Status report on KEKCC and Grid systems

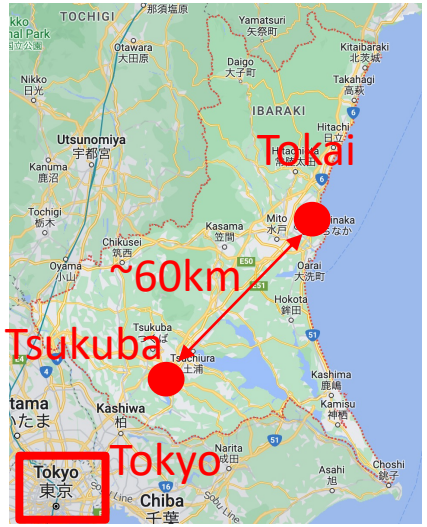
Tomoe Kishimoto

Computing Research Center, KEK

Mission of KEK

- KEK covers diverse field of accelerator based science making full use of the electron machines in Tsukuba and the proton machines in Tokai, e.g.)
- SuperKEKB/Belle II experiment
  - B-factory experiment to make precise measurements of weak interaction parameters and find new physics
- J-PARC
  - Hadron hall: Particle and nuclear physics with fixed target
  - Neutrino facility: Neutrino beamline for T2K experiment and upgrade program for Hyper-Kamiokande

Ibaraki prefecture



Tsukuba campus

SuperKEKB, PF/PF-AR
LINAC, ATF/STF

Tokai campus

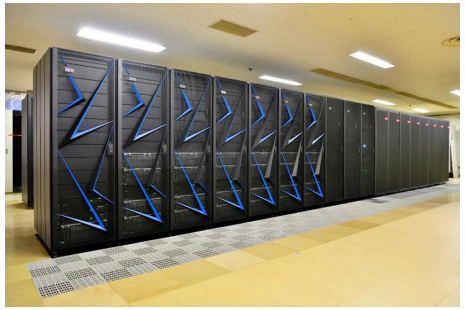
J-PARC (co-operated by KEK and Japan Atomic Agency (JAEA))

KEK Computing Research Center

- KEK Computing Research Center (CRC) provides computational resources and networks to support many research activities at KEK
  - Central Computer System (KEKCC) @ Tsukuba campus
  - Internet connections, campus networks (KEK-LAN, JLAN)
  - Supercomputer system
  - E-mail and Web systems, etc
- Research and Development
  - Geant4 for detector simulations
  - Distributed computing, etc

KEK CRC HP

CRC building



KEKCC

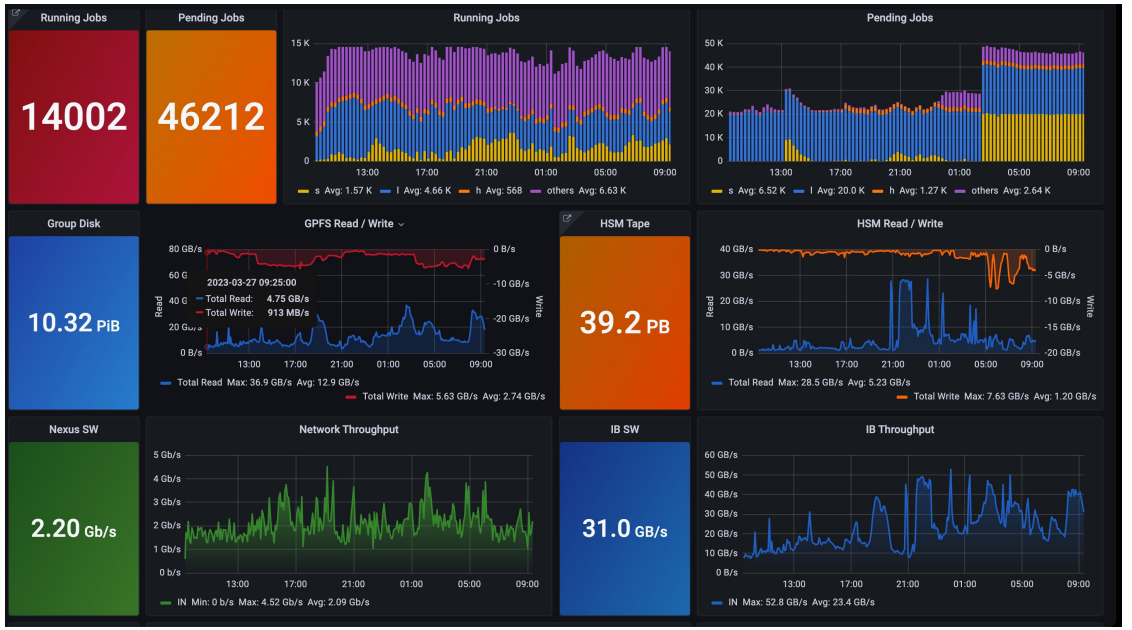
Supercomputer
(NEC SX-Aurora TSUBASA)

KEKCC in numbers

- KEKCC is a rental system replaced every 4-5 years
  - Current KEKCC started operations in Sep. 2020, the next procurement is ongoing
  - Linux cluster + storage system (GPFS/HSM)
- CPU: 15,200 cores
  - Intel Xeon Gold 6230 2.1 GHz, 380 nodes
- Memory: 87 TB
  - 4.8 GB/core (80%) + 9.6 GB/core (20%)
- Disk: 25.5 PB
  - 17 PB: GPFS for experimental groups
  - 8.5 PB: GPFS-HPSS interface (GHI) as an HSM cache
- Tape: 100 PB as maximum capacity

KEKCC
Central Computing System

Monitoring dashboard

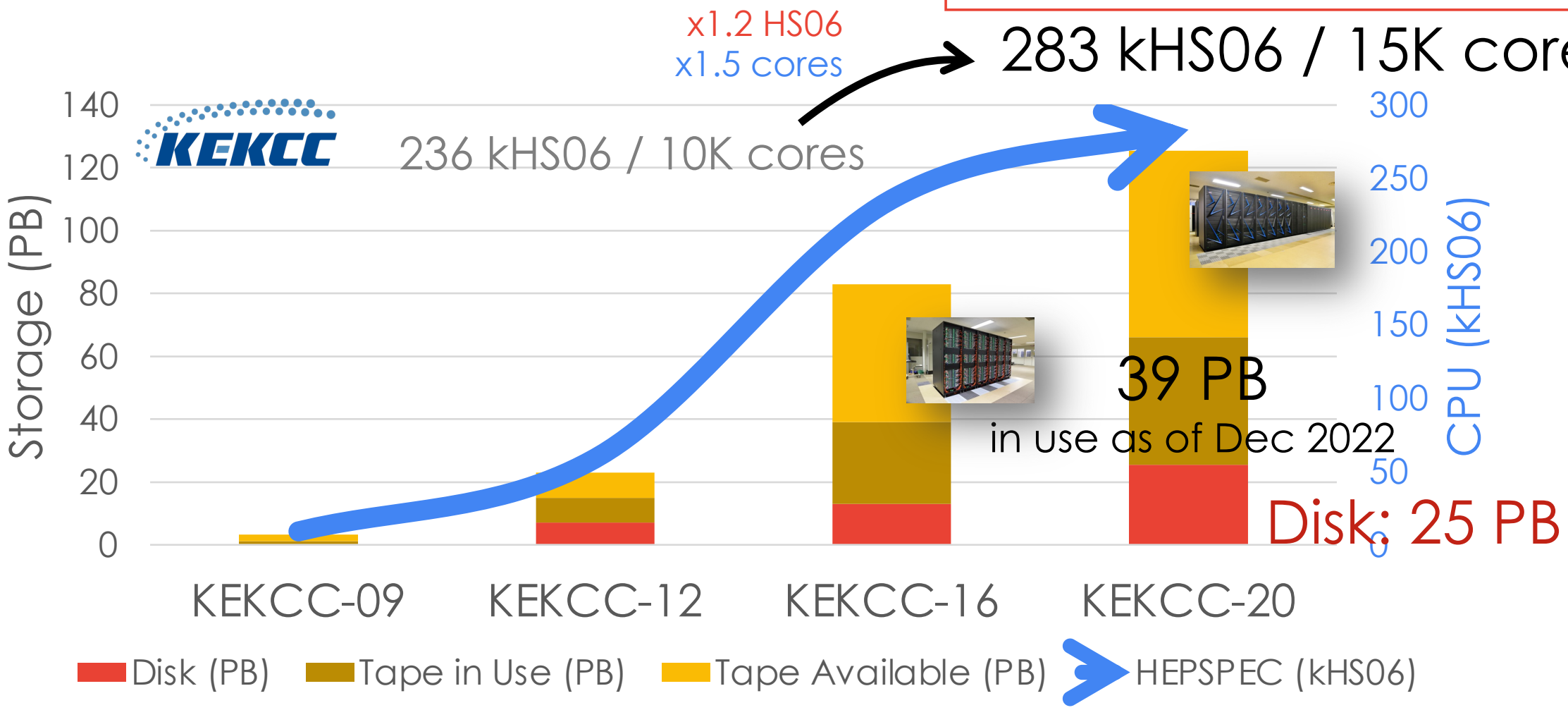


Site Scale Evolution

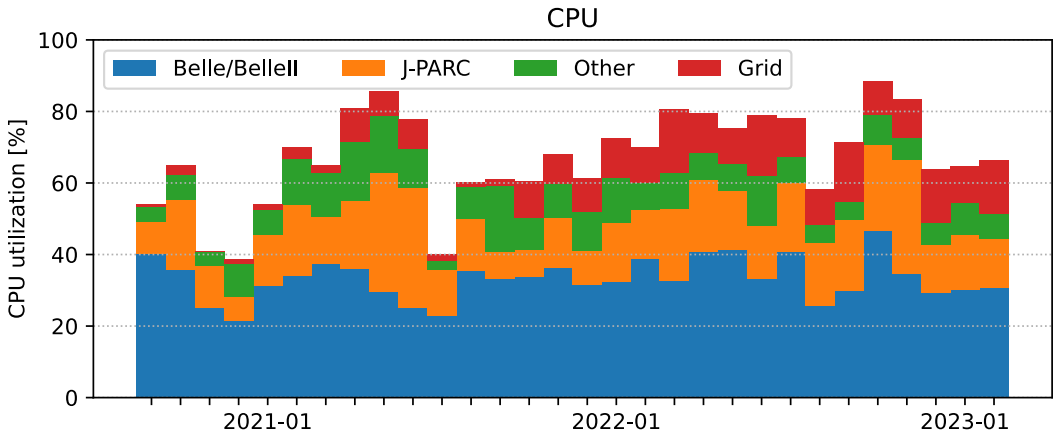
Resource History (Last 4-Gen)

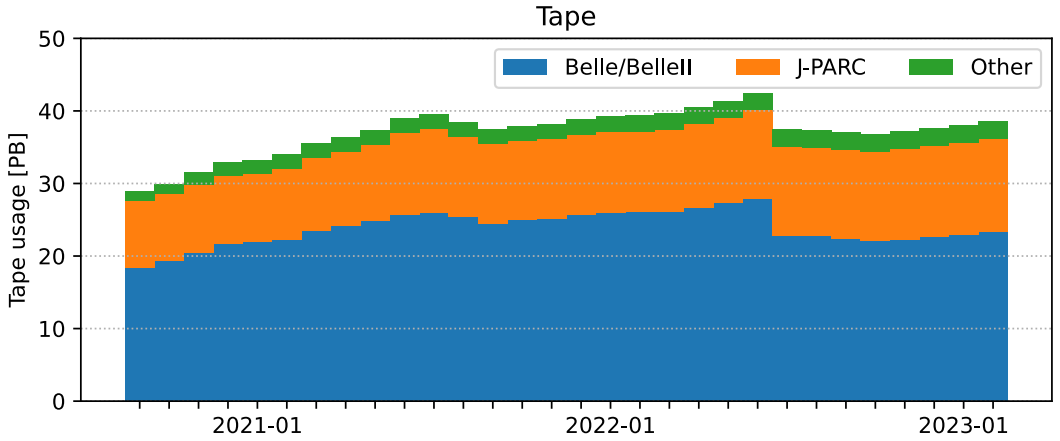
G. Iwai

283 kHS06 of CPU
25.5 PB of disk
Max 100 PB of tape capacity



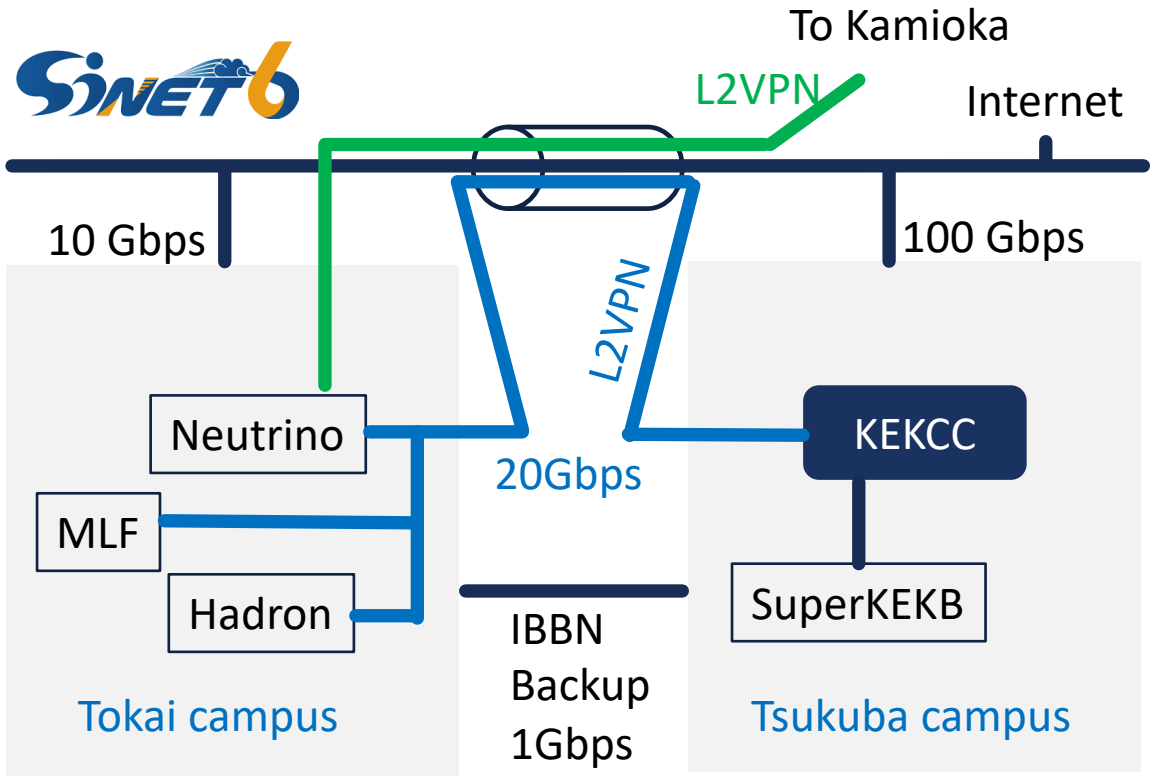
KEKCC resource utilization in the current system





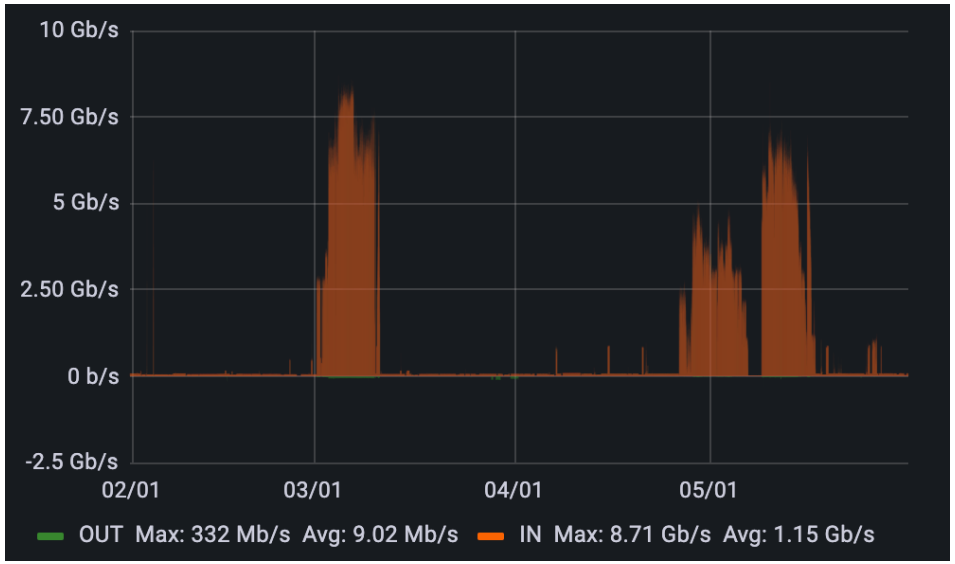
- CPU resources are well utilized
  - Belle II experiment (local batch system) is a dominant consumer
  - Grid jobs are increasing
- Tape usage reached 40 PB

Networks between campuses



IBBN: Ibaraki Broad Band Network
hosted by Ibaraki prefecture

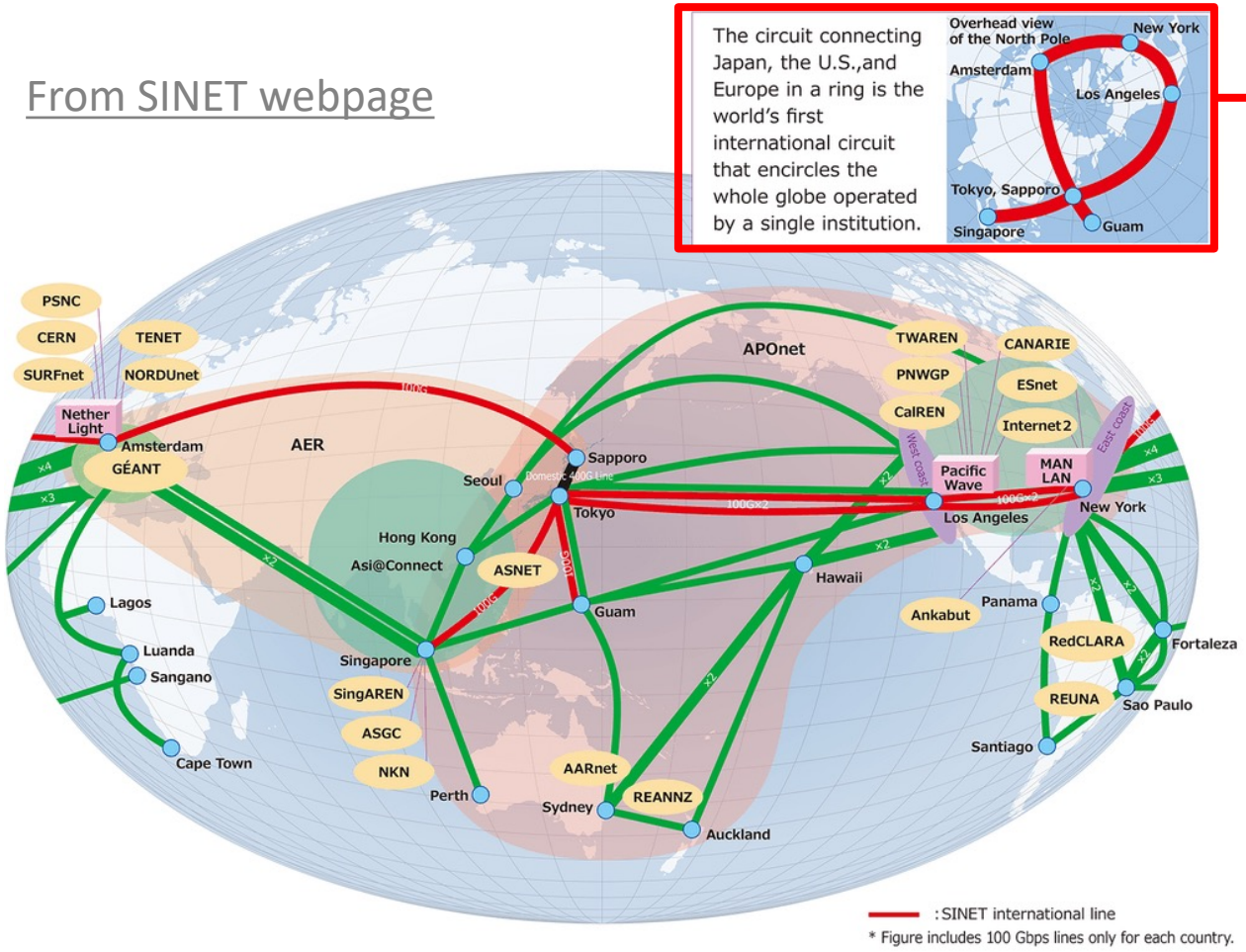
J-PARC (JLAN) ⇄ KEKCC



- Experimental data produced in J-PARC are transferred to KEKCC via SINET L2VPN

Global networks (SINET6)

From SINET webpage



- 100 Gpbs global ring
  - USA: Los Angeles and New York, 100Gbps x2
  - Europe: Amsterdam, 100Gbps
  - Asia: Singapore and Guam, each 100Gbps
- KEKCC connects to LHCONE (L3VPN) for BelleII data transfers with other sites
  - Shares VRF with ICEPP (ATLAS)

G. Iwai

Grid Services 2023

Both Belle II StoRM now on CentOS7

as Belle II dedicated

Service	OS	VM/Bare metal	Ethernet	IPv6	High Availability	Uninterruptable
StoRM (FE/BE)	CentOS7	Bare metal	10GE	✅	✅	✅
VOMS	CentOS7	VM on RHEL8	10GE	✅	✅ SIOS LifeKeeper	✅
LFC	RHEL6 + ELS	VM on RHEL8	10GE			
AMGA	7	Bare metal	10GE			
Top BDII	7	VM on RHEL8	10GE			
Site BDII	CentOS7	VM on RHEL8	10GE	✅	✅	✅
ARGUS		Bare metal	10GE	✅	✅	✅
FTS3		Bare metal	10GE	✅	✅	✅
ARC-CE	CentOS7	Bare metal	10GE	✅	✅	
GridFTP / **WebDAV**	CentOS7	Bare metal	40GE	✅	✅	✅
CVMFS Stratum Zero	CentOS7	Bare metal	10GE	✅	✅	
CVMFS Stratum One	CentOS7	Bare metal	10GE	✅	✅	
HTTP Proxy	CentOS7	Bare metal	10GE	✅	✅	

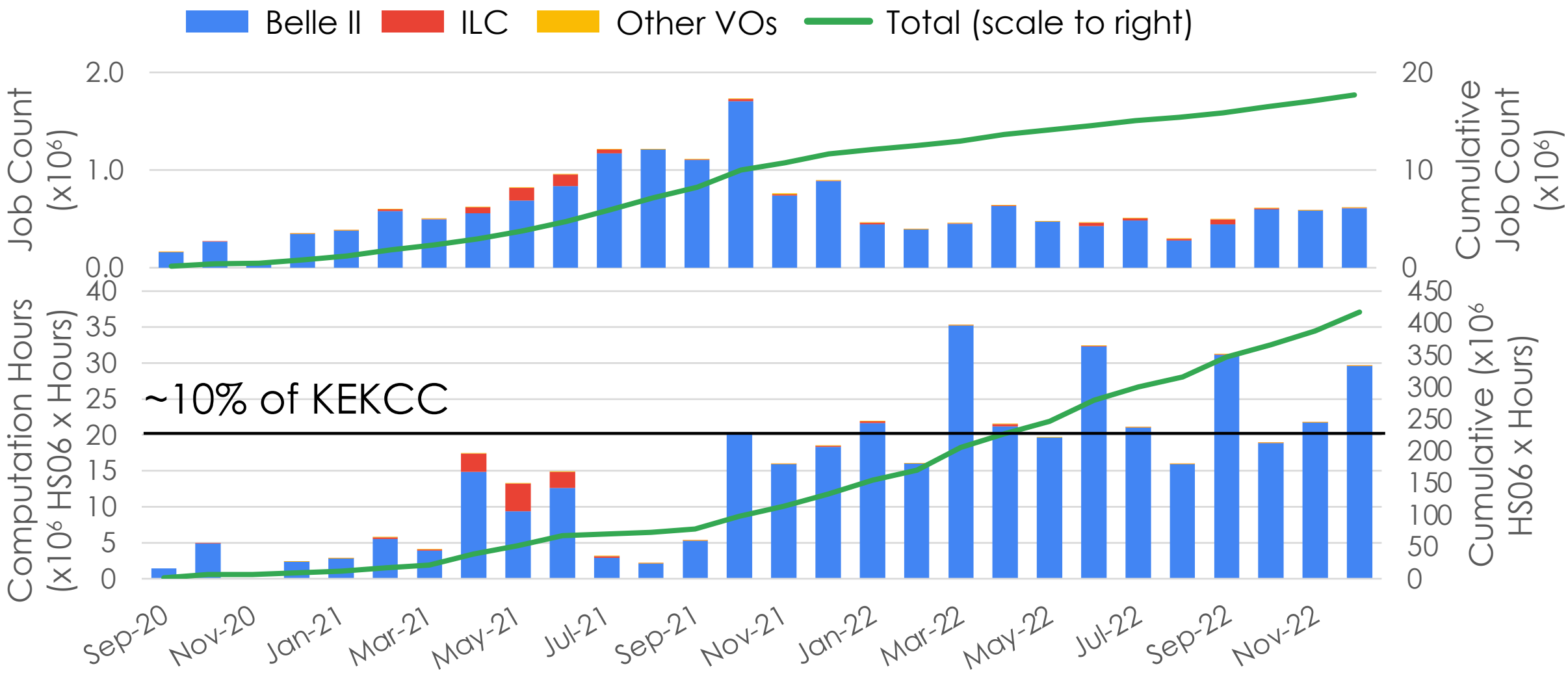
Decommissioned Dec 2021

Migrated and IPv6 ready Sep 2021

New ARC instances replaced Dec 2021

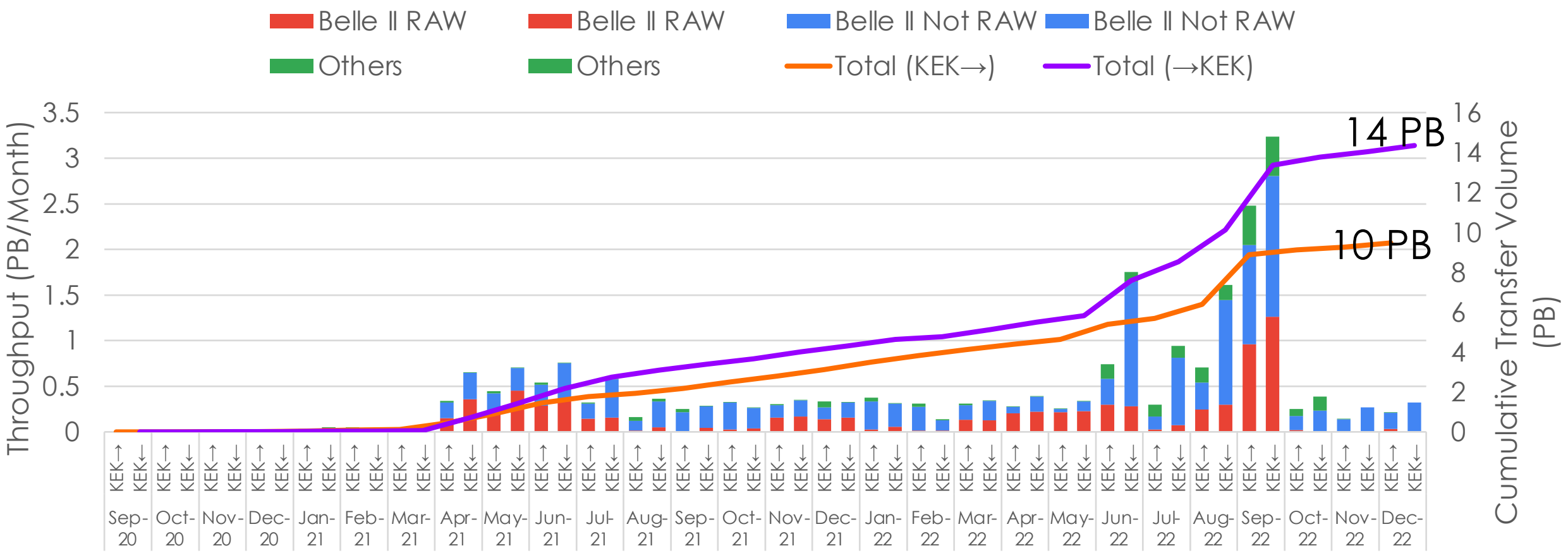
G. Iwai

Grid Jobs

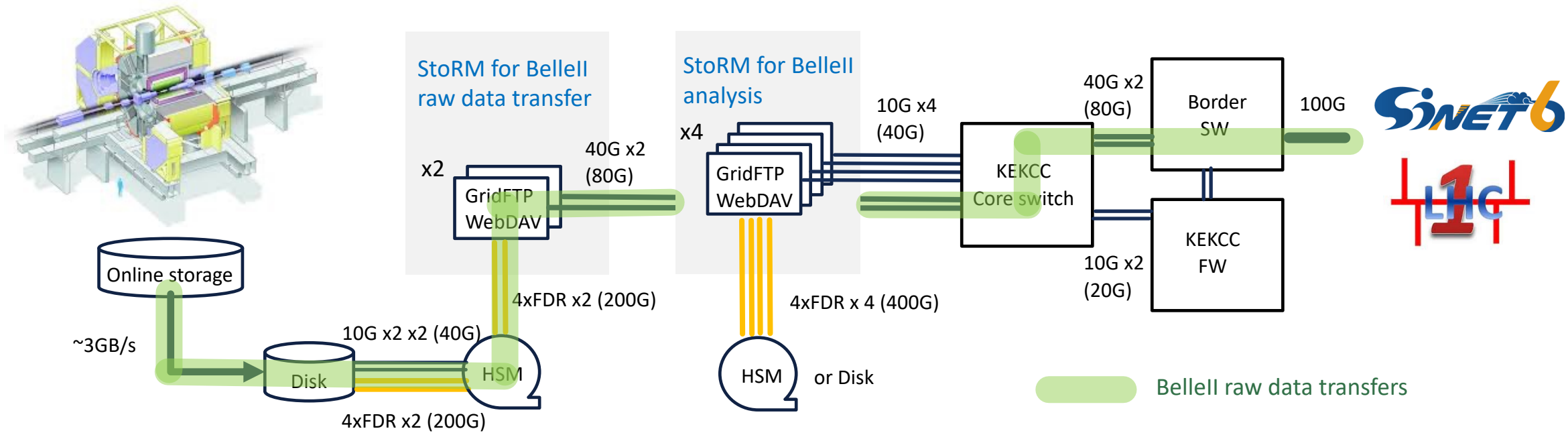


G. Iwai

Transfer Volume from/to StoRM (Not Including Internal Transfer)



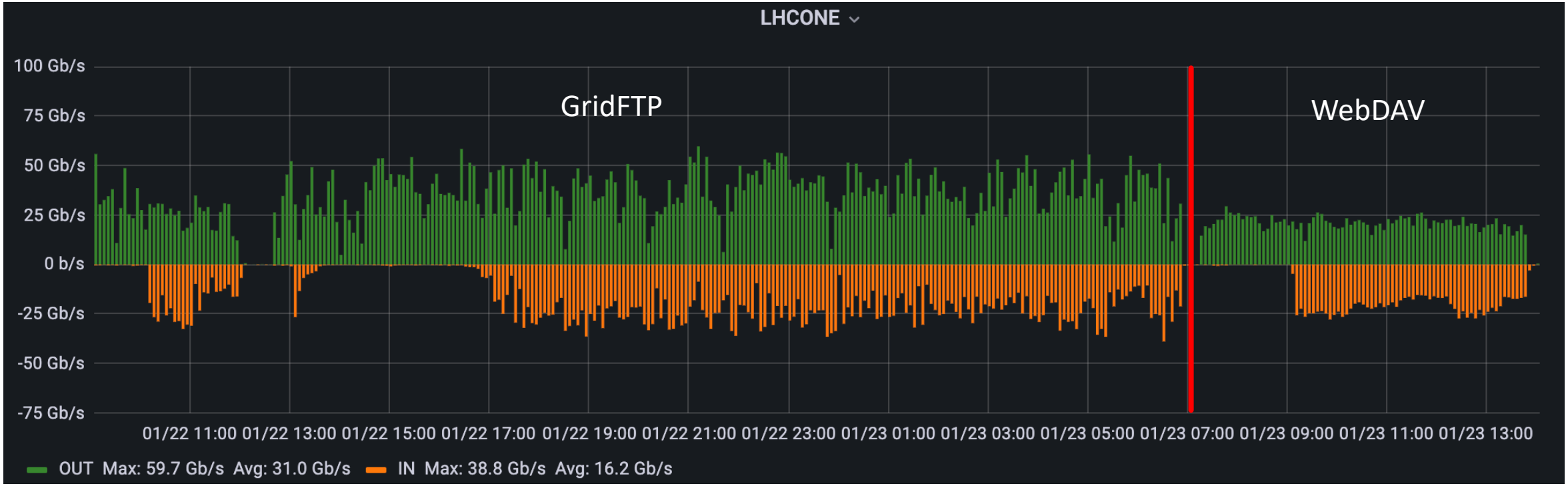
StoRM configuration for BelleII



- BelleII raw data transfers are one of main missions of Grid system
  - Separated StoRM instances from analysis activities and other VOs
  - Multiple StoRM instances to ensure the transfer capability (DNS round robin to select an instance)

Storm transfer performance

KEKCC ⇄ Raw data centers



- WebDAV degraded the throughput in our environment

WebDAV

- WebDAV transfers seem CPU intensive
  - Currently, two instances for Belle II raw data transfers
  - >75% CPU usages were observed

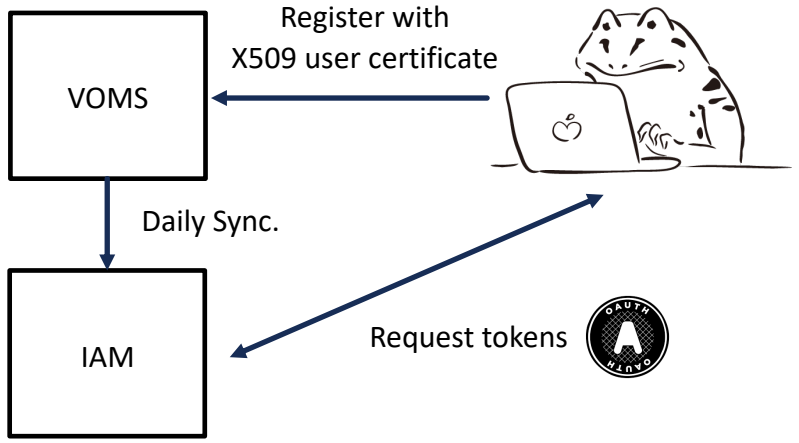
  → Maybe, better to increase transfer instances
- Load-balancing mechanism based on DNS round-robin seems a poor control

  → Considering using NGINX (redirect/reverse proxy) as a load-balancer

CPU usage of two transfer instances

Token migration

- IAM instances have been deployed to support token-based AuthN/Z for BelleII activities
  - User information is synchronized with VOMS
  - Currently, still pre-production mode with limited users
- Third Party Transfers (TPC) based on tokens have been confirmed using FTS+StoRM
  - Job submission tests using ARC-CE are ongoing
- Need to establish a registration procedure without X509 user certificate after terminating VOMS service



Summary

- KEKCC and Grid systems are well in operations
  - The next KEKCC procurement is ongoing and will launch the new system in September 2024
- WebDAV is a main protocol for BelleII raw data transfers between Grid sites
  - Considering to deploy NGINX for a better load balancing
- Toward getting ready for token-based environment
  - Confirmed: File transfers between StoRMs by using FTS
  - Not yet: Verify ARC-CE to work with tokens
  - Not yet: Establish registration procedure to IAM without X509 user certificate