東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

ATLAS EXPERIMENT

ICEPP Site Report

1st Nov. 2023

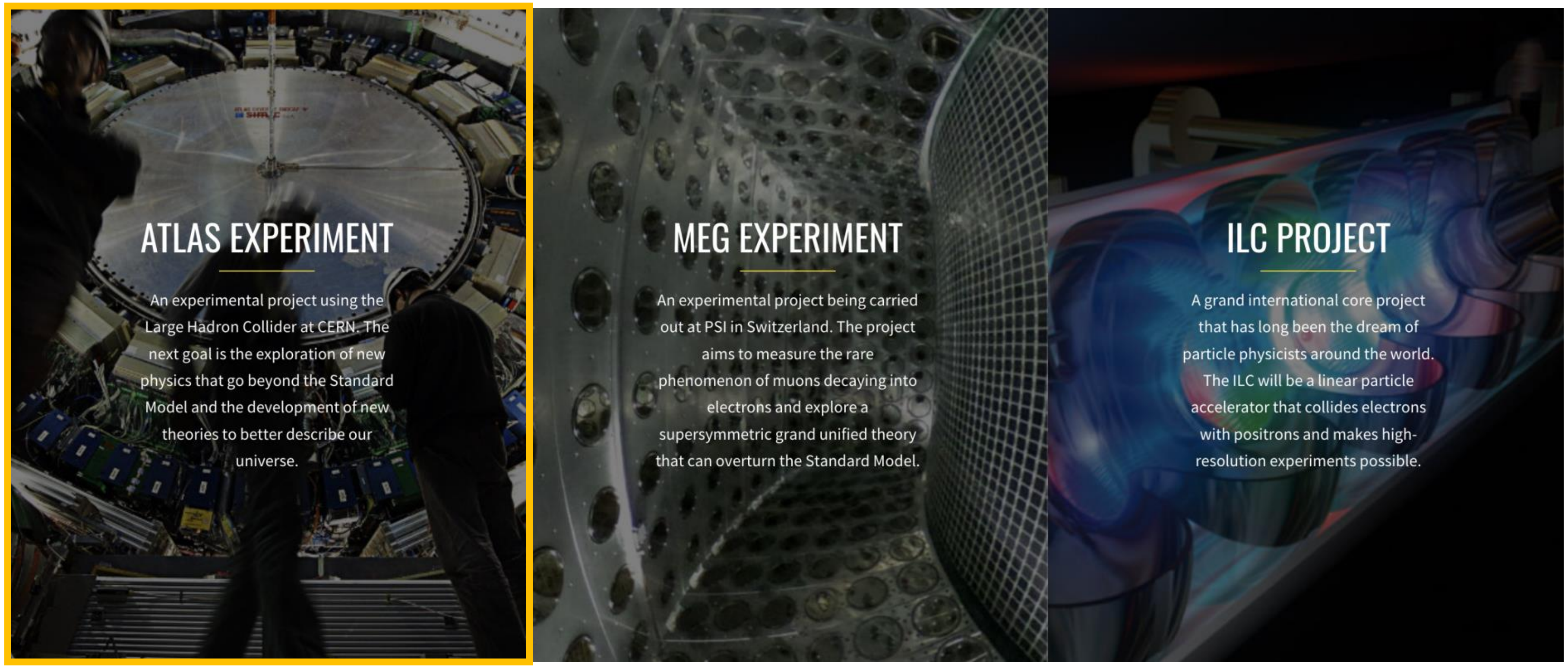
The 7th Asian Tier Center Forum (ATCF7)

Masahiko Saito, on behalf of the operation team

ICEPP, The University of Tokyo

International Center for Elementary Particle Physics (ICEPP)

Main projects at ICEPP



ICEPP
The University of Tokyo

ATLAS-Japan group

- 13 institutes and ~170 members (49 members from ICEPP)
- Contributes to a wide area of the experiment
  - muon triggers, silicon tracker, **Tier2 operation**

ATLAS Japan

➡ ICEPP operates **Tokyo Regional Analysis Center** for ATLAS/ATLAS-Japan

Tokyo Regional Analysis Center

- Support for ATLAS VO in WLCG (Tier2) and provide ATLAS-Japan dedicated resources

- Tier2 (WLCG) (focus on this presentation)
  - Worker nodes (ARC-CE/HTCondor): ~11k cores
  - Storage (DPM→dCache): ~15 PB

- Tier3 (ATLAS-Japan)
  - Interactive nodes: ~ 200 cores
  - Worker nodes (LSF→ HTCondor): ~ 1.7 kcores
  - Storage (GPFS): 3 PB
  - GPU resources: V100, T4

We have been providing large-scale computing resources to the ATLAS experiments and ATLAS Japan for over 15 years.

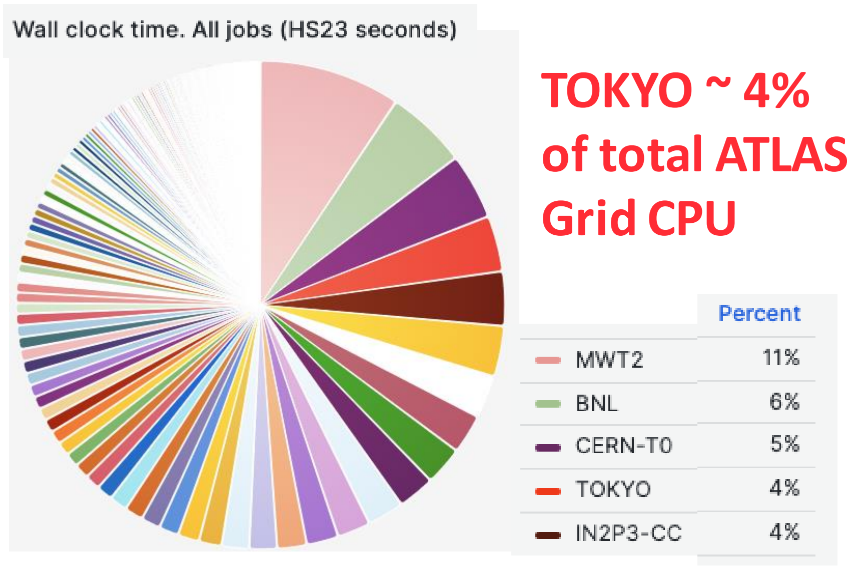
Outline

- Tokyo Regional Center Overview
- Recent Activities
  - Storage Middleware Migration: DPM to dCache
  - HTCondor 8.8.x to 10.x
- Miscellaneous
  - Monitoring System
  - Automation Tool

Overview of Tier2

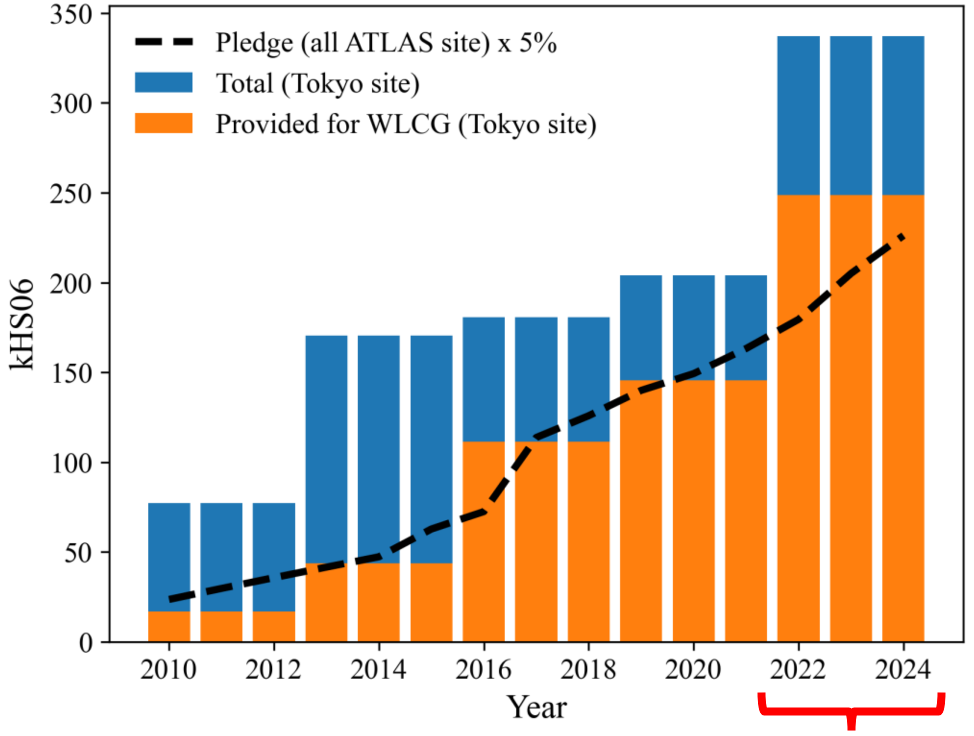
Provided resources for ATLAS as Tier2

- One of the biggest Tier2 sites
  - CPU: ~4% of total ATLAS resources
  - Disk: ~3% of total ATLAS resources
  - cf. ATLAS-Japan member ratio to author list is ~ 3%

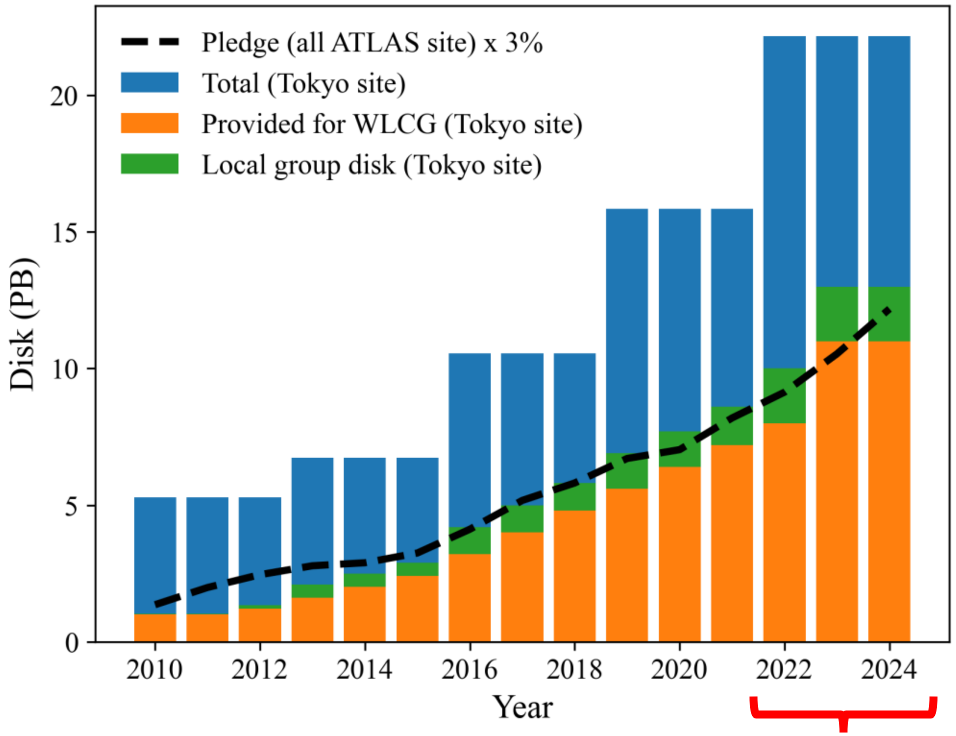


TOKYO ~ 4% of total ATLAS Grid CPU

Compute resources

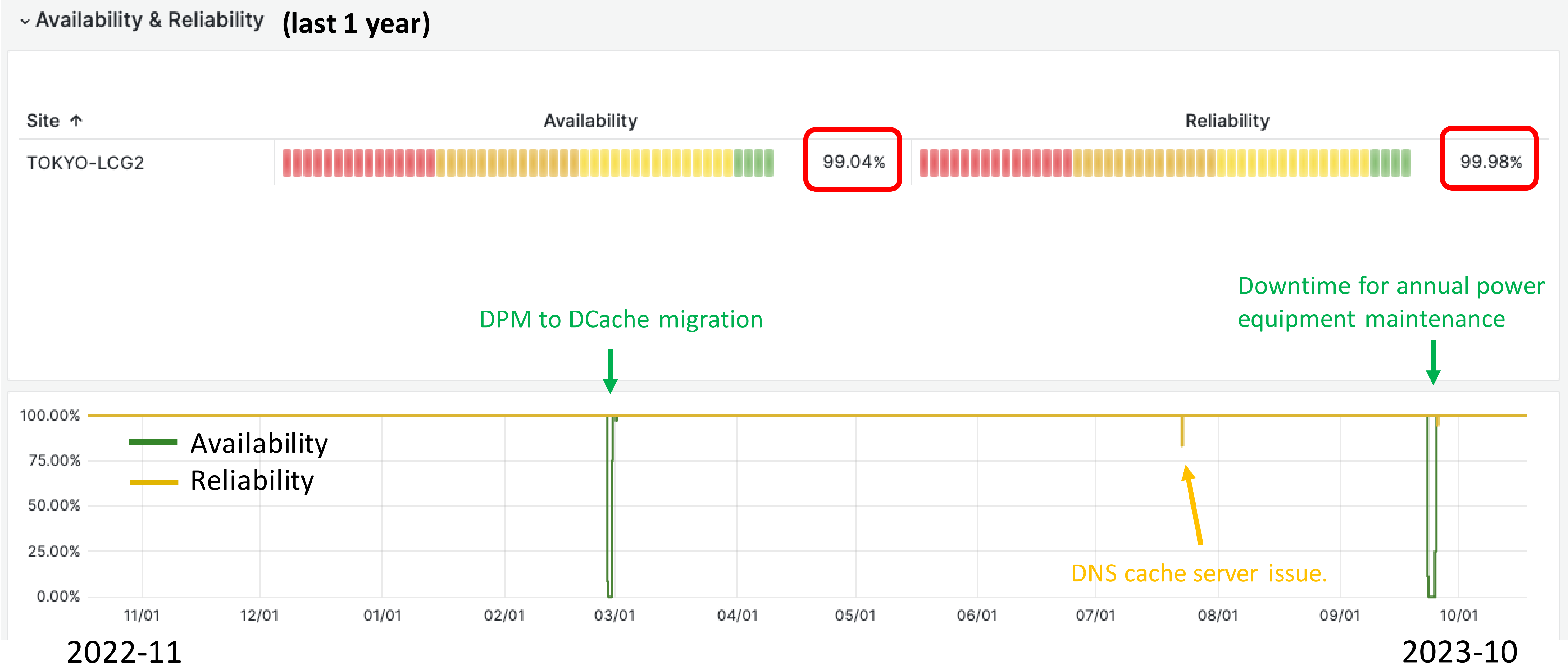


Storage resources



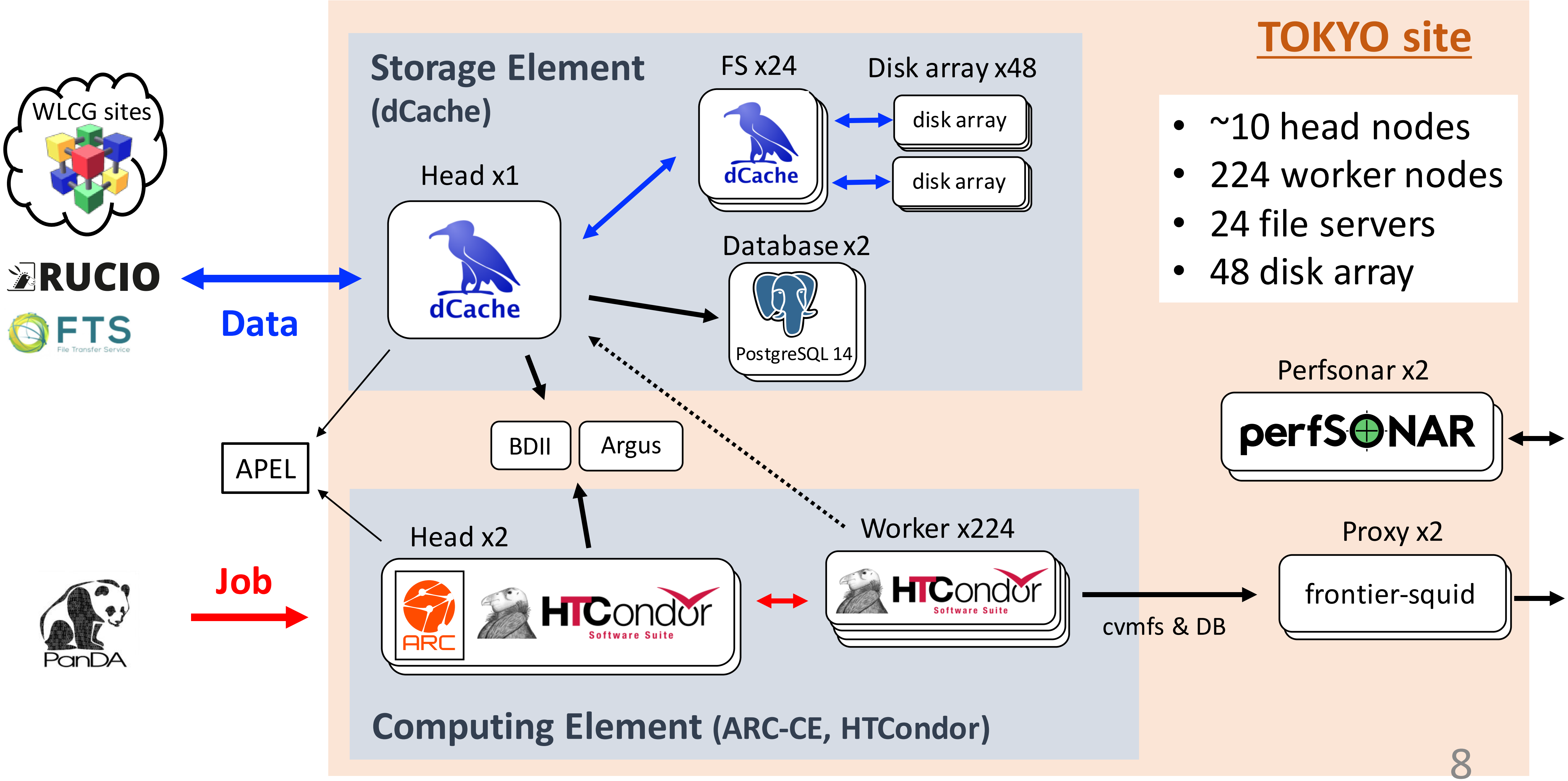
6th system: 2022 - 2024

Availability & Reliability



High availability (~99%) and **reliability** (~99.99%) operation

Overview of TOKYO Grid services



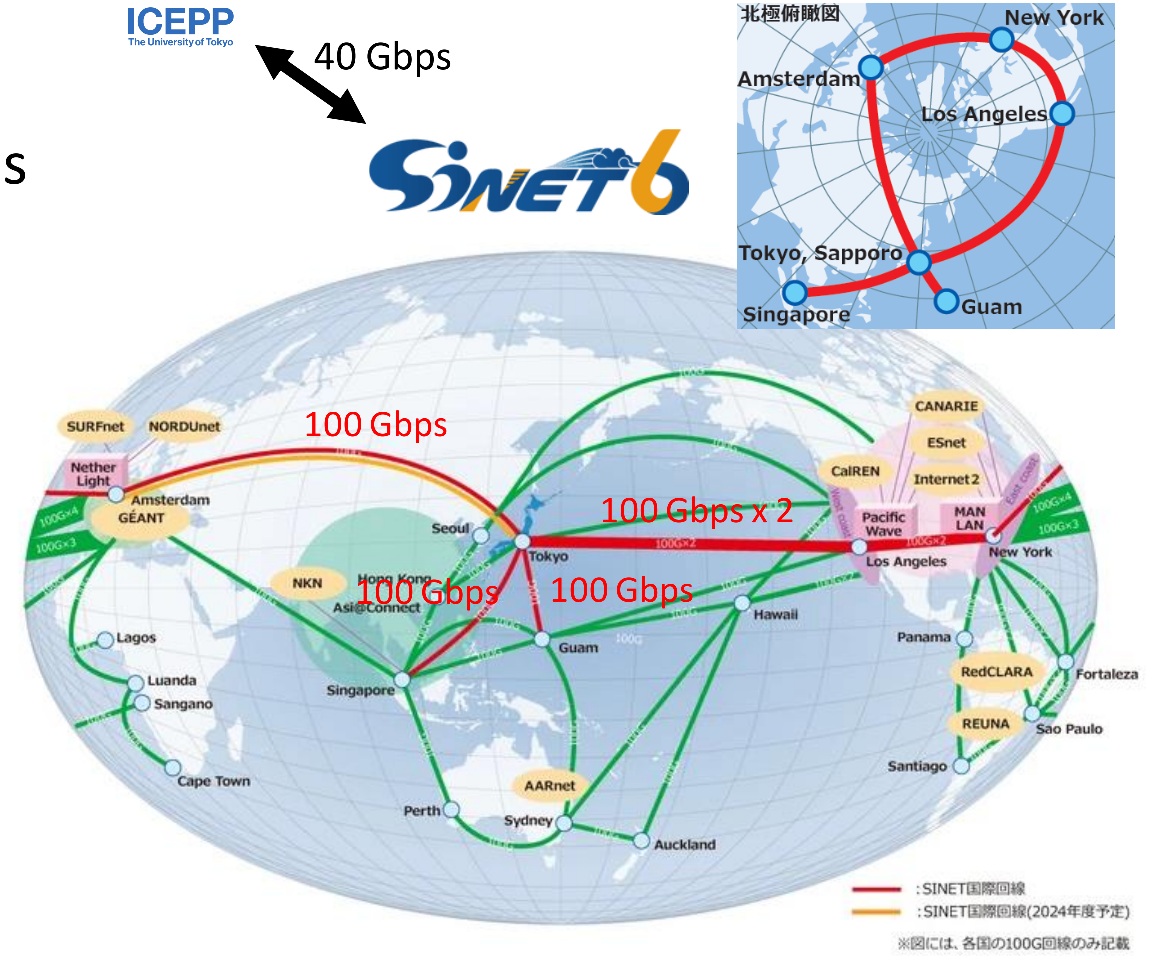
- ~10 head nodes
- 224 worker nodes
- 24 file servers
- 48 disk array

Network

TOKYO regional center

40 Gbps

- External (WAN):
  - Connected to SINET (and LHCONE) at 40 Gbps
    - Upgrade to 100 Gbps **delayed** due to procurement of network equipment.
    - Procurement has been completed. Testing is ongoing.
  - SINET6 international network
    - Upgrade of Tokyo – Amsterdam in Apr 2024
      - 100 Gbps → **100 Gbps x4**
      - Latency: 150ms → **200ms ~ 250ms** (via North America)



from SINET6 webpage

- Internal (LAN):
  - Core switch ← (25 GbE) → File server
  - Core switch ← (40 GbE) → Edge switch ← (10 GbE x 16) → Worker node x 16

Network (WAN)

(dCache file servers ↔ LHCONE/Internet)

- Data transfer volume: in(out)bound 30 (45) PB / year → **200 TB / day**
- Dominant transfer region is Europe, followed by North America.
- Transfer rate is sometimes **limited by the bandwidth** (40 Gbps). Expected to improve when upgraded to 100 Gbps.

Recent activity

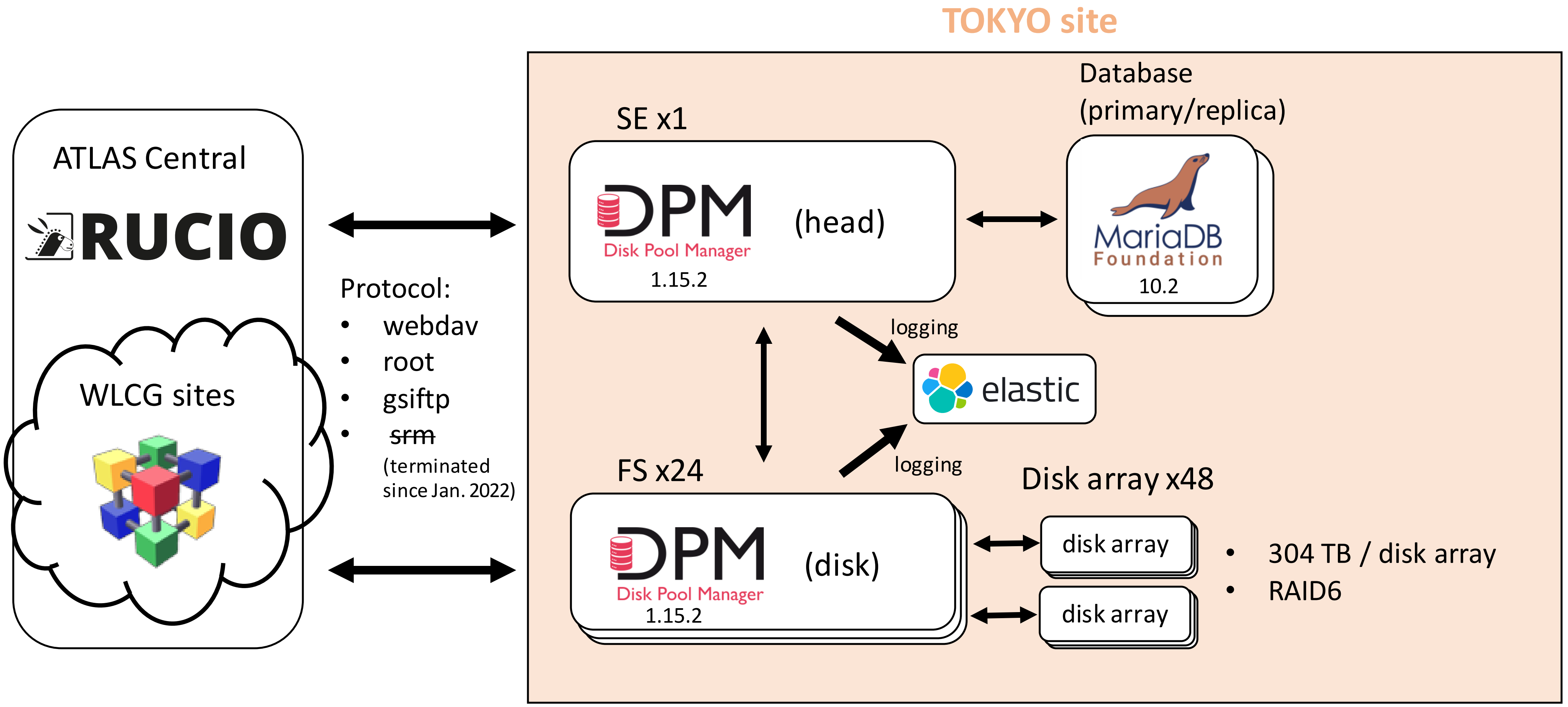
Storage element migration from DPM to dCache

- Tokyo site has been using **DPM** since the beginning (2006~)
  - Provided 40 TB in 2006, 200 TB in 2007, and now ~10 PB & 70 M objects
  - Probably one of the biggest DPM users

- DPM EOL is summer 2024, and EGI support was in June this year. Migration to other storage middleware was mandatory.

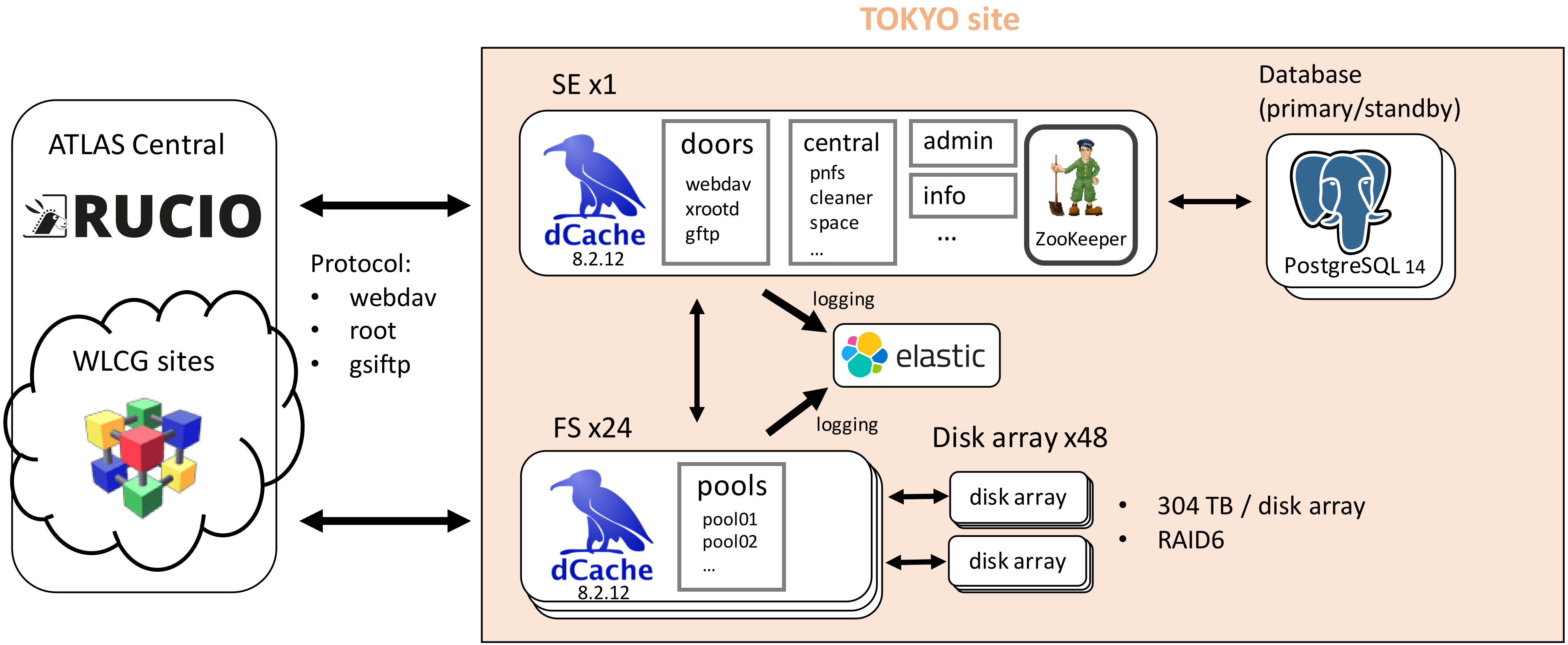
- Decided to move into dCache
  - We don’t need to copy files and to prepare additional (many) servers.
  - Several sites have already migrated from DPM to dCache.
    - Experience and knowledge has been accumulated.

Overview of Tokyo Tier2 storage element (SE)



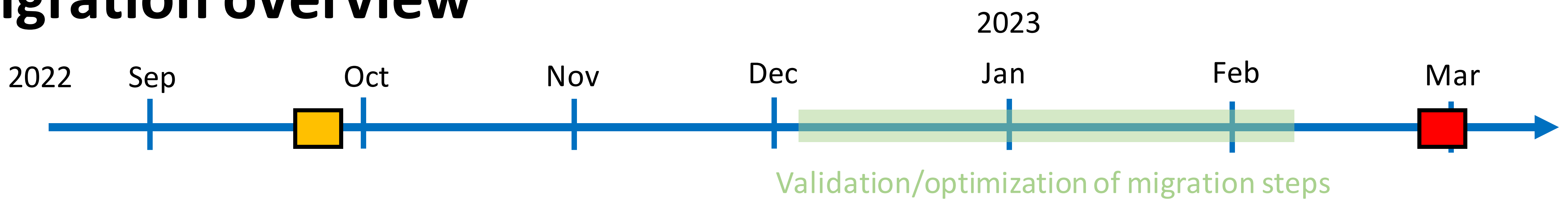
- Storage volume: 14.6 PB (provided 8 + 2 PB), 70M objects are stored
- Database size: 22 GB

Overview of Tokyo Tier2 storage element (SE)



- Use the same servers (head/disk/db) with DPM
- Transparent to end users except for SRR URL

Migration overview



<u>24 Sep – 26 Sep</u>

3 days downtime
(for annual power equipment maintenance)

- Updated DPM to v1.15.2 (latest version)
- Fixed DB inconsistency (~ 12 h)
  - lost/dark data, metadata inconsistencies, missing checksum etc.
  - Found many inconsistencies accumulated over 15 years
  - Fixed them after checking the type of inconsistency and the attributes of the target file.

<u>27 Feb – 1 Mar</u>

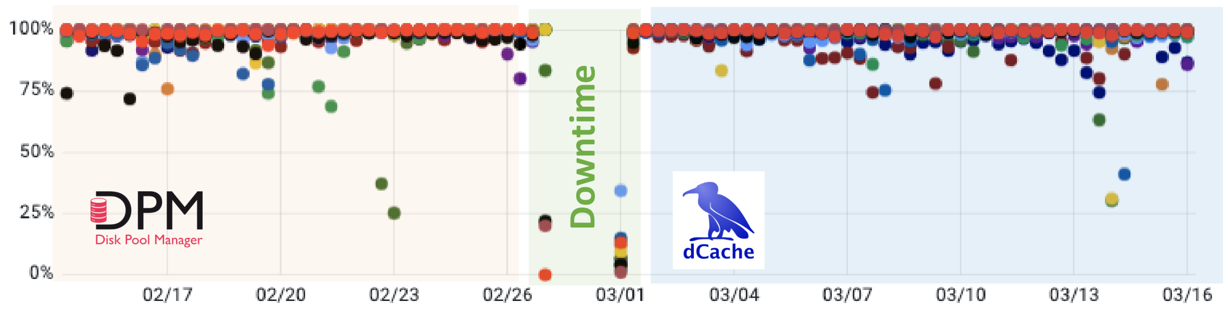
3 days downtime
for the migration
(actual ~55h)

- Migration steps
  - Stop DPM, and install dCache
  - RDBMS replacement (MariaDB to csv, csv to PostgreSQL)
    - incompatible table/schema between two middleware
  - Create hard links of physical files
    - different directory structures in file servers between two middleware
  - start dCache
- No critical issues happened.

Transfer efficiency/volume

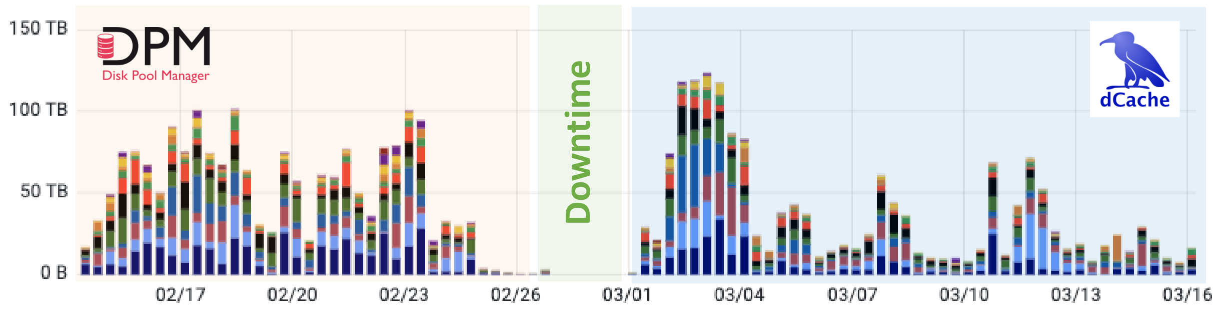
Transfers: Others → Tokyo

Transfer Efficiency



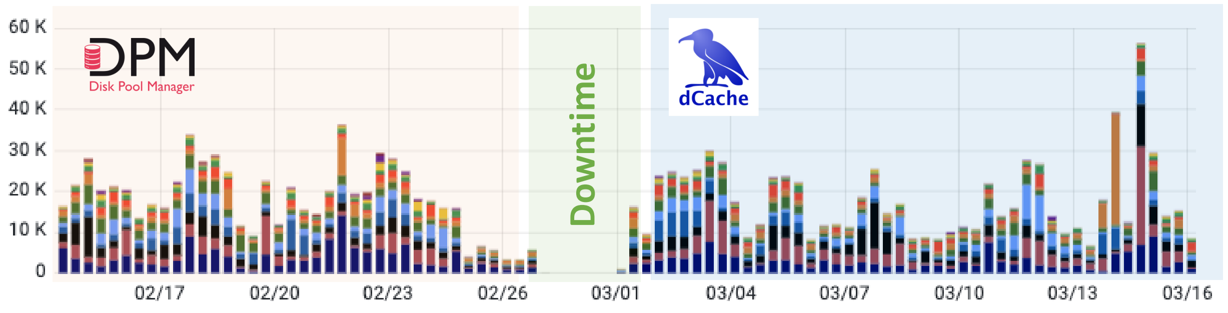
	min	max	avg ^	current
UK	6.71%	100%	94.4%	99.8%
TW	4.41%	100%	94.9%	100%
DE	4.16%	99.9%	95.4%	99.1%

Transfer Volume



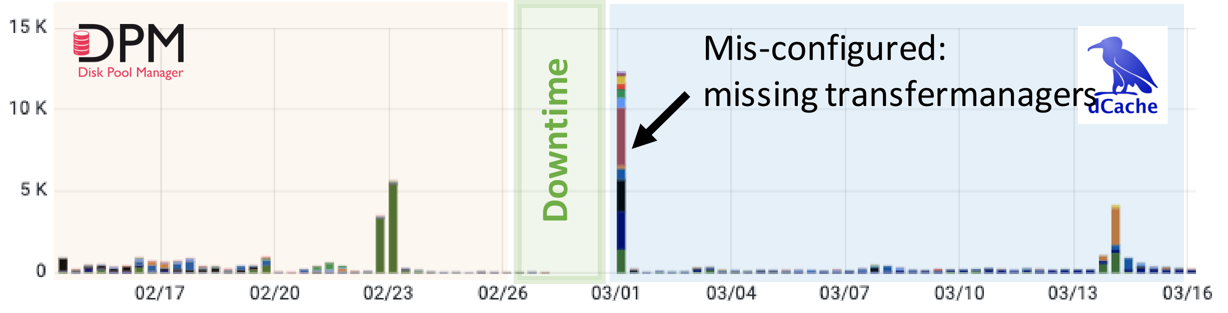
	avg	total ⌄
US	7.68 TB	699 TB
ND	4.90 TB	446 TB
CERN	4.77 TB	434 TB

Transfer Successes



	avg	total ⌄
US	3.04 K	277 K
CERN	2.36 K	214 K
DE	2.15 K	196 K

Transfer Failures



	avg	total ⌄
UK	163	14.8 K
US	97.5	8.87 K
DE	74.0	6.74 K

No issues for transfer

Events after migration (1)

- Mar 2023: Migration completed.

- Apr 2023: dCache pool services were restarting frequently.
  - This happened under heavy load.
  - The service automatically restarted within a few minutes. Not critical, but not good.
  - Cause: Out of memory
    - The JVM memory of the dCache pool service was too small (4 GiB).
    - A xrootd connection uses 8 MiB by default. 2000 connections per each FS are allowed. → 16 GiB is required.
  - Solution: Changed the dCache settings
    - dcache.java.memory.heap : 4 GiB → 8 GiB
    - dcache.java.memory.direct : 2 GiB → 8 GiB
    - pool.mover.xrootd.frame-size : 8 MiB → 2 MiB

- June 2023: The number of active transfers has reached the limit (500).
  - Once the upper limit is reached, no more new connection to the FS can be allowed.
  - Many zombie (TPC) processes were observed.
  - Killed such zombies & added alert to detect them.

Events after migration (2)

- July 2023: dCache stuck due to degradation of DNS cache servers
  - Our DNS cache servers consist of 2 servers (primary and secondary). If the primary server goes down, the secondary DNS server is automatically used.
  - However, after the primary DNS cache server went down, dCache performance has deteriorated significantly.
    - With the current settings there is a 5 second timeout for the fallback.
  - After changing the priority of the DNS cache servers, the problem was resolved.
  - We plan to change the DNS resolve system in the future.
- Sep 2023: Updated from 8.2.12 → 8.2.32 No big troubles. Smooth upgrade!
- Next:
  - Upgrade to 9.2.x release
  - Support WLCG token
  - OS Upgrade: CentOS7 → Alma9

HTCondor 8.8.x to 10.0.x

- Upgraded from HTCondor 8.8.x to 10.0.x
- Strong backward compatibility allows two versions (8.8 and 10.0) to co-exist.
- Also changed authentication methods from Password to IDTokens.
- We have two CE nodes and want to avoid downtime, rolling update was a bit complex.
  - Need to communicate with each other using the same authentication methods.

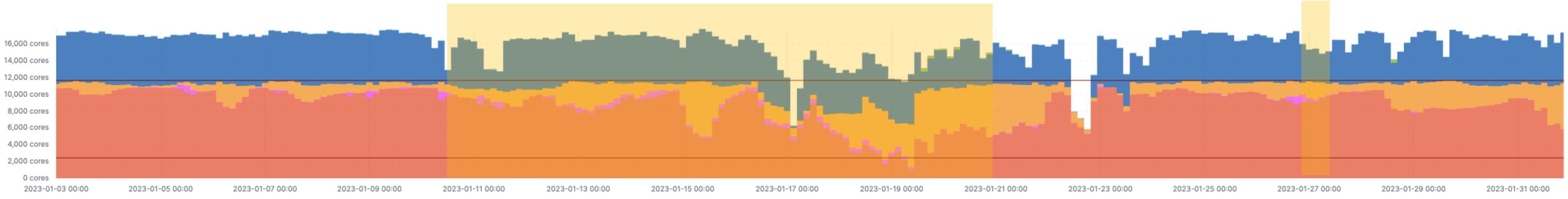
Worker nodes (EP)	HTC 8.8 Password auth.		HTC 8.8 Password auth.		HTC **10.0** Pass & **Token** auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Token auth.
Primary CE (AP, CM)	HTC 8.8 Password auth.	➡	HTC 8.8 Password auth.	➡	HTC 8.8 Password auth.	➡	HTC **10.0** **Token** auth.	➡	HTC 10.0 Token auth.
Secondary CE (AP, CM)	HTC 8.8 Password auth.		HTC **10.0** Pass & **Token** auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Token auth.

HTCondor 8.8.x to 10.0.x

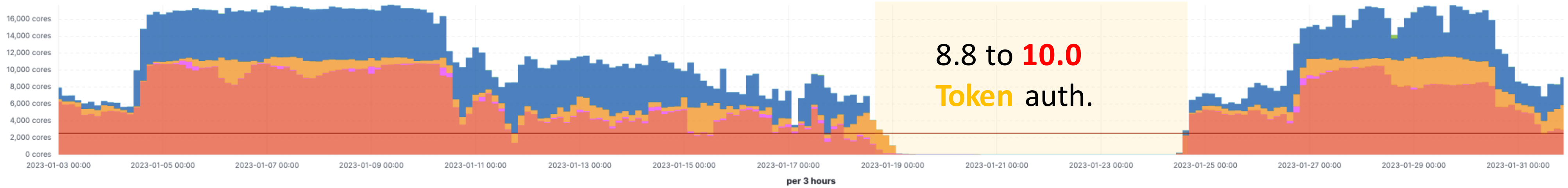
8.8 to 10.0
Pass & Token auth.

Remove Pass auth.
only Token auth.

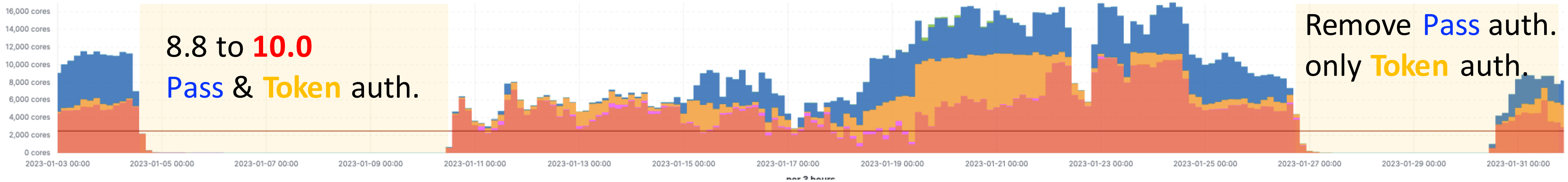
Total cores



Cores managed by
Primary CE



Cores managed by
Secondary CE



Worker nodes (EP)	HTC 8.8 Password auth.		HTC 8.8 Password auth.		HTC **10.0** Pass & **Token** auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Token auth.
Primary CE (AP, CM)	HTC 8.8 Password auth.	➡	HTC 8.8 Password auth.	➡	HTC 8.8 Password auth.	➡	HTC **10.0** **Token** auth.	➡	HTC 10.0 Token auth.
Secondary CE (AP, CM)	HTC 8.8 Password auth.		HTC **10.0** Pass & **Token** auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Pass & Token auth.		HTC 10.0 Token auth.

Miscellaneous

Monitoring and Automation tool

Monitoring system in Tokyo Tier2 site

- Prometheus/Grafana
  - Each server publishes its metrics via exporters. Prometheus server retrieves them.
  - Node info (CPU/Disk/Mem usage, etc.), power usage, network inbound/outbound, etc.
  - Alerts when something happens (e.g. full disk usage, high temperature, cvmfs status etc.)
    - Notification via email and slack. Some of problems are automatically recovered.

Prometheus

Grafana

- Elasticsearch/Kibana (plan to move to OpenSearch)
  - Each server sends its logs via filebeat to the elasticsearch server.
  - Syslogs, network logs (iftop at FS), HTCondor logs, dCache billing, etc.

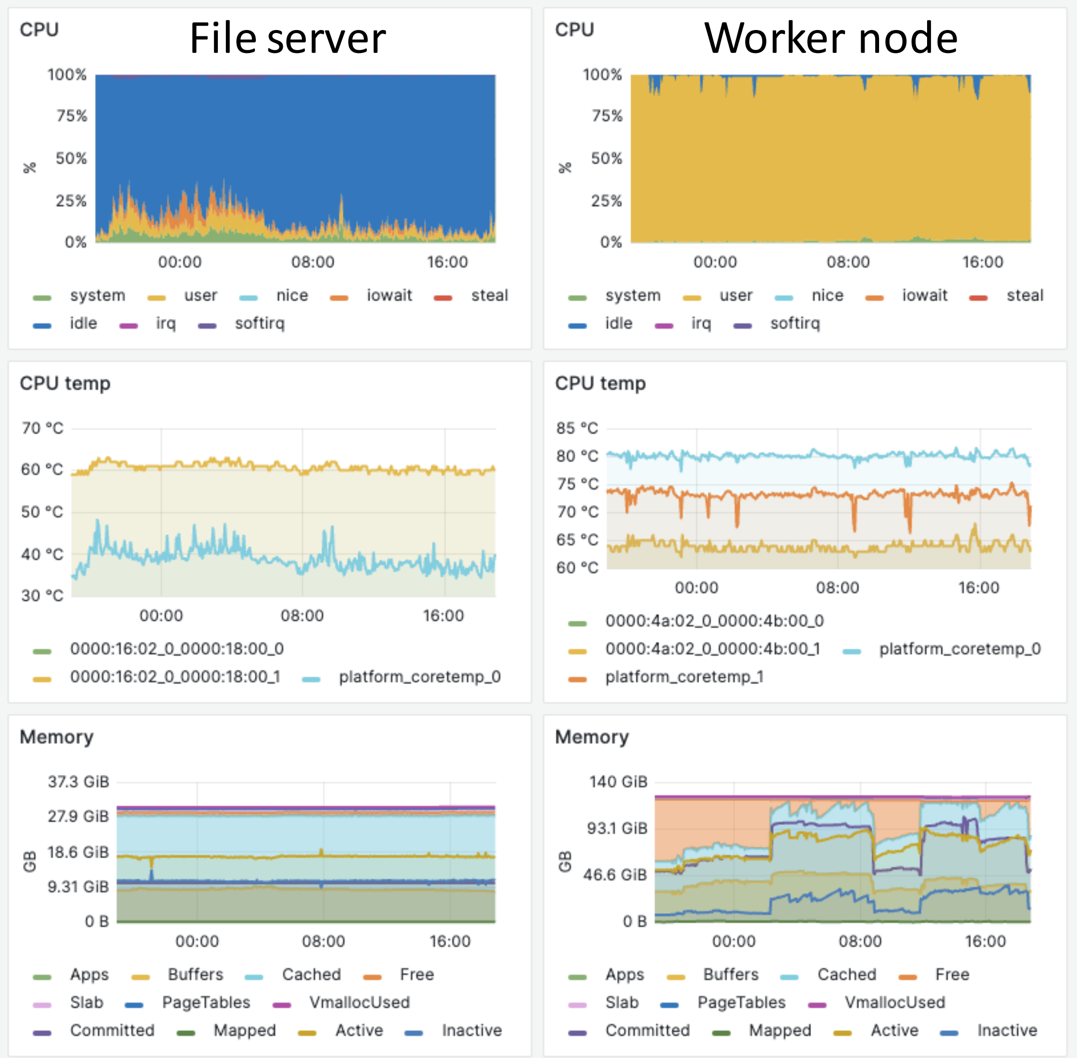
elasticsearch

Kibana

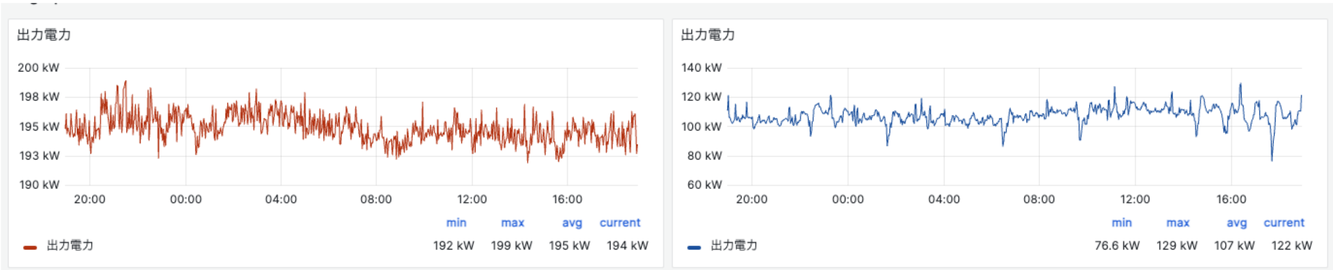
- Others:
  - Legacy cron scripts
  - Site-network information requested recently by WLCG

Monitoring: Prometheus/Grafana

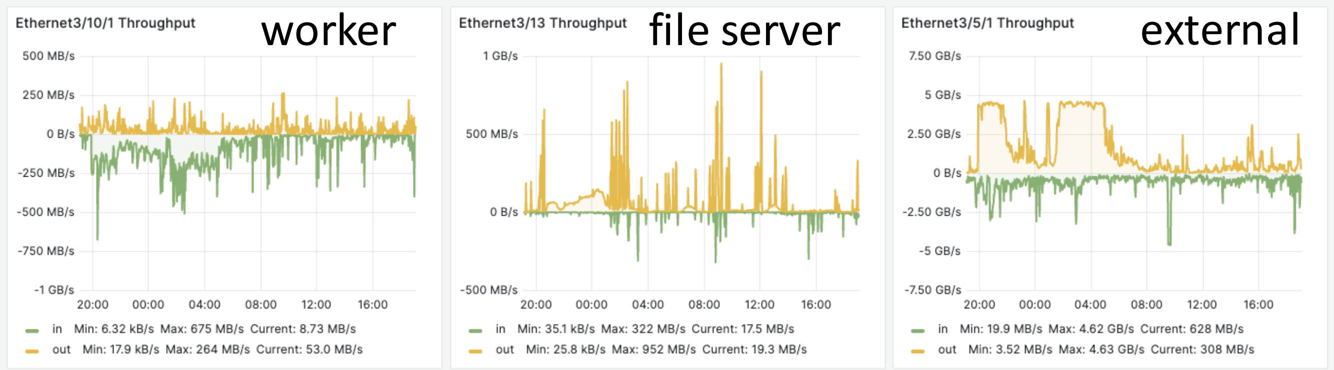
Node metrics



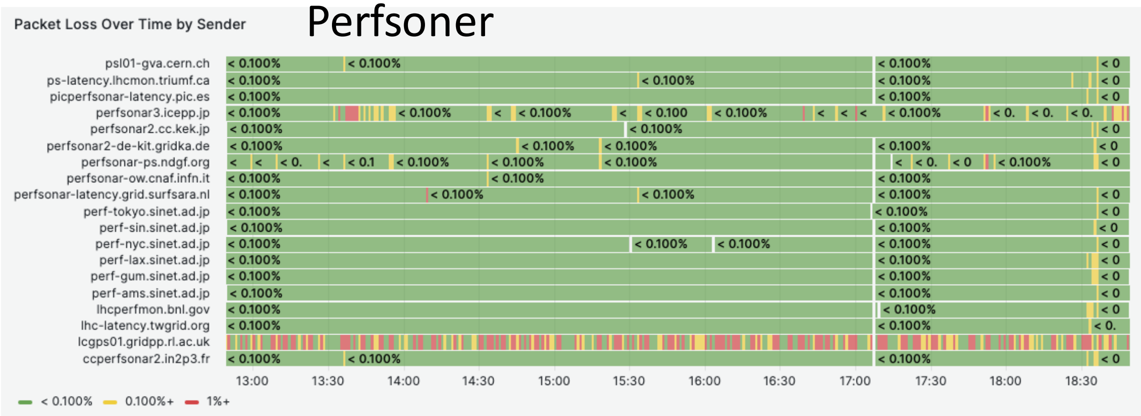
Power metrics



Network metrics

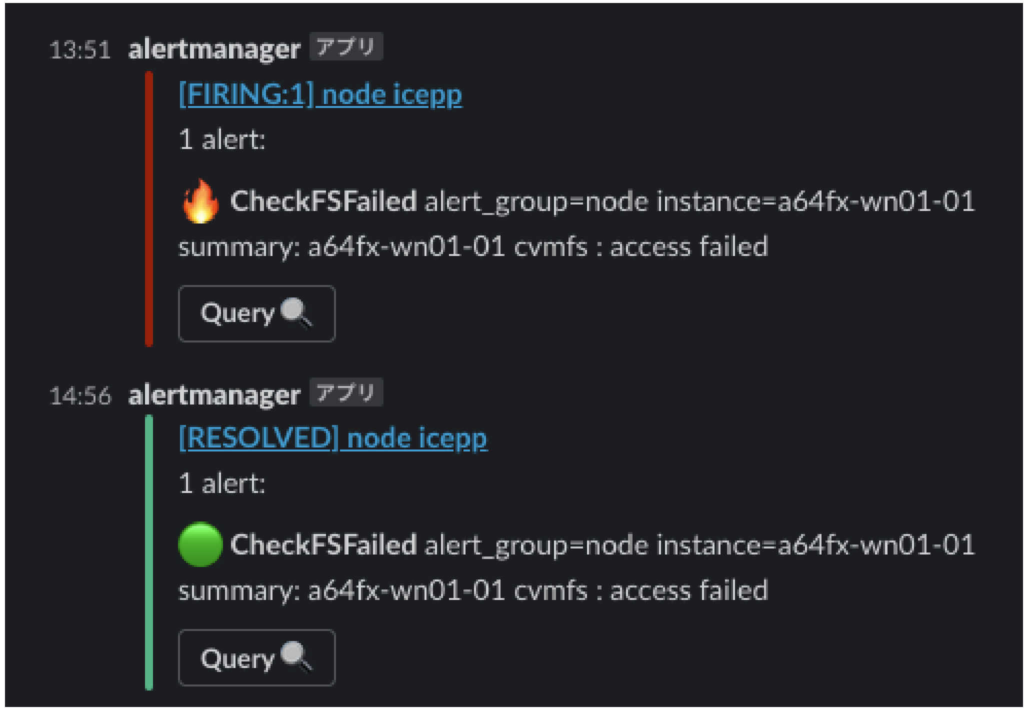


Perfsoner

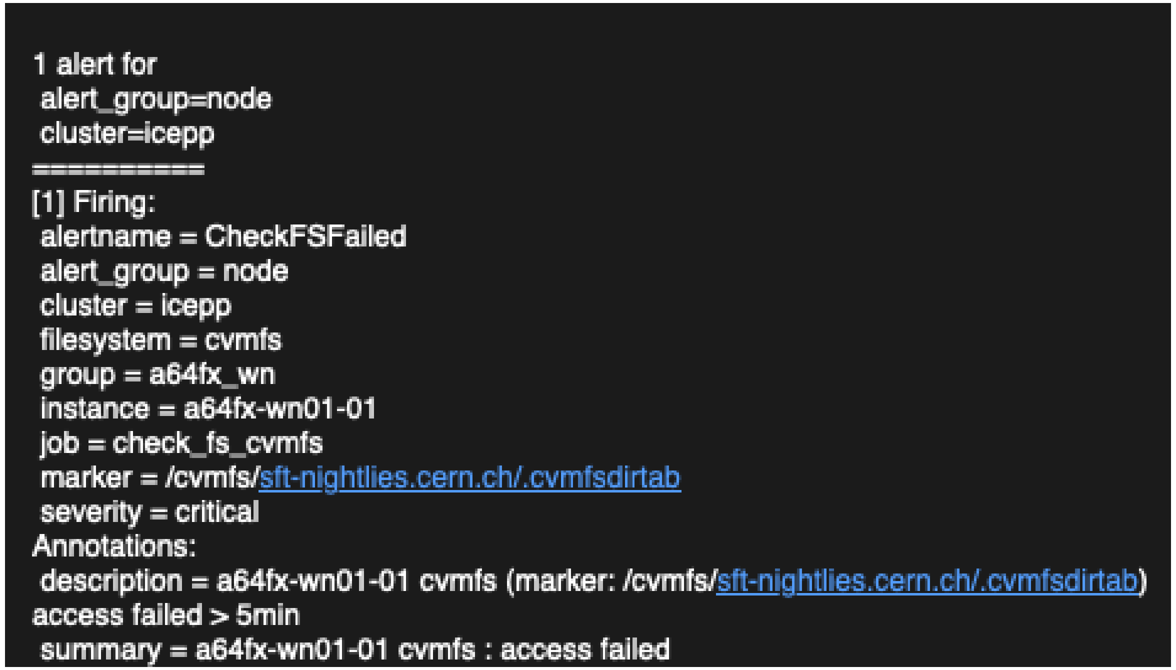


Monitoring: Alerting

Slack

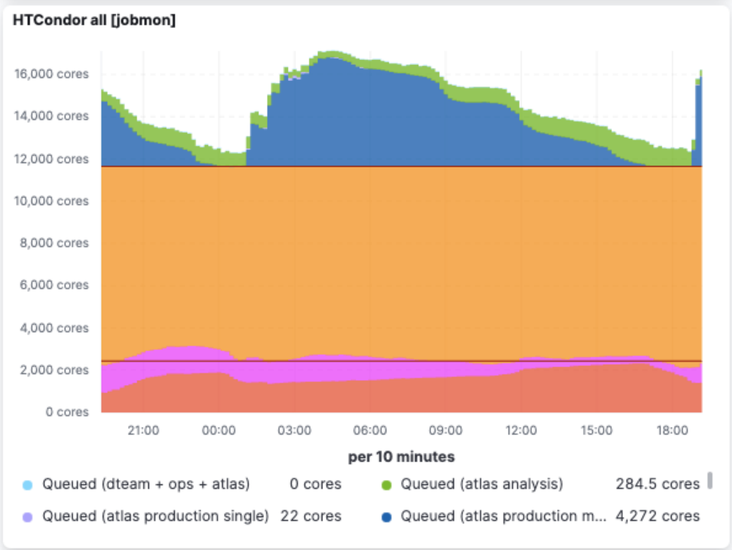


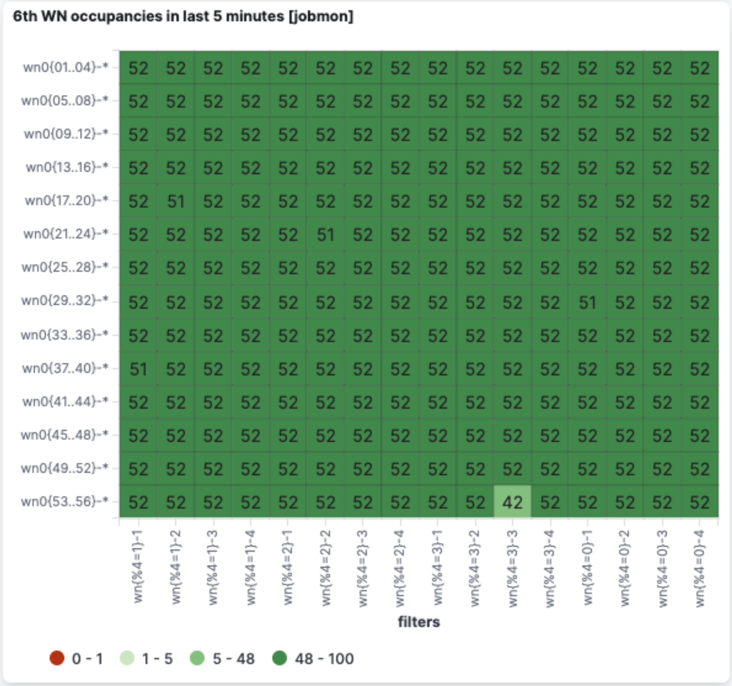
Email

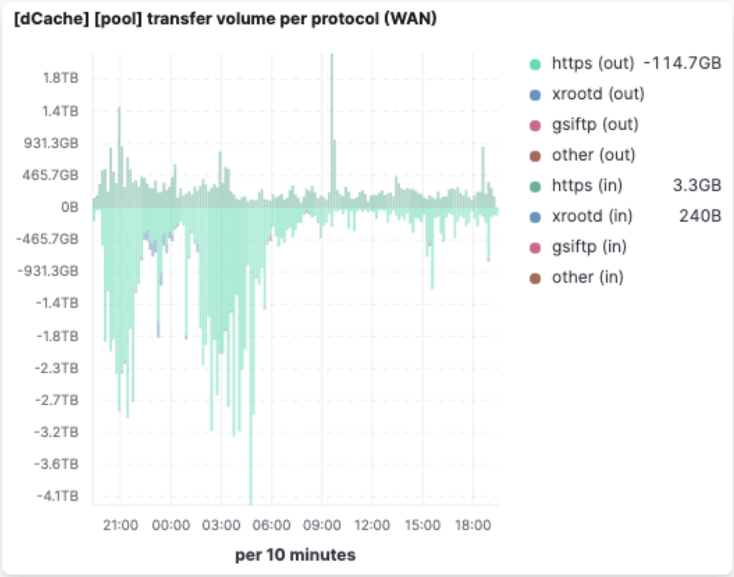


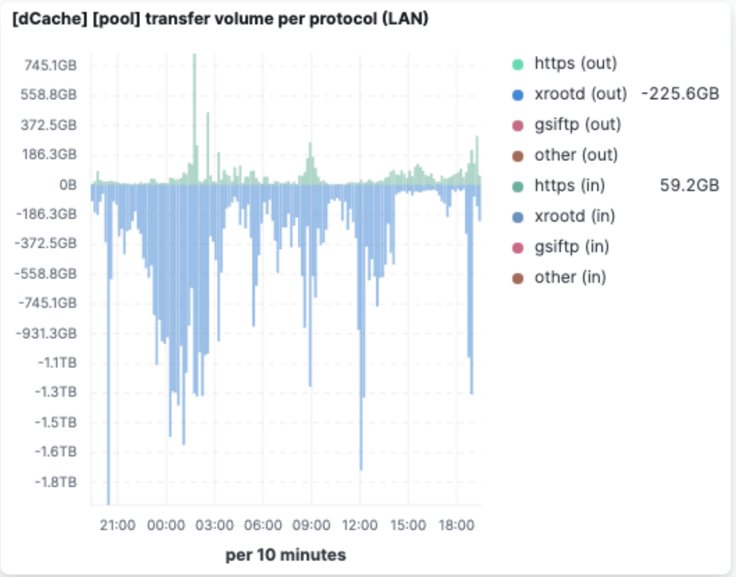
We can receive alerts as soon as a problem occurs and take action quickly.

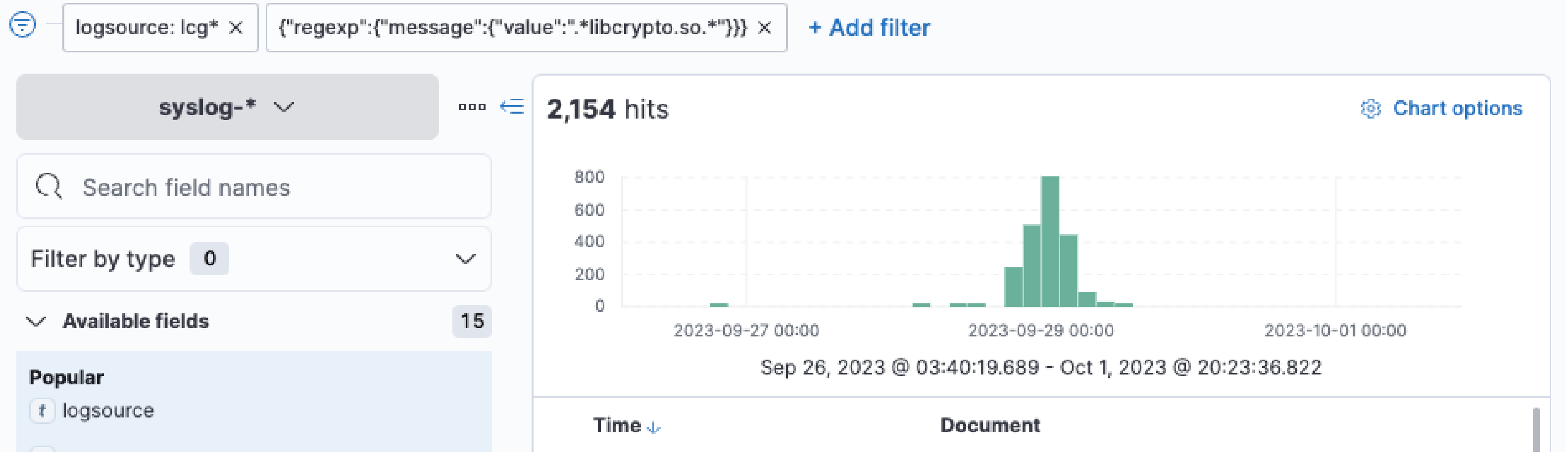
Monitoring: Elasticsearch/Kibana











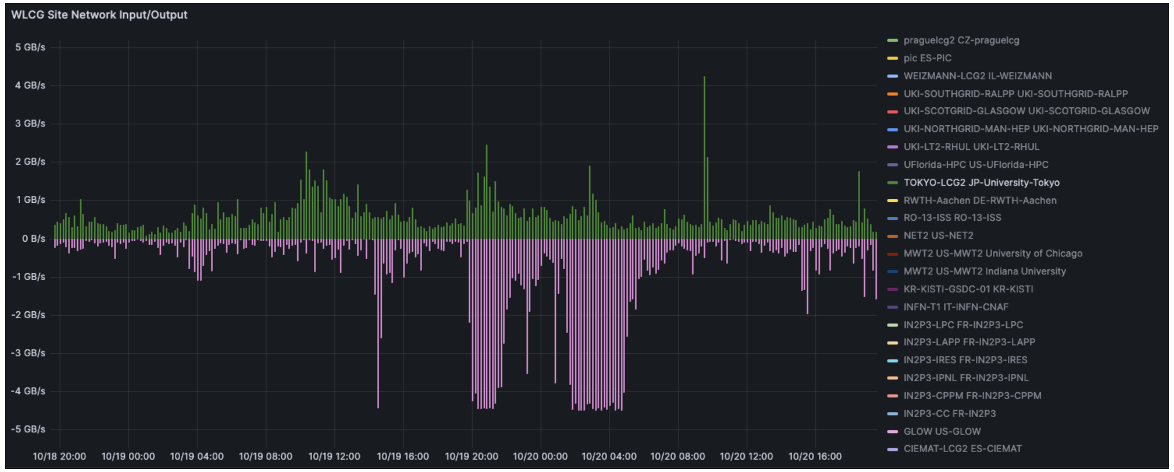
Monitoring: WLCG Site Network

- Published site-network information
  - https://www.icepp.jp/site-network-information/icepp-netmon.json
  - https://www.icepp.jp/site-network-information/ICEPP.html

icepp-netmon.json


```json
{
  "Description": "Network statistics for TOKYO-LCG2",
  "UpdatedLast": "2023-10-20T10:35:15.600539+00:00",
  "InBytesPerSec": 604442581.7994184,
  "OutBytesPerSec": 297742126.9578676,
  "UpdateInterval": "60 seconds",
  "MonitoredInterfaces": [
    "7050sx3-1_Port-Channel3"
  ]
}
```


MONIT Grafana



ICEPP.html

SITE Network Information

This page describes **ICEPP** (International Center for Elementary Particle Physics, The University of Tokyo) network information for WLCG use.

LAST UPDATE: 22-Aug-2023 12:00 JST

Network Overview

Network Description

ICEPP Tier2 Center connects to SINET6 through the Information Technology Center (ITC) of the University of Tokyo. There is a 40 Gbps link between the ICEPP Tier2 Center and the ITC, and a 400 Gbps link between the ITC and SINET6.

All external network traffic to/from the ICEPP Tier2 Center is routed through the L3 router at the ITC, where it is peered to LHCONE and others (non-LHCONE) via SINET6. A firewall is installed on the ICEPP side of the boundary between ICEPP and ICT.

Network Equipment Details

We use Arista 7050SX3 for the firewall and Arista 7504R for the backbone switch.

Network Monitoring

We monitor network traffic on our border router (Arista 7050SX3). It is available at https://www.icepp.jp/site-network-information/icepp-netmon.json

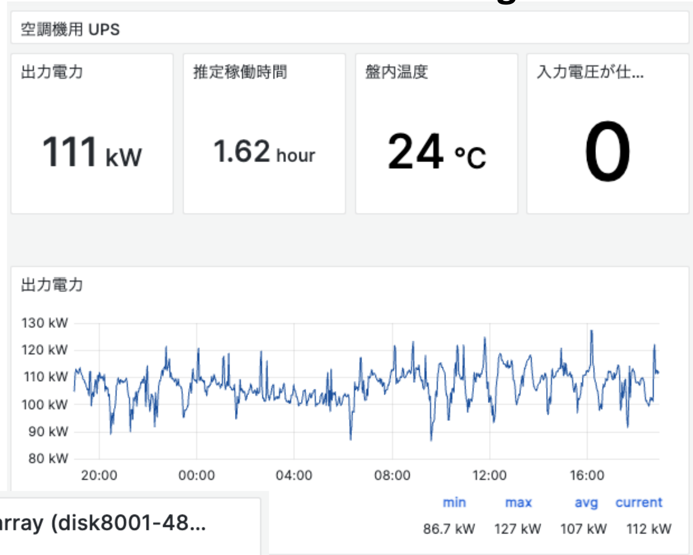
Network Monitoring Link Into CRIC

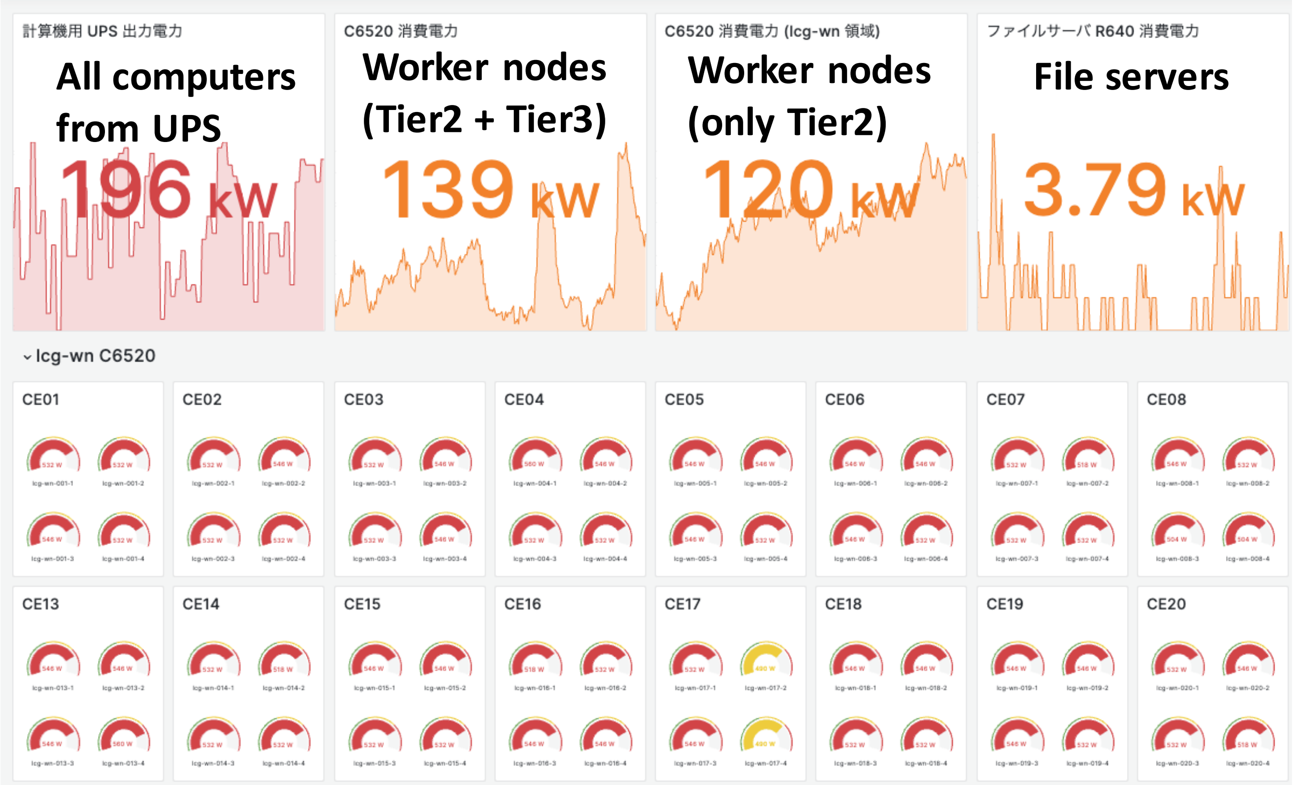
- https://wlcg-cric.cern.ch/core/netsite/detail/JP-University-Tokyo/

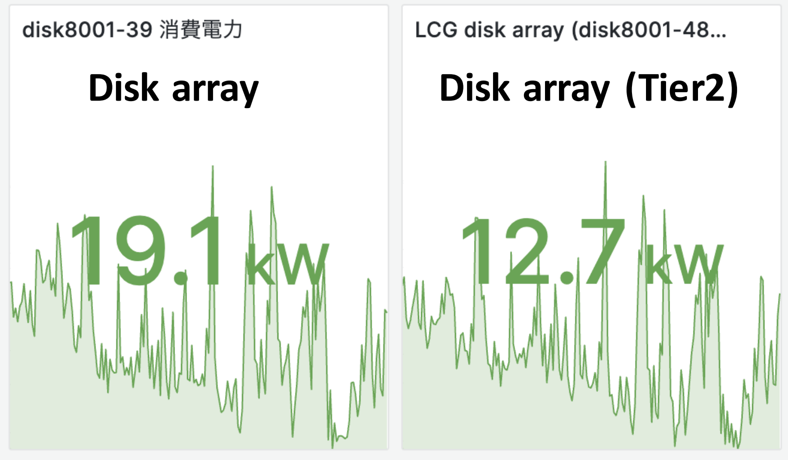
Power consumption monitor

- We monitor the power consumption of our computing clusters.
- We have recently added new monitors to collect information from each server.
- We use this monitoring for future R&D to reduce power consumption.
  - Current PUE ~ 1.56
  - Plan to change air conditioning which has been running since 2007

Air conditioning







Automation tool

- We use Ansible to automate middleware deployments.
  - Gradually moved away from Puppet since ~2020.
  - Ansible is easy to learn, agentless and supported by Red Hat.
- Most of LCG middleware roles have been ported from Puppet. Community developed roles are used for other common software.
- All deployments and configuration changes are done through Ansible. (Except for testing/debugging)
  - For the 2022 system migration, almost all deployments were done using Ansible.
- Do we publish our Ansible deployment code? A lot of work is needed to do this.
  - Improve documents (e.g. README), check if there is confidential information, check if a licence is correct.
  - Please wait for a while. We are positive about this.

Summary and plan

- ICEPP regional analysis center is operating stably.
- Contributes to ~5% CPU and ~3% Disk of ATLAS sites
- SE migration (DPM → dCache) was successfully completed.
- Near term upgrade plan
  - External network: 40 Gbps → 100 Gbps
  - CentOS7 → Alma9
  - Dual stack of worker nodes.
- Next hardware replacement
  - Tokyo site's hardware has been leased and replaced every 3 years so far.
  - Due to recent delays and increased of delivery times, we plan to extend the the lease by one year. In the future, we expect to replace the hardware every 4 or 5 years.