LAL and GRIF Site Report

Michel Jouvin
LAL, Orsay
jouvin@lal.in2p3.fr

HEPiX, BNL, October 2015

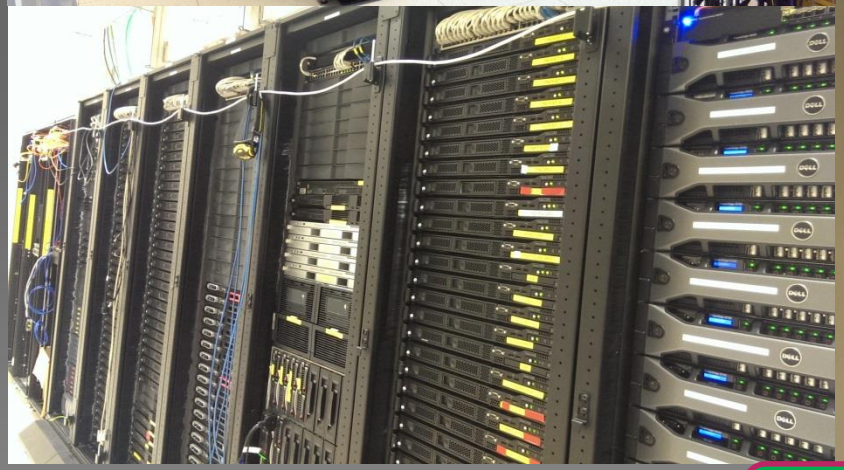
Disclaimer

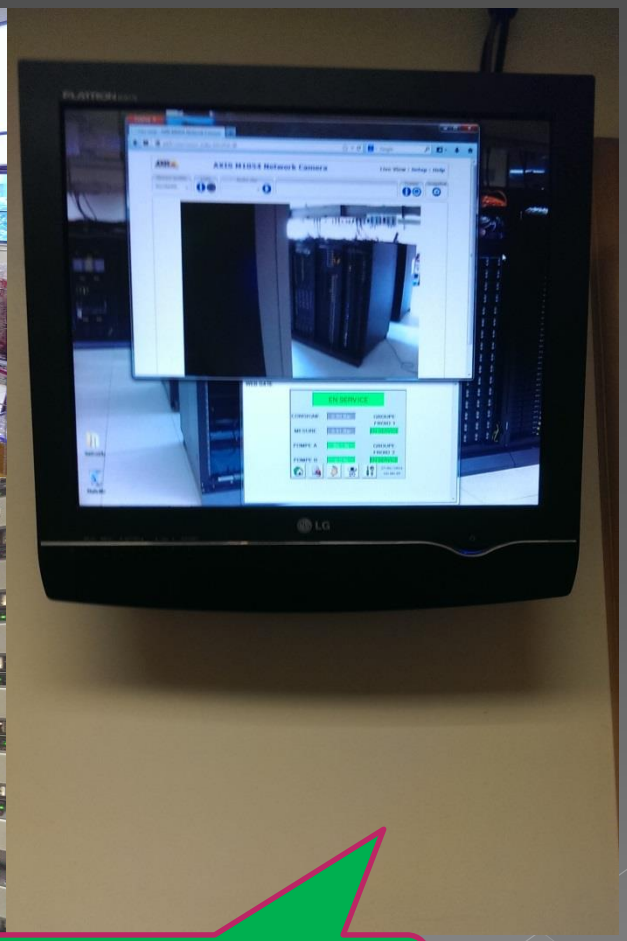
Except when noticed explictly, the information in these slides applies to LAL only

Infrastructure

- Infrastructure built 2 years ago running smoothly
  - Shared by several labs
  - Hosting 250 kW IT in 30 racks currently (400 kW, 30 racks max)
  - No issue so far
  - Chiller redundancy is a real improvement: not sensitive to chiller issues or maintenances
  - Cooling system based exclusively on rear door heat exchangers working very well: racks up to 20 kW without problems
    - ATOS racks with fans for hot air extraction
- Extension planned during next year: funding secured
  - Start of studies delayed by (minor) administrative problems
  - Goal: 900 kW IT in 50 racks, 3 chillers + 3 power circuits in N+1 configuration
  - Extensible to 1,5 MW/84 racks by adding chillers and power

Infrastructure





remote monitoring

Manpower

- GRIF: decreasing manpower, ~5-6 FTEs
  - Same number of people but with a decreasing fraction of their time dedicated to GRIF
- LAL: a better period ahead of us?
  - Just hired one more sysadmin for grid and cloud resources
    - Unfortunately another sysadmin announced he'll leave mid-December: replacement not yet clear…
  - 1 apprentice just started as a cloud resource sysadmin for 2 years
    - The person has already been a 1-year apprentice this year and proved to be very efficient!
    - Funded by H2020 CYCLONE project (testbed activities)
  - Several persons have left or leaving (retirement)
    - Not affecting server sysadmins but desktop and conference support
    - Replacement partly anticipated… but partly uncertain!

Hardware Update…

- A lot of (very) old HW due to the infrastructure problems during a few years… A challenge to renew it!
- Servers: standardized on dual-twin servers to optimize the footprint
  - Dell C6xxx + Dell FX2 mainly
  - Dell FX2: a blade-like chassis with a shared network switch and ability to chain the switch (with redundancy) with 10 GbE uplinks
    - Replacement of the top-of-the-rack switch with a more flexible oversubscription
  - Still 2-3 racks of IBM 3550 (8 years)…
  - Service virtualization (cloud) to use efficiently "big boxes"

… Hardware Update

- Storage: almost replaced 10-year old DDN arrays by Dell MD3xxx (iSCSI or SAS)
  - Should be completed by the end of the year
  - 100 TB per server
  - Proved to be performing well and very resilient/dependable
  - 20% more expensive than a storage-in-a-box solution: can we continue to afford it?
    - No other alternative solution identified yet: attractive one from SM turned out to be pricy…
    - Dell OEM from NetApp coming to an end in 1 year: new Dell solutions to be evaluated
  - Still have to replace HP-based storage (200 TB): not dependable, a lot of operational issues

OS Changes: Unix

- Business as usual… mainly SL6 systems currently
  - Still running a few SL5 machines but becoming marginal
- Started to deploy a few EL7 machines
  - Moving to CentOS: no problem seen, except the shorter support for minor versions (no errata after the release of the next minor)
    - Rolling version: problem with update of installation kernel/initrd
  - Currently mainly web servers but soon grid systems
- Moved to a unified authentication infrastructure based on Windows Active Directory (LDAP + Kerberos)
  - LAL was still relying on NIS…
  - Has been a lot of work to clear the inconsistencies between Unix and Windows accounts… but a useful cleanup!
  - Still have to convert web systems: impact on service ACLs
  - Developed an account creation tool as a Django application

OS Changes : Windows

- Windows 7 and 8 mainly
  - Installed with WDS, managed with Active Directory
  - WSUS for patching
- Still running quite a few WXP boxes
  - Often used for system controls, in particular local electronics development
  - (Slow) progress in discussion with users…
  - Restricting network connections… when possible
- Planning to start at W10 support soon
  - Be ready before the first boxes with W10 pre-installed arrive

Projects

- H2020 CYCLONE: enabling technologies for cloud federations
  - Strong focus on federated authentication and virtualized network services based on openNaas solution.
    - openNaas is the outcome of FP7 projects
  - Complementary (with some overlap!) to INDIGO Datacloud
  - LAL operating a testbed and doing some developments
- Irfu involved in H2020 INDIGO Datacloud with similar responsibilities
  - INDIGO: develop an integrated data/computing platform for scientific communities
    - SW components allowing execution of applications on Cloud and Grid as well as on HPC clusters.
  - OpenStack cloud with Ceph backend being deployed
  - Support some user communities (e.g. CTA) as early adopter of the new infrastructure and SW components available

GRIF Update

- A grid site in Paris region involving LAL, Irfu and 4 other labs
  - 10000 cores, 7 PB of disks
- 10th anniversary last Spring… already!
  - Despite some difficulties, a very successful (human) story
  - Unusal object for VOs but at the end they recognize the benefit
    - In particular in term of manpower and support
  - Plans for the next 10 years!
- Like others, have to live with tighter budgets and manpower
- Significant change in 2015: move to HTCondor for 3 subsites
  - 1 site also moved to an ARC CE
  - Saw the benefit of the collective work (and Quattor!): relatively easy upgrade after the first subsite
    - In fact 1 subsite moved to Puppet (GRIF/Irfu)

Cloud

- Cloud inherited from the StratusLab project turned into a University-wide (Univ. Paris Sud) resource
  - University funded last year the doubling of the resources
    - 800 cores, 100 TB
  - Goal: allow University users to assess if a cloud resource addresses their needs for scientific computing, experiment with cloud resource sharing between different (funding) groups
- Most (big) users need a job paradigm and don't want to become sysadmins
  - Evaluating offering a job interface to the cloud through HTCondor
- LAL is still considering virtualizing all its grid WNs
  - With HTCondor, in a much better position to do it
- Several GRIF labs using Proxmox for service virtualization

P2IO VirtualData…

- P2IO: a common initiative all HEP, NP and Astrophysics labs in Orsay to foster synergies between each others
  - Physique des 2 Infinis et des Origines: 8 labs, including LAL and Irfu
  - Make fundamental physics visible in the new Paris Saclay university
    - May be a funding opportunity…
  - Several technological platforms in P2IO: VirtualData = computing
- VirtualData goal: build a computing expertise network around a shared computing platform
  - 130 people involved in computing
    - Development (~75) and Operations (~55)
  - Shared computing platform hosted in 2 shared facilities to enable redundancy when needed
    - Use cloud expertise gained with StratusLab

… P2IO VirtualData

- Facility part of the project is now over
  - 2 datacenters: Orsay and Ecole Polytechnique (+ CEA/Irfu)
  - Common management of the facility is now well established, at least for the Orsay datacenter
    - GRIF served as a useful experience and model: 4 GRIF partners in VD
- Future challenges: run common services and foster synergy between developers
  - R&D work started around Ceph with all the 8 labs involved.
    - Struggling for funding a decent prototype (0.5 to 1 PB) with Ceph-tuned hardware: currently using decommissioned grid HW
    - Not yet managed to get good performances (1.5 GB/s with 6 servers)
  - Recently failed to get funded a development project around parallelization involving HEP (tracking) and Astrophysics (HPC)
    - Looking for new opportunities but something has started…