Source: LSST

Overview of LSST activities at CC-IN2P3

fabio hernandez

Contents

◦ Computing resources

◦ Software

◦ Ongoing and future R&D activities

◦ Today's topics

◦ Summary

Computing resources

Platform upgrade

○ Migration to CentOS 7

○ Available now

compute nodes for interactive work

compute nodes for MPI jobs

compute nodes for GPU

4 compute nodes in the batch farm

2 login nodes: `ssh cca7.in2p3.fr`

○ Please get prepared to run on CentOS 7

if there are missing tools, please submit a ticket: `https://cc-usersupport.in2p3.fr`

○ Detailed schedule of migration to be established in the next 3 months

current major stopper is LHC experiments

Resources

◦ Staged increase of capacity

◦ Batch farm

+240 compute nodes to be put in production by the end of April

configuration: 48 CPU cores, 144 GB RAM, 1 Gbps NIC, 3 TB local disk (raw)

◦ Storage

```
/sps/lsst
```

managed by GPFS

allocation: 130 TB

used: 93 TB

Resources (cont.)

- Storage (cont.)

  request for 2017: +260 TB

  delivery schedule:

      +100 TB — March 27th

      +30 TB — April

      +130 TB — second half 2017

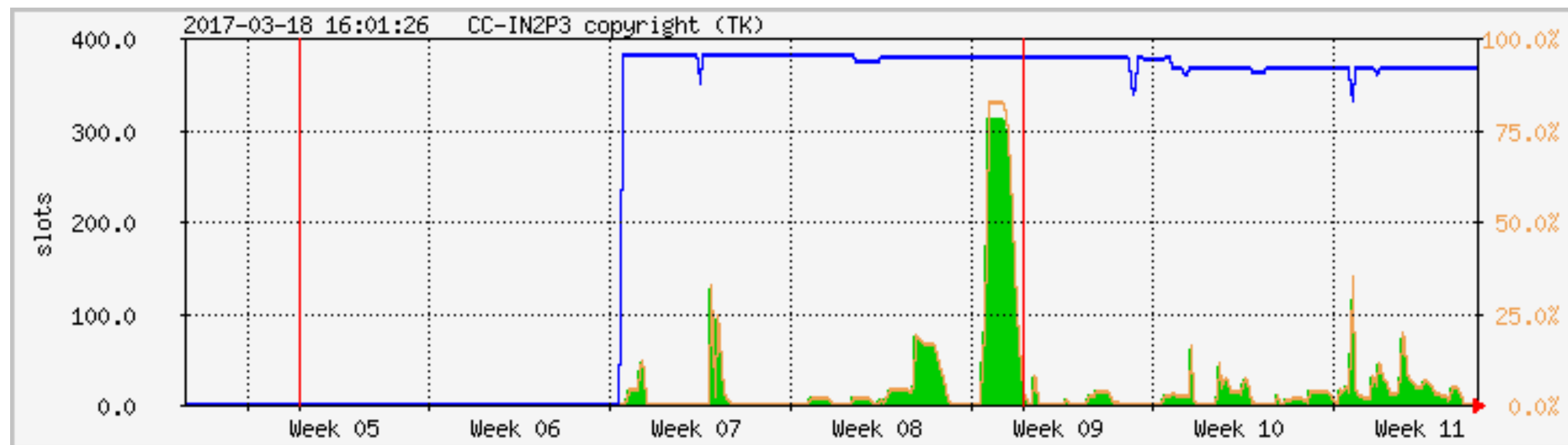- Ongoing tuning of the GPFS-managed areas used by LSST

  automatic redistribution of files among more file servers: expected increase in throughput

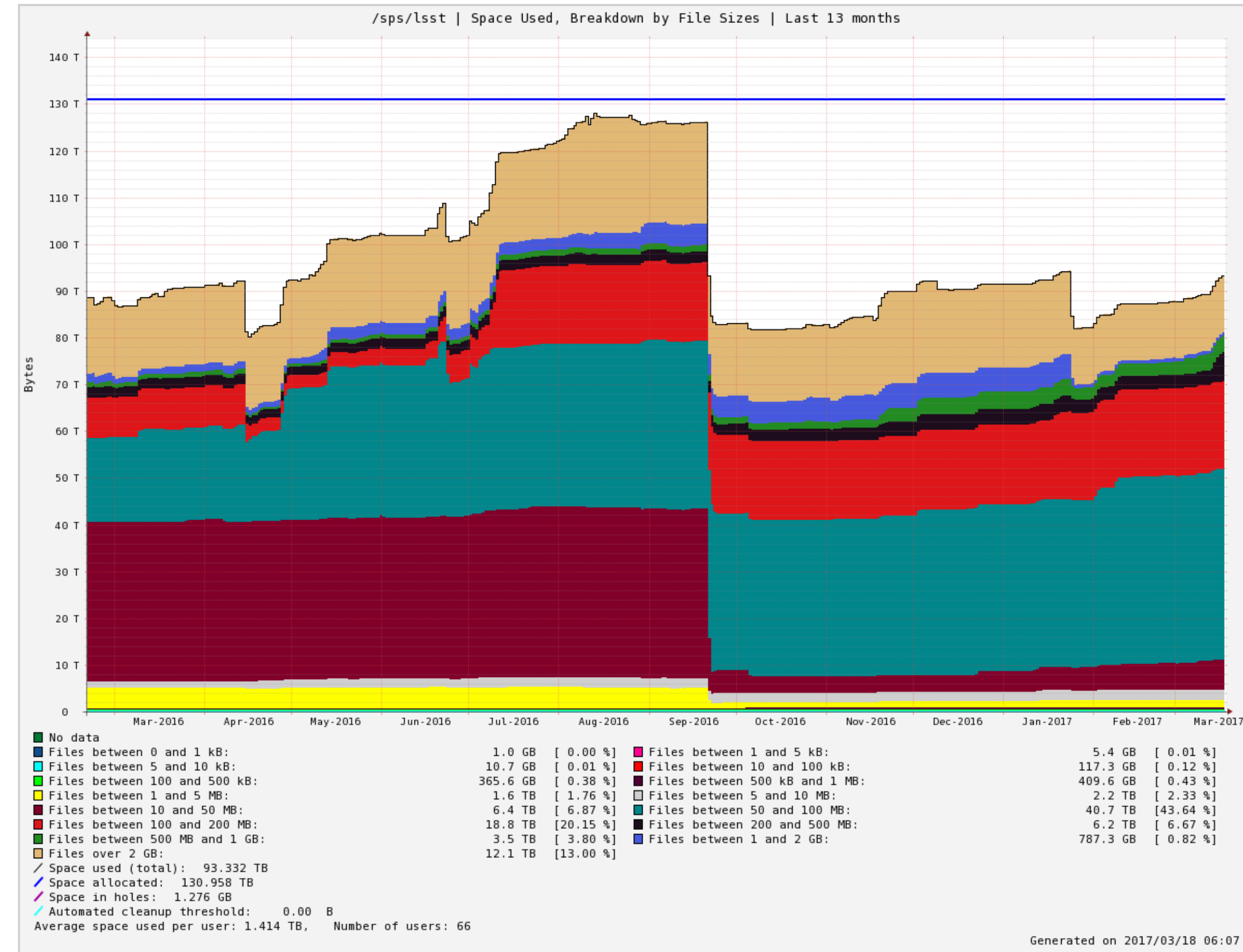 *automatic migration of files to different areas based on file usage and file size: transparent to users*

Resources (cont.)

/sps/lsst utilisation

/sps/lsst | Space Used, Breakdown by File Sizes | Last 13 months

CPU utilisation

Resources (cont.)

○ LSST data management plan to be prepared sometime in the future

○ Already 31 inactive LSST accounts (out of 100)

inactive means that password expired (6 months period) or no explicit account validity period

need involvement by czars

to avoid orphan data act before people leave the institute/labs: e.g. copy the data, modify ownership of the data, archive the data or remove the data

○ It helps if massive data recall from tape is performed in a coordinated way

Software

◦ LSST software framework (a.k.a. the stack) available via several channels

◦ In the cloud

binary distribution of <u>stable versions</u> delivered via CernVM-FS

accessible under: `/cvmfs/lsst.in2p3.fr/software`

built from source on CentOS 7 and OS X

works on other Linux distributions (e.g. Ubuntu)

CC-IN2P3 compute nodes running CentOS 7 are configured to also use `/cvmfs`

details: `https://github.com/airnandez/lsst-cvmfs`

Software (cont.)

○ At CC-IN2P3

installed under `/sps/lsst/software/lsst_distrib`

both <u>weeklies</u> and <u>stable versions</u>

built from sources on CentOS 7

example of usage:

```
$ source /sps/lsst/software/lsst_distrib/v13.0/loadLSST.bash
$ setup lsst_distrib
```

Software (cont.)

○ At CC-IN2P3 (cont.)

currently available under `/sps/lsst/software/lsst_distrib`

```
v13.0
w_2017_5
w_2017_6
w_2017_7
w_2017_8
w_2017_9
w_2017_10
w_2017_11
```

What is a sensible policy for keeping weeklies?

R&D activities

Ongoing R&D activities

○ Experimentation with big memory machines

two nodes, provided by Dell

   CPU: x2, 28 CPU cores, 56 execution threads

   RAM: 512 GB

   persistent local storage (data): 2 TB NVMe, 3 TB SSD, 10 TB HDD

   network: 10 Gbps

one node devoted to the Qserv team for experimenting with this machine as a head node hosting the object index

one node devoted to data processing activities, currently being used by N. Regnault et. al. for processing Subaru data

see Marc Betoules's presentation later today

Ongoing R&D activities (cont.)

○ Containerisation of steps of the data release processing workflow

use containers for packaging each step of the workflow

goals: ensure reproducibility, exploit data locality

to exploit data locality requires intelligent scheduling

preliminary results using Docker containers: packaging OK, scheduling not addressed yet

currently exploring QEMU-based packaging

Ongoing R&D activities (cont.)

○ Exploring HTTP-based **bulk file transfer** over high latency network link

optimise throughput, not latency

why HTTP: standard, future-proof, ubiquitous, programmability of both client and server, support for any relevant programming language, security mechanisms based on standards (TLS)

dedicated file transfer testbed deployed at CC-IN2P3

campaign of tests using 10 Gbps network link between CC-IN2P3 and NCSA (110 ms RTT)

encouraging results: HTTP seems a viable alternative to custom-built tools, for bulk data transfer

white paper with initial results delivered to PRACE 4IP project: in review

Ongoing R&D activities (cont.)

- Operations of Qserv development cluster

  50 nodes, 400 CPU cores, 800 GB of memory, 500 TB raw storage capacity

  developers located at SLAC

- Evaluating OpenStack Swift as an Internet-connected storage platform

  typical use cases: data archival, sharing of (big) big datasets

  details: `http://swift-in2p3.readthedocs.io`

  friendlier user interface would allow to appreciate the potential of this platform

- Prototyping a Python Notebook-based data analysis platform for LSST

  need to understand what the needs are

  see Bertrand Rigaud's presentation later today

Future activities

◦ Pursue ongoing R&D activities

notebook service, containerisation, data transfer

◦ New iteration of sizing and deployment plan of computing resources at CC-IN2P3 for data release processing needs

in coordination of update to size plan being pursued at NCSA

requires deeper understanding of the data processing workflows and assumptions

◦ Start prototyping LSST data backbone

exercise with spectrograph data: will require foundations of authorisation and authentication, data transfer, data archival, file metadata and file location services

CC-IN2P3 as archival location of selected data for disaster recovery

lead by NCSA

◦ Consolidate CernVM-FS-based software distribution platform

new hardware being prepared

new namespace consistent with local deployment: `/cvmfs/lsst.eu/sw/lsst_distrib/…`

Future activities (cont.)

○ Documentation platform

doc.lsst.fr

goal: to build the single entry point for documentation for LSST-France community

as a member of LSST-France,

> *how can I use CC-IN2P3 resources?*

> *where can I find the relevant datasets?*

> *where can I find the software?*

> *how can I use the LSST software framework?*

> *where can I find examples and best practice information?*

> *where can I get help?*

> *what are the communications channels?*

currently exploring a solution based on readthedocs.org, consistent with what is used by LSST

git-based workflow for writing the documentation, automatic rendering triggered by changes commit

allows for contributions by multiple writers

Future activities (cont.)

◦ Would training sessions be useful for current and future members of the LSST-France community?

◦ Two main topics foreseen, separated by several weeks
Python 3 (~2 days)

LSST software stack usage and best practices (~2 days)

◦ ~12 people

◦ Could be hosted by CC-IN2P3 (or elsewhere)

◦ Summer or fall

◦ Volunteer speaker(s) ?

Today's topics

Today's topics

- Big memory machines [Marc Betoule]

- HDF5 format [Guy Barrand]

- Reprocessing discussion [Dominique Boutigny]

- coffee break

- Using Qserv [Nicolas Chotard, Fabrice Jammes]

- Python Notebook service [Bertrand Rigaud]

- Spark [Christian Arnault]

- Virtual observatory [Hervé Wozniak]

- Data analysis center discussion [Dominique Boutigny]

Conclusions

Summary

◦ Computing resources requested for 2017 to be progressively deployed

◦ Several ongoing and foreseen activities

◦ Deployment plan of computing resources for data release processing to be revisited

◦ Start prototyping the LSST data backbone

◦ (Re)processing campaigns to come

useful to understand future LSST data processing activities