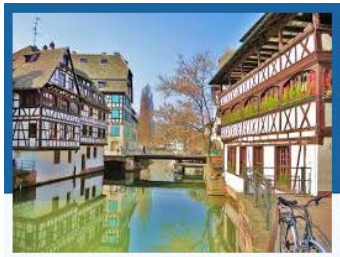


2017 Joint Workshop of the France-Korea (FKPPL) and France-Japan (TYL/FJPPL) Particle Physics Laboratories

10-12 mai 2017
IPHC Strasbourg (France)
Europe/Paris timezone

TYL-FJPPL: COMP_03
Computing platforms for future experiments
- status and perspectives -

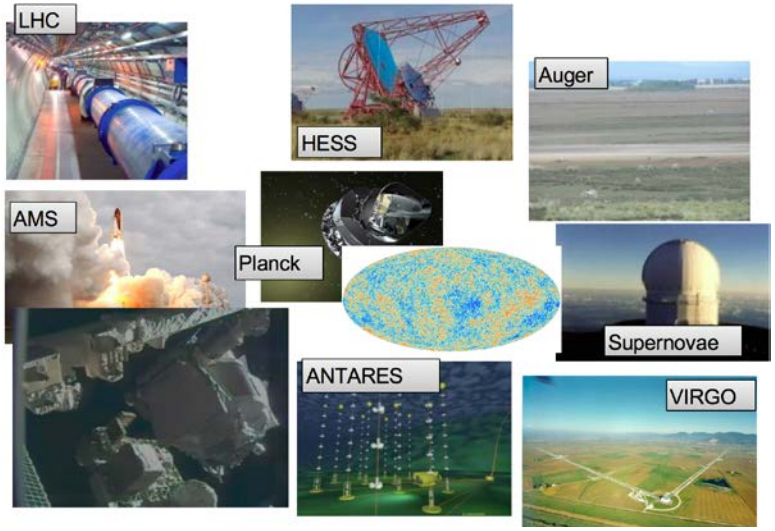
Tomoaki Nakamura
on behalf of

CC-IN2P3: F. Hernandez*, G. Rahal, L. Caillat-Vallet, M. Puel, V. Hamar, P-E. Macchi, **CENBG**: S. Incerti
KEK-CRC: T. Nakamura*, G. Iwai, H. Matsunaga, K. Murakami, T. Sasaki, S. Suzuki, W. Takase, Y. Watase
* Leader

Computing platforms for future experiments

Supporting a lot of ongoing/future experiments.

CC-IN2P3

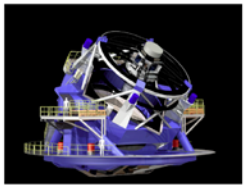


LSST

Whole dataset available at CC-IN2P3

50% of the processing by CC-IN2P3

other 50% by NCSA



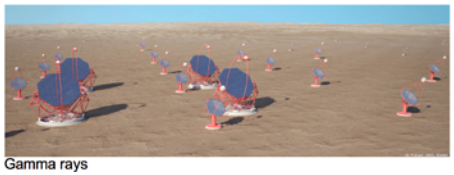
EUCLID

CC-IN2P3 is the French Data Center for processsing and data management

dark energy and dark matter

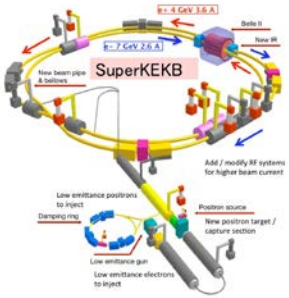
CTA

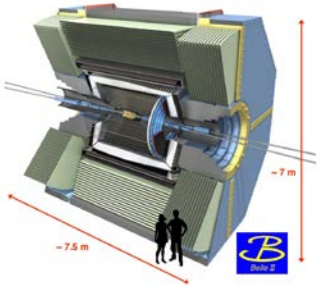
CC-IN2P3 should play a key role in the CTA data processing



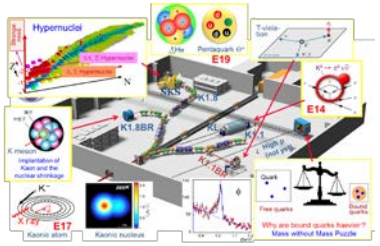
Gamma rays

KEK-CRC





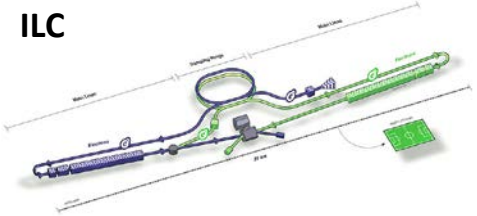
SuperKEKB/Belle II

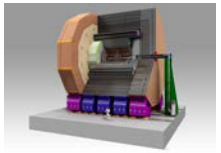


J-PARC

T2K

ILC





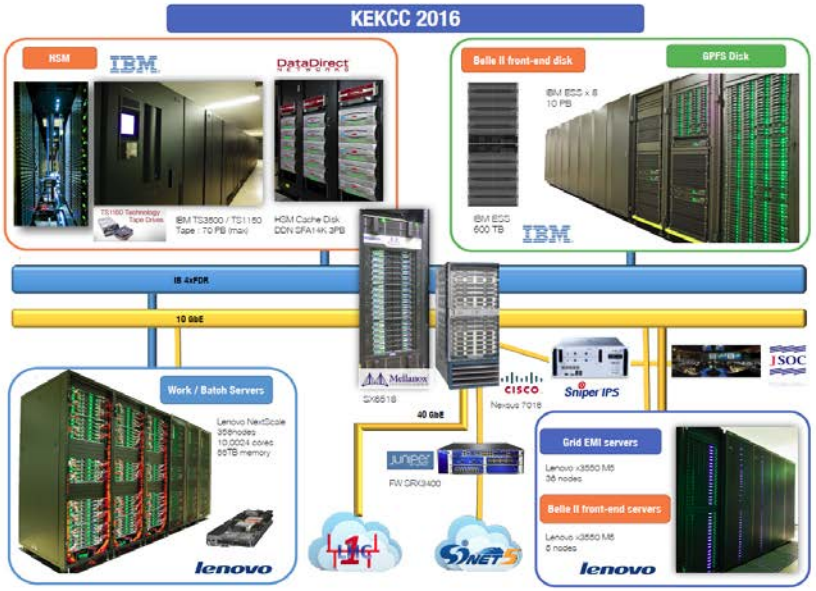
Computing infrastructure

- CC-IN2P3
  - System upgrade and resource enhancement are made gradually, year-by-year.
- KEK-CRC
  - System duration is 4~5 years.
  - Replace whole HW components at once.
- Exchange of experience is quite essential!
  - Type of CPUs, Storage system, Tape archiving system.

Upgrade of CC-IN2P3 in 2017

- HTC farm
  - 257 700 HS06 => 323 700 HS06 (***+25.6%***)*
  - Introduce CentOS7 (current is SL6)
  - Assessment of container use in computing (docker, shifter, singularity ?)
- Tape storage – HPSS
  - 44PB => 61PB (***+38.6%***)
  - DropT10K-C (5TB tape cartridge)
- Backup – IBM Spectrum Protect
  - Drop LTO4 tape cartridges and drives
- Disk storage – DCACHE, XROOTD, IRODS
  - 18,7 PB => 21,5 PB (***+15%***)
- Prepare to replace AFS
  - A part of the data to GPFS space
  - A part of the data to ISILON space (NFSv4 access)
- Network
  - Backbone evolution to prepare the arrival of 100Gbps external links
- Virtualization On Demand (aka cloud openstack)
  - See Mattieu PUEL's talk
- Many studies,
  - Container orchestration, CEPH, storage front-end and transfer gateway (LSST use case)

The new KEKCC has already been in production mode since the last September.



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

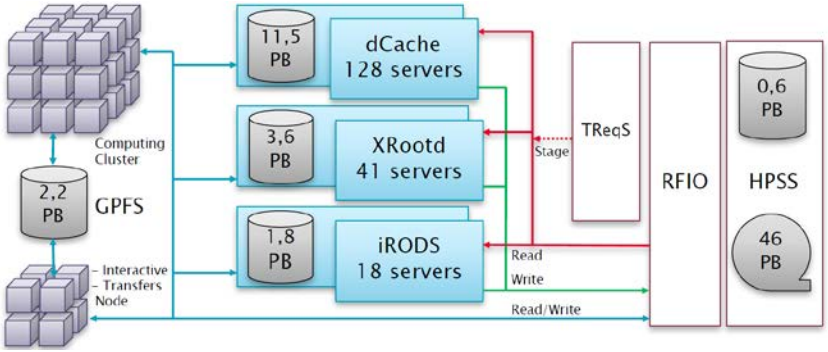
Tape : 70 PB (max cap.)
HSM data : 8.5 PB data, 170 M files, 5,000 tapes

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

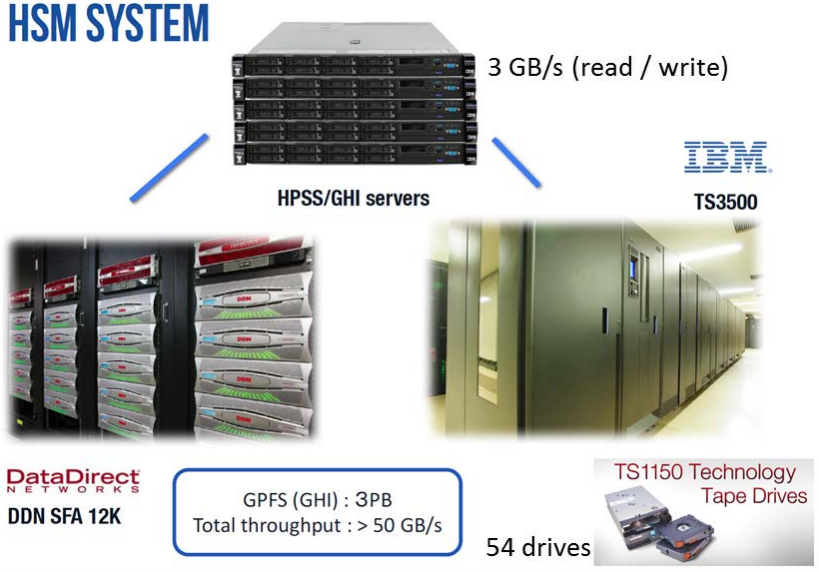
JOB scheduler : Platfrom LSF v9

Effective use of storage system

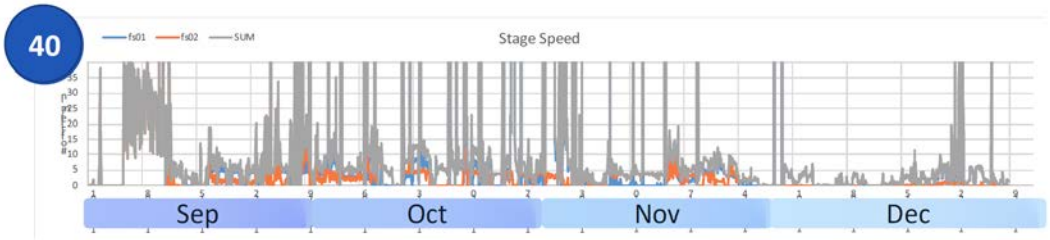
CC-IN2P3



KEK-CRC

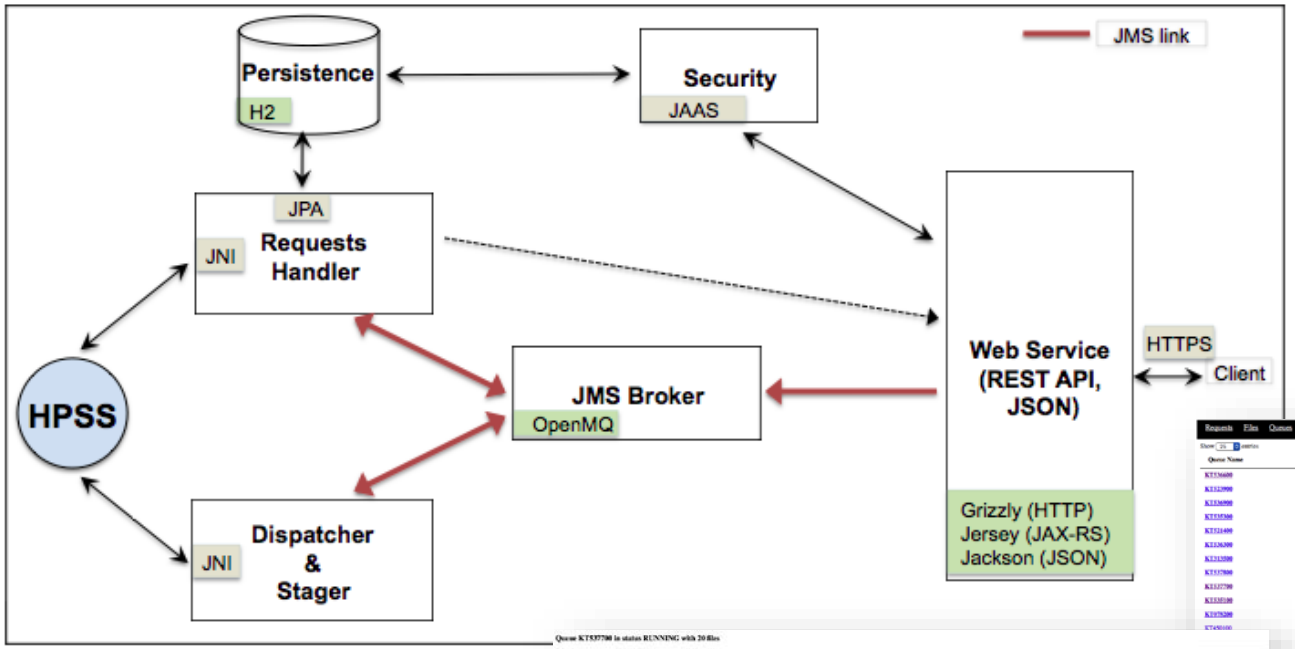


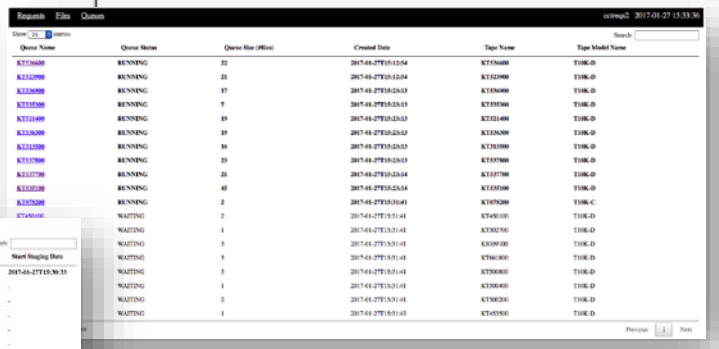
All data was not in the disk cache at the beginning of new system of KEK-CRC.
Spikes indicate manual staging (10K files/min.) But sometime, the staging performance was degraded.

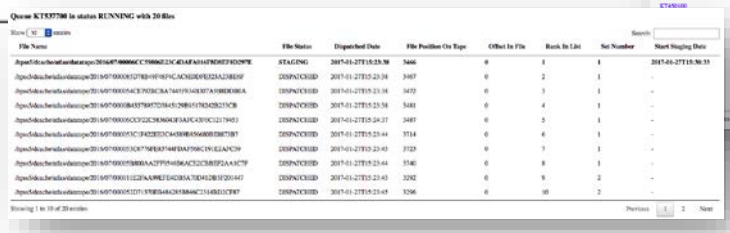


Tape request scheduling, TReqS

TReqS at CC-IN2P3







HPSS
High Performance Storage System

CC-IN2P3 and KEK-CRC has the same hieratical tape archive system (HPSS).

- Confliction between manual staging by administrator for high priority data and pileup of user request.
- It will be an issue again at the next system migration (3~4 years later) at KEK-CRC.

Solutions?: e.g. Tape-by-tape queuing and/or Periodical shuffling of queue by tape order in real-time

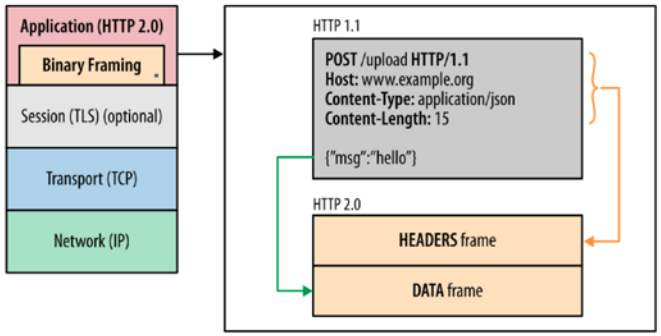
- CC-IN2P3 already has one of the solutions: Tape Request Scheduler (TReqS).
- It is quite attractive, KEK-CRC started the study of its applicability.

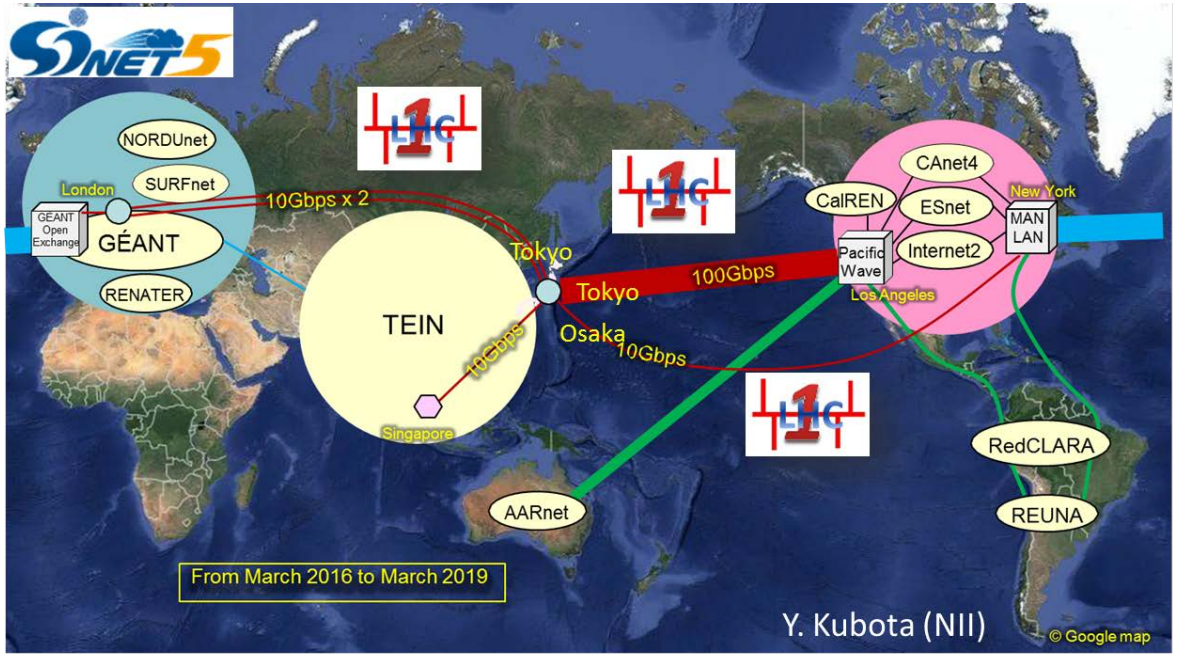
Long range network monitoring

- Connectivity of the international network is one of the most important component for the data and resource sharing.
- The bandwidth has been increased beyond the 10 Gbps order.

Is it sufficient? The answer is No.

- Need to optimize many parameters.
- CC-IN2P3 and KEK-CRC has already started the performance monitoring for the long range network by using 10G perfSONAR.
- CC-IN2P3 developed the new tools to check the applicability of the standard protocols i.e. HTTP and TLS instead of the conventional GridFTP.
- We will start the performance measurement for the long distance.





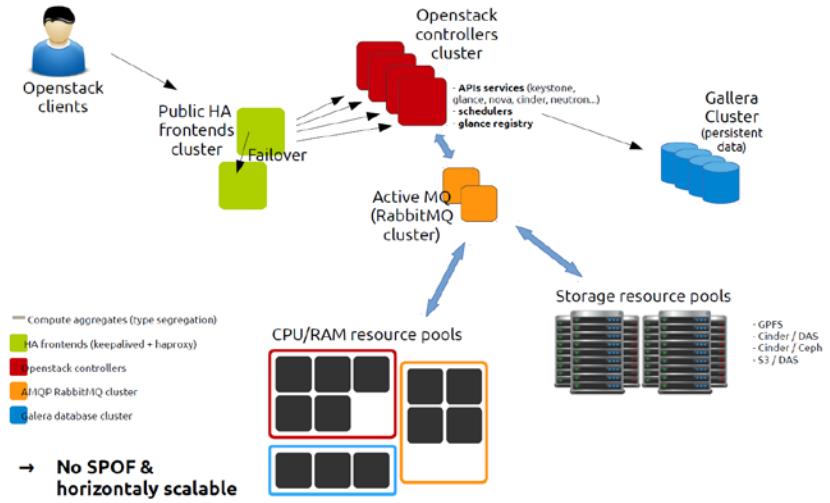
perfSONAR

Measurement between CC-IN2P3 and KEK-CRC with 10 Gbps NIC

Private cloud deployment with OpenStack

OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter. It consists of many open source software and tools.

CC-IN2P3



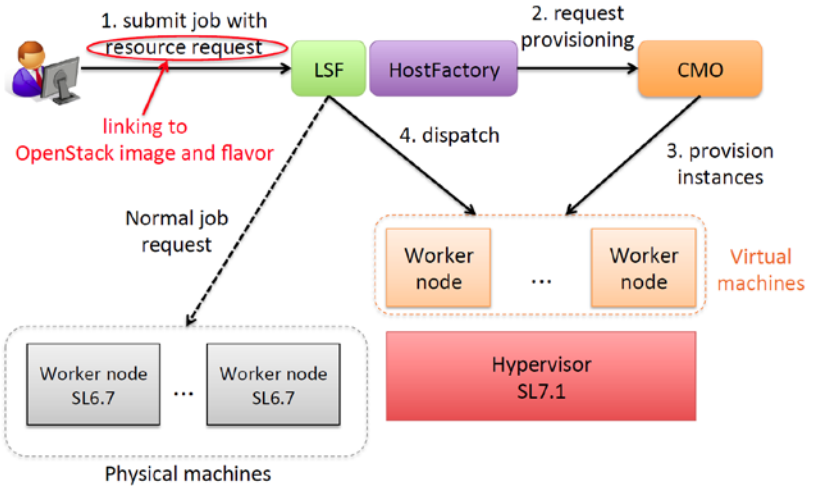
Compute clusters (may '16) :

	Cores	RAM	Storage
HA	544	3.6 TB	36 TB
R&D	288	1.2 TB	24 TB
Computing	1216	5.2 TB	40 TB

→ 2048 cores in total, 10TB RAM, 100 TB storage

KEK-CRC

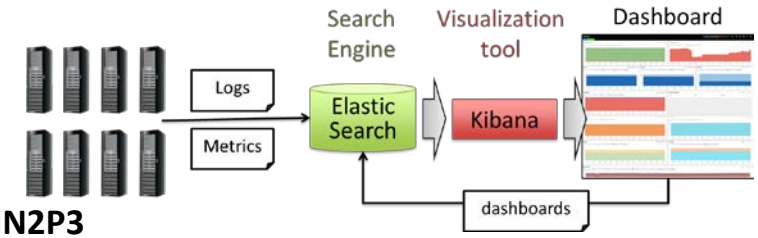
IBM Cloud Manager with OpenStack



Deployment schedule in 2017

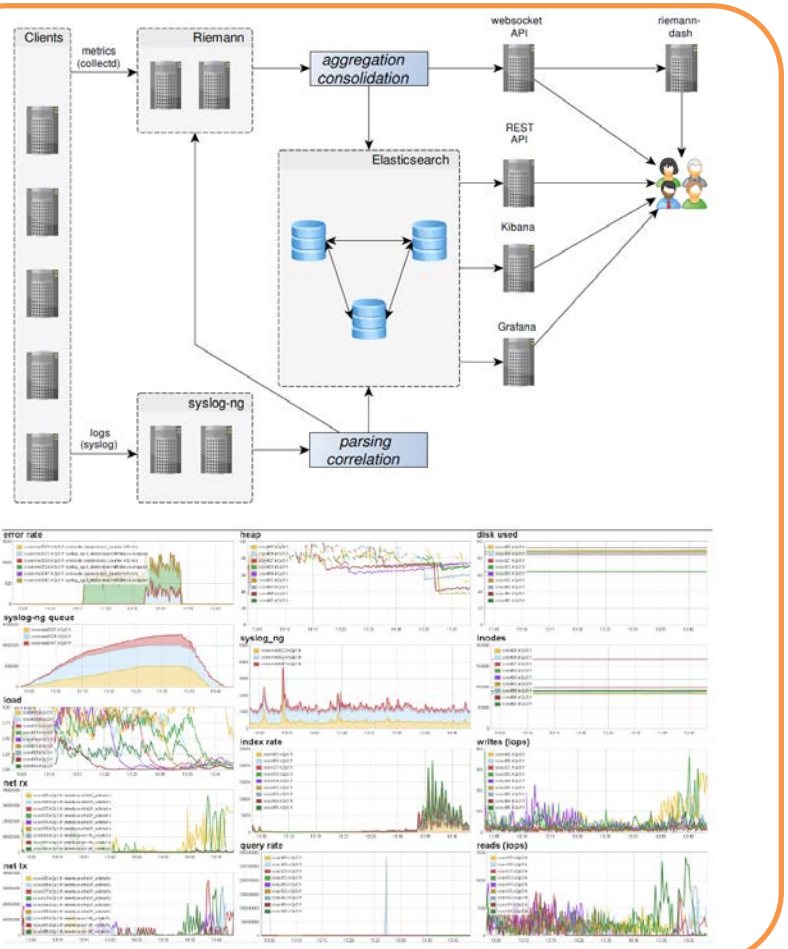
Feb.:	Evaluation
Mar. - May.:	Quality Assurance
Jun. - Jul.:	Define workflow
~Aug.:	Start testing by group manager and end user

System monitoring, Kibana and Elasticsearch

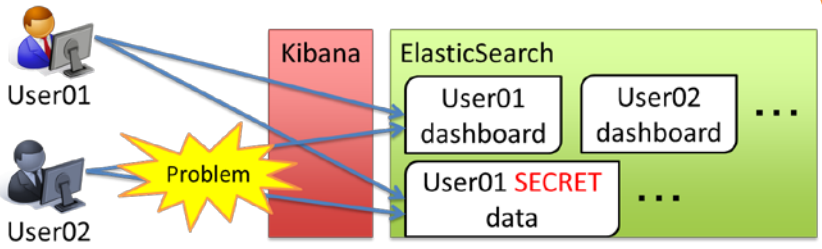


Combination of Kibana and Elasticsearch became the standard solution for system monitoring.

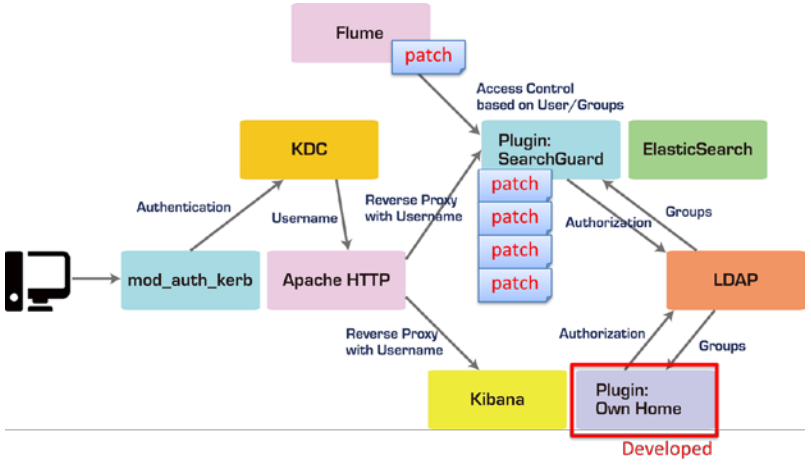
Developed at CC-IN2P3



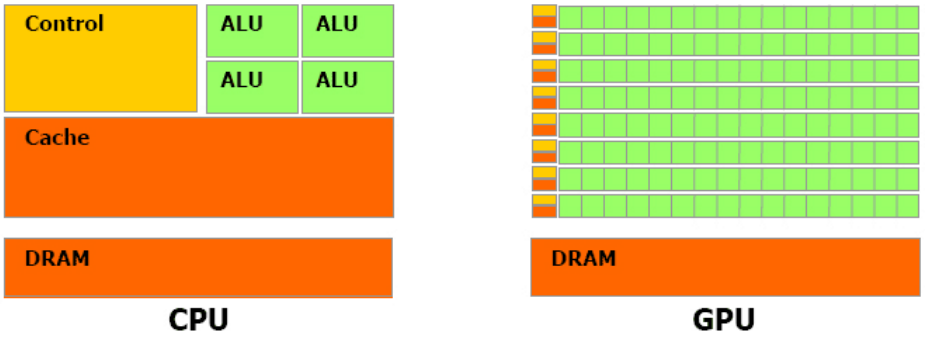
Developed at KEK-CRC



Authentication module for each users/groups



Study on GPU cluster, usability and performance

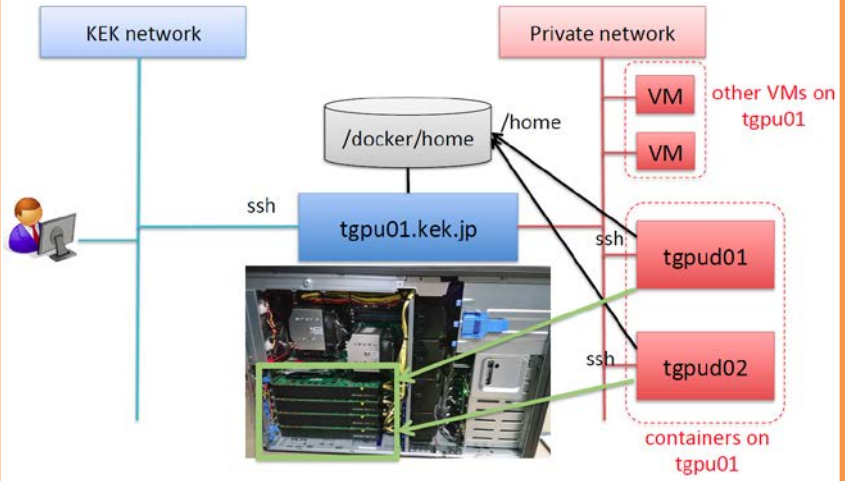


The performance of individual CPU core is saturated (clock frequency is not increased). Effective use of GPU is one of the candidates to increase the computing performance per cost for the next generation experiments.

Service in production at CC-IN2P3

- Servers:
  - 10 Dell C4130
    - 2 Xeon E5-2640v3 @2.6 Ghz (8 cores)
    - 128 GB RAM
    - SSD
    - 1 Gb/s NIC, 10 Gb/s projected
    - 2 Nvidia Tesla K80
      - ➡ 4 GPUs Nvidia GK210 per node, 12 GB DDR5 each
  - Cluster totaling 40 GPUs
- Interconnect:
  - InfiniBand QDR

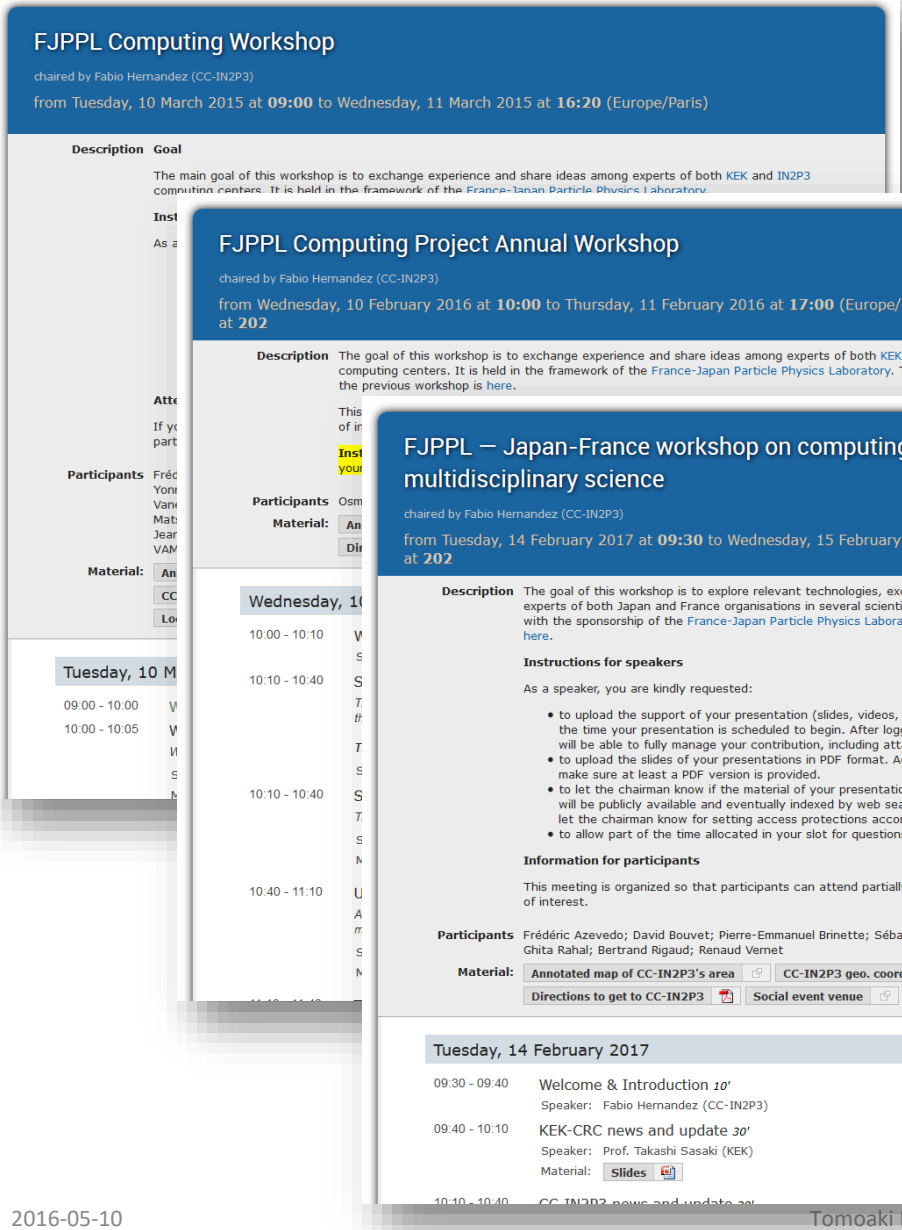
Under development at KEK-CRC



Joint with TYL-FJPPL: APP_01

Development of application software.

Annual meetings



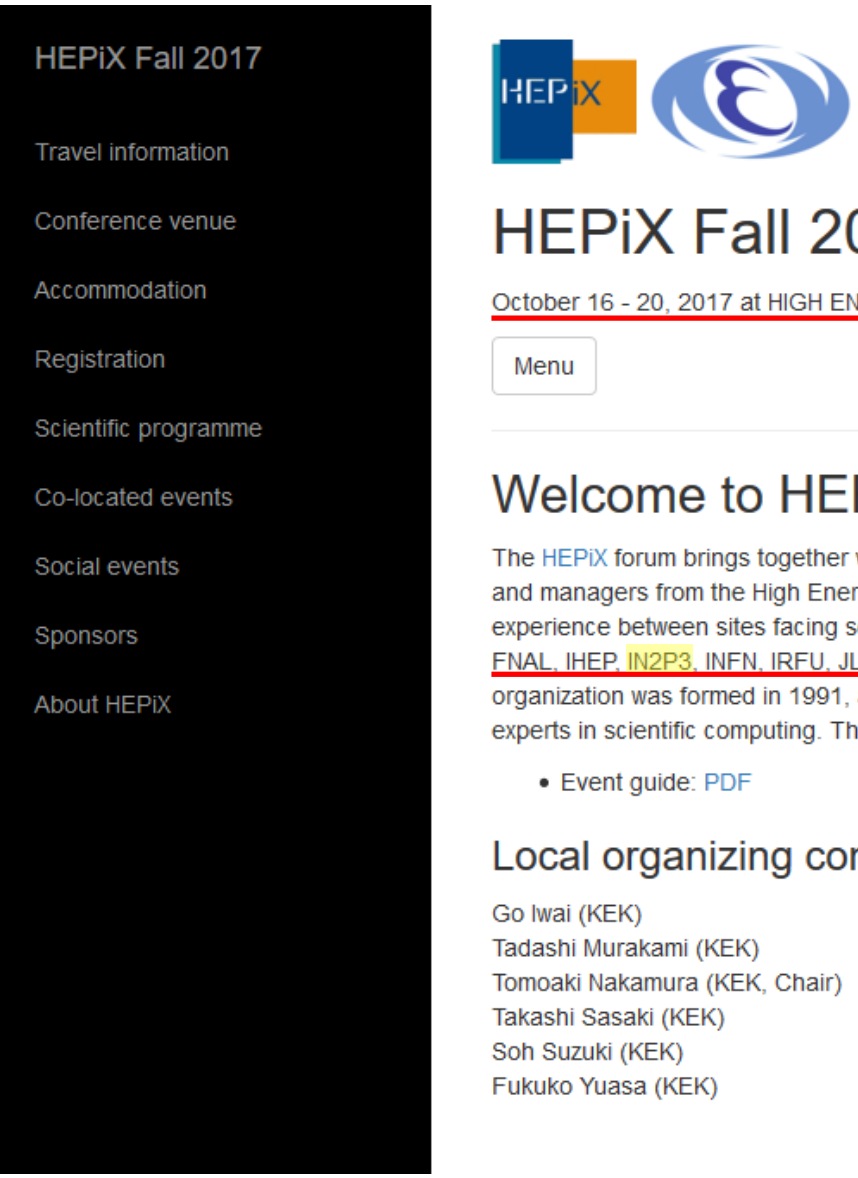
2015: 23 talks
https://indico.in2p3.fr/event/11289/

2016: 12 talks + hands on + discussion
https://indico.in2p3.fr/event/12701/

2017: 18 talks + discussion
https://indico.in2p3.fr/event/14157/

- Information exchange.
- Share experience.
- Common interest for both centers.
- Determine the specific subjects to collaborate.

HEPiX Fall 2017 Workshop



- Major conference among the IT division in HEP laboratories.
- Next conference will be held at KEK in Oct. 2017.
- Report progress covered by this collaborations as much as possible.

Summary

A lot of items on the novel computing technology for the current/future experiments are actively developed and deployed at both computing centers, CC-IN2P3 and KEK-CRC.

We work together on the common subjects and refer to each other by the technical interchange in the framework of TYL-FJPPL COMP_03 project.

Continuous collaboration is necessary for the further challenges!