



Experience on the operations at new KEKCC - KEK Site Report -

Tomoaki Nakamura
on behalf of KEKCC team

Computing Research Center
HIGH ENERGY ACCELERATOR RESEARCH ORGANIZATION, KEK



Recent topics from KEK



Press Release

The Belle II detector was "rolled-in" to the collision point of the SuperKEKB accelerator

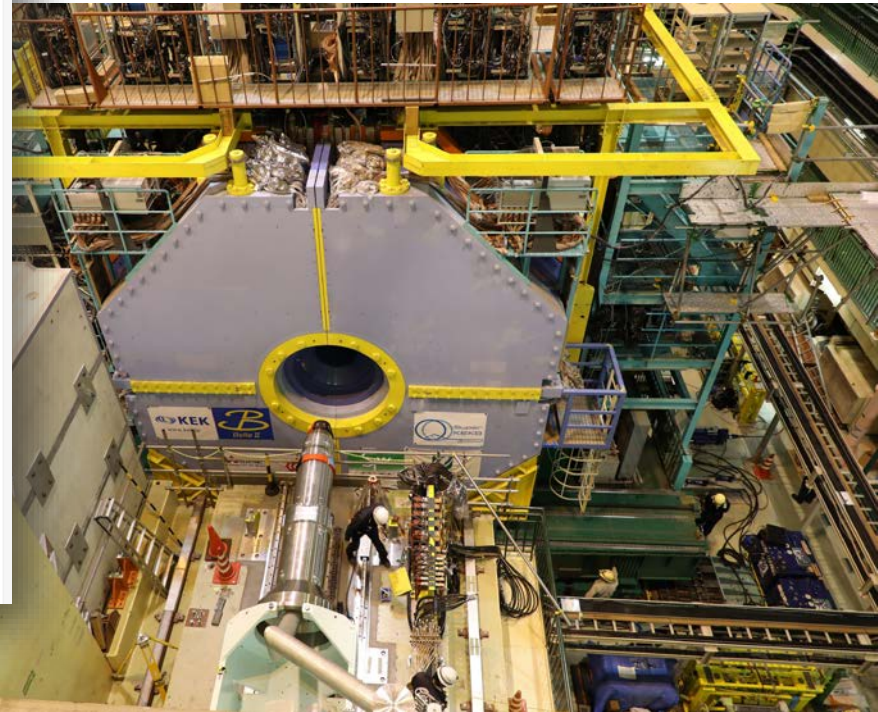
April 11th, 2017 The Belle II detector was "rolled-in" to the collision point of the SuperKEKB accelerator. The High Energy Accelerator Research Organization (KEK) successfully completed the "roll-in" of the Belle II detector on the afternoon of Tuesday April 11th. This is an important milestone for the international Belle II collaboration and the SuperKEKB accelerator. The Belle II experiment is an international collaboration with 23 member nations hosted by KEK in Tsukuba, Japan. Using a state-of-the-art experimental apparatus, Belle II explores the mysteries of the beginning of the universe. The Belle II detector precisely measures elementary particle interactions artificially created with the upgraded SuperKEKB accelerator. The term "roll-in" refers to the operation of moving the entire Belle II detector system, following the completion of the assembly and integration of the various components, from its assembly area to the beam collision point. The total weight of the Belle II detector is about 1400 tons. Belle II was slowly and carefully moved about 13 meters from the assembly point to the beam collision point. The Belle II detector and the SuperKEKB accelerator are now integrated. In the Belle II experiment, we will observe various elementary particles generated from high energy electron-positron collisions using the 8-meter tall Belle II detector consisting of seven types of subdetectors; we will investigate the various kinds of elementary particles that emerge from these collisions. The detector will provide measurements of the direction and momenta of the newly produced particles. Compared to the previous Belle experiment, Belle II will allow for the collection of much larger data samples with much improved measurement precision. More than 700 researchers from the 23 member countries and regions around the world will participate in the Belle II experiment. Their goal is to find a significant "deviation" from the Standard Model of particle physics and perhaps determine which of the many proposed new theories describes the world of elementary particles.

Apr. 11, 2017

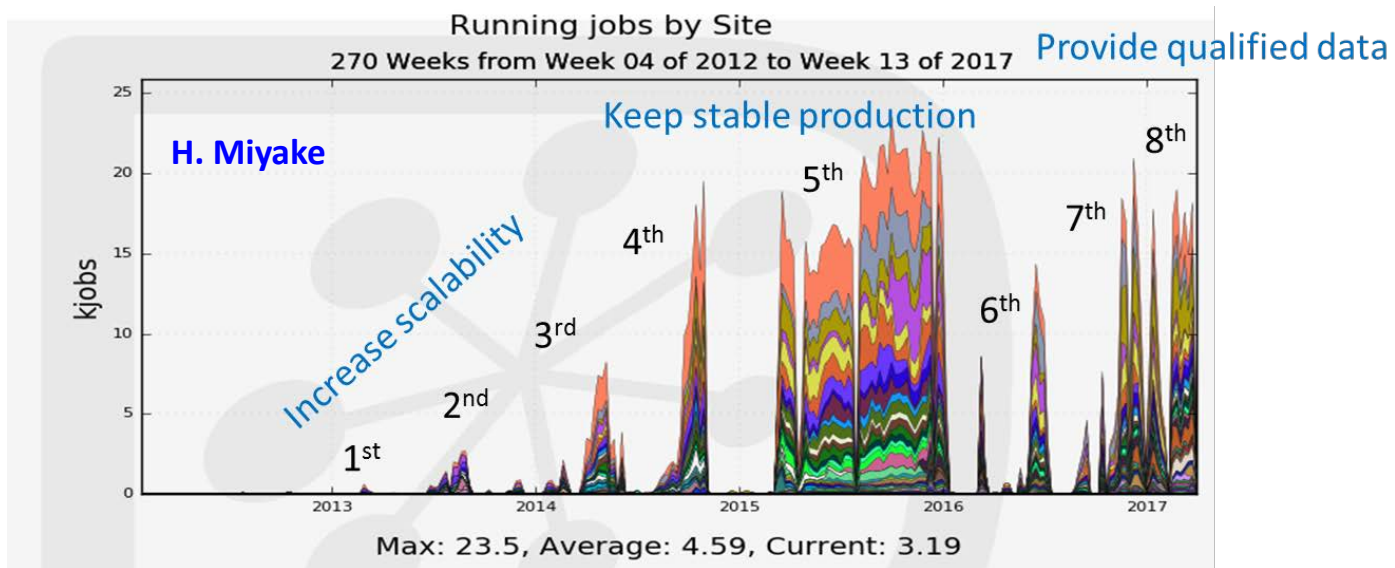
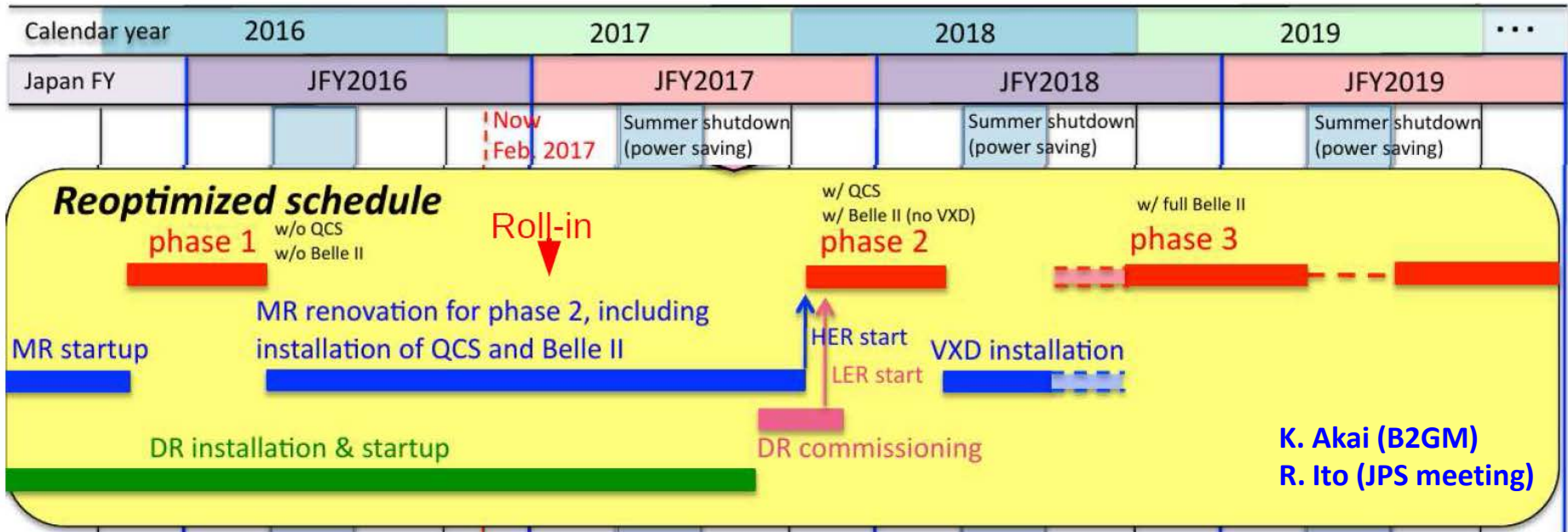
<http://www.kek.jp/en/NewsRoom/Release/20170411171500/>

Mar. 2016: Test operation of SuperKEKB
Successful storage of e^+e^- beam

Apr. 2014: Belle II detector was rolled-in



Belle II schedule and distributed computing

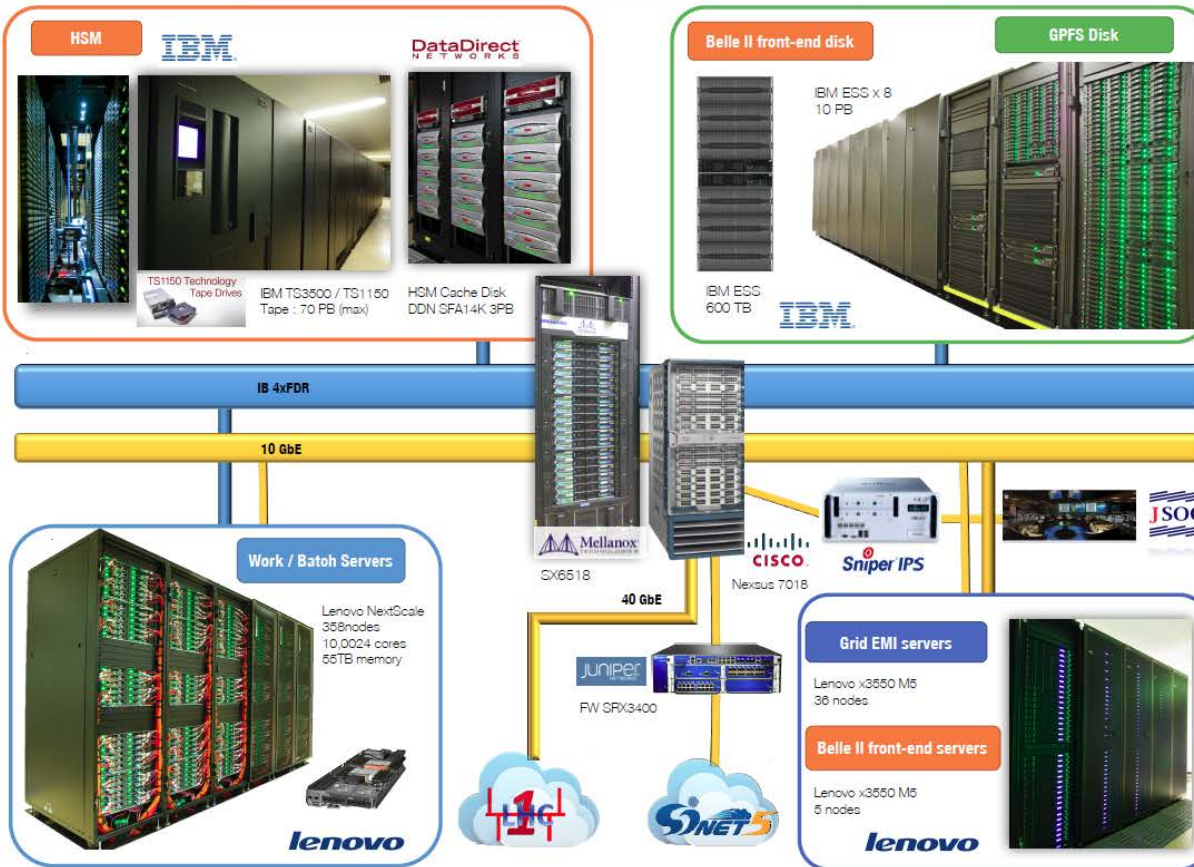


New KEK Central Computer System (KEKCC)



System has already been in production mode since the last September.

KEKCC 2016



SYSTEM RESOURCES

- CPU :** 10,024 cores
- ▣ Intel Xeon E5-2697v3 (2.6GHz, 14cores) x 2 358 nodes
 - ▣ 4GB/core (8,000 cores) / 8GB/core (2,000 cores) (for app. use)
 - ▣ 236 kHS06 / site

Disk : 10PB (GPFS) + 3PB (HSM cache)

Interconnect : IB 4xFDR

Tape : 70 PB (max cap.)
HSM data : 8.5 PB data, 170 M files, 5,000 tapes

Total throughput : 100 GB/s (Disk, GPFS), 50 GB/s (HSM, GHI)

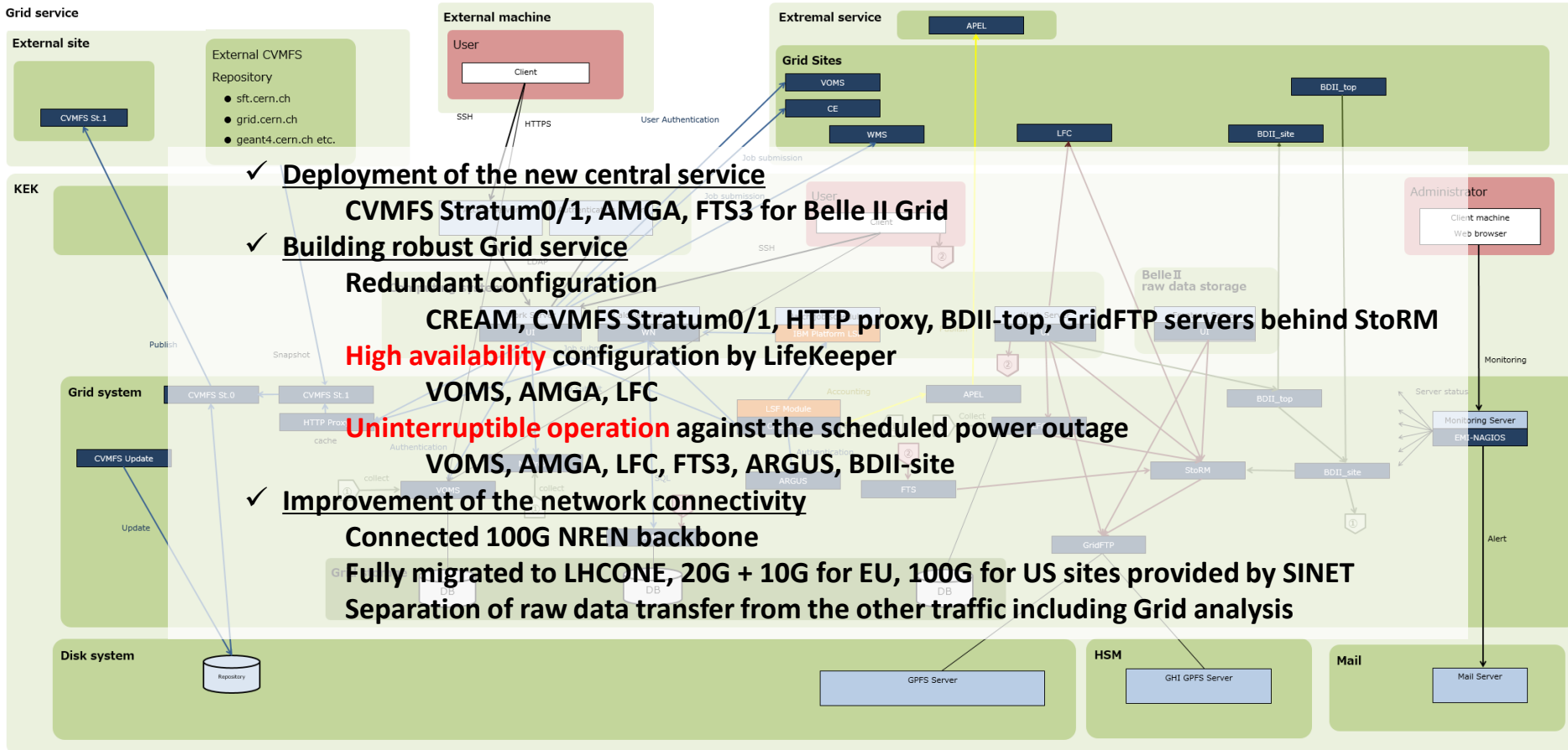
JOB scheduler : Platform LSF v9

K. Murakami (CHEP2016)

Upgrade of Grid system

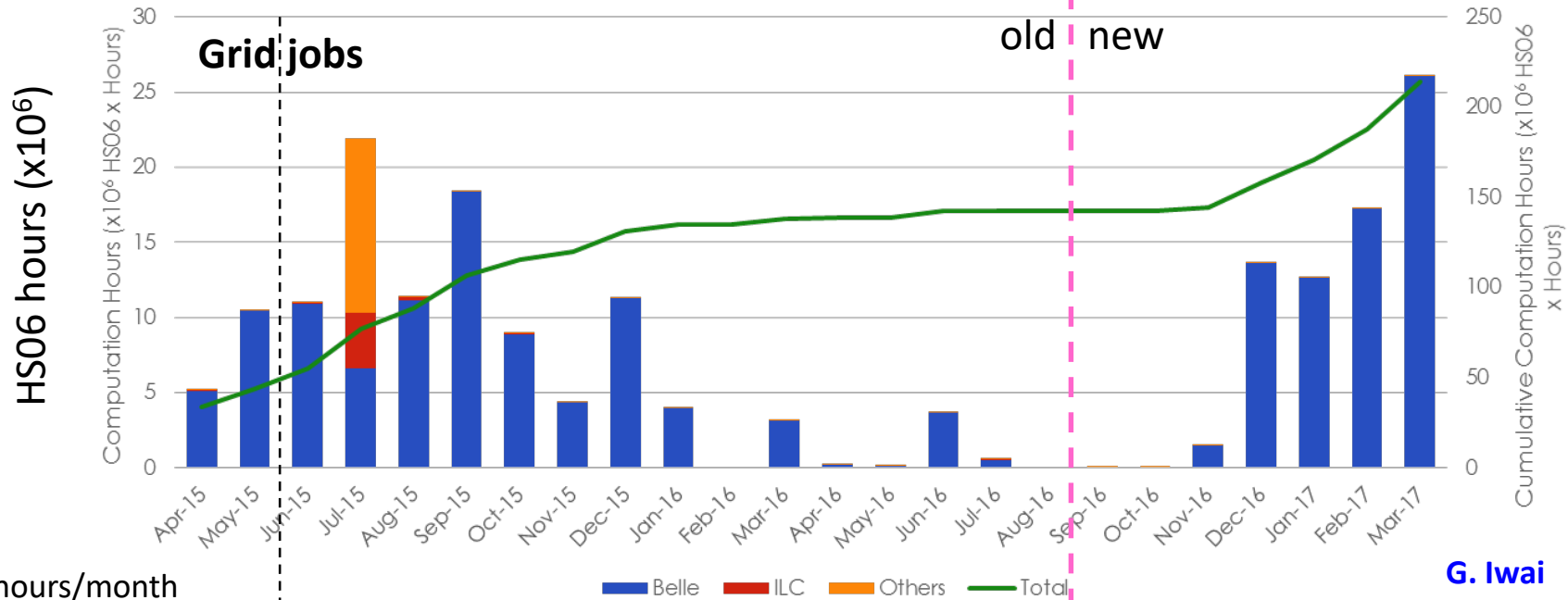


System has already been in production mode since the last September.



T. Nakamura (CHEP2016)

CPU consumption at new KEKCC



Old system:

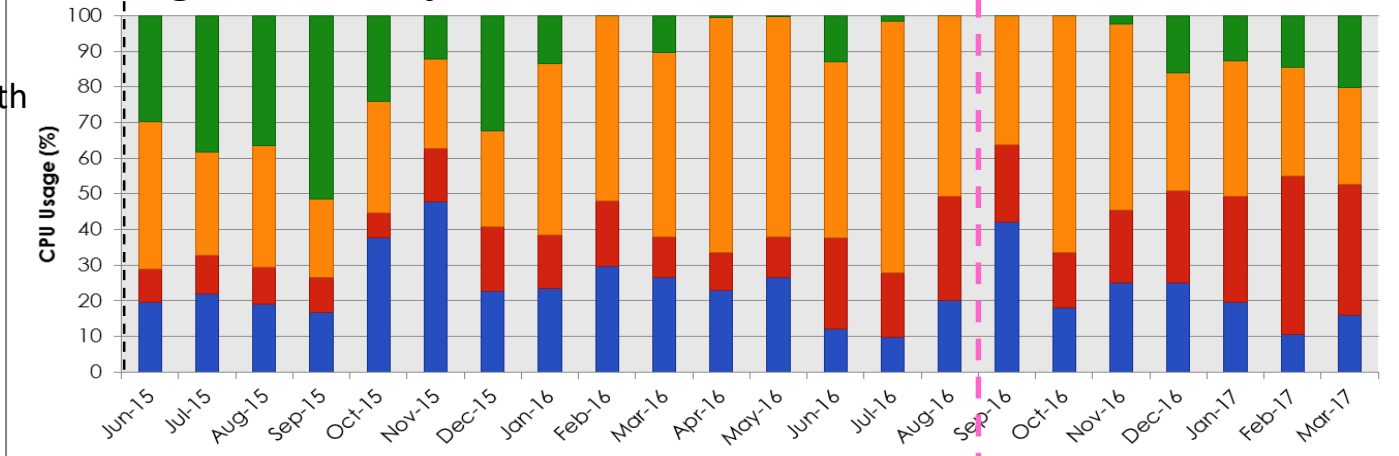
~40M HS06 hours/month
(14.7 HS06/core)

New System:

~168M HS06 hour /month
(23.5 HS06/core)

CPU cores: x2.5
HS06 score: x1.6
Performance : > x4.0

Including local batch jobs



G. Iwai

GHI performance after the system migration



HSM SYSTEM



3 GB/s (read / write)

HPSS/GHI servers



TS3500



DataDirect NETWORKS
DDN SFA 12K

GPFS (GHI) : 3PB
Total throughput : > 50 GB/s



54 drives

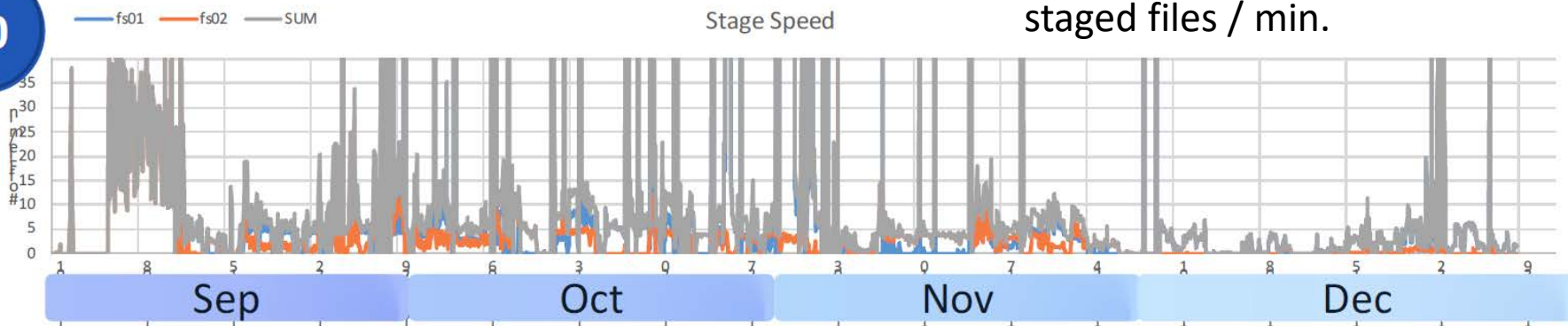
All data was not in the disk cache at the beginning of new system

Spikes indicate manual staging 10K files / min. But sometime, the staging performance was degraded.

Confliction between manual staging by administrator for high priority data and pileup of user request.

K. Murakami

40



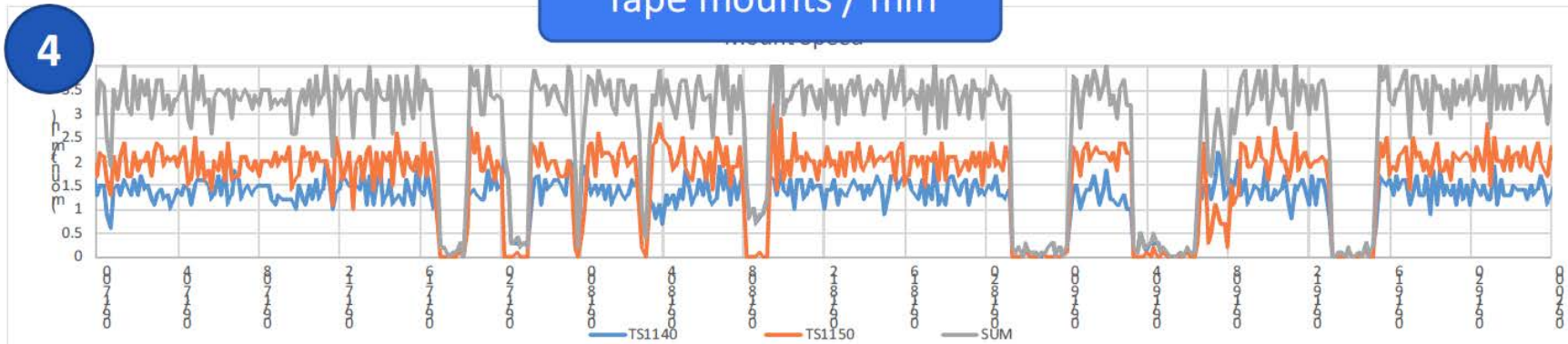
Data localization issue in tape



TS3500 Library accessor spec.: 15 sec/(un)mount = 4 tapes/min

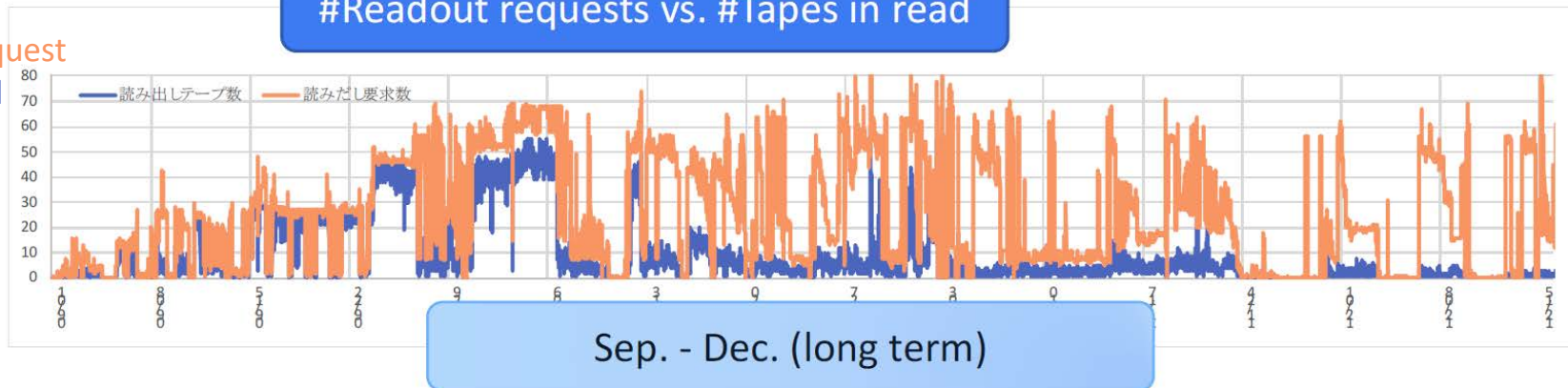
In worst case: File staging = ~4 files/min (continuous requests on different tape medias)

Tape mounts / min



#Readout requests vs. #Tapes in read

#Readout request
#Tape in read



It will be an issue again at the next system migration (3~4 years later.)

Solutions?:

Tape-by-tape queuing

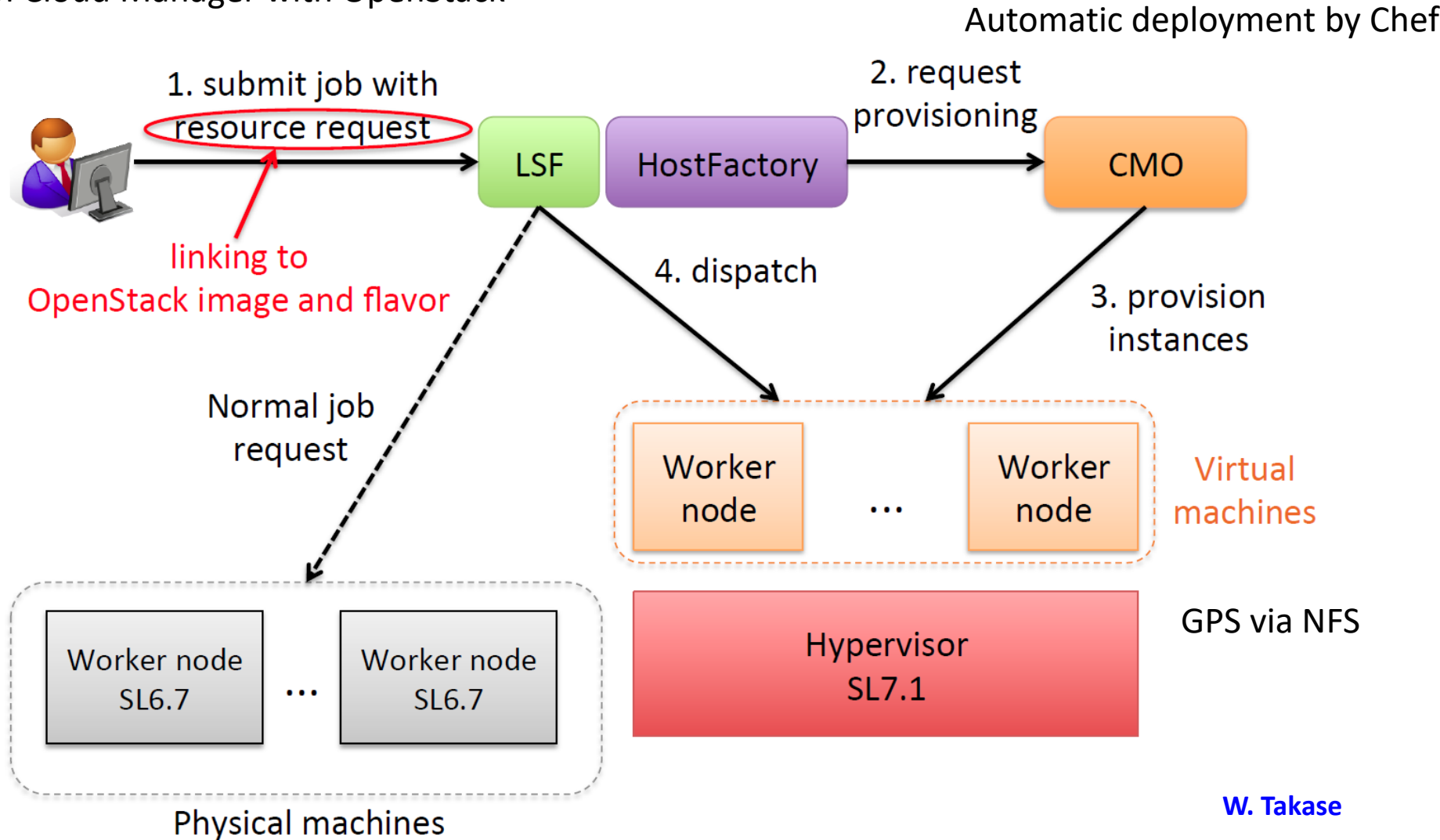
Periodical shuffling of queue by tape order in real-time

K. Murakami

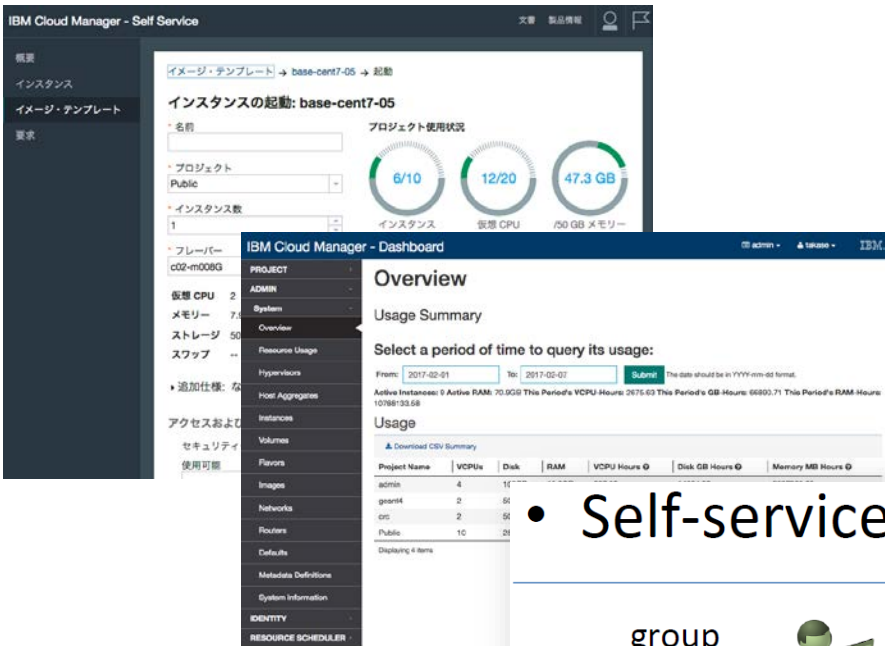
Private Cloud with OpenStack



IBM Cloud Manager with OpenStack



Schedule of private cloud deployment



(2017)

Feb.:

Mar. - May.:

Jun. - Jul.:

~Aug.:

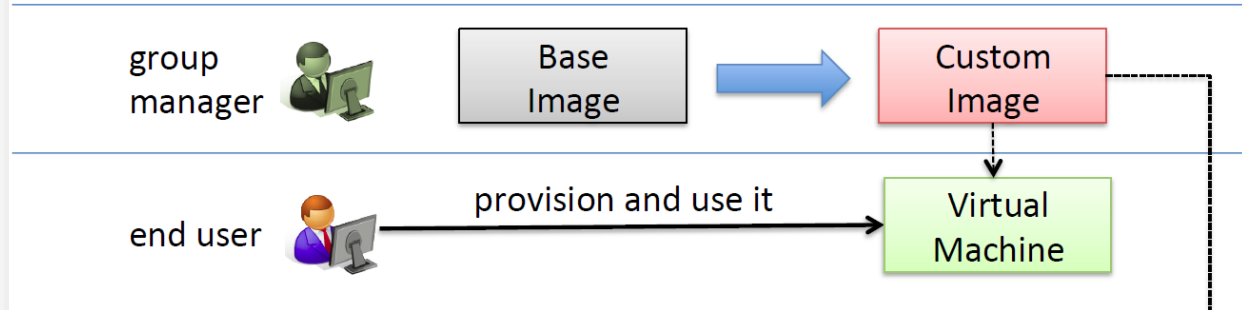
Evaluation

Quality Assurance

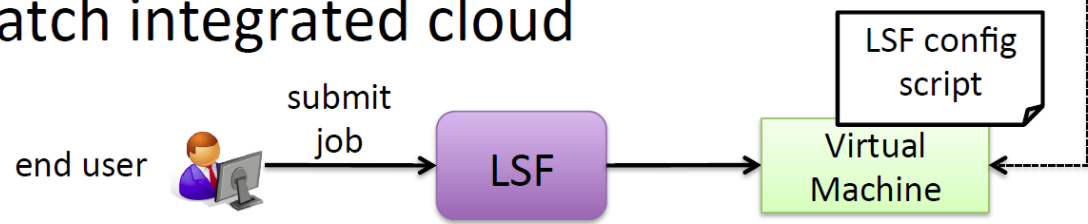
Define workflow

Start testing by group manager and end user

• Self-service



• Batch integrated cloud



W. Takase

Summary



The Belle II experiment will start the physics data taking from the beginning of 2018. This yields heavy load to the KEKCC and Grid service for the massive data production.

Many activities on the system development are ongoing for the optimization of the tape archiving system, HSM, and also for the future private cloud infrastructure at new the KEKCC.

Today, I could not introduce everything due to the time limitation. We expect that many things will be presented at the next HEPiX meeting at KEK.

We are looking forward to seeing you at the next HEPiX fall 2017.