

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

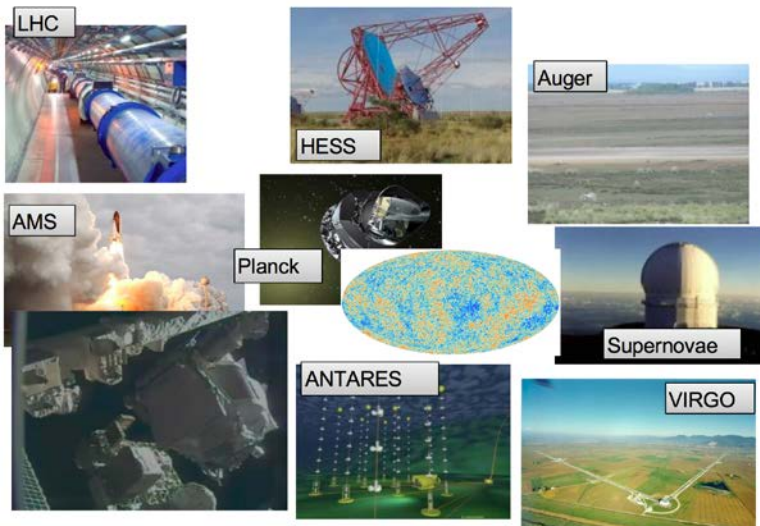
CC-IN2P3: [F. Hernandez](#)^{*}, G. Rahal, L. Caillat-Vallet, M. Puel, V. Hamar, P-E. Macchi, **CENBG:** S. Incerti
KEK-CRC: [T. Nakamura](#)^{*}, G. Iwai, H. Matsunaga, K. Murakami, T. Sasaki, S. Suzuki, W. Takase, Y. Watase
^{*} [Leader](#)



CC-IN2P3

Supporting a lot of ongoing/future experiments.

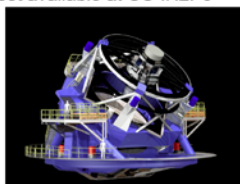
KEK-CRC



LSST

Whole dataset available at CC-IN2P3

50% of the processing by CC-IN2P3
other 50% by NCSA



EUCLID

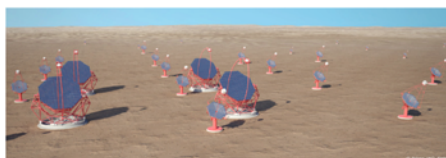
CC-IN2P3 is the French Data Center for processing and data management



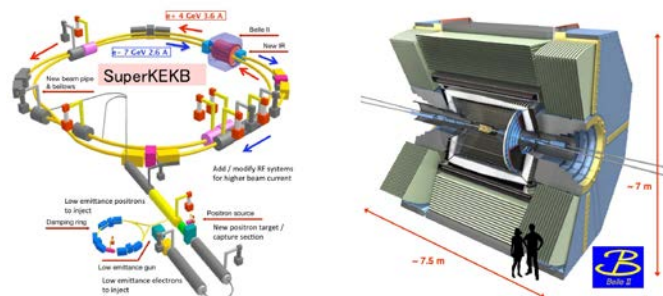
dark energy and dark matter

CTA

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



Gamma rays



SuperKEKB/Belle II

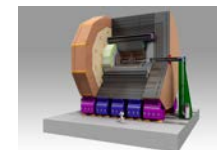
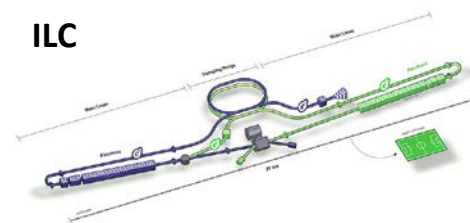


J-PARC



T2K

ILC



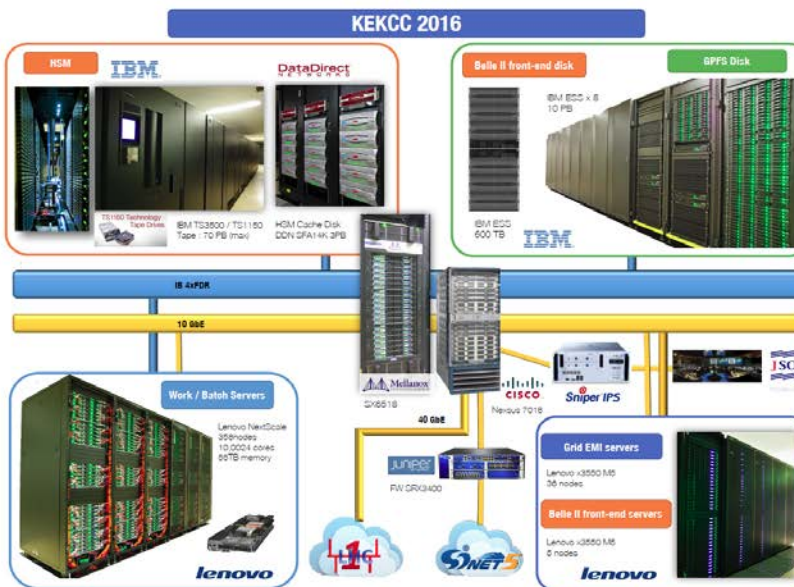


- **CC-IN2P3**
System upgrade and resource enhancement are made gradually, year-by-year.
- **KEK-CRC**
System duration is 4~5 years.
Replace whole HW components at once.
- **Exchange of experience is quite essential!**
Type of CPUs, Storage system, Tape archiving system.

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.

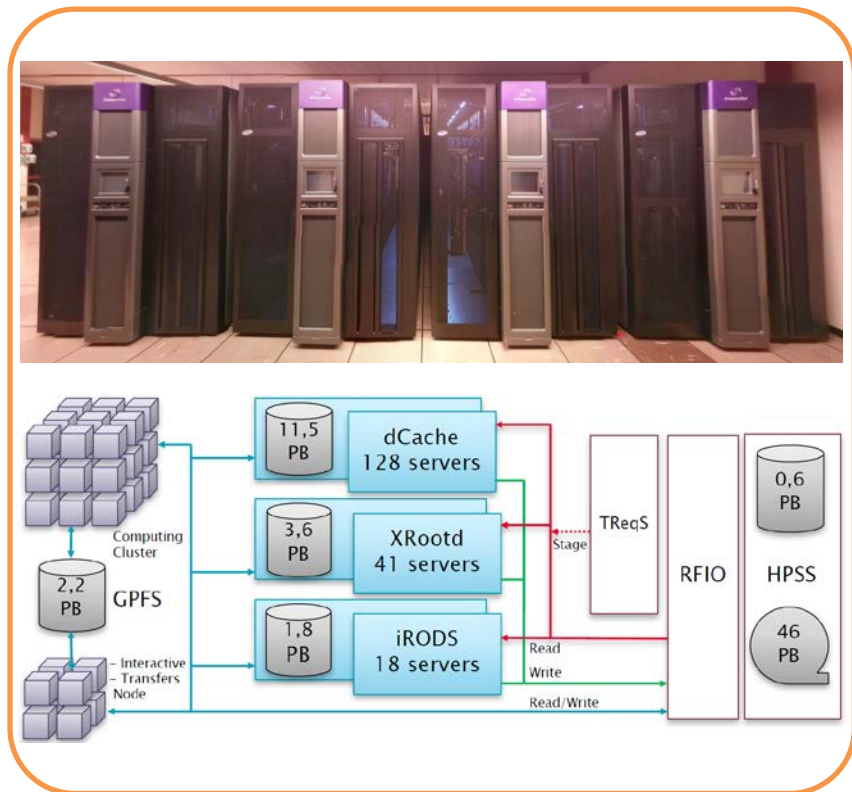


SYSTEM RESOURCES

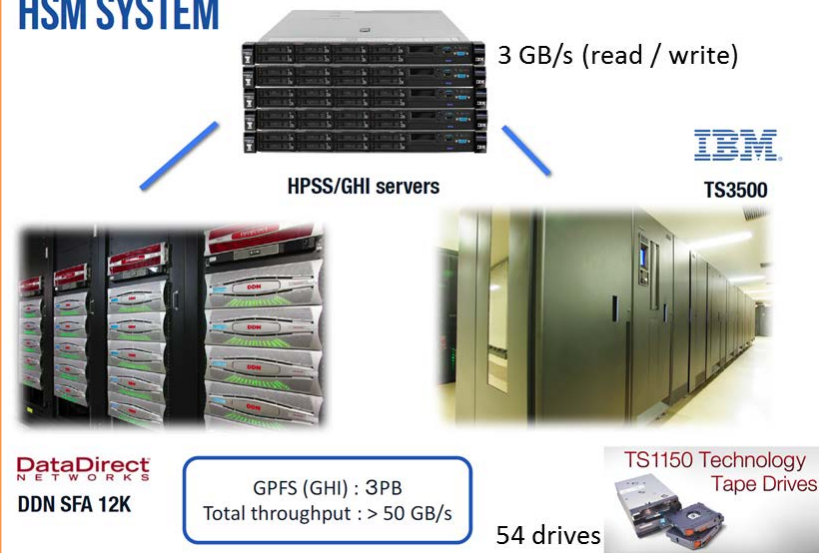
- CPU** : 10,024 cores
 - Intel Xeon E5-2697v3 (2.6GHz, 14cores) x 2 358 nodes
 - 4GB/core (8,000 cores) / 8GB/core (2,000 cores) (for app. use)
 - 236 kHS06 / site
- Disk** : 10PB (GPFS) + 3PB (HSM cache)
- Interconnect** : IB 4xPDR
- Tape** : 70 PB (max cap.)
- HSM data** : 8.5 PB data, 170 M files, 5,000 tapes
- Total throughput** : 100 GB/s (Disk, GPFS), 50 GB/s (HSM, GHI)
- JOB scheduler** : Platform LSF v9

CC-IN2P3

KEK-CRC

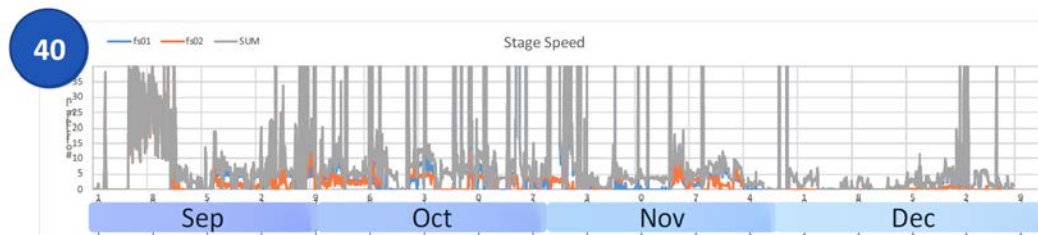


HSM SYSTEM

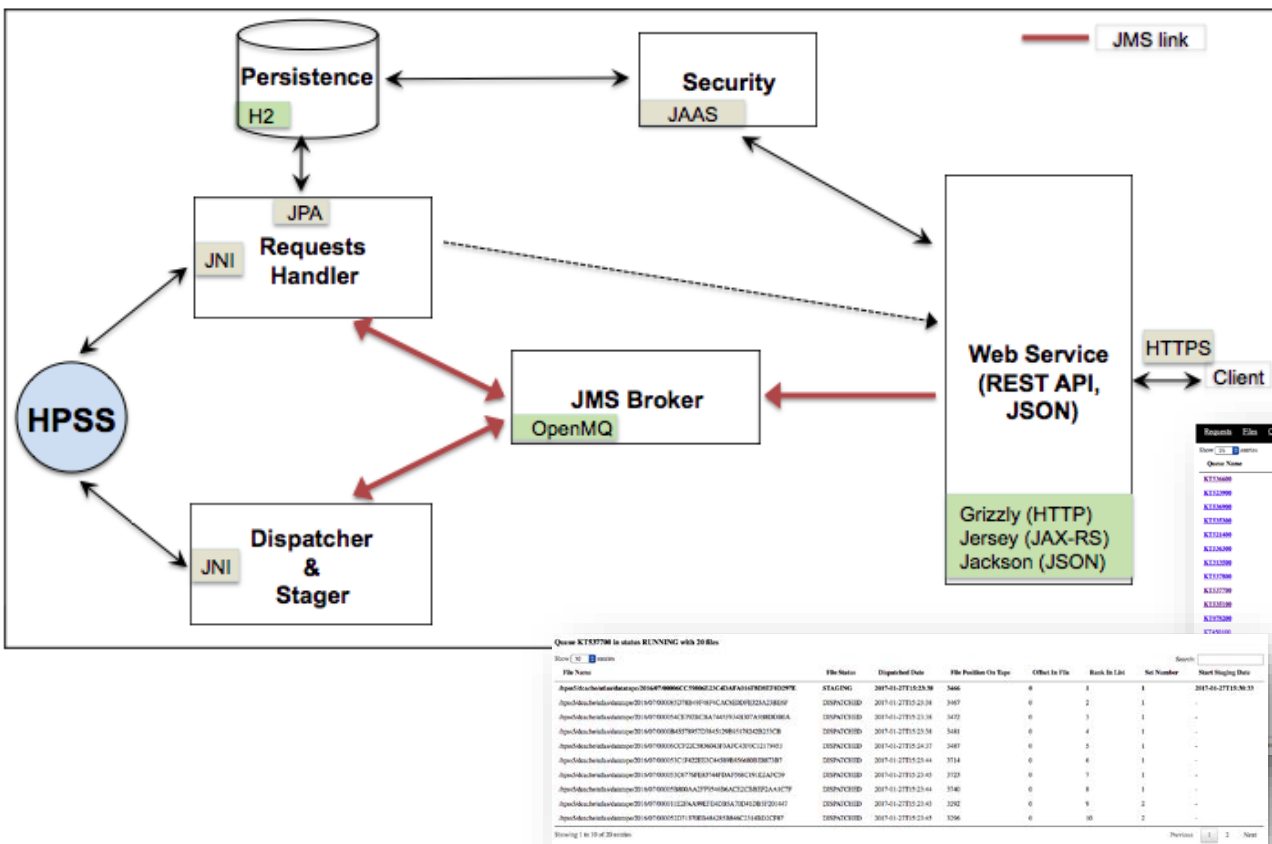


All data was not in the disk cache at the beginning of new system of KEK-CRC.

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



TReqS at CC-IN2P3

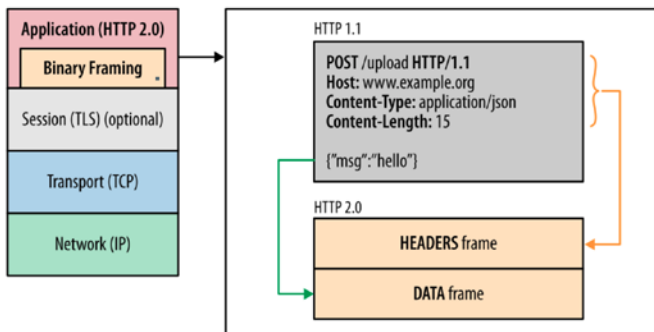
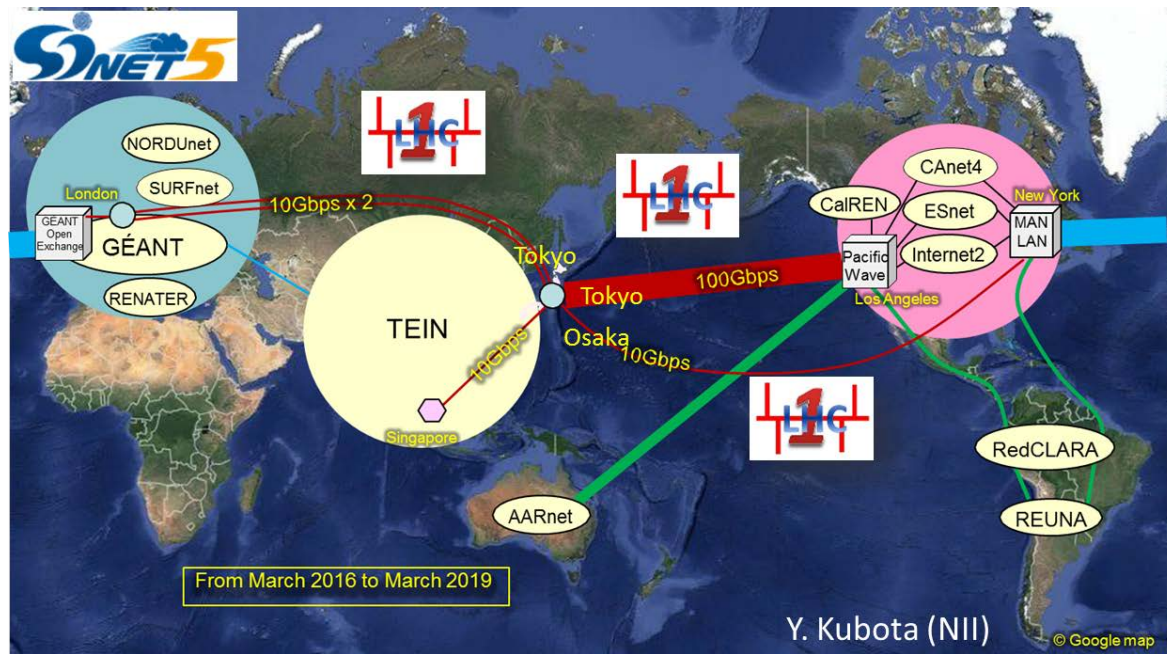


CC-IN2P3 and KEK-CRC has the same hieratical tape archive system (HPSS).

Queue Name	Queue Status	Queue Size (Files)	Created Date	Type Name	Tape Model Name
KT11000	RENNING	22	2017-01-27 15:14:54	KT11000	TIK-D
KT11000	RENNING	23	2017-01-27 15:14:54	KT11000	TIK-D
KT11000	RENNING	17	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	9	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	10	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	10	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	16	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	33	2017-01-27 15:14:53	KT11000	TIK-D
KT11000	RENNING	20	2017-01-27 15:14:54	KT11000	TIK-D
KT11000	RENNING	40	2017-01-27 15:14:54	KT11000	TIK-D
KT11000	RENNING	2	2017-01-27 15:14:41	KT11000	TIK-C
KT11000	WAITING	2	2017-01-27 15:14:41	KT11000	TIK-D
KT11000	WAITING	1	2017-01-27 15:14:41	KT11000	TIK-D
KT11000	WAITING	3	2017-01-27 15:14:41	KT11000	TIK-D
KT11000	WAITING	3	2017-01-27 15:14:41	KT11000	TIK-D
KT11000	WAITING	1	2017-01-27 15:14:41	KT11000	TIK-D
KT11000	WAITING	2	2017-01-27 15:14:41	KT11000	TIK-D

- 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.

- 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

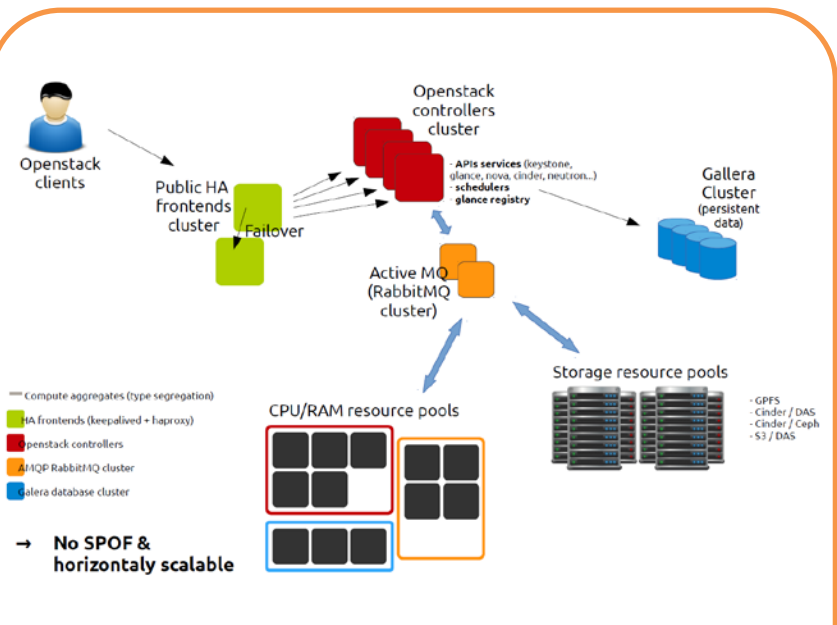




OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter. It consists of many open source software and tools.

CC-IN2P3

KEK-CRC

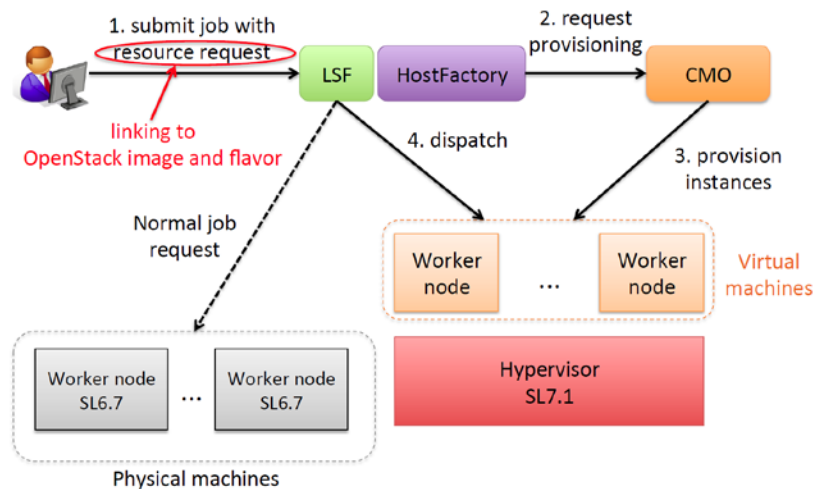


Compute clusters (may '16) :

	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

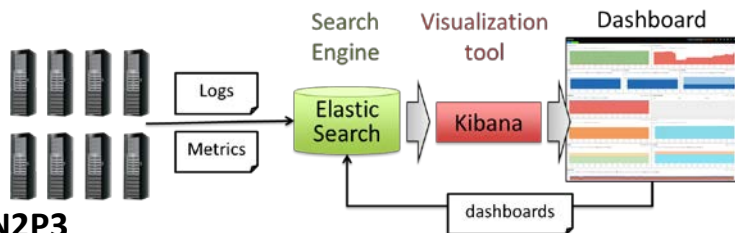
→ 2048 cores in total, 10TB RAM, 100 TB storage

IBM Cloud Manager with OpenStack



Deployment schedule in 2017

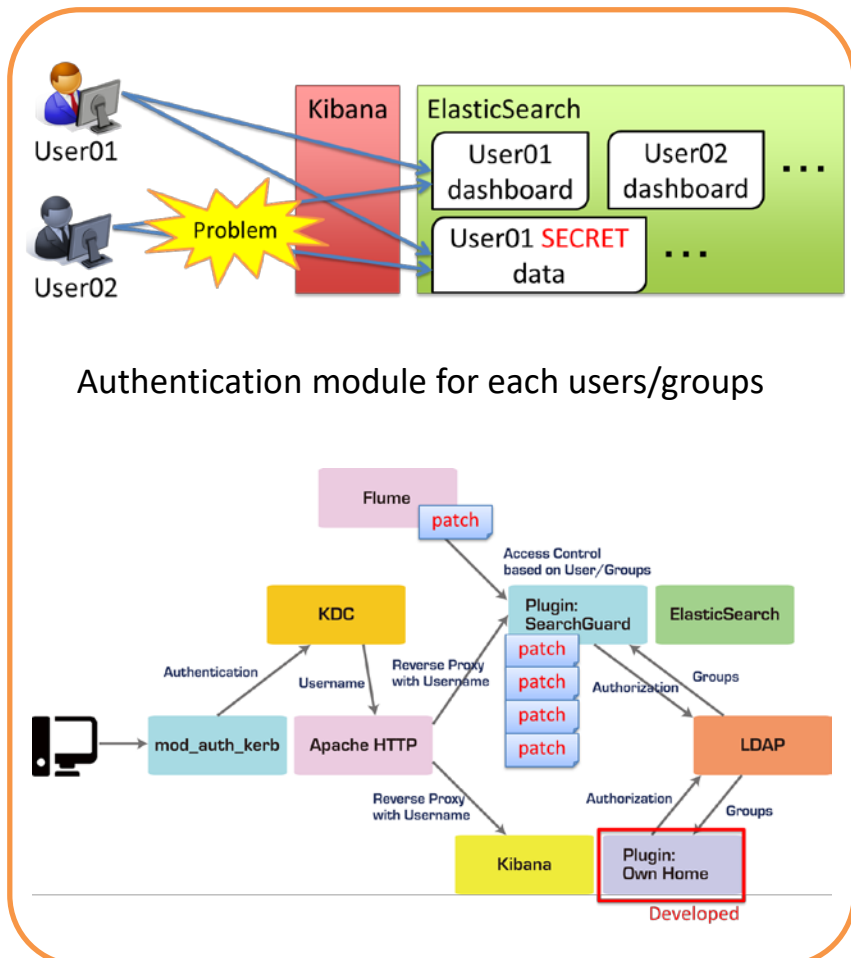
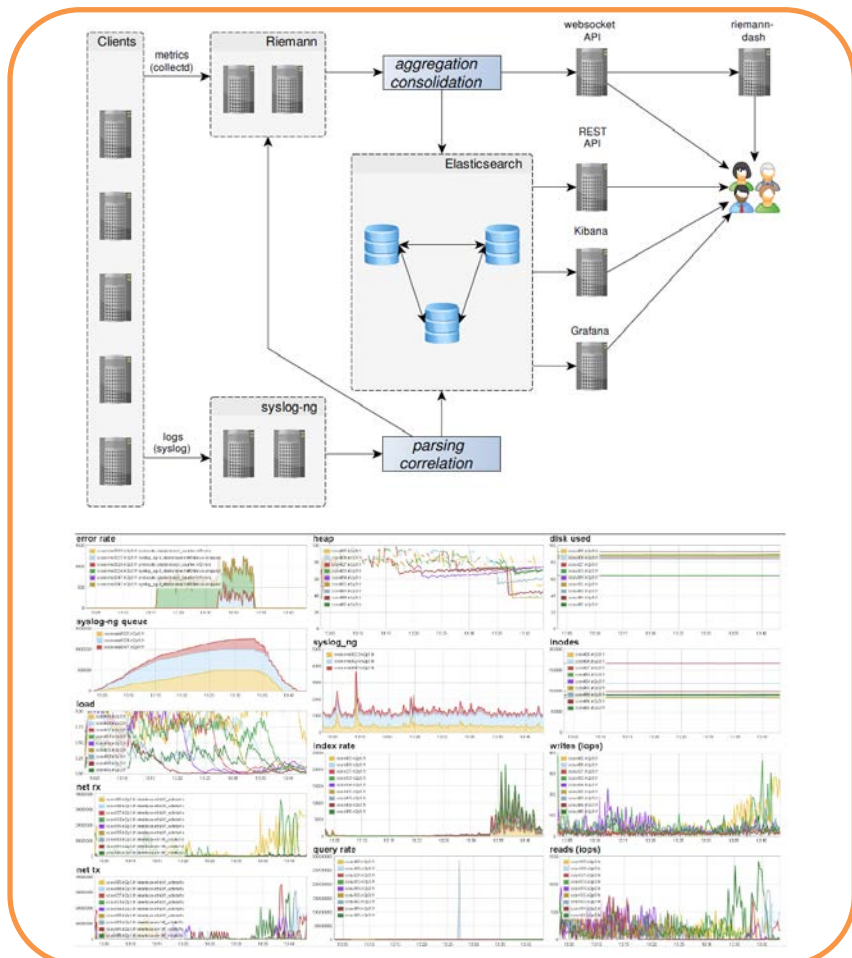
Feb.:	Evaluation
Mar. - May.:	Quality Assurance
Jun. - Jul.:	Define workflow
~Aug.:	Start testing by group manager and end user



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

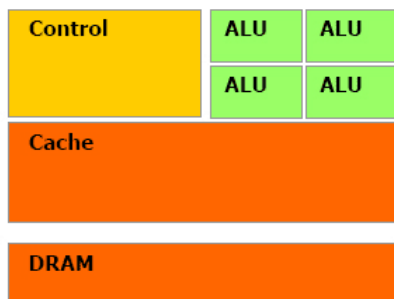
Developed at CC-IN2P3

Developed at KEK-CRC

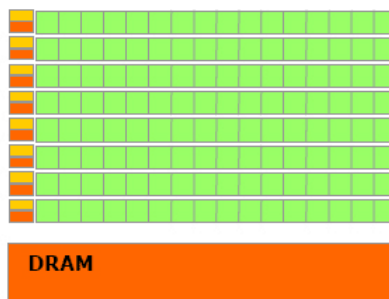


Authentication module for each users/groups

Developed



CPU



GPU

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:

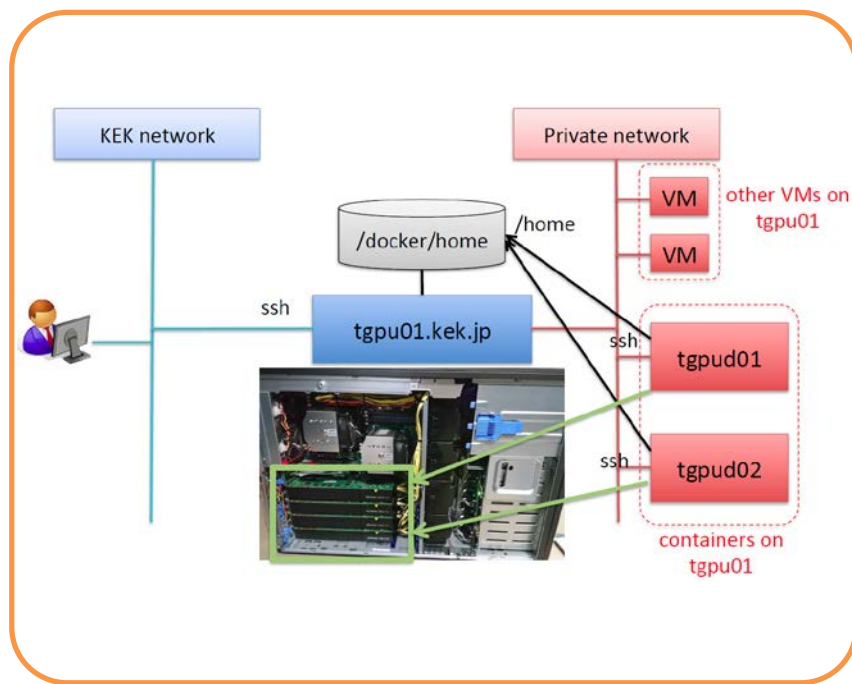
- 10 Dell C4130
 - 2 Xeon E5-2640v3 @2.6 Ghz (8 cores)
 - 128 GB RAM
 - SSD
 - 1 Gb/s NIC, 10 Gb/s projected
 - 2 Nvidia Tesla K80
 - → 4 GPUs Nvidia GK210 per node, 12 GB DDR5 each
- Cluster totaling 40 GPUs



Interconnect:

- InfiniBand QDR

Under development at KEK-CRC



Joint with TYL-FJPPL: APP_01
Development of application software.

FJPL Computing Workshop

chaired by Fabio Hernandez (CC-IN2P3)

from Tuesday, 10 March 2015 at 09:00 to Wednesday, 11 March 2015 at 16:20 (Europe/Paris)

Description Goal

The main goal of this workshop is to exchange experience and share ideas among experts of both KEK and IN2P3 computing centers. It is held in the framework of the France-Japan Particle Physics Laboratory.

Inst

As a

FJPL Computing Project Annual Workshop

chaired by Fabio Hernandez (CC-IN2P3)

from Wednesday, 10 February 2016 at 10:00 to Thursday, 11 February 2016 at 17:00 (Europe/Paris) at 202

Description The goal of this workshop is to exchange experience and share ideas among experts of both KEK and IN2P3 computing centers. It is held in the framework of the France-Japan Particle Physics Laboratory. The agenda of the previous workshop is here.

Atte

If yo

part

Participants

Fréd

Voni

Vani

Mat

Jear

VAM

Material:

An

CC

Lo

Wednesday, 10

10:00 - 10:10

10:10 - 10:40

10:10 - 10:40

10:10 - 10:40

10:40 - 11:10

Tuesday, 10 M

09:00 - 10:00

10:00 - 10:05

FJPL – Japan-France workshop on computing technologies for multidisciplinary science

chaired by Fabio Hernandez (CC-IN2P3)

from Tuesday, 14 February 2017 at 09:30 to Wednesday, 15 February 2017 at 22:30 (Europe/Paris) at 202

Description The goal of this workshop is to explore relevant technologies, exchange experience and share ideas among experts of both Japan and France organisations in several scientific domains. It is organized in the framework of and with the sponsorship of the France-Japan Particle Physics Laboratory. The agenda of the previous workshop is here.

Instructions for speakers

As a speaker, you are kindly requested:

- to upload the support of your presentation (slides, videos, documents, ...) on time, that is, not later than the time your presentation is scheduled to begin. After logging into Indico with your individual identifier, you will be able to fully manage your contribution, including attaching material to it.
- to upload the slides of your presentations in PDF format. Additional formats are also accepted but please make sure at least a PDF version is provided.
- to let the chairman know if the material of your presentation is to be protected. The agenda of the meeting will be publicly available and eventually indexed by web search engines. If this is a problem for you, please let the chairman know for setting access protections accordingly.
- to allow part of the time allocated in your slot for questions and comments from the audience

Information for participants

This meeting is organized so that participants can attend partially and should feel free to attend only their talks of interest.

Participants Frédéric Azevedo; David Bouvet; Pierre-Emmanuel Brinette; Sébastien Gadrat; Vanessa Hamar; Rachid Lemrani; Ghita Rahal; Bertrand Rigaud; Renaud Vernet

Material: [Annotated map of CC-IN2P3's area](#) [CC-IN2P3 geo. coords - N 45.782709 , E 4.865241](#)

[Directions to get to CC-IN2P3](#) [Social event venue](#) [Social event venue map](#)

[Go to day ▾](#)

Tuesday, 14 February 2017

09:30 - 09:40 Welcome & Introduction 10'

Speaker: Fabio Hernandez (CC-IN2P3)

09:40 - 10:10 KEK-CRC news and update 30'

Speaker: Prof. Takashi Sasaki (KEK)

Material: [Slides](#)

10:10 - 10:40 CC-IN2P3 news and update 30'

Tomoaki Nakamura, KEK-CRC

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

Travel information

Conference venue

Accommodation

Registration

Scientific programme

Co-located events

Social events

Sponsors

About HEPiX



<http://hepixon-fall-2017.kek.jp/>

HEPiX Fall 2017 Workshop

October 16 - 20, 2017 at HIGH ENERGY ACCELERATOR RESEARCH ORGANIZATION, KEK, Tsukuba, Japan

Menu

Welcome to HEPiX Fall 2017 Workshop

The HEPiX forum brings together worldwide Information Technology staff, including system administrators, system engineers, and managers from the High Energy Physics and Nuclear Physics laboratories and institutes, to foster a learning and sharing experience between sites facing scientific computing and data challenges. Participating sites include BNL, CERN, DESY, FNAL, IHEP, IN2P3, INFN, IRFU, JLAB, KEK, LBNL, NDGF, NIKHEF, PIC, RAL, SLAC, TRIUMF, and many others. The HEPiX organization was formed in 1991, and its semi-annual meetings are an excellent source of information and sharing for IT experts in scientific computing. The workshop is hosted by HIGH ENERGY ACCELERATOR RESEARCH ORGANIZATION, KEK.

- Event guide: [PDF](#)

Local organizing committee

Go Iwai (KEK)
 Tadashi Murakami (KEK)
 Tomoaki Nakamura (KEK, Chair)
 Takashi Sasaki (KEK)
 Soh Suzuki (KEK)
 Fukuko Yuasa (KEK)

- 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.

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!