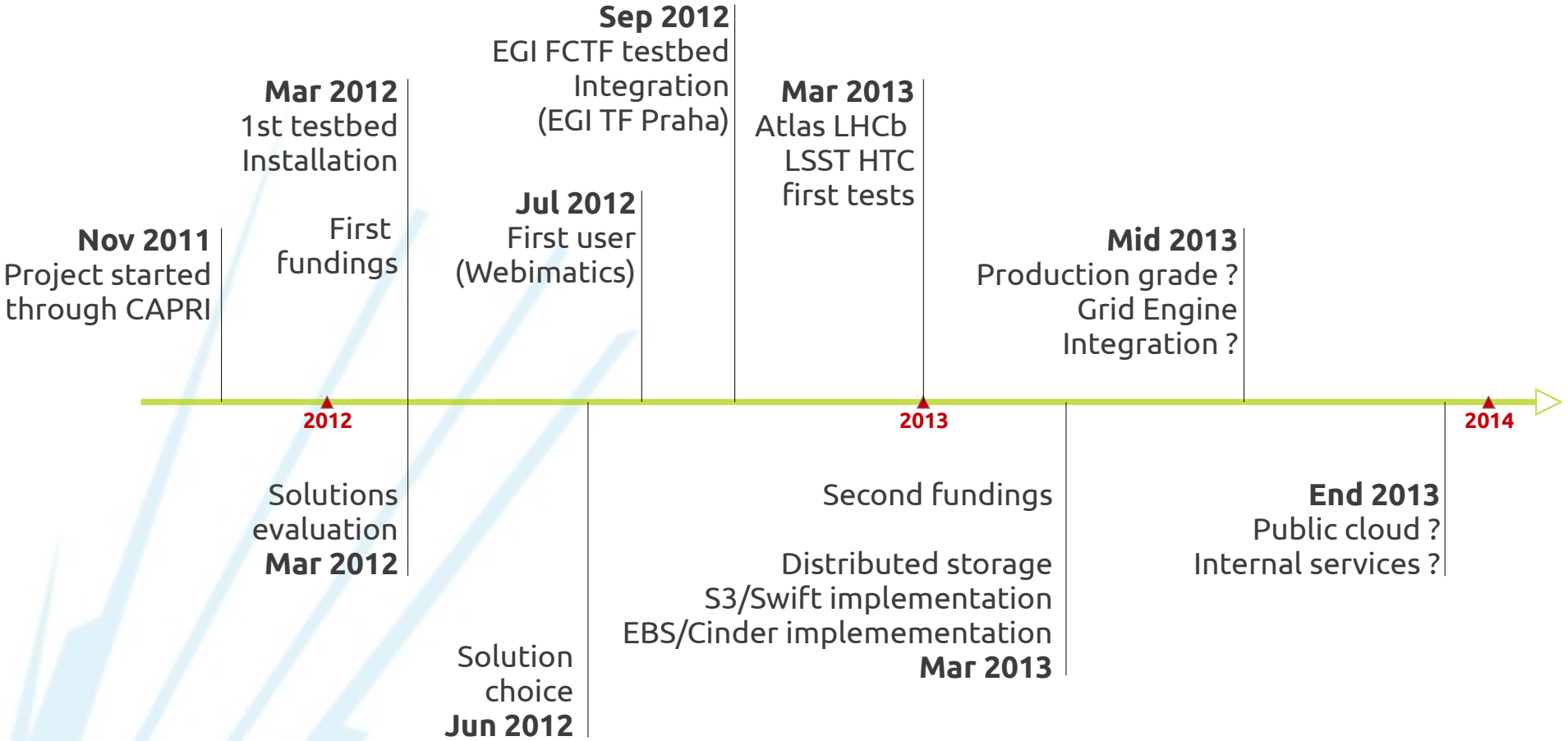
CC-IN2P3 Openstack status

CERN / IN2P3-CC Openstack teams meeting

Mattieu Puel – Mar 2013

Timeline


Nov 2011
Project started through CAPRI
Mar 2012
1st testbed Installation
First fundings
Sep 2012
EGI FCTF testbed Integration (EGI TF Praha)
Jul 2012
First user (Webimatics)
Mar 2013
Atlas LHCb LSST HTC first tests
Mid 2013
Production grade ? Grid Engine Integration ?
End 2013
Public cloud ? Internal services ?
2012
2013
2014
Solutions evaluation Mar 2012
Solution choice Jun 2012
Second fundings
Distributed storage S3/Swift implementation EBS/Cinder implemementation Mar 2013

Current platform (hardware)

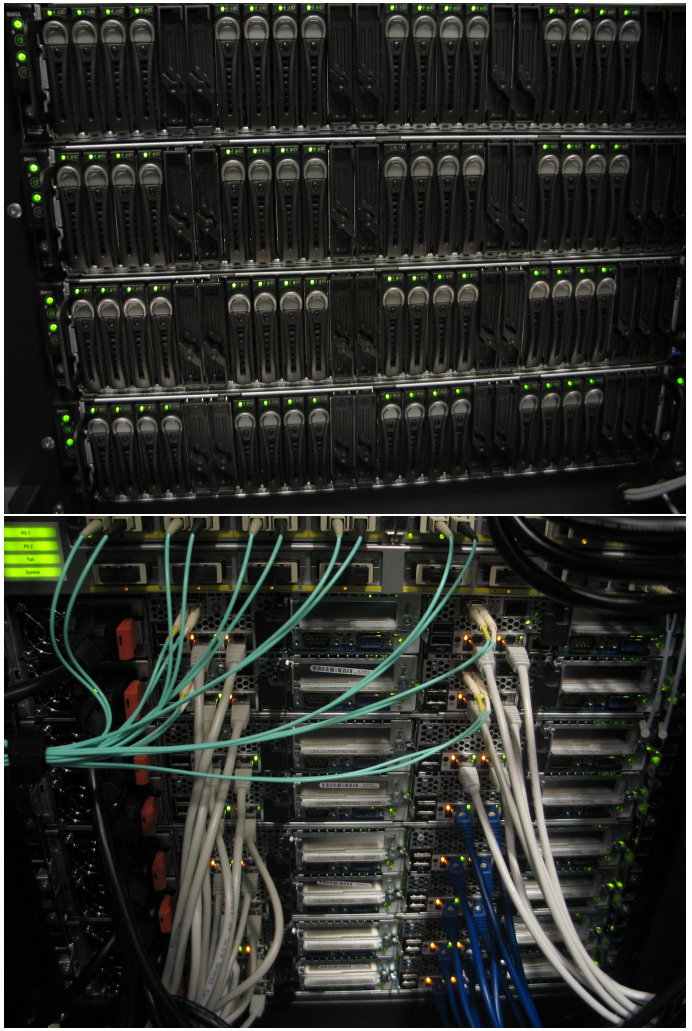
Nova

16 DELL Poweredge C6100 hosts :
- 2 Xeon 24cores X5675 @ 3.07GHz
- 96GB RAM
- 2TB raid 10 local storage (4 SAS 7.2krpm)

Total infrastructure :
- 400 cores
- 1.5TB RAM
- 32TB disk
- 32Gbps network

→ **~400 VMs m1.medium (1c, 4GB RAM, 50GB disk)**

10Gbps NICS with NPAR/SR-IOv technologies



Current platform (hardware)

Swift

3 DELL Poweredge R720xd :
- 10 SAS NL 7.2krpm 931GB drives
- 10 Gbps NIC

→ 22TB net capacity (220 users provided 100GB)

Cinder

6 DELL Poweredge R720xd :
- 10 SAS NL 7.2krpm 931GB drives (capacity)
- 10 SAS 15krpm 558GB drives (performance)
- 10 Gbps NIC

→ 30TB net capacity (300 100GB volumes)

Backend for /instances

10 DELL Poweredge R720xd :
- 10 SAS NL 7.2krpm 931GB drives (capacity)
- 10 SAS 15krpm 558GB drives (performance)
- 10 Gbps NIC

→ 37TB net capacity (~1.2k VMs)

EMC Isilon filer

→ 100TB net capacity (usable for both Cinder and /instances)

Current platform (software)

Operational :

Keystone
Glance
Nova
Horizon

Deploying :

Cinder
Swift

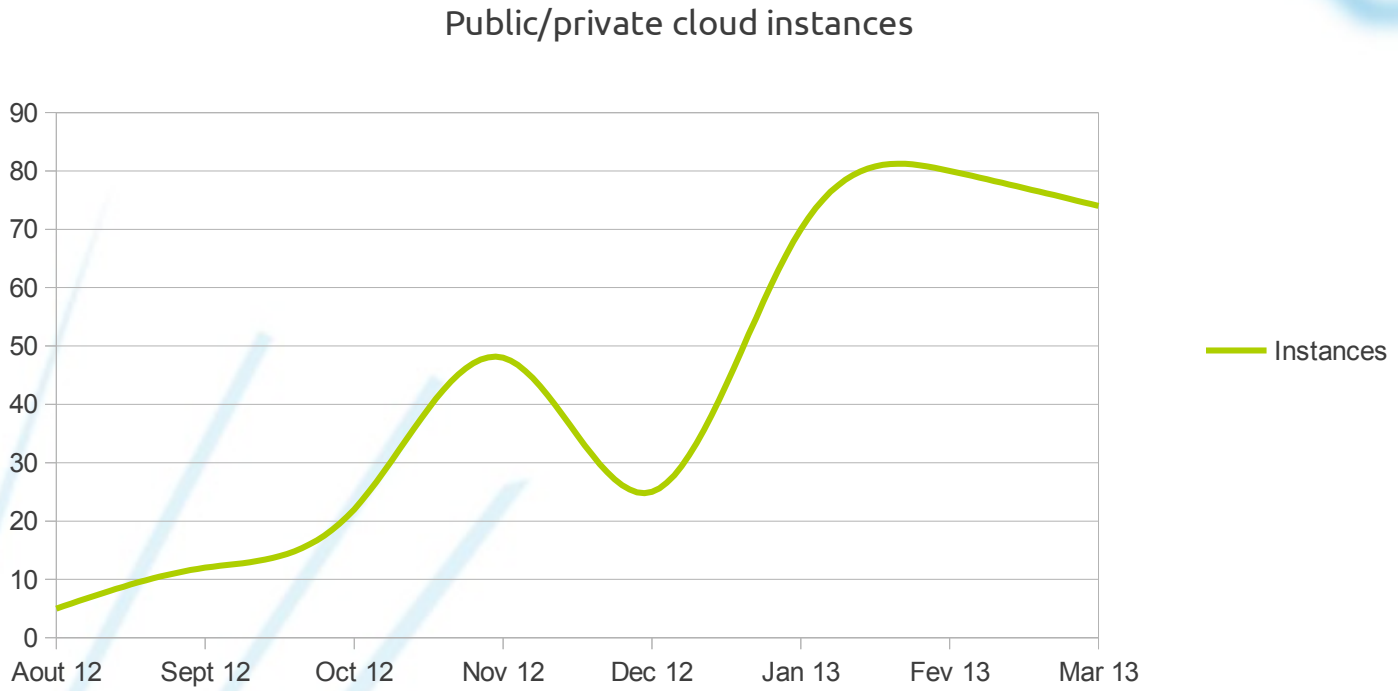
Interest in :

Quantum
Ceilometer

Features

Floating IPs
VLAN mode
GPFS backed Glance
VNC console

Platform usage



In regard with :

- ~180 infrastructure services physical hosts
- ~110 infrastructure services Vms
- ~700 physical worker nodes
- ~1400 total servers

Use cases

HTC corporate worker node

- Support for multiple specific environments (SL5/6, python dists, …)
- Full GE integration required (transparent to the users)

HTC user designed worker node (6 groups)

- DIRAC
- HEP use cases (CMS, Atlas, LHCb, Integral)
- Astroparticles (LSST, Euclid)
- EGI FCTF testbed
- **2 options:** w/o scheduler

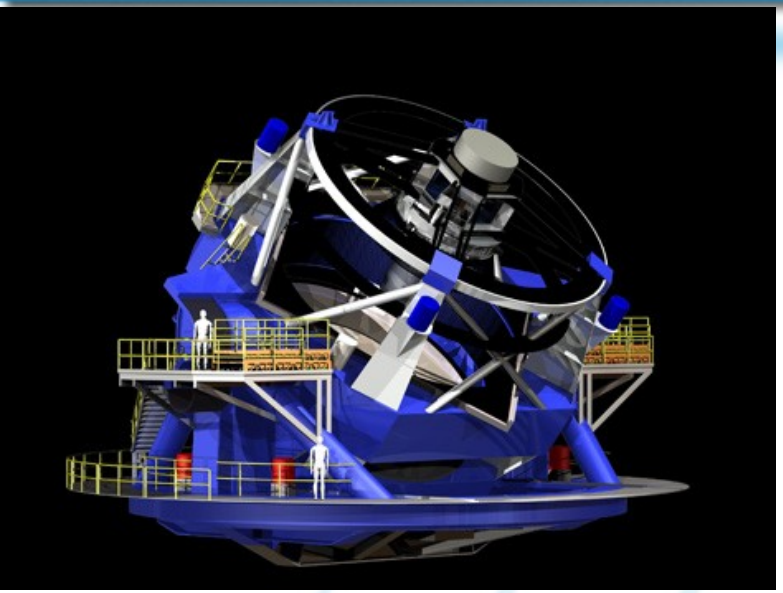
Public cloud (11 groups)

- IRT Bioaster
- Webimatics
- Academic/institutional (IN2P3...)
- Continuous integration (Jenkins)

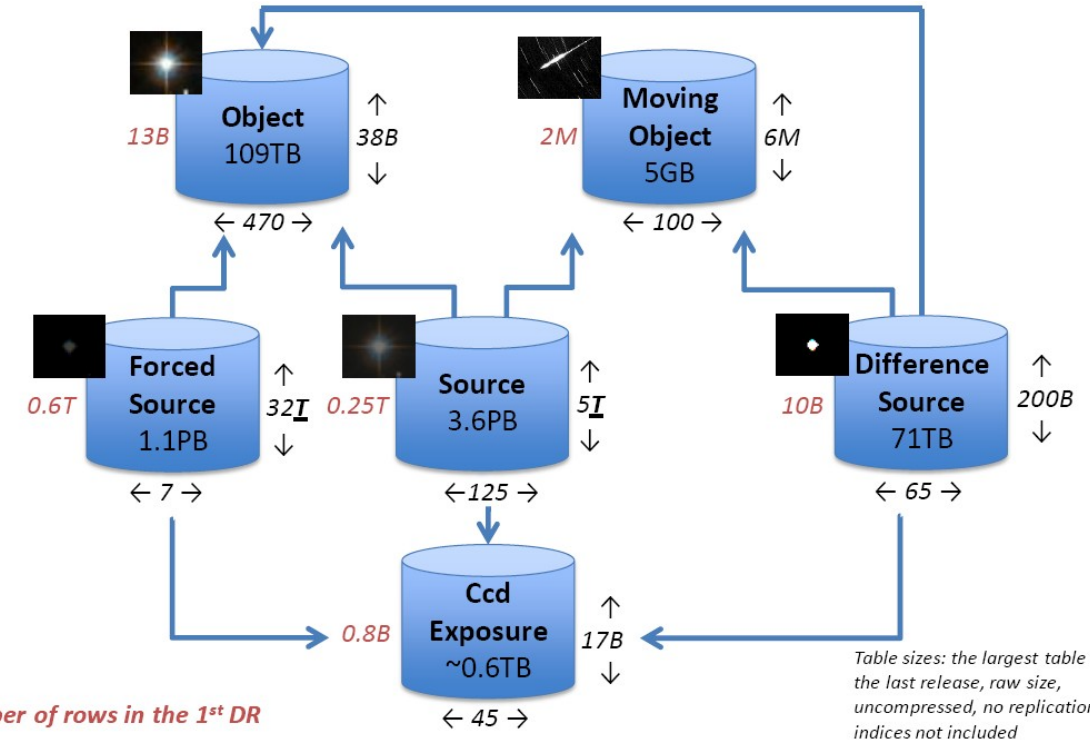
Private cloud (2 groups)

- eTRIKS (European Translational Information and Knowledge Management Services)
- Production internal services (migrate VMware vSphere services ?)
- Test servers

LSST (Large Synoptic Survey Telescope)



3.2 Gpixels – 2 images every 40s
→ ~15 TB of data every night



The whole visible sky scanned 1000 times during the lifetime of the project (10 years).

IN2P3-CC will process 50% of LSST data (the other 50% will be processed at NCSA). Catalog generation : 40k physical cores (2012 definition) in 2020 – 200k cores in 2029.

The whole LSST dataset will be available at IN2P3-CC.

The result of the processing (catalogs) will be stored in a gigantic distributed database system.

Computing

Tests

- Recycling of 70 DELL Poweredge 1950 systems (560 cores, 1.1TB RAM total capacity)
- Finish GPFS/Isilon testing for /instances storage
- Localising scratch space with distributed /instances (patch)
- Atlas/CMS/LHCb interested in testing/using ressources with CernVM
- LSST envisaging to start from scratch using cloud ressources
- Deployment of WNs is manual

Computing

Still required for real production :

- Need for a proper batch system integration, candidates :
  - Vcluster (Fermilab/KISTI)
  - Unicloud
  - Home made...
- Validate use cases (LSST, LHC experiments)
- Integration of former batch worker nodes, two constraints worked around :
  - Localise scratch space (easy before Folsom, now requires a patch)
  - Addition of distributed storage because of a lack of local space
  - Segregate Service / HTC VMs

Federation aspects

EGI Fedcloud Taskforce

- 14 RPs: (BSC, CC-IN2P3, CESGA, CESNET, Cyfronet, FZ Jülich, GRIF, GRNET, GWDG, INFN, KTH, SARA, TCD, OeRC, STFC, SZTAKI)
- Current work/choices:

resources advertisement	**BDII**
unified authN/authZ	**VOMS**
unified interface	**OCCI**
placement policies	**ø**
images sharing	**Stratuslab Marketplace HVWG tools (vmcatcher)**
aggregated accounting	**APEL/SSM**
brokering	**~~CompatibleOne~~ → ø**

Federation aspects

EGI Fedcloud Taskforce

+	Many resource providers Heterogeneous CMP technologies HEP, EGI related VOMS authentication Accounting
-	No placement policies No brokering Single interface supported (OCCI) Image deployment

Main issues, still investigated

Networking

- Secgroup model : let end users manage ports opening ??

- IPV4 public addressing limitations → floating Ips → breaks VLAN isolation. Let's move to IPv6 then.
- Multi host !

Storage

- Backend storage and former WNs recycling
- Ceph and SL6

Active/Future work

Active work

- Cinder implementation
- Swift implementation
- HTC productions (LSST, LHC experiments, …)
- Enable former WNs recycling (GPFS validation)
- France Grilles (French NGI) federated cloud (with LAL, IRIT, …)
- EGI FCTF

Future work

- Production grade (Nova, Cinder, Swift)
- Unicloud implementation
- Public cloud opening (define the economic model)
- IPv6 networking redesign
- Internal services provisionning (aim at replacing VMware vSphere)