GRIF - LPNHE
Site Report

Victor Mendoza
http://grif.fr

LCG France, Sites Meeting
CC-IN2P3, Lyon
November 22, 2010

Contents

- Site report
  - Tools
  - Resources
  - Old Server Room
  - New Server Room
  - Relocation Plans
  - Problems After Move
  - Infrastructure
  - Server Room Pictures
    - Procurements
    - Local Tools
    - Local User Support
    - Virtualization
    - Troubles
    - Plans
- GRIF-LPNHE vs BDII v5
- Plans

GRIF

- Tier 2 LHC
  - 6 locations
  - APC, IPNO, IRFU-CEA, LAL, LLR, LPNHE
  - 6 CE, 6 SE
- Pooling and consolidating
  - IT resources: failover, round-robin, sharing
  - Human resources: ~20 people
    - December 1, 2010: 2 CDD → CDI
- GRIFOPN: 10Gbps private network

GRIF's Tools

- Quattor – SVN: installation, configuration
  1020 servers
- Trac: documentation, tracking
- Jabber: private chat
- Accounting: BIRT
- Monitoring: Nagios
- Security: Pakiti
- Collaborative tools

GRIF-LPNHE's Resources

Computing	GRIF	LPNHE
Computing: SPEC06	60K	5.8K
Computing: CPUs	1.4K	166
Computing: cores	5.5K	620
Pledges 2010: SPEC06	35K	4.4K
Pledges 2011: SPEC06	58.3K	6.4K

Storage	GRIF	LPNHE
DPM: TB (·)	2.2K	367
Pledges 2010: TB	2.24K	401
Pledges 2011: TB	2.96K	619

- Native xroot storage not accounted

LPNHE's Old Server Room

- Housed in an asbestos rich building

- Used to share server room with the rest of IT

- A "Grid" server room was built in 2006-2007

- Was sized taking planned move into account

- Investment was limited due to impending move

- The move was delayed by more than 2 years

- At the end we had 20kW of hardware off due to power and cooling limitations

LPNHE's New Server Room

- Requirements were set "in stone" in 2005
- No modifications were accepted: 2005 - 2010
  - Raised floor
  - End of aisle cooling
- No access rights
- First visit: 2010 Q1
  - Very low ceiling
  - Unknown load bearing capacity
  - Water leaks
  - No power from campus emergency generators
- Some "critical" (light) modifications were accepted

LPNHE's Relocation Plans

- All risk analysis, validation, certification was ignored by "higher-ups"

- Tried to minimize downtime

- All hardware had to be cleaned (asbestos)

- Despite the requested modifications, move was not delayed

- Very late access, started installing racks and cabling only one week before move

- The actual move: 2 day operation

LPNHE's Problems After Move

- Electrical, tripped breakers

- Chilled water cuts

- UPS in bypass

- Access, security

- Several false fire alarms

- Server room is still under construction
  - Almost no specifications or documentation
  - Delivery date: 2011 Q2

LPNHE's Infrastructure

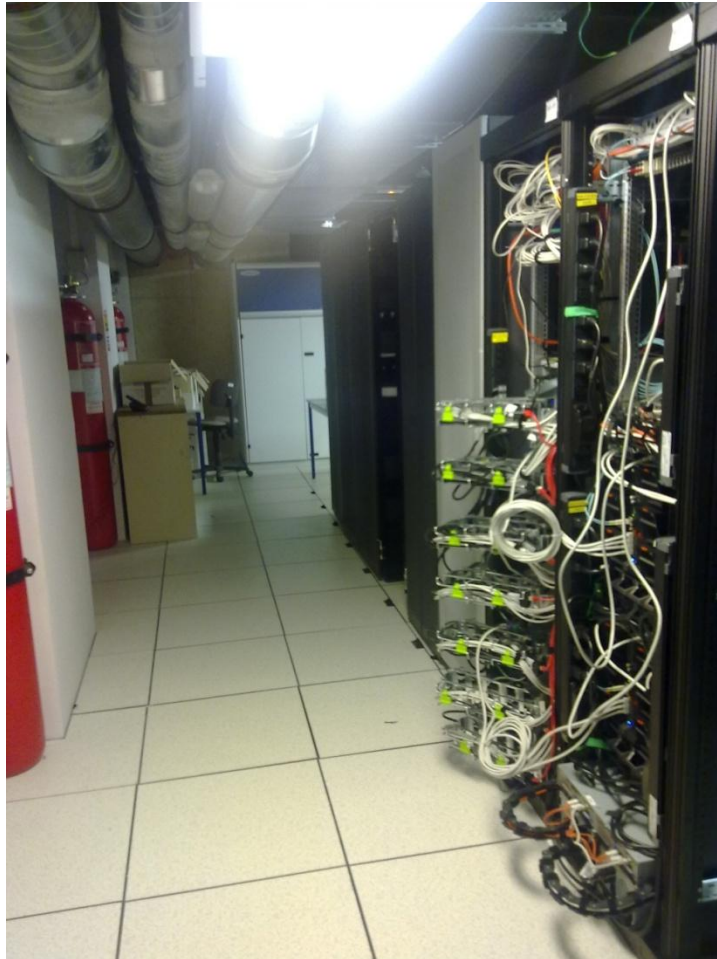|	Old Server Room	New Server Room
Network	4 x 1Gbps	10 Gbps
Power	80 kW	200 kW
Cooling	35 kW	200 kW
UPS	100 kW	200 kW
Racks	6	12
Fire detection	Minimal	Multiple, redundant
Fire suppression	Manual	Automatic, 300 m^3 Argon

LPNHE's
Server Room Pictures

LPNHE's
Server Room Pictures

LPNHE's Procurements

- Storage: 240 TB
  - DDN 6620
  - 2 I/O servers
- Computing: 2.5K SPEC-06
  - Dell PE M1000e
  - 16 Dell PE M610
- Network:
  - Core extension switch: 24 10G fiber ports
  - Top of rack switch: 48 port, 10G fiber uplink

LPNHE's Local Tools

- Cacti
  - Nice rrdtool based graphs
  - Very painful to manage (mouse clicks)
- Automatic Shutdown System
  - Scripted
  - Reacts to ambient temperature
  - Can react to UPS state
  - Staged:
    - Closes queues
    - Shuts down idle WNs
    - Shuts down all WNs
    - Shuts down the rest
  - Storage not powered off

LPNHE's
Local User Support

- Very limited due to lack of manpower

- Mostly for ATLAS users
    - Large backlog of requests

- Plan to push user support for non HEP VOs

- Documentation: seem to find ourselves always waiting for some calmness

LPNHE's Virtualization

- VMware in production
  - But no Grid related servers
- KVM
  - Just test servers for now
  - Hosts and guests managed through Quattor
  - No location awareness
  - Shared FS: GPFS
  - Very happy with results

LPNHE's Troubles

- kswapd taking 100% of CPU on WNs
- BIOS disabling RAM sticks (Dell PE M610)
  - already replaced 32 sticks
  - waiting for delivery of next batch
- Chassis network outage (Dell PE M1000e)
- Reduced available slots due to small HD size on older WNs

GRIF-LPNHE vs BDII v5

- GGUS ticket: [54678](#)

- BDII v5 performance degrades with time

- Using openldap-2.4 helps

- Keeping the LDAP DB in "tmpfs" also helps

- Problem remains unsolved
  - blamed on how openldap handles indexes

- After gLite-3.2 update 19
  - performance degrade even faster (at first?)

LPNHE's Plans

- Test kernel patch for 100% kswapd problem

- Split LCG CE

- Add a Cream CE

- Virtualization: Test StratusLab

- Cacti for GRIF through XML

GRIF's Plans

- More NGI (France Grilles) responsibilities:
  - Monitoring
  - Site certification infrastructure
  - Security
  - Early deployment
- Test CVMFS as a replacement for NFS in software areas