allocated the computer resources, and execution is controlled by the job scheduler function and the job resource management function. [End-user's Guide]

Job ACL function

The job ACL function sets the standard and maximum values of the resource limit values for each group and user. Jobs submitted without specified values are set with the standard values of this job ACL. [End-user's Guide, Administrator's Guide for Job Management]

Job execution environment

The software environment for executing job. [Overview, End-user's Guide, Administrator's Guide for Job Management]

Job execution environment customizing function

The job execution environment customizing function is a function to switch the software environment (job execution environment) in which jobs are executed according to user's specification. [End-user's Guide, Administrator's Guide for Job Management]

Job information notification API

The job information notification API is an API provided by the scheduler plugin function. Notifies job-related information to a program created by an administrator that performs unique job operation processing when the job state changes. [Administrator's Guide for Job Management, API user's Guide for Job Information Notification API]

Job manager exit function

The job manager exit function is a hook for the job manager function. Execute the exit process (exit function) created by the administrator at the time of judging whether to allow acceptance or execution of the job. [Overview, Administrator's Guide for Job Management]

Job manager function

The job manager function is part of the job operation management function. It manages jobs. [Overview, Administrator's Guide for Job Management]

Job master node

One of the nodes allocated to the job, on which the job script process runs. [End-user's Guide, End-user's Guide for Master-Worker Job, Command Reference]

Job model

A job model is a classification by job structure. Examples include normal jobs, bulk jobs, step jobs, master-worker jobs, and workflow jobs.

Job operation management function

The job operation management function controls job execution and manages the computer resources required for jobs. [Overview, End-user's Guide, Administrator's Guide for Job Management]

Job Operation Software

The Job Operation Software is platform software for managing such a large-scale computing system and for managing and controlling execution of programs in the system. [Overview]

Job power estimate function

The job power estimate function estimates power consumption by a submitted job on a compute node. It is part of the power cap scheduling function.

Before a submitted job is scheduled, this function provides an advance estimate of the compute node power consumption required for executing the job. [Administrator's Guide for Job Management, Administrator's Guide for Power Management]

Job resource manager exit function

The job resource manager exit function is a hook for the job resource management function. Perform processing (exit script) specified by the administrator, such as before allocating resources for a job or after releasing resources for a job. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

Job resource manager function

The job resource management function is part of the job operation management function. It manages the computer resources assigned to jobs. [Overview, Administrator's Guide for Job Management]

This function can be categorized in the following three types of functions based on the internal role.

- Master resource management function
This function collectively manages the computer resources in the entire system.
- Sub resource management function
This function enables communication between the master resource management function and node resource management function.
- Node resource management function
This function manages computer resources in the node on which it is running.

Job scheduler exit function

The job scheduler exit function is a hook for the job scheduler function. Execute the exit process (exit function) created by the administrator when the job is scheduled. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

Job scheduler function

The job scheduler function is part of the job operation function. It selects the computer resources to be assigned to jobs and coordinates the job execution sequence in cooperation with the job resource management function. [Overview, Administrator's Guide for Job Management]

Job script

A job script is a shell script with the written program execution procedure. The Job Operation Software executes a job based on this job script. [Overview, End-user's Guide]

Job selection policy

The job selection policy is evaluation criteria for determining the job execution priority. It is set by the administrator. [End-user's Guide, Administrator's Guide for Job Management]

Job slave nodes

Nodes other than the job master node allocated to the job. [End-user's Guide, End-user's Guide for Master-Worker Job, Command Reference]

Job statistical information function

The job statistical information function outputs statistical information such the start and end times, CPU time, memory utilization, etc. You can use this statistical information to bill users and monitor the use status. [End-user's Guide, Administrator's Guide for Job Management]

Job suspend function

The job suspend function stops a running job while maintaining its status. It is effective in reducing power consumption. [End-user's Guide, Administrator's Guide for Job Management]

Job type

A job type is a classification by job execution format. Examples include batch jobs and interactive jobs. [End-user's Guide]

KVM mode

KVM is an abbreviation for Kernel-based Virtual Machine. KVM mode is a job execution environment. It executes jobs on KVM, the software that enables you to use a virtualization environment on Linux. Other job execution environments are "normal mode", "Docker mode", and "Mckernel mode." [End-user's Guide, Administrator's Guide for Job Management]

Large page

A memory area extended to a larger size than a normal page. An available size for a large page is 2MiB. [End-user's Guide for HPC Extensions]

Large page library [FX]

A library for using large page by programs (a.out) constructed with -Klargepage option when translating using Development Studio. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

Layered storage system [FX]

The layered storage system combines LLIO and FEFS. These features enable high-speed and large-capacity file systems. [Overview, End-user's Guide]

LLIO (Lightweight Layered IO-Accelerator) [FX]

LLIO is a high-performance file system using high-speed flash memory. Combined with FEFS, it provides layered storage with LLIO as first- layer storage and FEFS as second-layer storage system. [Overview, End-user's Guide]

Location LED [FX]

LED to be installed in BoB of FX server, and to light to show BoB to be maintained when hardware is maintained. [Administrator's Guide for System Management]

Log management function

The log management function collects and manages the system logs. [Overview, Administrator's Guide for System Management]

Login node (LN: Login node)

Node used by the end user to create applications and request the Job Operation Software to execute jobs. [Overview]

Main unit/Shelf [FX]

Unit of minimum composition of FX server system. It is composed of 3BoB. Sometimes, this is called shelf. [Overview]

Maintenance mode

Maintenance mode is a state that enables node maintenance. Examples include hardware maintenance mode and software maintenance mode. Maintenance mode isolates the nodes to which no job is assigned. Moreover, hardware maintenance mode is used for power off, and software maintenance mode is used to stop services. [Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Management network

Network used to control services and communicate information related to the operations of the Job Operation Software. [Overview, Setup Guide]

McKernel mode

One of the job execution environments. Mckernel mode is a job execution environment. Jobs can execute on McKernel isolated from the host OS. Other job execution environments are "normal mode", "Docker mode", and "KVM mode." [End-user's Guide, Administrator's Guide for Job Management]

MDS node (MDS: Metadata Server)

Node for storing and managing the FEFS metadata. [Overview]

Mellanox HCA

HCA card of Mellanox. When InfiniBand is used, it uses it. [Setup Guide, Administrator's Guide for System Management, Administrator's Guide for Maintenance]

Mesh mode [FX]

Mesh mode is one method of placing nodes at Tofu coordinates for node allocated jobs. The smallest unit of node allocation for jobs is one node. Allocated nodes are mutually adjacent in the Tofu coordinate space. Other methods of placing nodes at Tofu coordinates for node allocated jobs include "torus mode" and "non-contiguous mode." [End-user's Guide, Administrator's Guide for Job Management]

MGS node (MGS: Management Server)

Node for managing the file system (FEFS) configuration information. [Overview]

MPMD (Multiple Program Multiple Data)

MPMD is a programming model for parallelization. It uses multiple different programs for process sharing. [End-user's Guide]

MPS

Abbreviation for Multi-Process Service. MPS is a function for effectively using GPU compute resources. [Administrator's Guide for Job Management]

Multinode job

A multinode job is a job that requires multiple nodes for execution. In contrast, a job that requires only one node for execution is called a "single node job." [End-user's Guide]

Multiuse cluster

A group of nodes that can be used for any purpose. A Multiuse cluster can perform power control and status monitoring independently of a compute cluster or storage cluster. [Overview]

Multiuse node

Node used for any purpose. [Overview]

Node

A node is the operating range of one OS. In some models, one node corresponds to a main unit chassis in a 1:1 relationship. In other models, a single chassis contains multiple nodes. [Overview]

Node allocated job

A node allocated job is a job to which node resources are allocated in units of physical nodes. A node allocated job is also a node-exclusive job. In contrast to a node allocated job, the job in which the node resources are allocated in units of virtual nodes is called "virtual node allocated job." [End-user's Guide]

Node group

A node group is the unit of a set of nodes for managing compute nodes. It is monitored by a compute cluster sub management node. [Overview]

Node selection method [PG]

A node selection method is the method of selecting the nodes to allocate to jobs. The method specifies which of the following has priority: unused nodes or nodes already in use. [End-user's Guide]

Node temporary area

A Node temporary area is a local storage area that can be used by the individual compute nodes allocated to the job. This area is on the first-layer storage for fast access. The contents are deleted when the job ends. [Overview, End-user's Guide]

Node type

Node type is an attribute indicating the role of a node. [Overview]

Node-exclusive job

A node-exclusive job is a job executed by monopolizing a node allocated to it. A node-exclusive job is also a node allocated job. In contrast to node-exclusive jobs, jobs that execute on virtual nodes (jobs that the division of physical nodes among multiple jobs) are called "node-sharing job." [End-user's Guide]

Node-sharing job

A node-sharing job is a job executed with a virtual node. This involves the division of physical nodes among multiple jobs. A node allocated job is also a virtual node allocated job. In contrast to a node-shared job, a job that executes while monopolizing the physical node to which it is allocated is called a "node-exclusive job." [End-user's Guide]

Non-contiguous mode [FX]

Non-contiguous mode is one method of placing nodes at Tofu coordinates for node allocated jobs. The smallest unit of node allocation for jobs is one node. Nodes are allocated so that they are mutually adjacent in the Tofu coordinate space as much as possible. Other methods of placing nodes at Tofu coordinates for node allocated jobs include "torus mode" and "mesh mode." [End-user's Guide, Administrator's Guide for Job Management]

Normal mode

Normal mode is a job execution environment. Jobs are executed in the standard OS environment (Linux) installed on the system. Other job execution environments are "Docker mode", "Mckernel mode", and "KVM mode." [End-user's Guide, Administrator's Guide for Job Management]

Normal page

Normal page refers to the memory page of the system standard size. The size of the memory area called page is differs by OS or hardware. FX server OS uses 64KiB memory pages. [End-user's Guide for HPC Extensions]

NUMA allocation policy

NUMA allocation policy consists of rules on the allocation of jobs to NUMA nodes. There is a method for allocating a job so that it fits in a NUMA node, and a method for allocating a job so that it spans NUMA nodes. [End-user's Guide, Administrator's Guide for Job Management]

NUMA node

In NUMA (Non-Uniform Memory Access) architecture, in which the cost of accessing memory shared by multiple CPU cores is not uniform, this is the CPU core group in which the memory access cost is the same. [End-user's Guide, Administrator's Guide for Job Management]

NVIDIA Container Toolkit

Tools for using GPUs from Docker containers. [Setup Guide, Administrator's Guide for Job Management]

OBP

Abbreviation of Open Boot PROM. Firmware that is executed immediately when the system is powered on, and boot the operating system. [Administrator's Guide for System Management]

Operational support function

The operational support function supports the operation of the system administrator such as the batch command execution. [Overview, Administrator's Guide for System Management]

OSS node (OSS: Object Storage Server)

Node for storing and managing file data for FEFS. [Overview]

Paging

One of the virtual memory management methods. This method divides the memory area into a certain size called page and manages it by assigning a virtual address separately from the physical address. [End-user's Guide for HPC Extensions]

Parallel execution environment

The parallel execution environment is a mechanism for executing parallel jobs by making multiple compute nodes look like a single computer to the job. [Overview, Administrator's Guide for Job Management]

Parallel job

A parallel job is a job for executing a parallel processing program using multiple threads and/or multiple processes. In contrast, a job that executes a single-thread or single-process program is called a "sequential job." [End-user's Guide]

Parser

A parser is information that holds optional specifications for a command in the command API. [API user's Guide for Command API]

PCI card

The one to connect CPU (microprocessor) and peripherals, etc. chiefly by one of standards of data transmission road (bus) where it connects between devices in computer. [Administrator's Guide for Maintenance]

Power API

The Power API is a library for measuring and controlling power based on the Sandia Power API provided by the Job Operation Software. You can use the Power API to change hardware (CPU frequency, etc.) and measure power.

The Sandia Power API (Power Application Programming Interface) defines a library interface proposed by Sandia National Laboratories to measure and control power. [Overview, Administrator's Guide for Power Management, API user's Guide for Power API]

Power cap scheduling function

The power cap scheduling function is a job scheduling function that treats system power consumption as a resource. Based on power consumption estimates from the job power estimate function, this function performs job allocation and determines job execution times in such a way that the set allowable power is not exceeded. [Administrator's Guide for Job Management, Administrator's Guide for Power Management]

Power control driver/library [FX]

A driver/library to make Power API available in the FX server. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

Power control function

The power control function is part of the system management function. This function powers on and powers off the system. It can power on and off an entire cluster or a node in the cluster. [Administrator's Guide for System Management]

Power group

A power group is a unit for power measurement. A collection of arbitrary compute nodes and external devices can be defined as a power group. [Administrator's Guide for Power Management]

Power knob

A power knob is an individual performance limiting function for hardware that can be dynamically controlled by the software. Use this function to control node power consumption. [Administrator's Guide for Power Management, API user's Guide for Power API]

Power management function

The power management function monitors the power of system-related devices and manages the power of the entire system as a resource. [Overview, Administrator's Guide for Power Management]

Power saving function

The power save function reduces power consumption of compute nodes to which jobs are not allocated. [Overview, Administrator's Guide for Power Management]

Priority control of allocated nodes [PG]

Priority control of allocated nodes is a method of setting priority on each node and allocating nodes to jobs according to that priority. [End-user's Guide, Administrator's Guide for Job Management]

Prologue and epilogue function

The prologue and epilogue functions a hook for the job manager function. Executes administrator-created processes (prologue script and epilogue script) before and after the job execution. [End-user's Guide, Administrator's Guide for Job Management]

Prorogue script, Epilogue script

Prologue and epilogue scripts are exit processes for the prologue and epilogue function. These are shell scripts that you create. These are performed by the job manager function. Prorogue scripts are executed before the job runs, and epilogue scripts are executed after the job ends. [Administrator's Guide for Job Management, Administrator's Guide for Job Operation Manager Hook]

PSU (Power Supply Unit)

PSU is a device to supply electric power to each part such as hard disks and mother boards. [Administrator's Guide for Maintenance]

QDR (Quad Data Rate)

One of the transfer rate of IB. QDR is an abbreviation for Quad Data Rate. [Administrator's Guide for System Management]

Rack [FX]

Special case equipped with 384 nodes or less (24= BoB and =8 shelves) in FX server. It is sometimes referred to as "main unit rack." [Overview]

Rank

A rank is a concept equivalent to a process ID in an MPI program. [End-user's Guide]

Rank number

A rank number is a number that is set in the order of rank generation. [End-user's Guide]

Relay node

Node to support an installer node in a large-scale environment. Since this node has the repository replication function and TFIP server function, it balances a load when installing an OS or when applying a patch. An example of a relay node is the compute cluster sub management node for the Job Operation Software. [Overview, Setup Guide, Administrator's Guide for System Management]

Resource group

A job can be executed in a resource group, which is a range of computer resources within a resource unit. A resource group can be seen as the smallest job operation unit formed by finely dividing the job operation within a resource unit. [Overview, Administrator's Guide for Job Management]

Resource unit

A resource unit is a unit formed by logically dividing nodes within a cluster. You can start the job scheduler function for this unit. Resource units are also used with the purpose of assigning dedicated resources to specific users and groups. [Overview, Administrator's Guide for Job Management]

Response Information

The response information is the information which holds the request result for the job operation management function from the command API. [API user's Guide for Command API]

Resume

Resume restarts a stopped job while retaining the status of a running job. Stopping a running job while retaining its state is called "suspend." [End-user's Guide, Administrator's Guide for Job Management]

Rolling update

Partial maintenance while continuing the job operation in the clusters of some compute nodes without stopping the system or the entire cluster when applying a package. [Administrator's Guide for Maintenance]

Scheduler API

The scheduler API is an API provided by the scheduler plugin function. An original scheduling algorithm can be incorporated by replacing a part of the processing unit of the job scheduler with a module (shared library) created by the job operation administrator. [Administrator's Guide for Job Management, API user's Guide for Scheduler API]

Scheduler plugin function

The scheduler plugin function is a function that incorporates an original scheduling algorithm created by the job operation administrator into the job scheduler to replace the scheduling algorithm of the job operation management function with it. [Administrator's Guide for Job Management, API user's Guide for Scheduler API]

SDI specification

The SDI specification is a method of selecting and utilizing the execution mode deployed on the system in the job execution environment customizing function. In contrast, the method using the job execution environment prepared by the user is called "UDI specification." [End-user's Guide, Administrator's Guide for Job Management]

Second-layer storage

A file system that can be shared by login nodes and each compute node using the parallel distributed file system FEFS. [Overview, End-user's Guide]

Sector cache driver/library [FX]

A hardware function that divides process cache into two or more sectors and enables the application to use them arbitrarily. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

Sequential job

A sequential job is a job that executes a single-thread or single-process program. In contrast, a job for executing a parallel processing program using multiple threads and/or multiple processes is called a "parallel job." [End-user's Guide]

ServerView Suite

ServerView Suite is a suite of integrated systems management tools that help ensure the stable operation of PRIMERGY servers. Reduce operational management and deployment workload. The ServerView Suite includes the SVIM (ServerView Installation Manager), ServerView Agent, and ServerView RAID Manager components that set and install the OS and other user management and server management software. [Overview, Setup Guide, Administrator's Guide for System Management]

Shared temporary area

A shared temporary area is a storage area that can be shared by the compute nodes allocated to the job. This area is on the first-layer storage for fast access. The contents are deleted when the job ends. [Overview, End-user's Guide]

shared_disk

Directory /var/opt/FJSVtcs/shared_disk on the system management node and the compute cluster management node, to store the logs, the dump files, and management information related to job operation by the Job Operation Software. If the system management node and compute cluster management node are redundant configuration, shared_disk mounts and utilizes only the active node. [Setup Guide, Administrator's Guide for System Management]

Shelf [FX]

Unit of minimum composition of FX server. It is composed of 3BoB. [Overview, Administrator's Guide for System Management]

Single node job

A single node job is a job that requires only one node for execution. In contrast, a job that requires multiple nodes for execution is called a "multinode job." [End-user's Guide]

Singularity

Container virtualization software designed for HPC. The Job Operation Software can execute Singularity as a job in Docker mode when configured for Singularity. [Setup Guide, Administrator's Guide for Job Management]

SIO group [FX]

Compute node group that uses the same storage I/O node in compute cluster. [Overview]

Software environment check function

The software environment check function checks the Job Operation Software status in the system. [Overview, Administrator's Guide for System Management]

Step job

A step job is a group of jobs that have an execution order or dependency relationship. Each job is called a sub job. [End-user's Guide]

Storage cluster

A group of nodes providing second-layer storage (FEFS) to compute clusters. [Overview]

Storage cluster management node (SCM: Storage Cluster Management node)

Node that manages the configuration and monitors the services in a storage cluster. [Overview]

Storage I/O node (SIO: Storage I/O node) [FX]

I/O node responsible for input/output for first-layer storage in the FX server. [Overview]

Sub job

Sub job is the processing unit of job in the bulk job and step job. [End-user's Guide]

Suspend

Suspend is a job suspend function. Stops a running job while retaining its state. While the job is suspended, the resources assigned to the job are not released. Restarting a suspended job is called "resume." [End-user's Guide, Administrator's Guide for Job Management]

Suspend period

The suspend period is the time from the start of suspend to the completion of the resume. [End-user's Guide, Administrator's Guide for Job Management]

System control function

The system control function controls to start or stop a node in the system. [Overview, Administrator's Guide for System Management]

System maintenance function

The system maintenance function supports hardware maintenance and software maintenance. [Overview, Administrator's Guide for System Management]

System management function

The system management function manages system configuration information and monitors the status of services. [Overview, Administrator's Guide for System Management]

System management node (SMM: System Management node)

Node for controlling power to start/stop a cluster and monitoring the nodes and services in the cluster. [Overview]

System monitoring function

The system monitoring function monitors the hardware or software status. [Overview, Administrator's Guide for System Management]

System power data collection function

The system power data collection function measures the power consumption by system, job, or system-related device, collects the data, and supports visualization of this information. [Overview, Administrator's Guide for Power Management]

Technical Computing Suite

Technical Computing Suite is a general term for system software that provides operational functions for exascale computer systems, and an environment to use applications. [Overview]

It consists of three products:

- Job Operation Software
- File system (LLIO, FEFS)
- Development Studio

TLB (Translation Look-aside Buffer)

An address translation buffer (Virtual memory address to physical memory address translation table used by the program.) [End-user's Guide for HPC Extensions]

Tofu coordinates [FX]

The Tofu coordinate is the coordinate that indicates the placement location in Tofu units on the FX server. It is expressed in six dimensional coordinates along the X, Y, Z, a, b, and c axes. [Overview, End-user's Guide, Administrator's Guide for Job Management]

Tofu interconnect D [FX]

A high-speed network connecting nodes to one another in the FX server. In the Job Operation Software, the network that connects compute nodes to one another is called a compute network. It is sometimes referred to as "Tofu Interconnect." [Overview]

Tofu port [FX]

Joint of node and Tofu interconnect D. [Administrator's Guide for System Management]

Tofu route [FX]

Communication route to execute job by multiple nodes in FX server. [Administrator's Guide for System Management]

Tofu unit [FX]

A Tofu unit is a unit for managing nodes in the FX server. A Tofu unit is configured with 12 adjacent nodes connected by the Tofu interconnect. [Overview, End-user's Guide, Administrator's Guide for Job Management]

TofuD driver [FX]

A driver to make Tofu interconnect D available in the FX server. [Overview, End-user's Guide for HPC Extensions, Administrator's Guide for HPC Extensions]

Torus mode [FX]

Torus mode is one method of placing nodes at Tofu coordinates for node allocated jobs. The smallest node allocation unit for jobs is the Tofu unit (12 nodes). Allocated nodes are mutually adjacent in the Tofu coordinate space. Other methods of placing nodes at Tofu coordinates for node allocated jobs include "mesh mode" and "non-contiguous mode." [End-user's Guide, Administrator's Guide for Job Management]

UDI specification

The UDI specification is a method to use the job execution environment prepared by the user in the job execution environment customizing function. In contrast, the method of selecting and using the execution mode deployed on the system is called "SDI Specification." [End-user's Guide, Administrator's Guide for Job Management]

User definition service

Services monitored by the system monitoring function that have been set as monitored by the administrator. The system administrator can also detect abnormalities in important services other than Job Operation Software by registering a script to monitor an arbitrary service in the system monitoring function. [Administrator's Guide for System Management]

Virtual node

A virtual node is the unit of resource allocation. It is the concept of grouping CPU cores and memory together. [End-user's Guide, Administrator's Guide for Job Management]

Virtual node allocated job

A virtual node allocated job is a job to which node resources are allocated in units of virtual nodes. Multiple virtual nodes can exist on a node. Therefore, when a single node is shared by multiple jobs, it is referred to as "node-sharing job." In addition. In contrast to a virtual node allocated job, the job in which the node resources are allocated in units of physical nodes is called the "node allocated job." [End-user's Guide, Administrator's Guide for Job Management]

Virtual node placement policy [PG]

A virtual node placement policy is the concept of how virtual nodes are placed as opposed to physical nodes. [End-user's Guide, Administrator's Guide for Job Management]

Warm reboot

Operation to only restart an OS without powering off the hardware in restarting a node. In contrast, powering off the hardware and restarting a node is called "cold reboot." [Administrator's Guide for System Management]

Workflow job

A workflow job is a group of jobs with a defined execution flow (conditional branching and repetition). Specifically, it indicates not a job but a job control method. [End-user's Guide]