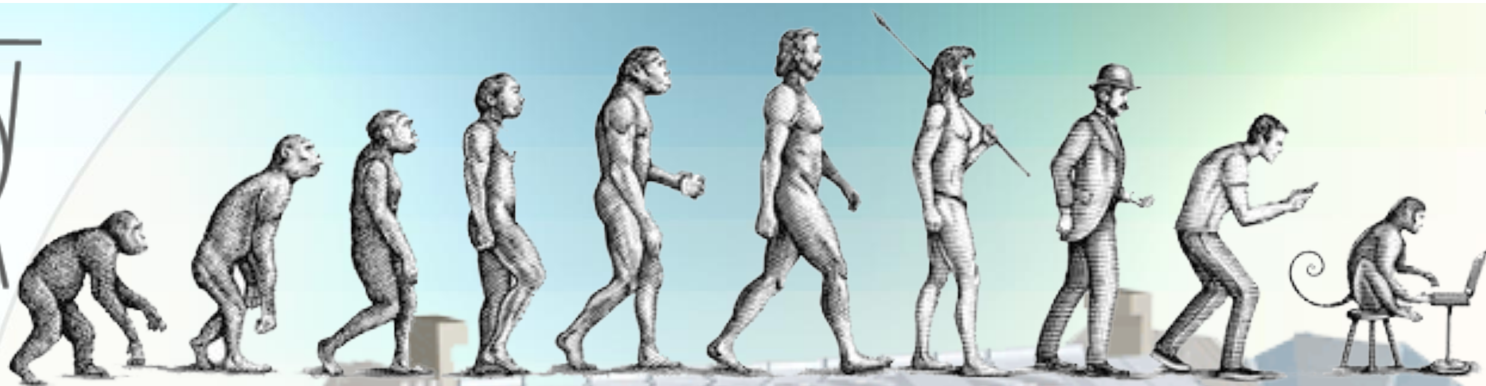


ATLAS TDAQ System Administration: an overview and evolution

ISGC 2013

Christopher Jon Lee

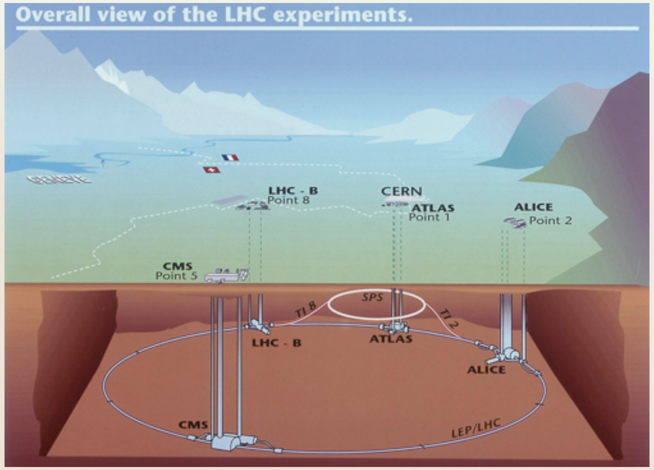
University of Johannesburg, South Africa
CERN

for and on behalf of the

ATLAS TDAQ SysAdmin team.

Introduction
LHC & ATLAS

- Large Hadron Collider, an accelerator, ~100m underground
- 27 kilometres in circumference
- Protons are accelerated in opposite directions at > 99.99 c
- Smashing together in the center of ATLAS, one of 7 experiments
- 600 million collisions per second, The collisions are measured in Luminosity or “number of particle per sec per square cm”, and are recorded by the Trigger and Data Acquisition system
- After 3 years of LHC runs, Long Shutdown 1 (LS1)



Introduction
Trigger and Data Acquisition

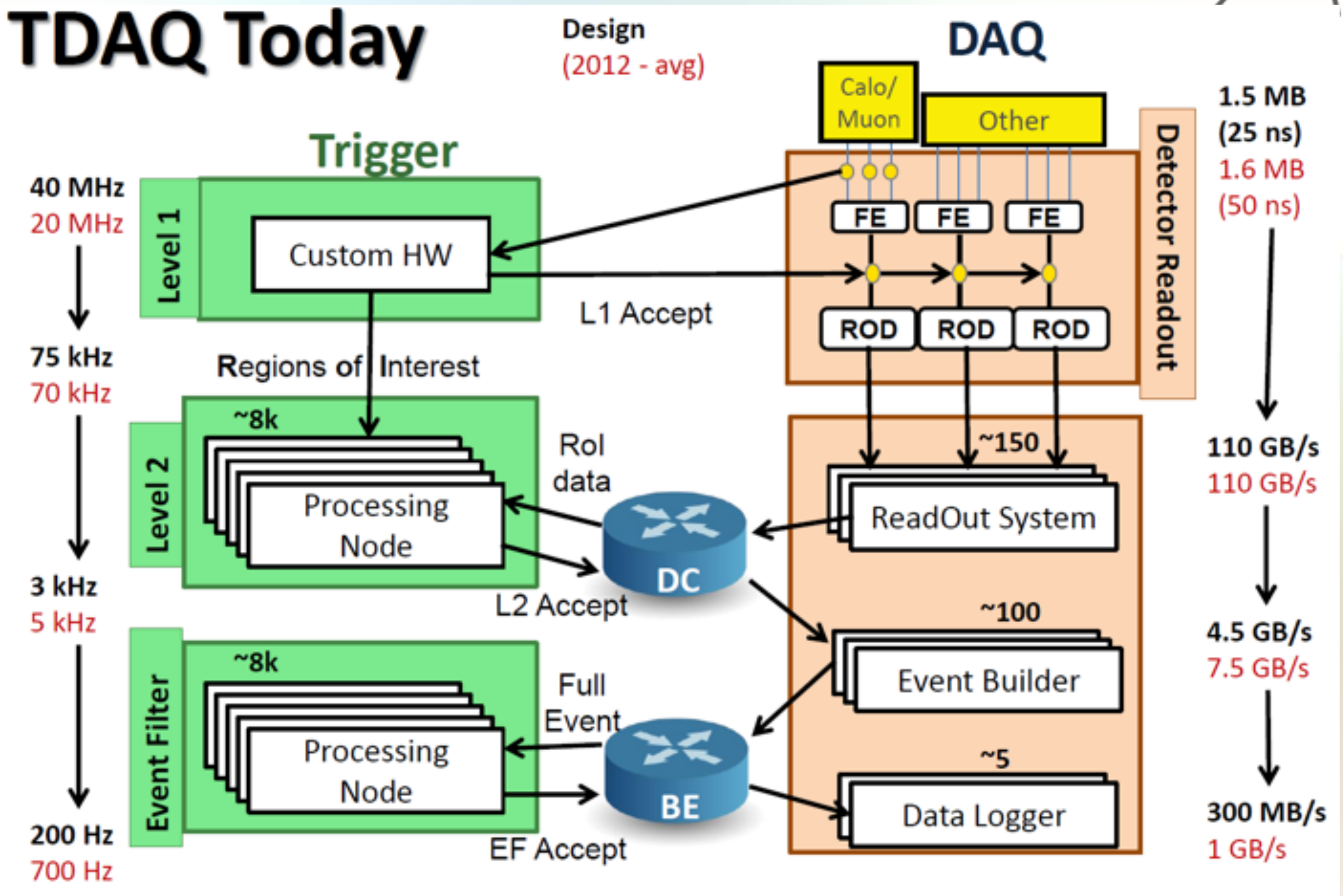
- Large online computing farm, used to process the data readout from ~100 million channels
- Ancillary functions (monitoring, control, etc.)
- Point 1
  - Approximately 100m underground, in close proximity to the detector (USA15)
  - On the surface near to the ATLAS Control Centre (P1, SDX1 & SCR)
- In General Public Network (GPN)
  - Laboratory for software development, prior to implementation into P1, recently been commissioned (TestBed)

The Racks

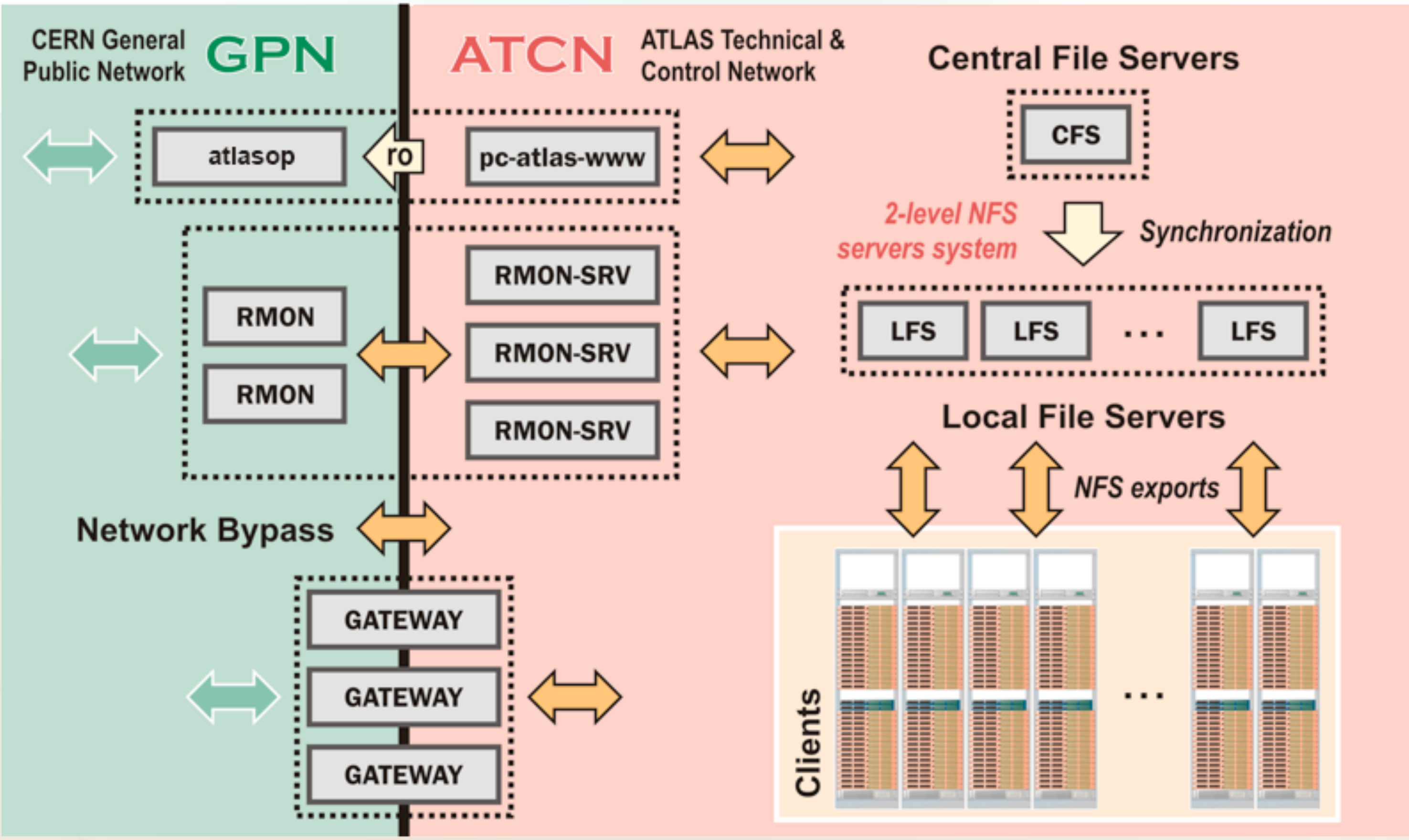
- USA15:
  - 220 racks deployed over 3 floors
  - 2009, ~70% filled, 1MW   2013, > 90% filled, 1.29MW*
  - 2.5MW of cooling can be provided
- SDX1:
  - 120 racks deployed over 2 floors
  - 2009, ~50% filled, 385 kW* 2013, ~91% filled, 709 kW*
- TestBed:
  - 22 Racks over 1 floor
  - 68% filled, 60kW of power. 11 Racks by TDAQ
  - 100 kW of Cooling can be provided
- Possibly an increase up to 90 racks (~4000 hosts) to deal with future luminosities

*estimated at ~6.5 kW/rack

TDAQ Design Overview



ATLAS Point1 Functional Layout



Redundancy

- Cant afford downtown & miss out on luminosity provided by the LHC
- Two centralised UPS lines with diesel backup generators
- Independent UPS lines available to SDX1 for mission critical equipment
- Redundant configuration for critical services
  - DNS/DHCP/NTP
  - LDAP
  - Domain Controller
  - LFS'
- NetAPP serves the most critical NFS and CIFS areas
  - 84 HDD's in RAID 6, dual head system with redundant fibre channels
  - Can survive the loss of one head, any one FC link, one FC interface on a shelf - just not of a whole shelf
- CERN's Tivoli for critical data
- SVN

Centralised and Distributed Storage Systems

- CFS
  - CFS nodes: user home directories, DAQ software distribution. Nagios RRD files, CIFS/NFS exports
  - Node configuration area
  - DAQ configuration area
  - Exports from the private to public webserver
  - Dedicated storage area shared among the gateways
- LFS
  - Synced from the CFS
  - Nagios Server
  - Exports to the clients

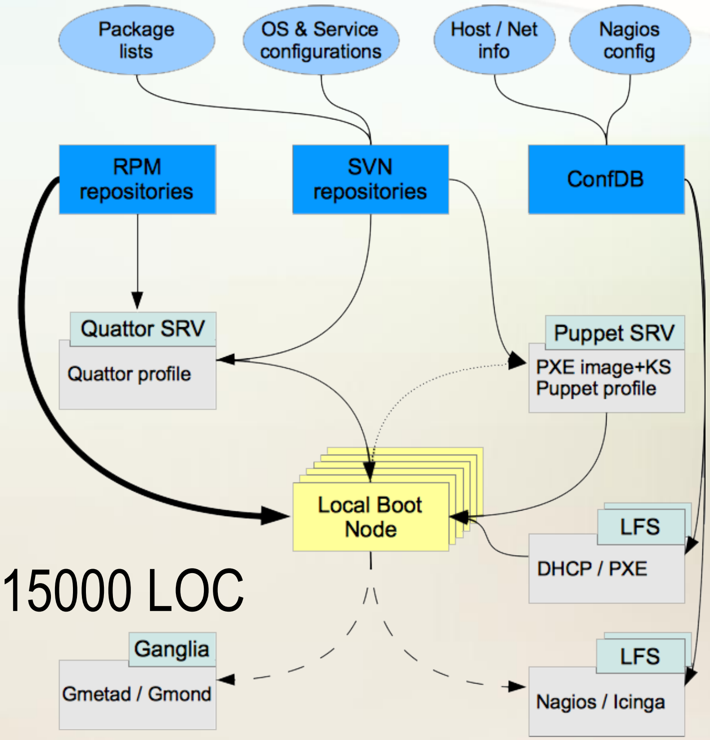
Clients
Local Boot

- Provisioning by PXE + Kickstart
  - DHCP + PXE provided by an LFS from ConfDB* information
  - Template-based kickstart files
- Quattor
  - CERN standard Configuration Management Tool
  - Production system, managing 237 hosts in the Online Farm
  - about 3500 LOCs in PAN (more dense code, less coverage)
  - Tight control on installed packages
  - Lack of flexibility for complex configuration/services dependencies
  - Multiple languages for implementing modules

*See slide 14

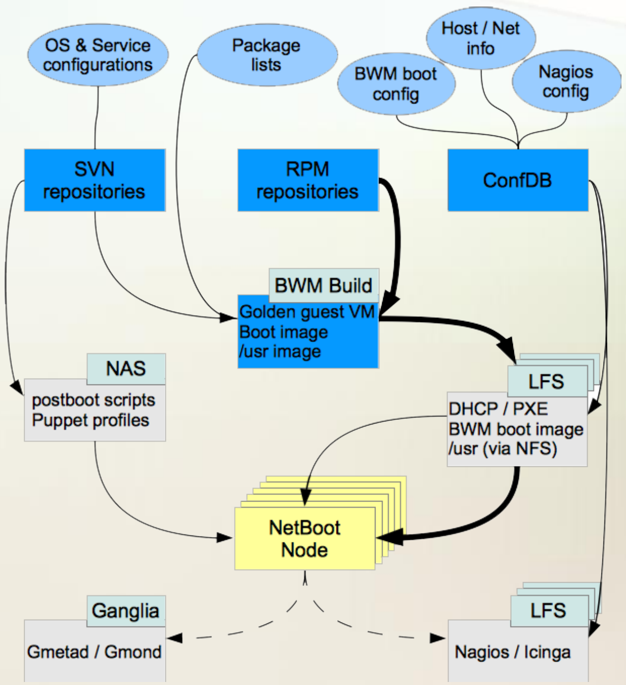
Clients
The Puppet Master

- Puppet
  - Widespread industry adoption, active development
  - Full featured, highly flexible
  - Gentler learning curve
  - Focus on consistence and idempotence
  - In production, manages exclusively 25 complex servers
  - Complements Quattor
- Migrate to Puppet and SLC6
  - our manifest code base has grown to ~15000 LOC
- Puppet is being adopted by CERN IT



Clients
Look Mom.. No disk

- The more components one has in a system, the greater the risk of failure, So… reduce any components that are not “needed”
- In ATLAS, extensive use of PCs with no operating system on disk
- “NetBooted” via PXE
- Advantages:
  - Ease of Maintenance
  - Reproducibility on a large scale
  - Reduced "installation" times
- Disadvantages:
  - requires ad-hoc development and support
  - not suitable for running servers



Clients
Netbooted Nodes

- ~ 2350 Nodes booted Scientific Linux CERN5 OS via PXE
- 75 Local File Servers (LFS), provide DHCP, PXE, TFTP, for booting, & `/usr` read-only directory via NFS
- Configuration of DHCP, PXE and boot parameters are provided by our in house built ConfDB*
- Boot With Me Tool (BWM)
  - Generates PXE boot images (kernel + RAMdisk root) and `/usr`
  - Uses reference SLC5 VM image as source
- BWM post-boot script system
  - Hierarchy of shell scripts, configures servers, disk, and NFS mounts.
  - Store on a central Network-Attached Storage, executed by the client
- Subversion repository is used to track changes of the BWM image creation configuration

*See slide 14

Clients
Puppet again..

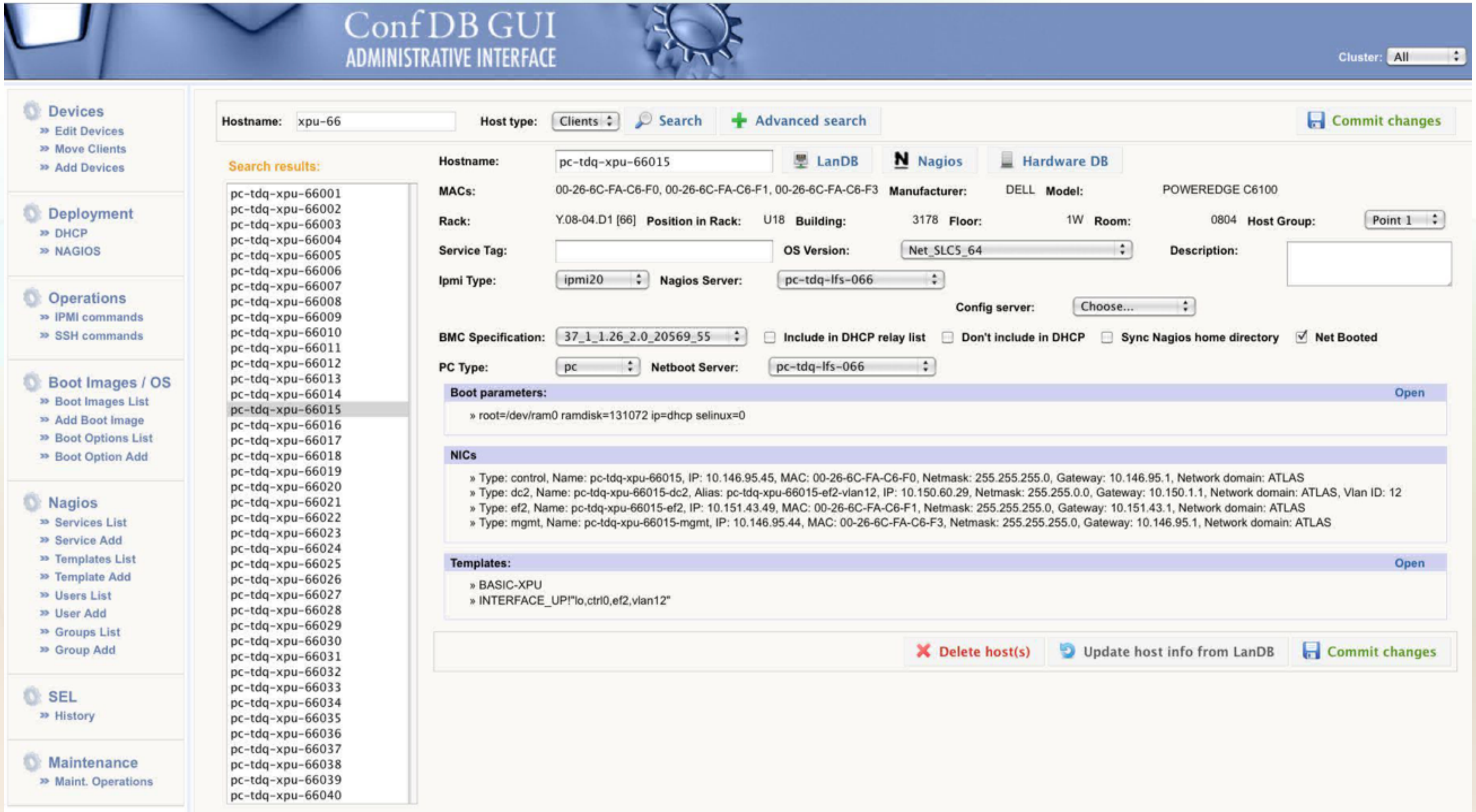
- Problems with BWM post-boot
  - No Permanent changes to a working system
  - Needs to be rebooted
  - Bash Scripts
- Move net-booted to puppet as well
  - Manifests hosted on CFS
  - No Daemon
  - Cron job that runs puppet and mails if problems on host
  - No need to reboot node for changes to take affect

- Minor problems on low resource machines and TDAQ /ROS drivers

Tools
Configuration Database

- Management of large number of NetBooted nodes is far from trivial.
- ConfDB was developed
  - based on a MySQL database
  - web based GUI
- Keep records of all the changes and system configurations
- Greatly help to speed up routine tasks:
  - Registering NetBooted clients, with data extracted from CERN's Network Database (LanDB)
  - Configuring boot options
  - Client to LFS assignment and migration
  - Deploying DHPC and NAGIOS configurations
  - Issuing IPMI and system commands on multiple nods in parallel

Tools
ConfDB UI



Tools
OS Repository Management

- Simple custom-made system for managing "time-frozen" snapshots of CERN package repositories
- Sufficient functionality through 2012, controlled upgrades, (theoretical) rollback capability
- We want more flexible functionality and easier management
  - Partial upgrades, client status reporting etc.
  - May adopt a third-party open source tool, e.g. Pulp
- Rollback/versionlock are possible in principle but not easy Puppet +yum does not offer the same detailed control as Quattor+SPMA

Monitoring
Nagios

- Nagios has been used since 2007
- Primarily monitors:
  - the health status of the OS
  - the hardware, selected services and network components
  - provides alerting for critical events
  - Significant development effort into hardware status via IPMI
- Separate Nagios server instance on each of the 80 LFS nodes
- Feeding data to a central storage area to a single MySQL Cluster
- A web-based interface was developed in-house
- Configurations integrated in ConfDB
- Over 5 years, machines monitored by NAGIOS increased from ~1500 to ~3000 in Point 1

Monitoring

GROUPS	TOTAL	ONLINE	OFFLINE	BROKEN	RESERVED
⊞ Gateways	6	6	0	0	0
⊞ WebServers	2	2	0	0	0
⊞ FileServers	2	2	0	0	0
⊞ DNS	2	2	0	0	0
⊞ CFS	1	1	0	0	0
⊞ LDAP	4	4	0	0	0
⊞ MYSQL	5	5	0	0	0
⊞ VH	2	2	0	0	0
⊞ ACR	128	119	0	0	9
⊞ SCR	49	33	16	0	0
⊟ TDQ	2186	2168	2	9	7
⊞ LFS	74	73	0	0	1
⊞ ONL	33	33	0	0	0
⊞ AMS	7	7	0	0	0
⊞ MON	32	32	0	0	0
⊞ GMON	6	5	0	1	0
⊞ ROS	157	156	1	0	0
⊞ SFI	48	48	0	0	0
⊞ SFO	9	9	0	0	0
⊞ DC	12	12	0	0	0
⊞ L2SV	8	8	0	0	0
⊞ XPU	1208	1196	0	8	4
⊞ EF	448	448	0	0	0
⊞ PRESERIES	131	129	1	0	1
⊞ RMON SRVs	3	3	0	0	0
⊞ NET-MON	7	7	0	0	0
⊞ SYS	3	2	0	0	1
⊞ SBC	161	153	5	2	1
⊞ PUB	14	12	2	0	0
⊞ DCS	102	98	4	0	0
⊞ MU-CALSRV	2	2	0	0	0
⊞ SWITCH	100	100	0	0	0
⊞ OTHERS	156	132	21	0	3
TOTAL	2926	2841	50	11	20

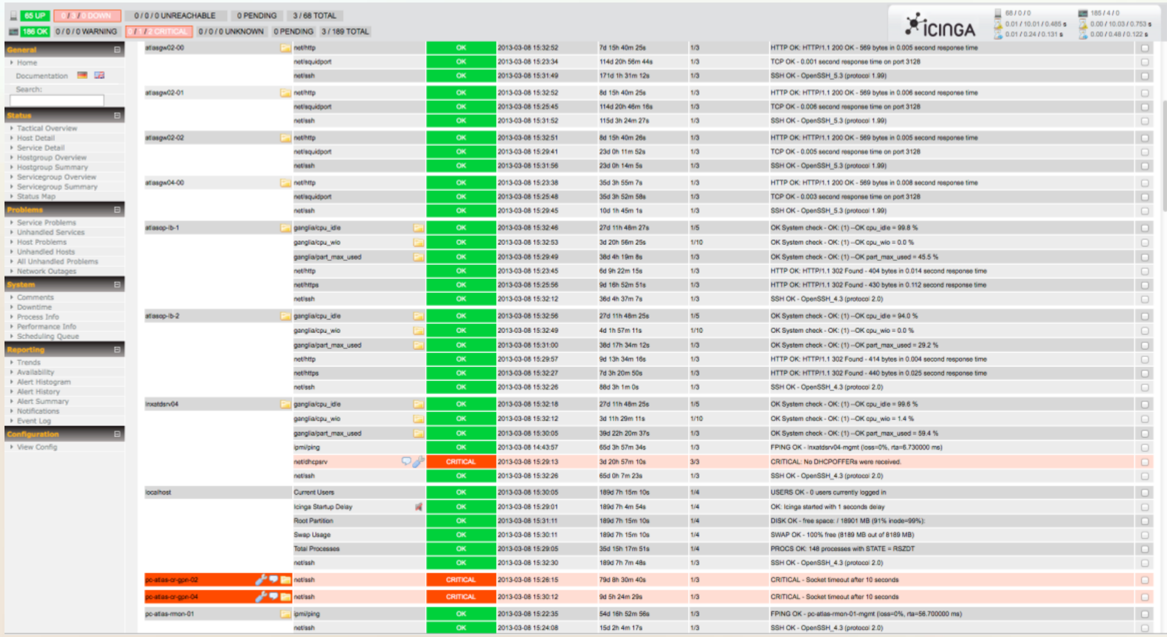
Monitoring
Ganglia

- 2011, we introduced Ganglia for performance monitoring and trending
- We are currently evaluating on a smaller scale system
- Used primarily for special purpose nodes
  - local boot nodes
  - Monitoring
  - Online sub-farms
- Single central server
- Round-Robin Database (RRD) caching deamon used for performance
- Ganglia is used by many Grid facilities
- Integration with Icinga

ICINGA and Gearman

Monitoring

- Icinga+Gearman provide active checks with distributed scheduling
- Alerting
- Icinga+Gearman adopted by CMS
- Icinga can reuse Nagios Plugins, and much of Nagios configuration



Security
Remote Access Subsystems

- Security at Point 1 is of utmost importance
- Access to the ATLAS Technical and Control Network (ATCN) is highly restricted
- only allowed via one of the following gateway systems:
  - ATLAS Point 1, Allowing expert users access to their restricted machines via SSH or SCP protocols
  - ATLAS Remote Monitoring System provides the graphical terminal services required for organising the remote participation in the ATLAS sub-detector monitoring shifts
  - ATLAS DCS Windows Terminal Servers, allowing DCS experts access to the supervisory control and data acquisition (SCADA) systems in ATLAS
- Host and network based accounting, security monitoring and intrusion prevention systems are configured on all the gateways

Security
Role Based Access Control

- Own user database in the form of an LDAP server based on OpenLDAP software
- Standalone but for consistency it is synchronized with IT
- User based authentication, NOT group based authentication
- Slave Windows Domain Controller using the CERN NICE credentials.
- Local service accounts, authentication (passwords) in LDAP
- Role Based Access Control (RBAC) authorization system
- 150 unique roles in a hierarchical structure and~1600 Users

- BUT..

  Access ONLY authorised by ShiftLeader, to on-call experts, during shifts

Virtualisation

- More in an "opportunistic" fashion than as a full battle plan
- Machines currently running as VM's
  - Gateways
  - Few Windows services
  - Core Nagios servers
  - Public Nodes
  - Domain Controllers
  - Development Web servers
  - Puppet & Quattor
- Xen on Gateways (Bastion Hosts) KVM with SLC6 for everything else
- Currently :
  - 35 machines in Point 1
  - 11 in GPN
- Expected:
  - Additional ~100 DCS
  - ~1500 for Sim@P1

Sim @ P1

- HLT Farm would be mostly idle during Long shutdown 1 (LS1)
- Use it for ATLAS simulations jobs considerable fraction (>10%) of ATLAS Grid
- Support from BNL during the setup phases
- Support from NetAdmin for dedicated network
- Preserve security of ATCN by isolating VMs and VLANs
- Tests are ongoing in TDAQ Lab4 Test Bed

#	Type	CPU Cores: non-HT / HT	Memory	Local disk
341	Dell PE 1950	8 / 8 (*)	16 GB	80 GB
320	Dell PE 6100	8 / 16	24 GB	250 GB
832	Dell PE 6100	12 / 24	24 GB	250 GB
1493	**Total**	**15272 / 27816**	**33 TB**	**315 TB**

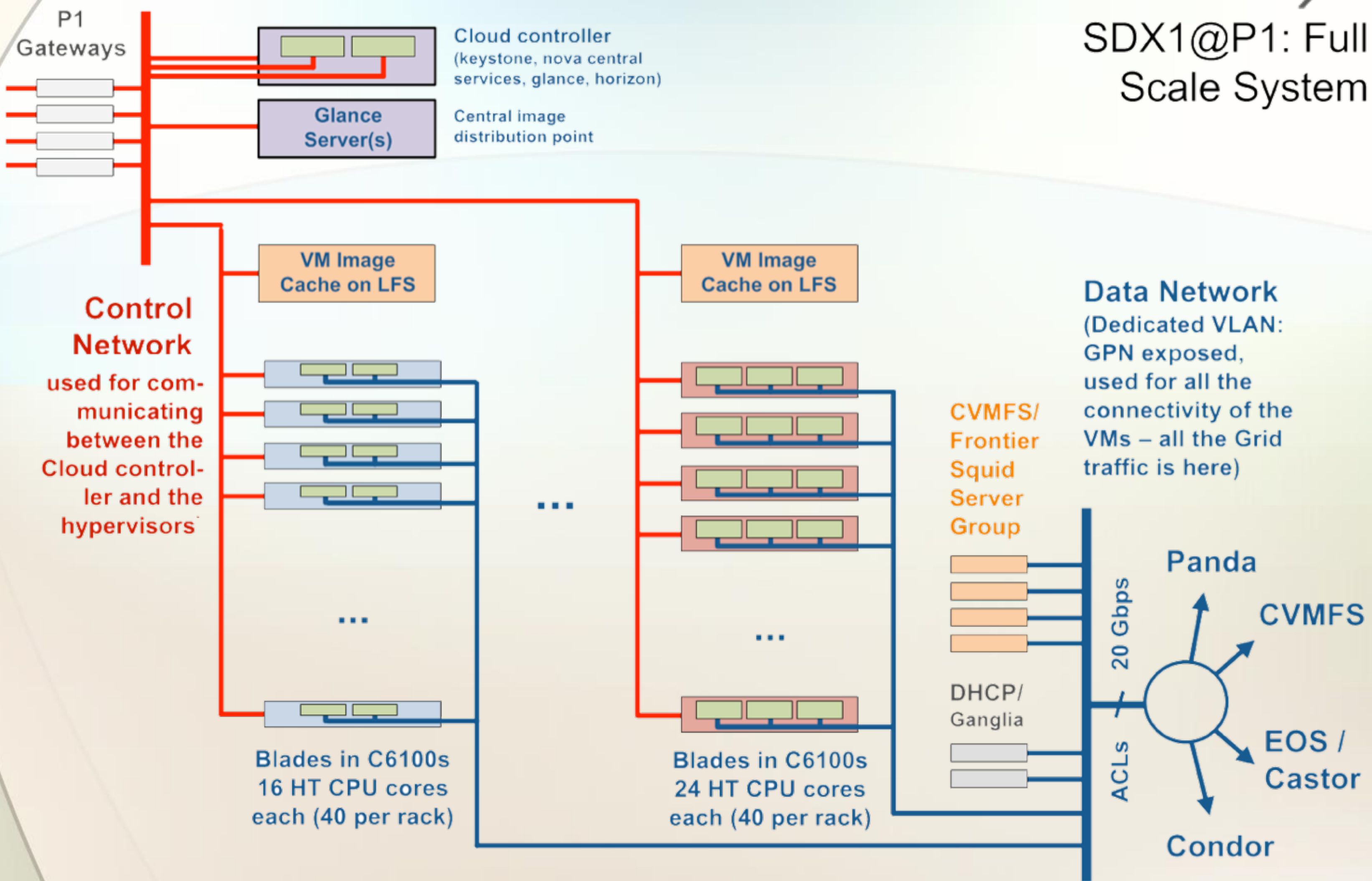
(*) HT not enabled

Sim @ P1
The Plan

- *OpenSTACK,* as an overlay infrastructure
  - Provides necessary management of VM resources
  - Support & control of physical hosts remain with TDAQ
  - <u>Delegate VM Farm support to Offline Operations</u>
- Easy to quickly switch from HLT $\leftrightarrow$ GRID
  - During LS1: monthly full-scale test of TDAQ sw upgrade
  - Can be used in the future also during short LHC stop
- BNL, CERN IT, CMS
  - Sharing experiences
  - Support if needed
  - ATLAS is already, successfully using BNL cloud resources

Sim @ P1

SDX1@P1: Full Scale System



Conclusion

- 3 years of LHC run point to a good, linear scalability of the LFS-based architecture, OK overall, but to be kept under control
- LFS' have the performance for serving more than one rack of clients, BUT:
  - Network locality decreases load on control network
  - More clients would be affected in case of failure of an LFS
- Netboot limited functionality, great flexibility in OS migration
- Localboot, used to be harder to configure, Puppet changing that
- Overall, the current architecture is sound and performing well, and the possible changes would increase complexity and not give drastic improvements
- We have decided to focus on trying to simplify and streamline the system where possible, making it more robust and maintainable

Backups / Spares

SFO

- Sub Farm Output
- Historically (2003), 200 Hz average output rate
- Defined as a compromise between physics goals and yearly data production volume and therefore offline storage requirements ... at least this is what the elders say ...
- 10 years later, in 2012, data storage and network capabilities allowed to run at an average of 600 Hz
- For 2015+, the proposal is to run at 1kHz average output
- The offline community has still however to acknowledge this new target