News from CC-IN2P3

Renaud Vernet – 18 juin 2012

Outline

- General presentation
- Cloud perspectives
- News on hardware and middleware

- Status
  - Prepare downtime
  - Current issues

The CC-IN2P3

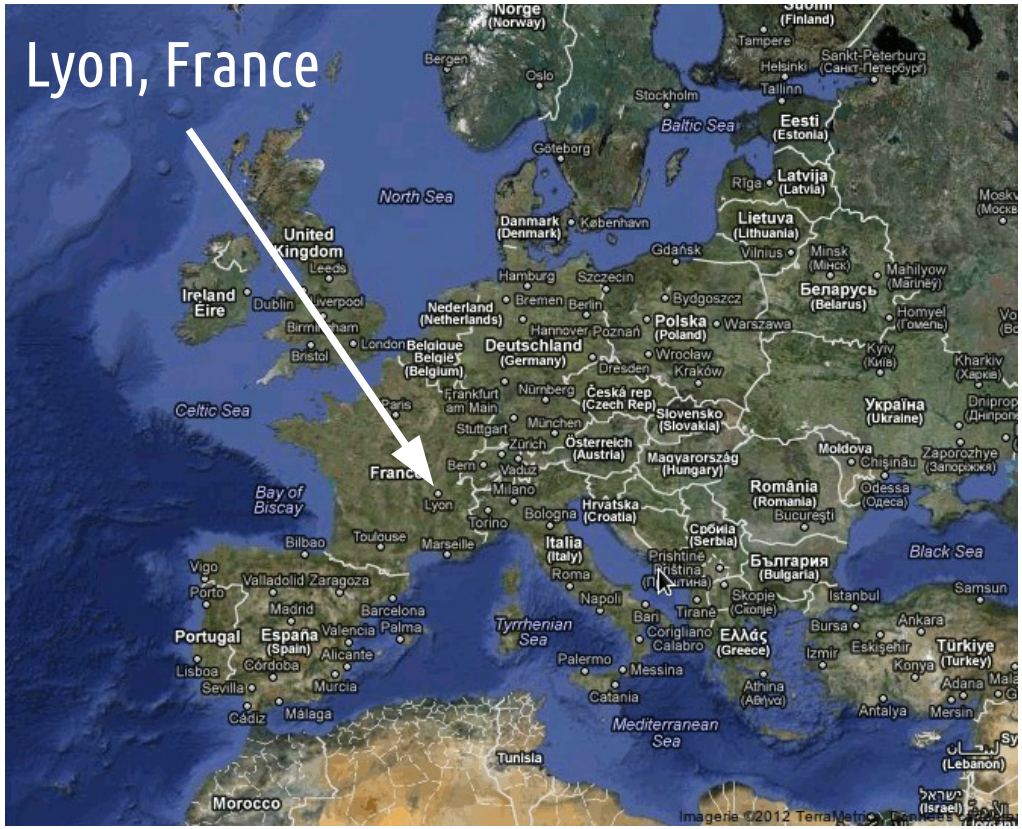
- IN2P3
  - National Institute for Nuclear and Particle Physics
- **C**omputing **C**enter originally created for physics purposes (nuclear & HEP)

→ The core of our activities

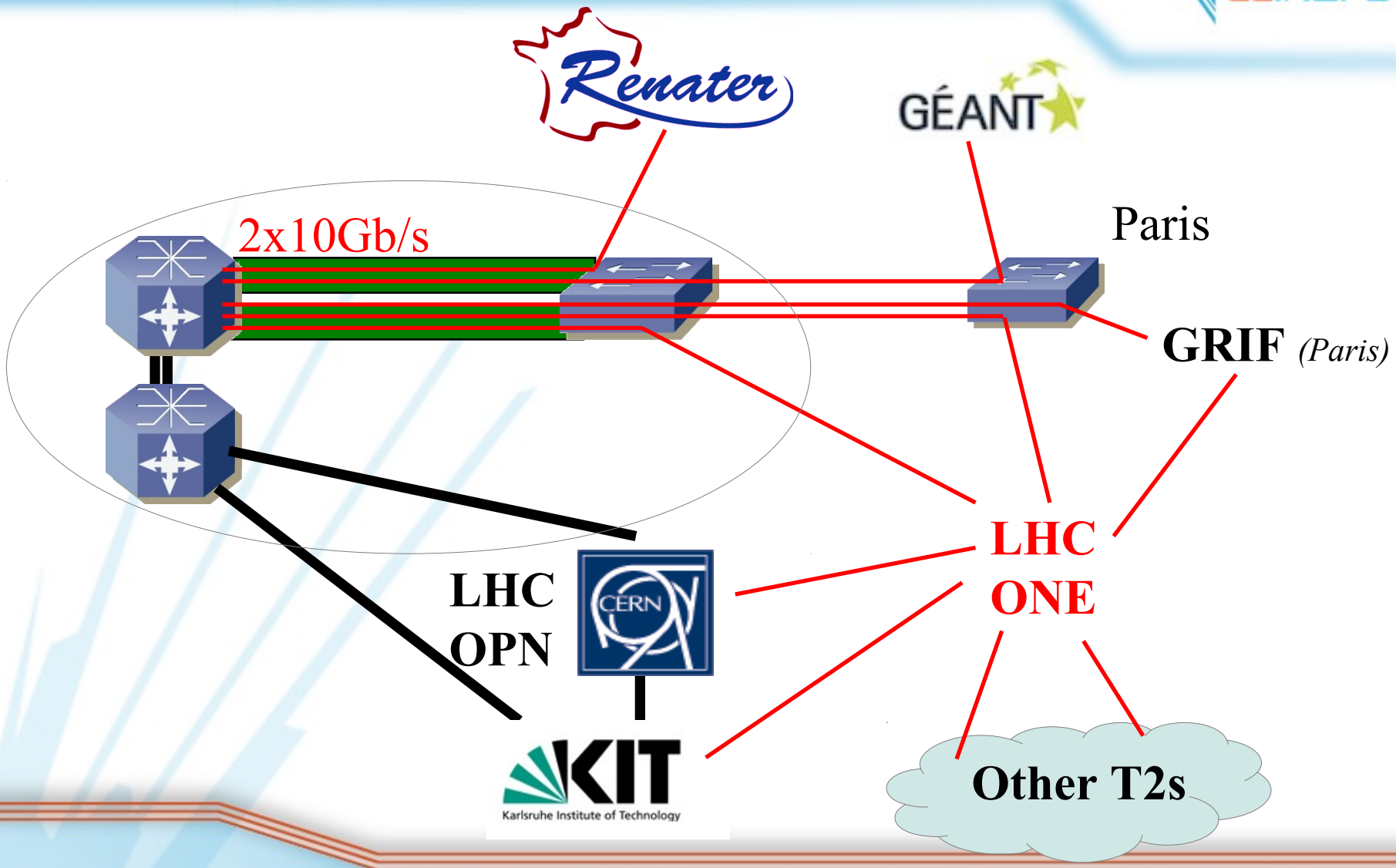
- Now serving also non-physics fields
  - Biology, medical imaging, neurosciences, humanities...
- Other activities
  - Grid development (portals)
  - Web hosting service (education purposes...)
- ~80 persons at CC
  - Some of them (user support group, 10 people) work in close contact with users/experiments

The CC-IN2P3

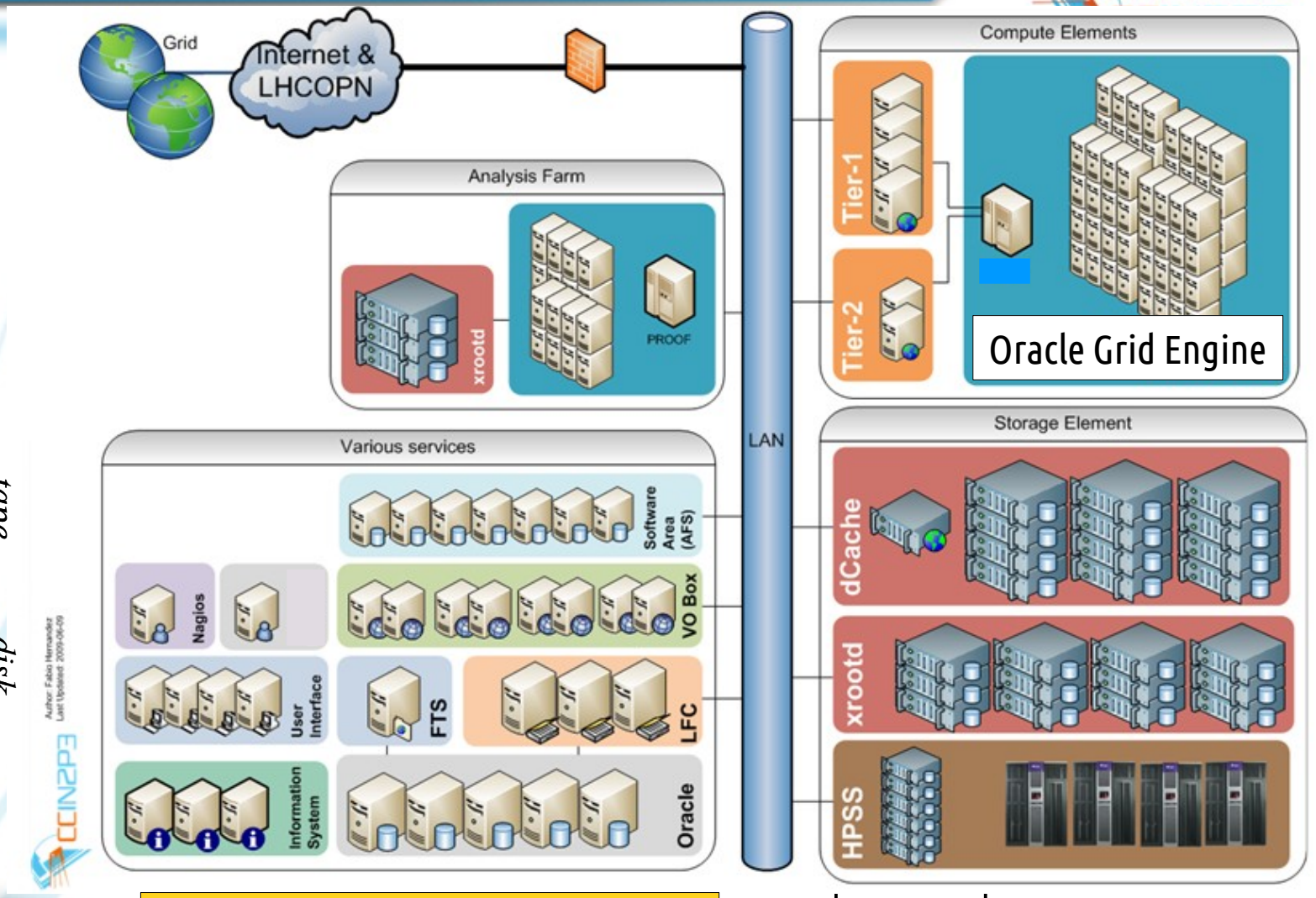
- Located in Lyon
  - ~150km from Geneva
- Hardware partners
  - Dell, Oracle, IBM, Cisco
- Network connectivity
  - LHC OPN
  - LHC ONE
  - Renater



Network


Renater
GÉANT
2x10Gb/s
Paris
GRIF (Paris)
LHC
ONE
LHC
OPN
CERN
KIT
Karlsruhe Institute of Technology
Other T2s

Resources overview



17 PB HPSS
1 PB TSM

8.2 PB dCache
1.9 PB irods
1.7 PB xrootd
1.2 PB GPFS
0.3 PB SRB

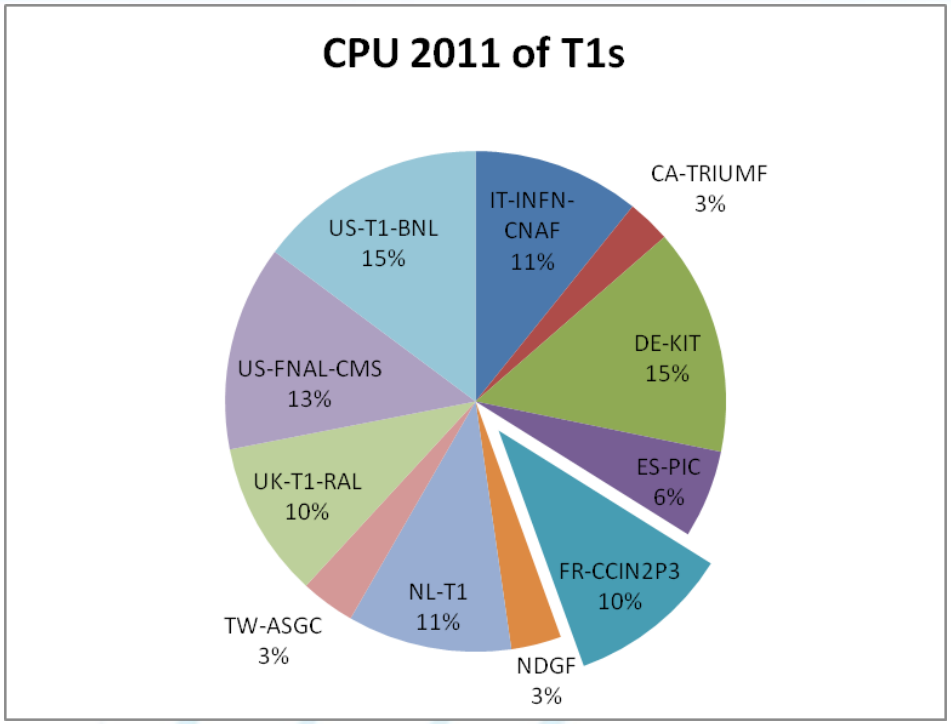
tape

disk

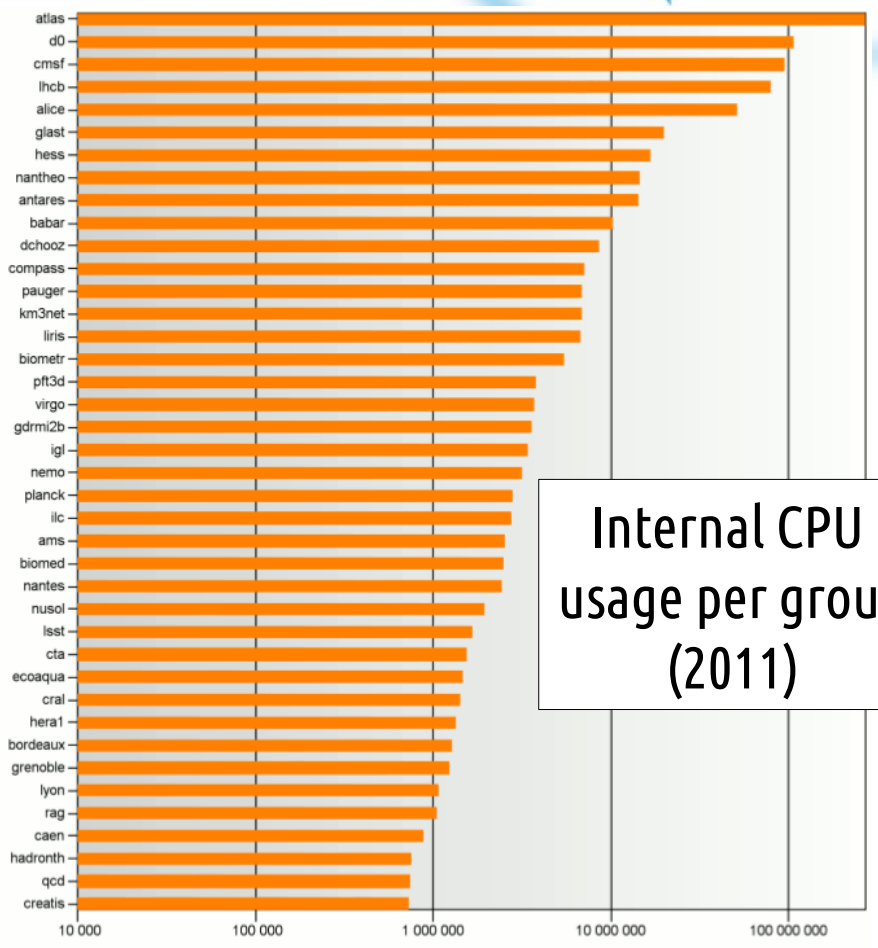
+ GPFS

~16k workpoints (essentially SL5)
~30PB storage (tape+disk)

CPU views



Share in WLCG



Internal CPU usage per group (2011)

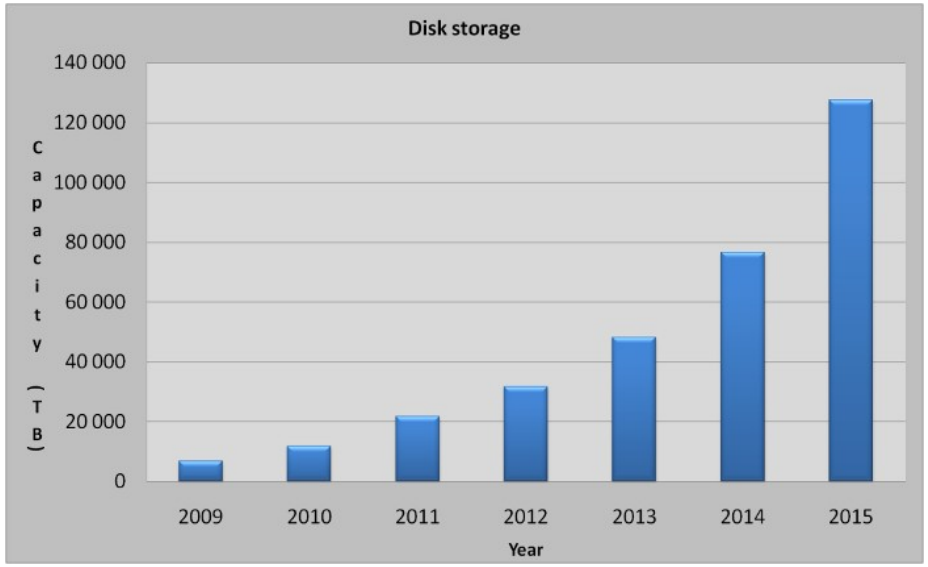
(Log scale)

HS06.h

Extension of CC-IN2P3

- Previous architecture of CC-IN2P3 could not face the upcoming growth of computing resources
  - Space, cooling reasons
- A new building was built to cope with these needs
  - Twice more space
  - New cooling technologies
  - Environmental aspects

A big design phase, a big investment for a big project :
unique opportunities in scientific data processing



Extension of CC-IN2P3

Cloud at CC-IN2P3

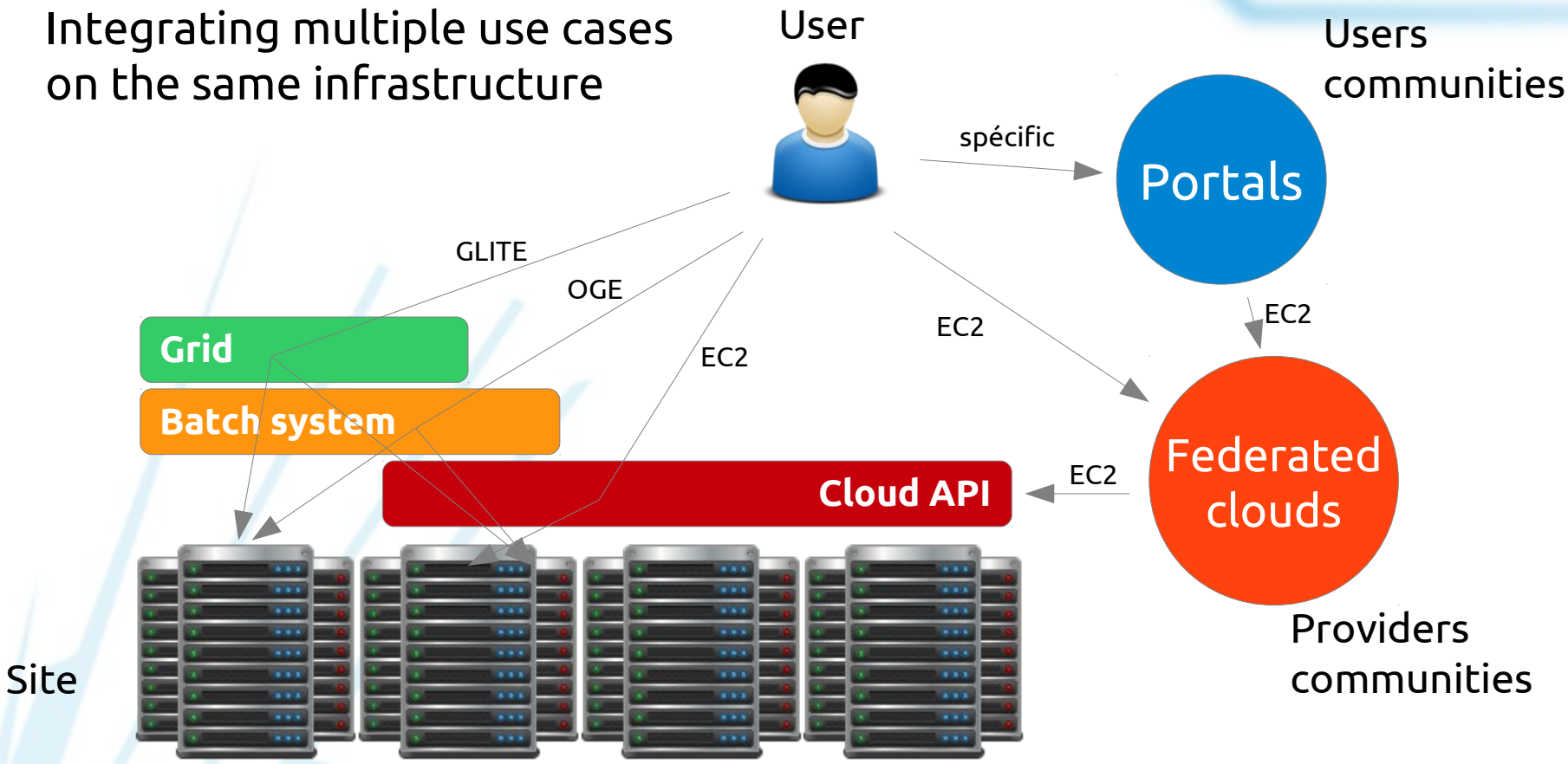
- Main goals
  - Build an academic community cloud, integrated to wider federations
  - HTC use case is no option
- Motivation & user needs
  - Users need flexibility (elasticity)
  - Another way of achieving distributed computing
  - Satisfying new use cases (servers on demand...)
- Steps
  - Offer IAAS ressources through generic interfaces (EC2/OCCI)
  - Integrate national/european/worldwide academic federations

Cloud at CC-IN2P3

- Work in progress
  - Evaluation of existing/upcoming technologies
    - Proprietary : IBM, Dell/Canonical, VMWare, Oracle, RedHat
    - OSS : OpenNebula, Openstack, Nimbus...
  - Reuse experience
  - Identify hot spots
    - Security, storage, performance, networking
- And then
  - Open to new communities
    - Other scientific fields, academic institutes, industry ?
  - High level tools for users (to PAAS/SAAS)
  - Branch to the batch system
  - How will the users follow these new technologies ?

Computing and cloud interfaces

Integrating multiple use cases on the same infrastructure



Hardware news

- CPU
  - We soon take off PowerEdge 1950 & Dataplex
  - New machines C6220 soon
    - 32 cores, 3 GB mem/core, 1 TB disk
    - Better cpu/power ratio, larger scratch area and memory per core
- GPU
  - Dell test machines ordered
  - ALICE is going to be a user
  - Other VOs ?

Dedicated infrastructure to be provided very soon
(Dell, chassis with 4 GPU blades)

- Tape
  - All new data written to T10K-B (1 TB) and T10K-C (5 TB) tapes
  - No more use of T10K-A (500 GB) tapes

Middleware migrations

- gLite 3.2 is retiring... supports ends this year
- EMI1 is now the reference
- EMI2 available since last May
  - Not (or very little) tested by sites
- CCIN2P3 is migrating to EMI1
  - CE, BDII, UI...
- EMI2 migration not planned yet
- gLExec
  - Deployed
  - CMS sees problem
  - We will ask ATLAS support to check

Fall downtime

- CCIN2P3 is going to be connected to a second power line
- Intervention is being planned together with the French electricity provider (EDF)
- A 2-3 day downtime will occur most likely next Fall
  - Date unknown yet
- As it is a 'long' downtime, we will keep the most critical services online
  - Power will be provided by the local generator
- Nothing critical for most LHC VOs but ATLAS
- ATLAS :
  - FTS, AMI, Top & site bdii
  - If possible
    - 1 Frontier server (RAL failover may not bear te load)
    - Pilot factory (unsured by CERN otherwise)

Current issues : tape usage

- 2.5 PB tape is allocated to ATLAS at IN2P3
- Some is used by local users
  - Some users use it quite extensively
- Actions have to be taken with the users
  - Tape is not a Tier-3 service

Current issues : network/transfers

- LHC ONE
- Performances
  - Change of ethernet card bounding algorithm
  - Java version on Solaris machines
    - Prevent us from upgrading
- Connection with China issues
- See Jerome/Eric talks

Current issues : Frontier/squid and GOC DB

- ATLAS requires Frontier/Squid service to belong to the GOC DB
  - Currently managed by ATLAS support at IN2P3
    - Not an service officially maintained by our site
- Our position as a site
  - Not yet a full WLCG service
    - no official technical support, security (?)
  - Officials have to decide whether we take the responsibility of managing this additonal service, officially
  - Otherwise, should be done at the vobox level

Notes, ideas & remarks

- Memory usage
  - WLCG pledges are 2 GB per core
  - Most VOs use 4 GB (or more)
    - CPU disponibility is therefore affected : naively twice less jobs (!)
  - Virtual memory calculation
    - What is it based on ? Account for shared memory ? Batch system dependence ?
- Accounting
  - Groups objectives expressed in terms of consumed CPU time on the batch system
    - → low cpu efficiency software penalises everybody
  - My personal view : we should base accounting on # jobs (wall time basis)
    - VO should be responsible on its CPU use
    - Analysis is usually rather I/O bound than CPU bound

That's all, folks

Corrupted files

- LHCb seems a fraction (0.03%) of corrupted files on our dCache
  - Checksum in LFC different from Chimera
  - Checksum verification not activated during transfers
- ATLAS not affected
  - Checksum verification activated
  - We do see transfer failures, but files not registered
- Thorough investigations on many suspects
  - LFC, network, worker node, middleware, dCache pool...
- Ticket marked as unsolved
- We asked LHCb to activate the checksum verification