Activity Report from AIST ~1st anniversary of ABCI~

Ryousei Takano

National Institute of Advanced Industrial Science and Technology (AIST), Japan

Energy Efficient HPC State of the Practice Kobe Meeting, 4 August 2019

ABCI: The World's First Large-Scale Open AI Infrastructure

東京大学 THE UNIVERSITY OF TOKYO　AIST

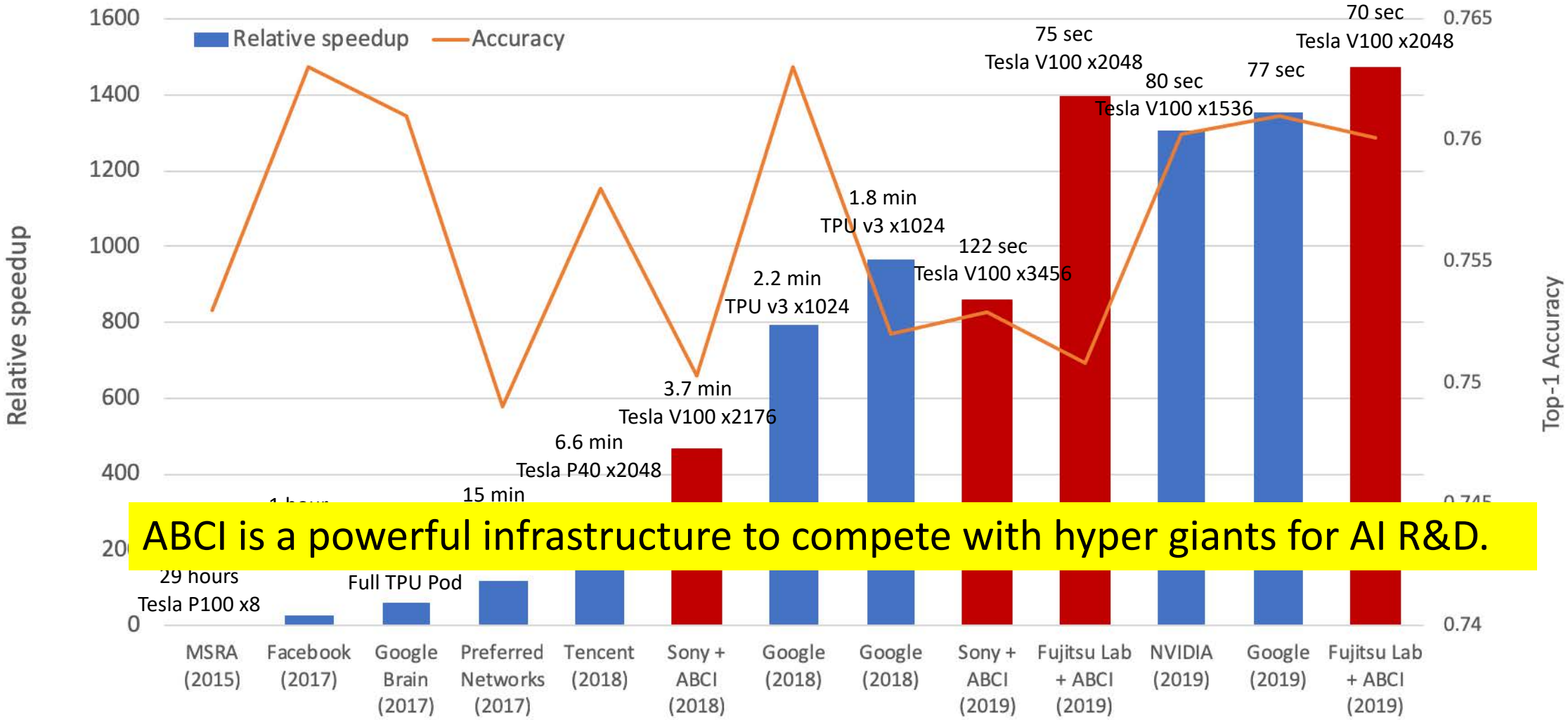
Univ. Tokyo / Kashiwa II Campus

ABCI AI Bridging Cloud Infrastructure

- World Top-Level compute and data process capability
- **Open, Public, and Dedicated** infrastructure for Al & Big Data Algorithms, Software, and Applications
- **Open Innovation Platform** to accelerate joint academic-industry R&D for AI

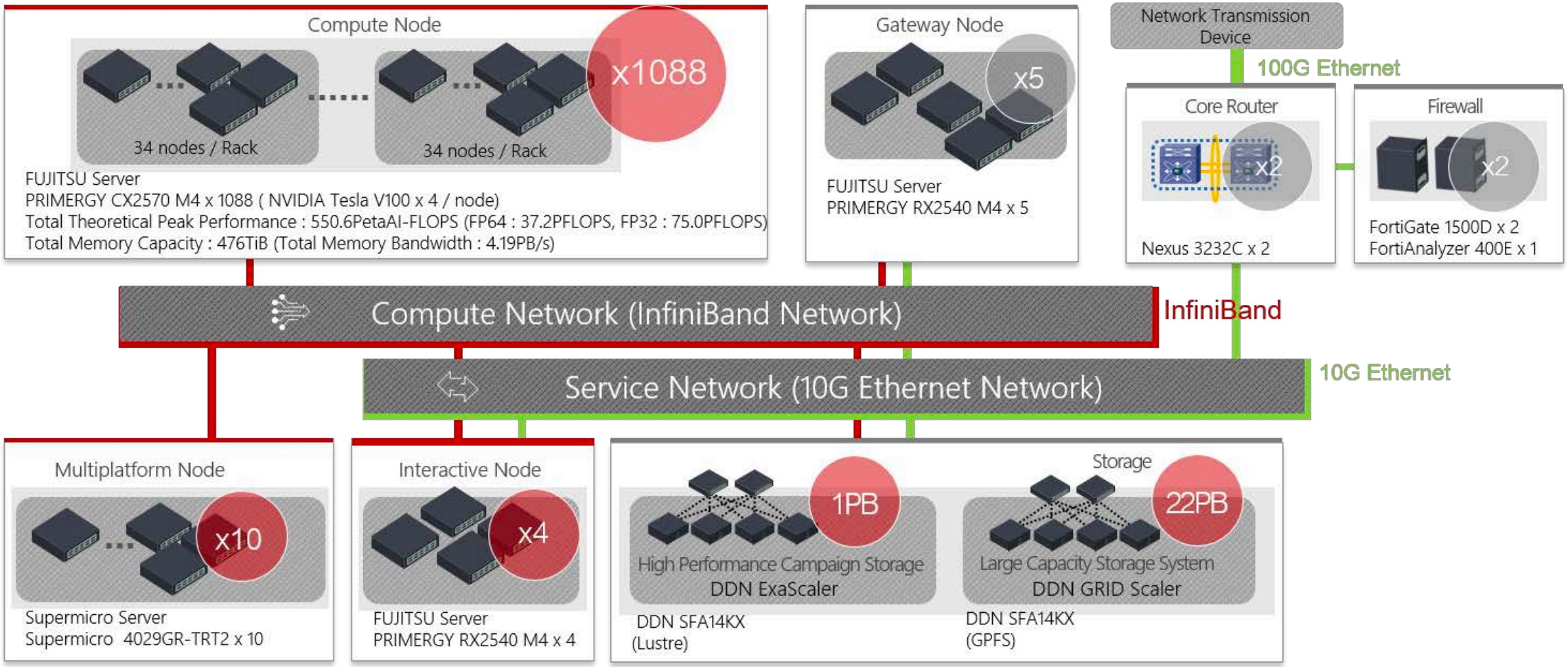
Peak Performance:
550 PFlops (FP16)
37 PFlops (FP64)
Effective Performance:
19.88 PFlops (#8 in TOP500)
14.423 GFlops/W (#3 in GREEN500)
509 GFlops (#5 in HPCG)
Power Usage: < 2.3 MW
Average PUE: < 1.1 (Estimated)

Speed up of ImageNet-1k/Resnet Training



ABCI is a powerful infrastructure to compete with hyper giants for AI R&D.

ABCI: System Overview



ABCI: Computing System

Chips
(GPU, CPU)

Tesla V100

Xeon Skylake-SP

GPU:
7.8 TFlops(FP64)
125 TFlops(FP16)

CPU:
1.53 TFlops(FP64)
3.07 TFlops(FP32)

NVIDIA Tesla V100
(16GB SMX2)

Compute Node
(4GPUs, 2CPUs)

PRIMERGY
CX2570 M4

34.2 TFlops(FP64)
506 TFlops(FP16)

~3.72 TB/s MEM BW

384 GiB MEM
200 Gbps NW BW
1.6TB NVMe SSD

Chassis
(2 Compute Nodes)

PRIMERGY
CX400 M4

68.5 PFlops(FP64)
1.01 PFlops(FP16)

Rack
(17 Chassis)

1.16 PFlops(FP64)
17.2 PFlops(FP16)

~131TB/s MEM BW

Full Bisection BW within Rack
70kW Max

System
(32 Racks)

1088 Compute Nodes
4352 GPUs

37.2 PFlops(FP64)
0.55 EFlops(FP16)

~4.19 PB/s MEM BW

1/3 of Oversubscription BW
2.3MW

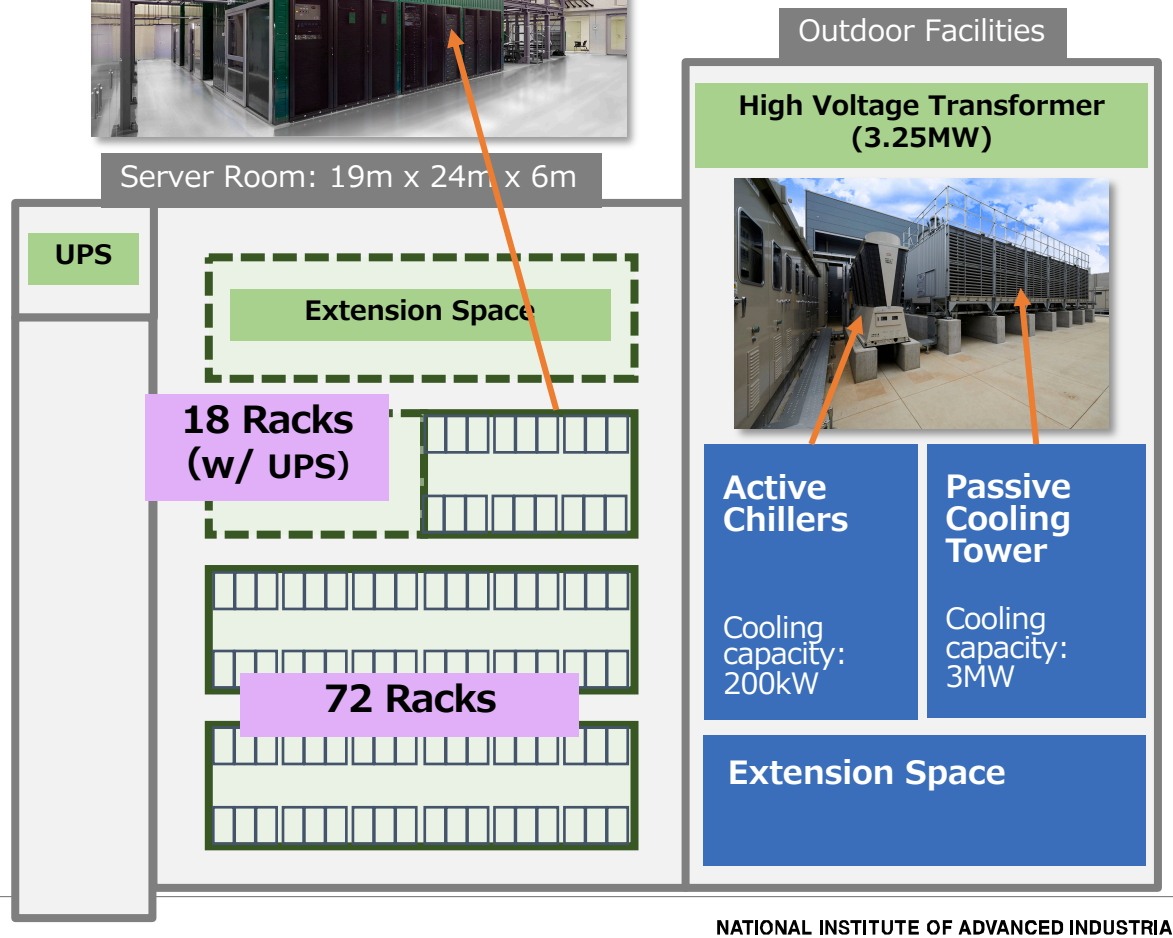
ABCI achieves ultra-dense packaged rack

*June 2019

	TSUBAME3.0	ABCI	Summit	TPU 3.0 Pod
Organization	Tokyo Tech	AIST	ORNL	Google
Start of operation	2017	2018	2018	2018
Number of nodes	540	1088	4608	unknown
Throughput Processor (TP)	NVIDIA Tesla P100	NVIDIA Tesla V100	NVIDIA Tesla V100	TPU 3.0
Number of TP	2160	4352	27648	unknown
Theoretical Perf. (FP64)	12.2 PF	37.2 PF	200 PF	unknown
Theoretical Perf. (DL)	47.2 PF	550 PF	3.3 EF	100 PF / Pod
TOP500*	25	8	1	unknown
GREEN500*	5	3	2	unknown
#Nodes / Rack	36	34	16	unknown
#GPU / Rack	144	136	96	unknown
kW / Rack	64.8 kW	67.33 kW	45-55 kW (est.)	unknown
DL Perf. / Rack	3.1 PF	17 PF	12 PF	12.5 PF (100 PF / 8 rack)

ABCI Datacenter

"Commoditizing supercomputer cooling technologies to Cloud (70kW/rack)"

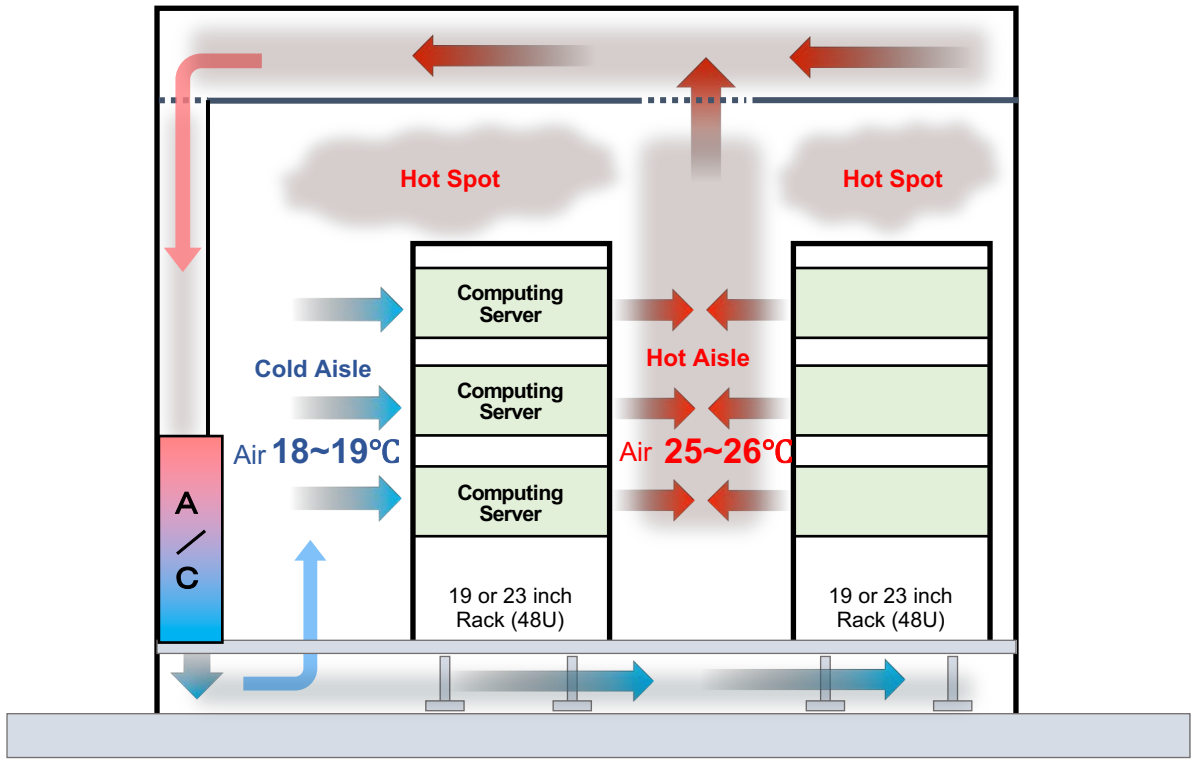


- **Single floor, cost effective building**
- **Hard concrete floor 2t/m² weight tolerance for racks and cooling pods**
- **Number of Racks**
  - Initial: 90 (ABCI uses 41 racks)
  - Max: 144
- **Power capacity: 3.25 MW**
  - ABCI uses 2.3MW max
- **Cooling capacity: 3.2MW**
  - 70kW/rack: 60kW water+10kW air

Free cooling with water/air hybrid cooling system

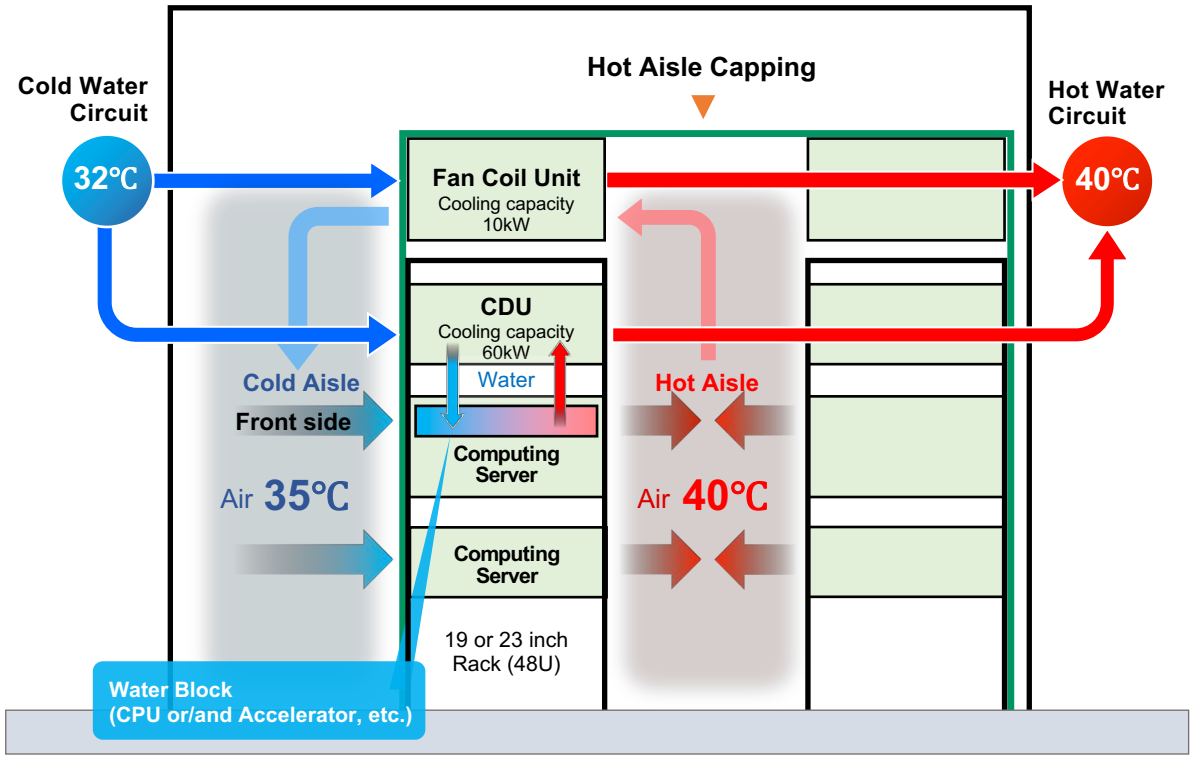
■ Conventional Datacenter Design

Average PUE : ≈ 2.0



■ ABCI Datacenter Design

Average PUE : ~ 1.1

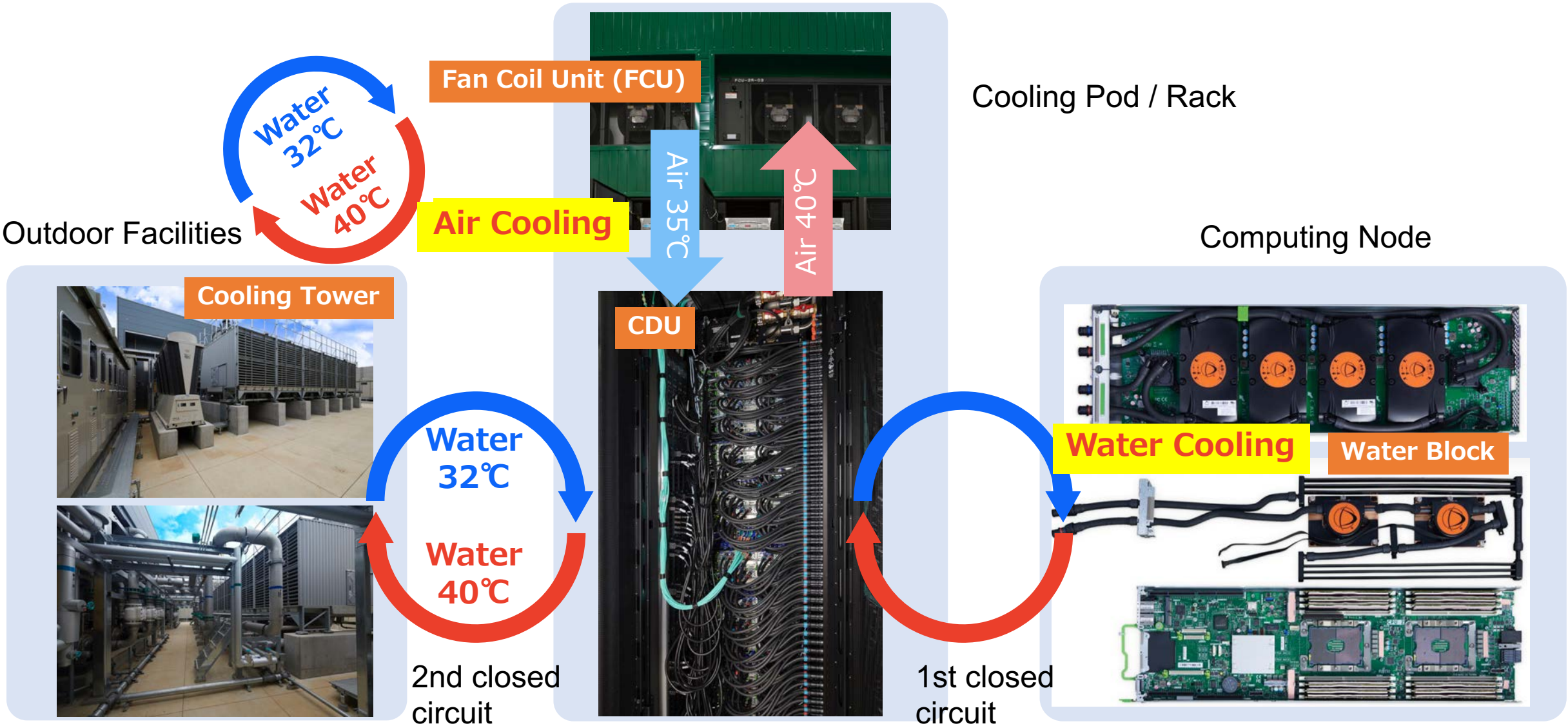


Power Usage Effectiveness (PUE) = $\frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$

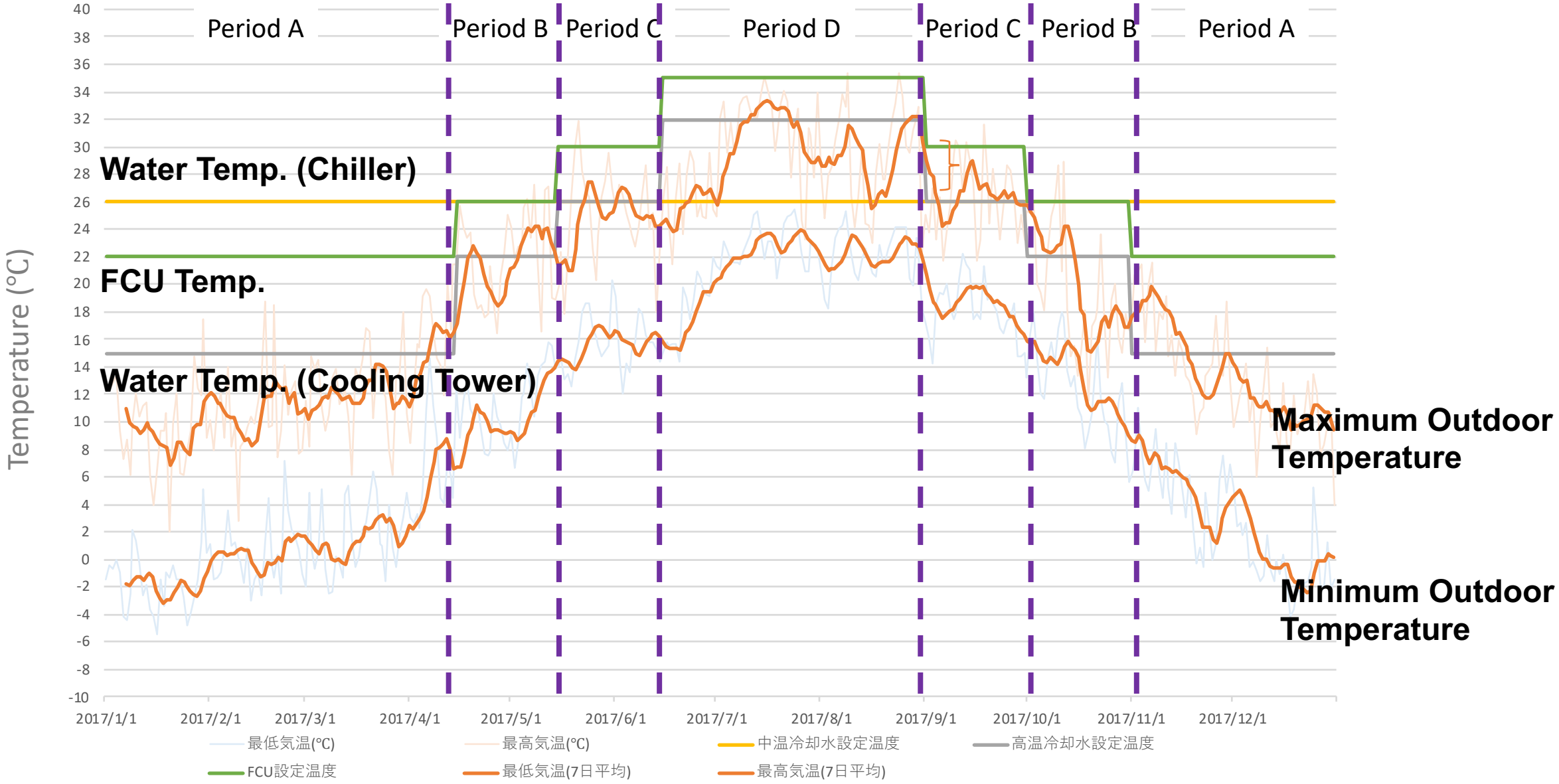
Configuration in Summer

(Lower the better. 2.0 Standard, 1.4 Good, 1.0 Best)

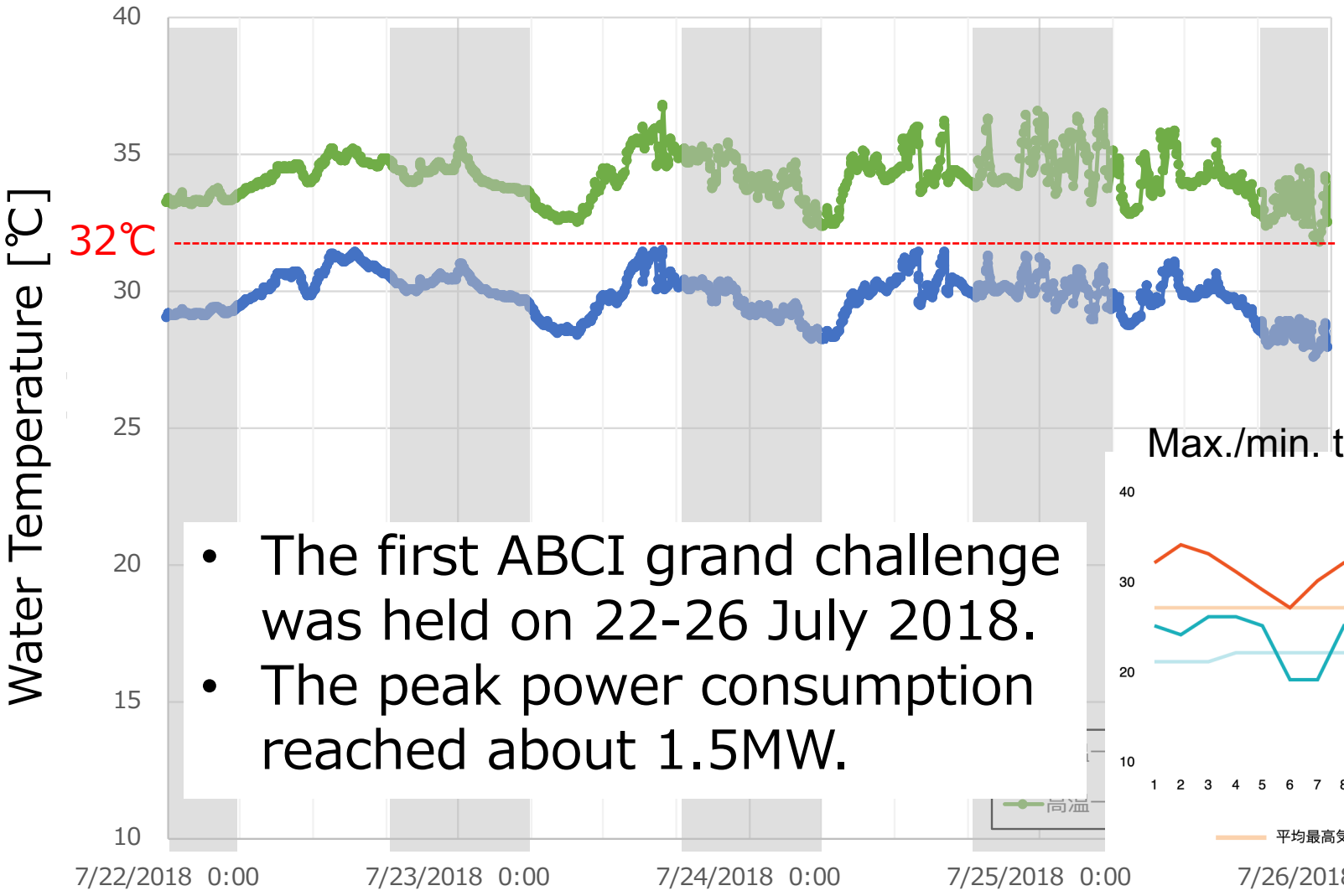
Free cooling with water/air hybrid cooling system



Cooling System Parameters



100% free cooling is possible in high summer

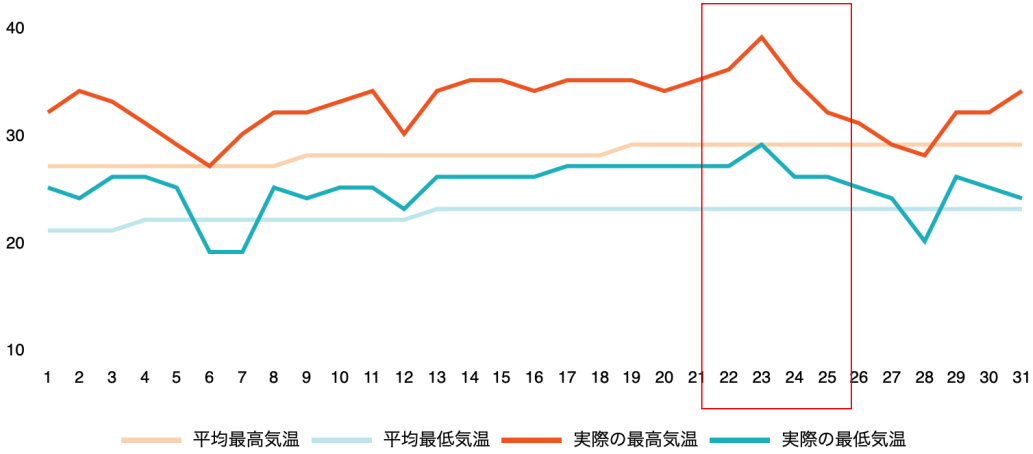


Cooling Water (From ABCI): <40℃

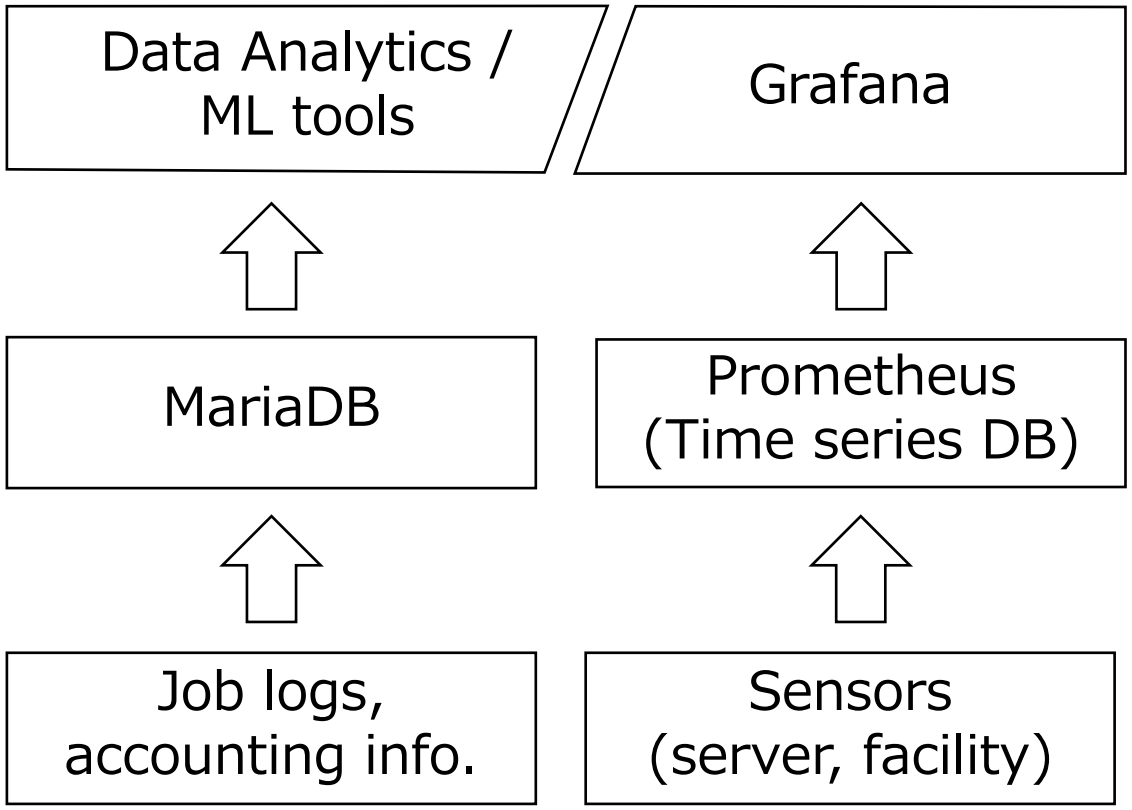
Cooling Water (To ABCI): <32℃

- The first ABCI grand challenge was held on 22-26 July 2018.
- The peak power consumption reached about 1.5MW.

Max./min. temperature in Kashiwa in July 2018



Monitoring Mechanism



ABCI dashboard through Grafana

Collecting Sensor Data

Location	#items	Sampling rate	Sensors
Server	873	1 min.~1 hour	CPU/GPU usage, memory usage, I/O usage, temperature, power consumption, etc
IB/Ethernet Switch	6245	1 min.~1 hour	
Job scheduler	26	-	Job ID, account ID, group ID, job type, resource type, walltime, application info., etc
AI data center	408	1 sec.	Temperature (hot/cold aisle), rack inlet air temperature, humidity, power consumption
			FCU inlet/outlet temperature, power consumption, status
			Water temperature, volume of water flow, power consumption of pump, status, CDU
Site	8	1 min.	Temperature, humidity, wind speed/direction, rainfall

Toward Autonomous HPC Datacenter

Develop a framework for optimizing the operation of HPC data center by leveraging ML/DL technologies [w/ RIKEN and Tokyo Tech]

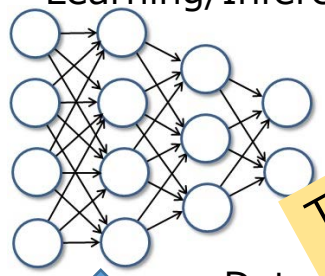
World leading supercomputing systems for big data / AI

TSUBAME 3.0

Data-driven operation

- Reduce power consumption
- Improve resource utilization
- Reduce HW maintenance fee using failure prediction

Learning/Inference

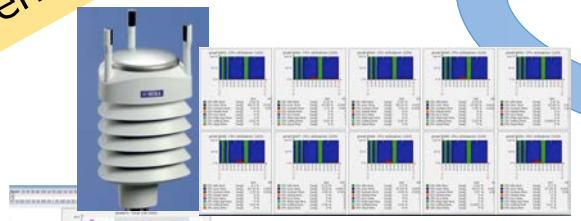


Train and apply parameters for improving the operation

Data for learning

Data analytics

Data center generates huge amount of sensor data



Monitoring

Monitoring data (job log and facility sensor data)

Data Store

ABCI Construction Timeline

AI Datacenter

2016.11 2017.1 2017.8 2018.1

Proposal & Contract → Design → Construction

Datacenter Facility

2017.5 .7 .12 2018.3

Proposal & Contract → Design → Installation

ABCI HW/SW

2016.8 .10 2017.8 .10 2018.3 .6

Conceptual Design → Spec & Benchmark Definition → Proposal & Contract → Detailed Design → Installation & Benchmark → End of 2018.7

Production

Test Op. → Operation

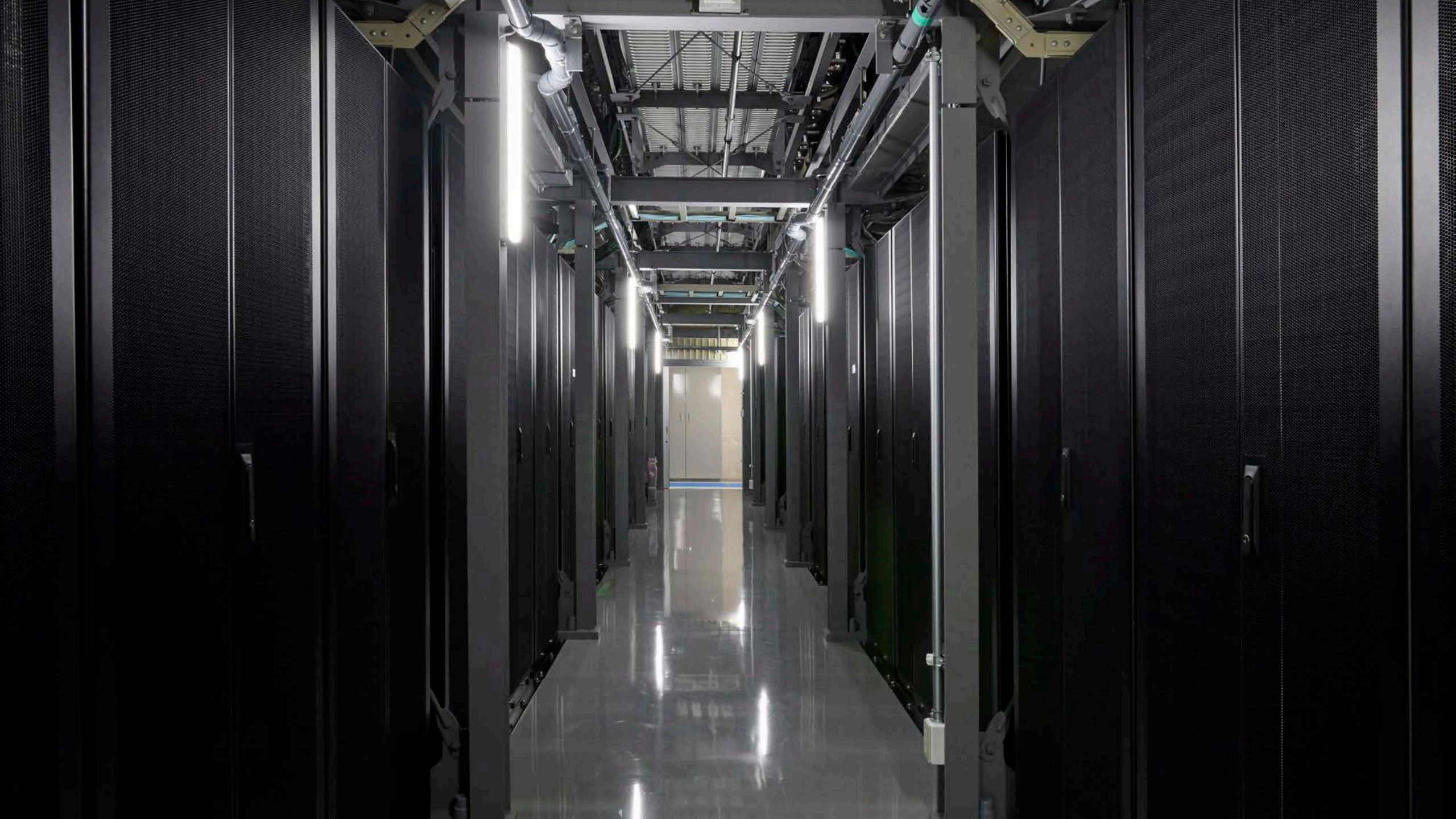

January 16, 2017


October 30, 2017


January 23, 2018


March 31, 2018
RC－1
空冷チラー


FCU-2R-01



Thank you for your attention!

More Information is available from https://abci.ai/