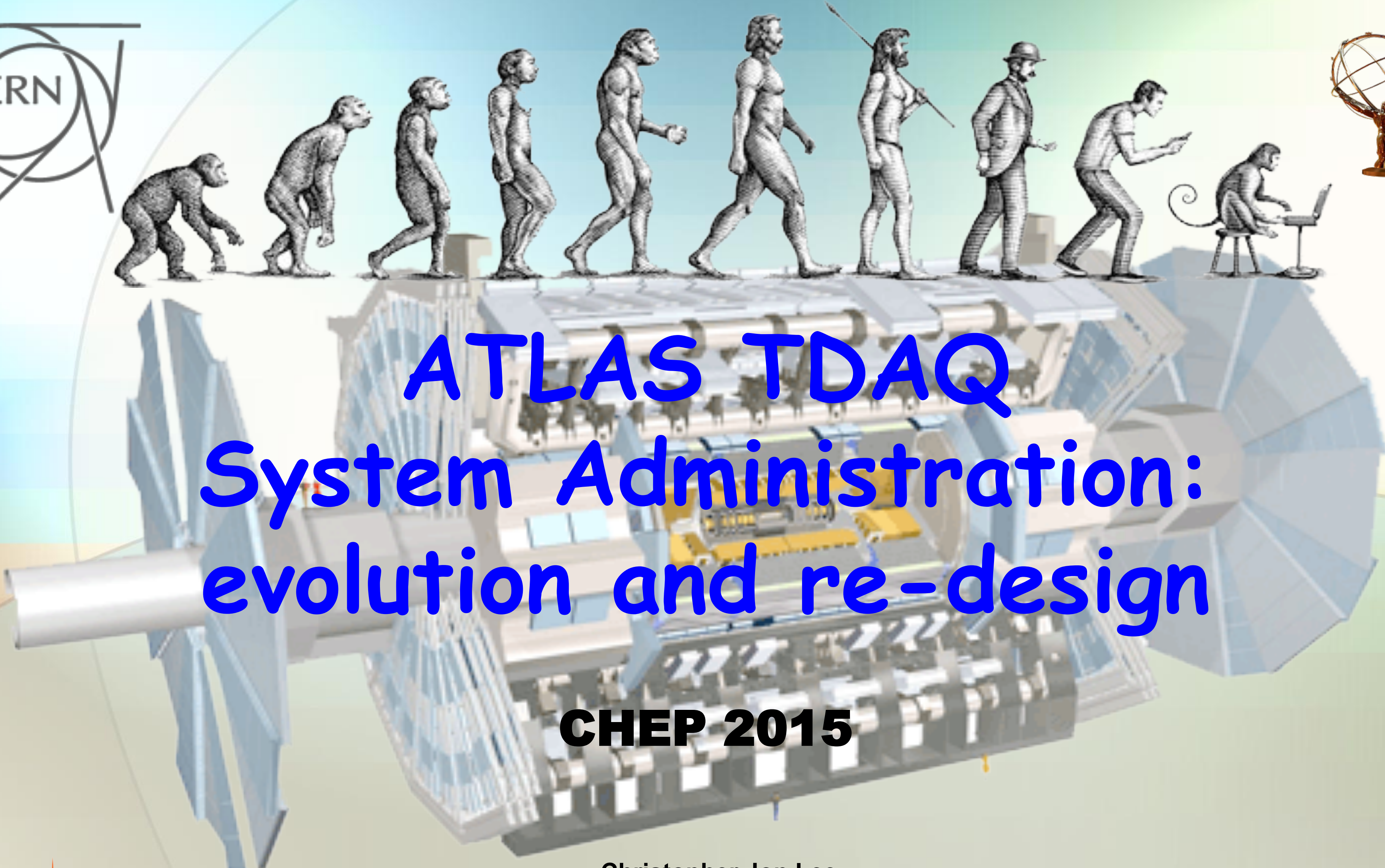


ATLAS TDAQ System Administration: evolution and re-design

CHEP 2015

Christopher Jon Lee

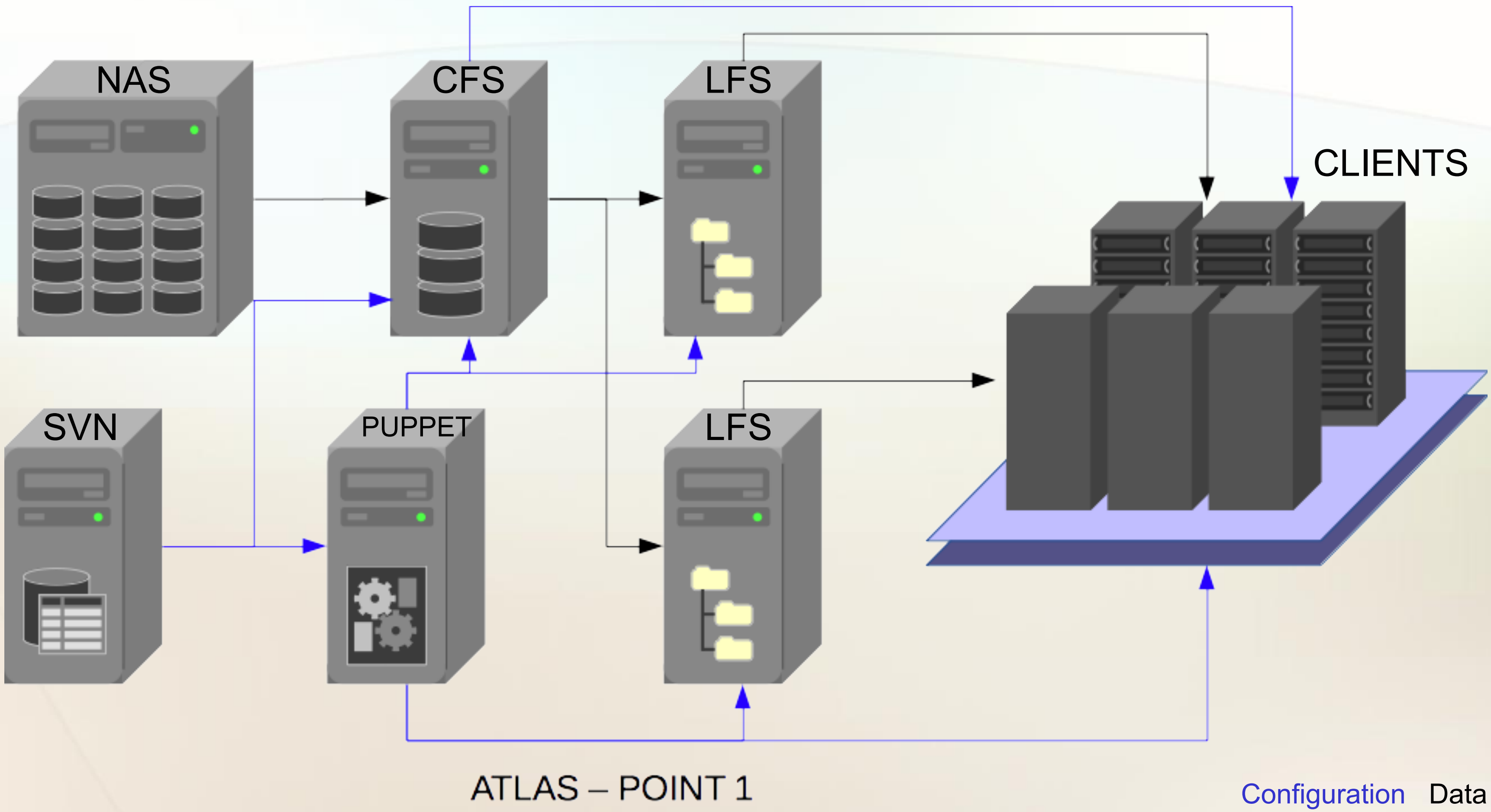
University of Johannesburg, South Africa
CERN

for and on behalf of the

ATLAS TDAQ SysAdmin team

Overview

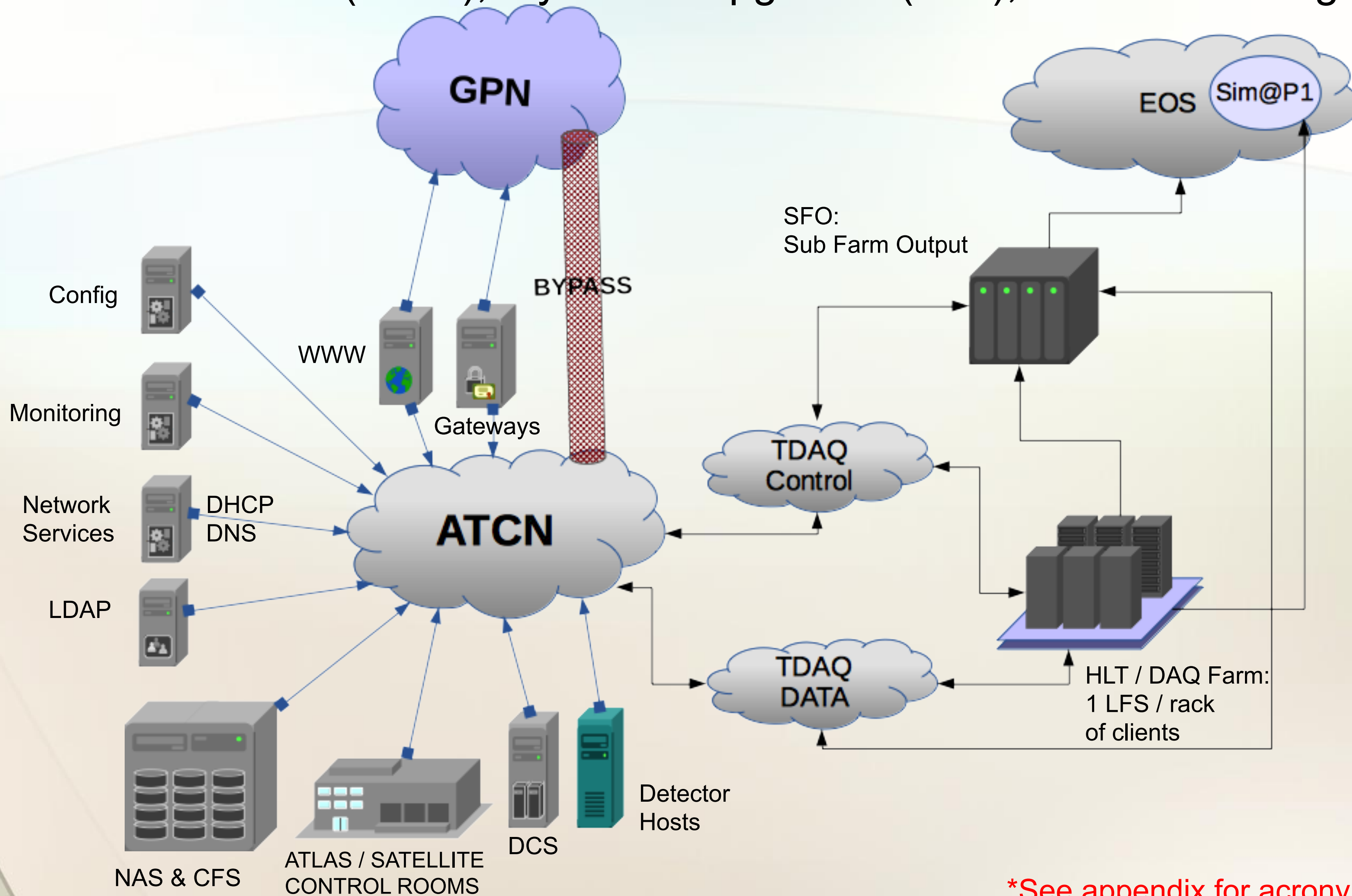
After 3 years of LHC beam (Run1), 2 years of upgrades (LS1), Restart has begun…



*See appendix for acronyms

Overview

After 3 years of LHC beam (Run1), 2 years of upgrades (LS1), Restart has begun…



*See appendix for acronyms

OS Upgrade

- Scientific Linux CERN
  - Only supported Linux OS in use
  - Full support direct from on-site experts
  - All Linux machines are now running SLC6
    - will remain the OS (Major) version for Run 2
- Windows OS
  - Used by Detector Control System (DCS) for **one** specific application
  - SLC6 hosts Windows VM
  - Windows VM is managed by DCS

- During beam, NO changes are made to the running system

Local boot vs Net boot

- 664 Local boot – Servers, DCS, TDAQ/SysAdmin Infrastructure
- 2392 Net boot – HLT Farm, Read Out Systems, Single Board Computers, etc
- Local boot (*Standard installation with boot from disk*)
  - provisioning by PXE + Kickstart + puppet labs
  - DHCP + PXE provided by an LFS from Configuration Database (See slides below)
  - template-based kickstart files
- "Net Boot" via PXE
  - 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
- Each reboot, is essentially a fresh clean OS

<u>Advantages:</u>

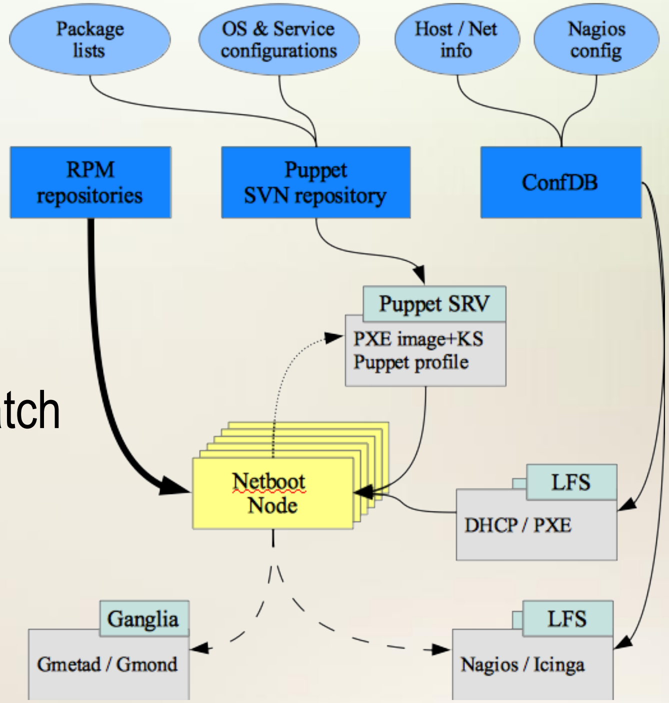
- ease of maintenance
- reproducibility on a large scale
- reduced HW replacement times

<u>Disadvantages:</u>

- requires ad-hoc development and support
- not suitable for running servers
- less flexible

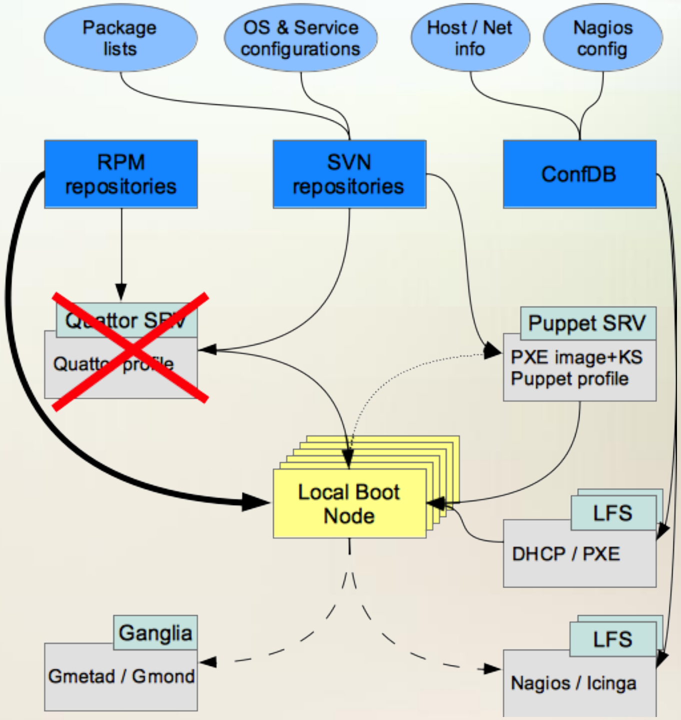
Netbooted – Redesign for SLC6

- A completely new netbooting system compared to SLC 4 & 5
  - based on NFSroot and customized to our needs
    - only R/W areas are kept in RAM (e.g.: /etc, /var, …)
    - "bind-mounts" overlaid on R/O NFS mount of / from the LFS
    - gives the users more free RAM for running their apps
  - Image created by puppet labs in a chrooted environment
  - **NO** "Golden Image"
    - always able to rebuild from versioned config
- Support for old hardware
  - 32 bit non-PAE kernel provided and maintained by CERN IT (on a best effort basis)
  - ELF image for non-PXE clients requires private patch of mknbi package



Configuration Management Systems

- With such a large cluster of machines CMS's are the only way to sanely control what happens on machines in a large farm
- Quattor was retired as it was becoming obsolescent
- puppet labs chosen over Chef and CF Engine
  - previous experience by existing Sys Admins
  - during LS1 CERN IT also adopted puppet labs
  - WLCG applications mostly puppetised
- All systems are Puppet controlled
  - local boot : Daemon run, every 30 min
  - net boot: `puppet --apply` via hourly cron job
  - no need to reboot in order to apply simple configuration changes
- Very similar system between the two
  - code re-usability
  - easier to maintain



ConfDB GUI ADMINISTRATIVE INTERFACE version2

- ConfDB is our core Configuration DataBase
  - PHP based web UI, Python for utilities and REST API
  - configuration "state" database of all systems
    - DHCP details
    - ICINGA checks
    - operational status
    - boot type / parameters
  - entire system other than ConfDB is maintained by "code"
  - interface between CERN IT Databases and puppet labs
  - web UI provides various tools to ease cluster management – ssh, IPMI etc
- Included an OS release system for SLC6+
  - different release versions can run on different machines
  - useful for testing new versions and/or revert changes in case of problems
- More functionality added to REST API, used by puppet and other tools e.g.
  - network interface configuration (including bonding)
  - getting machine status in puppet labs and configure it accordingly (e.g. TDAQ or Sim@P1*)

* See slide 9

Virtualisation

- Mostly CORE and TEST systems, NOT for DAQ/HLT
- No cloud-like approach
  - no shared storage, privileged simplicity
- Instead of a single redundant system, rely on multiplicity of systems

✓ 50 VMs in Point 1 ✓ ~100 DCS	✓ 51 VMs in GPN ✓ ~2700 Sim@P1 *

- Currently running as VM's
  - gateways
  - DCS Windows services
  - icinga
  - public Nodes
  - domain controllers
  - development web servers
  - puppet labs
  - LDAP
- For new VH hardware, tested and really happy with the results of FlashCache
  - huge improvement on disk IO vs cost

* See next slide

Simulation @ P1

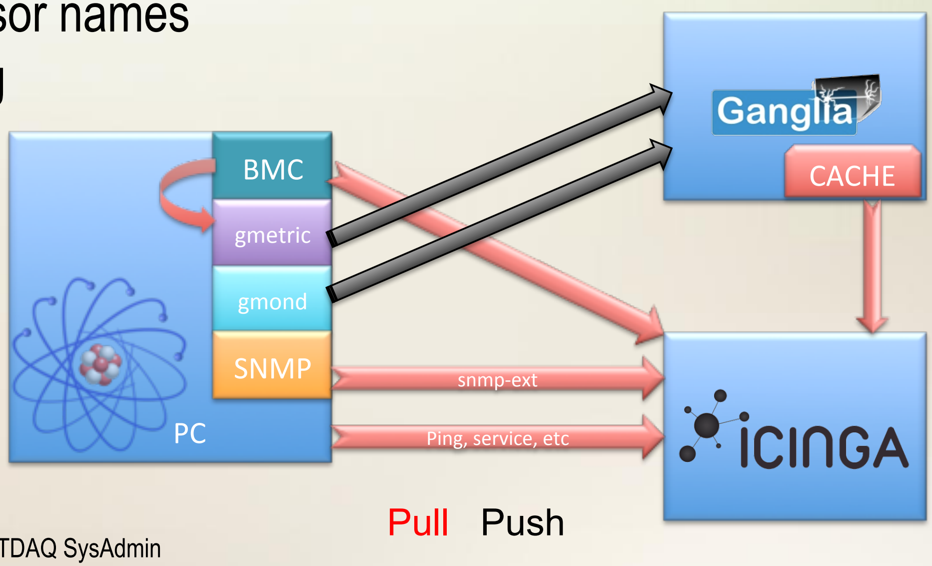
- Sim@P1 is the opportunistic use of the existing TDAQ HLT farm as a grid cluster
  - allow non-utilised resources to be exploited for ATLAS Prod jobs
- Virtual machines on top of the HLT machines, acting as computing nodes interconnected through a Virtual LAN on the data network
  - VMs and VLANs isolate the offline computing nodes network
    - no interferences with ATCN & DCS
    - security
- Communicate with the outside world (GPN) via a logically separated link to CERN IT
  - ACLs are in place for allowing only traffic towards the needed CERN services (Castor / EOS, Condor, etc.)
- More than 1300 nodes of the HLT farm are now able to run Monte Carlo, high CPU intensive jobs
  - EVGEN and SIMULATION only, by design
- Produced 1.7 billion Monte Carlo events since 1st Jan 2014
  - limited by downtimes during LS1 (cooling, power interventions etc)
- Switching between states is controlled by ATLAS control room
- For more information, see our poster on the topic
  - "Design, Results, Evolution and Status of ATLAS simulation in Point 1". Poster by Franco Brasolin

ICINGA & Ganglia

- Ganglia as collector for performance/health metrics, high scalability with rrdcache
- Ganglia-web provides advanced user interface to historical data in RRDs
- Icinga replaces Nagios
  - provides active checks and alerting
  - can use Ganglia data
  - can reuse Nagios plugins and much of Nagios configuration

HW monitoring via IPMI

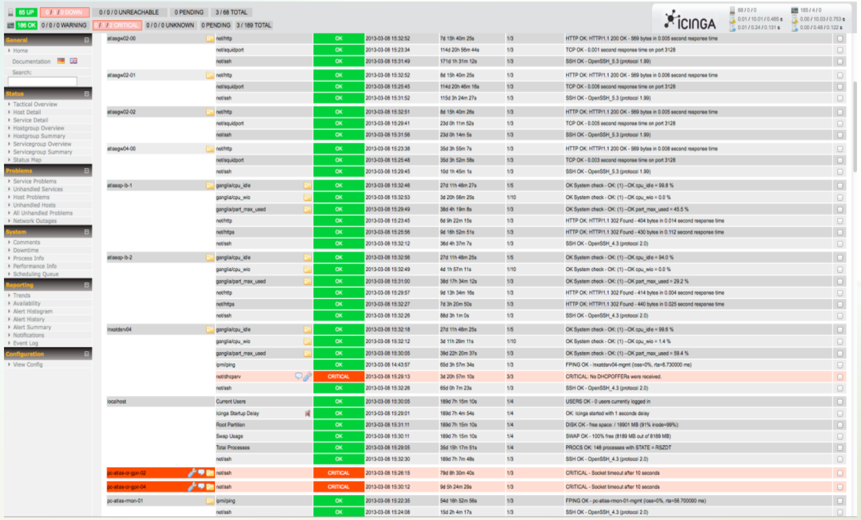
- Complete rewrite during LS1, work still in progress
- New version based on OpenIPMI, previous based on IPMItool
  - unique sensor ID's vs (unstable) sensor names
  - better performance with SDR caching
- Local readout fed to Ganglia
- ICINGA monitors SEL, Sensor OK state, specific values via Ganglia
- IPMI varies with vendor, type, version… always catching up



Monitoring & logging implementation

ICINGA

- One VM for core systems: ~570 nodes, ~5000 checks
- One PC for farm systems: ~2200 nodes, ~31000 checks
- Users receiving status notifications
- Testing Icinga 2 for distributed scheduling, configuration and performance

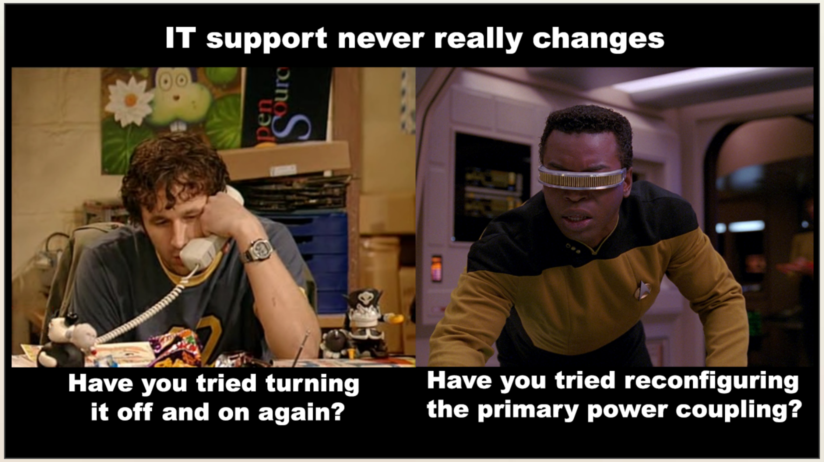


System Logs management

- Rsyslog on all machines, also as collector for remote (replaces syslog-ng)
- Remote logging to LFS or central syslog servers
  - net boot: 2 day retention (local), 30 day on LFS
  - local boot: 30 day retention
  - exposed systems: 12 weeks
- Remote logging for security critical servers to CERN IT
- Investigating central collection & analysis tools: Splunk, ELSA, logstash + elastic search

Summary

- LS1 was anything but a "shutdown" for our Team
- We have streamlined and improved the Point1 system
  - many more tasks are now fully automated, with very little human intervention needed
- Puppet, cleaner and easier to maintain
- Monitoring much more comprehensive than before
  - provides many more checks, still rapidly evolving
- Still Investigating Open LDAP (2.4.39) issues with opening/closing connections.
- Hoping Run 2 will provide us some "quiet time" to clean up



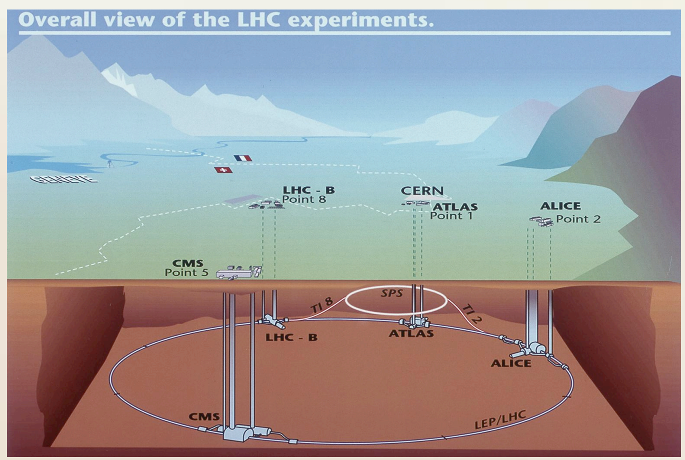
Glossary

ATCN: Atlas Technical and Control Network
CFS: Central File Server
DCS: Detector Control Systems
GPN: General Public Network
HLT: High Level Trigger
LFS: Local File Server
LHC: Large Hadron Collider
LS1: Long Shutdown 1
NAS: Network Attached Storage
PXE: Preboot eXecution Environment
ROS: Read Out System
SBC: Single Board Computer
SLC: Scientific Linux: Cern edition
SVN: Subversion
TDAQ: Trigger and Data Acquisition
WLCG: Worldwide LHC Computing Grid

BACKUPS / SPARE

Introduction
LHC & ATLAS

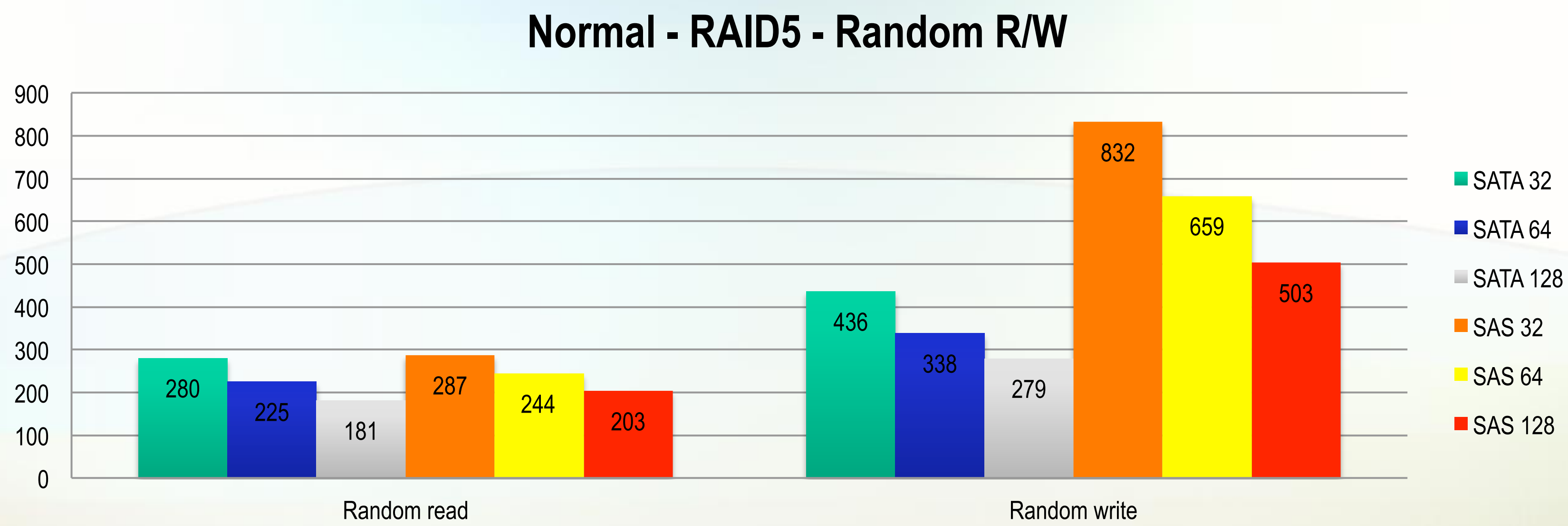
- Large Hadron Collider, an accelerator, ~100m underground
- 27 km in circumference
- Protons are accelerated in opposite directions at 4 TeV
- Smashing together in the center of ATLAS, one of 7 experiments
- 600 million collisions per second
- Data about these collisions are recorded by the Trigger and Data Acquisition system

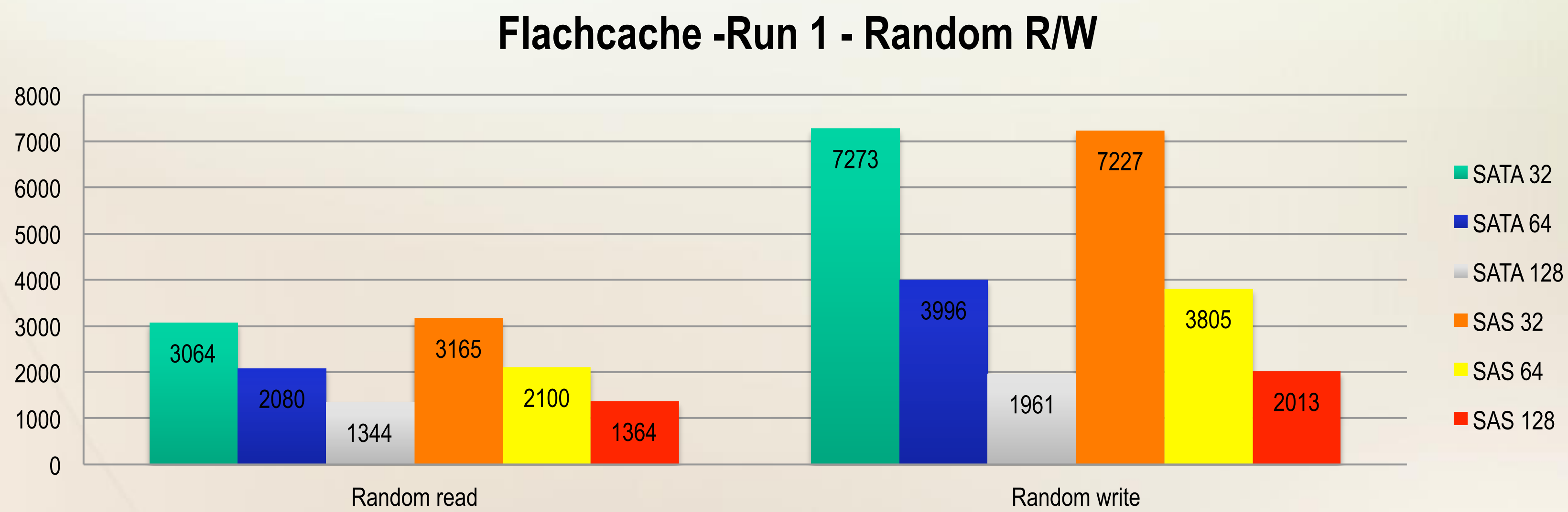


“Private networks” now managed by IT

- Private networks for us:
  - isolated Networks within the ATCN
    - high voltage power supplies
    - ATCA / VME crates
- Use of unregistered private networks is against CERN IT security policies
- Unmanaged switches are not supported anymore (= no spares)
- Integration with CERN IT
  - standardised system
  - no RFC1918 networks
  - all devices on the network registered and traceable
  - increased security
  - increased management
  - 4 hour piquet-like support from CERN IT

FlashCache tests





Sim@p1 -> TDAQ Mode

- Switching from control room shifter through a WEB gui when a LHC beam stop longer than 24h is foreseen
- Fast and automated:
  - from sim@p1 to TDAQ ~ 12 min
  - from TDAQ to sim@p1 ~ 1h
  - emergency switch from sim@p1 to TDAQ: 100 s

Monitoring system upgrade

Nagios v2 + Custom UI

- Old production system: stable, but complex
- Multiple standalone Nagios servers
- Central configuration from ConfDB
- Central storage in MySQL/cluster and RRDs
- High I/O load on MySQL and NetApp can become a bottleneck
- Custom WebUI requires maintenance
- Nagios v2 is obsolete

Nagios v3 + Custom UI

- With the LFS SLC6 migration complete, this was put into production
- One single MySQL server replaced the 4 machines cluster used for storing Nagios data
  - better performance
  - easier maintenance

Atlas Control Room

- Completion of the plan for PCoverIP migration
  (Desk remote technology over TCP/IP)
  - KVM (keyboard, video, mouse) from SDX1 to ACR over network
  - each machine has 1 or 2 PCoverIP cards (depending on number of screens)
  - each desk has 1 terminal client
- A joint collaboration between OPM and TDAQ NetAdmins+SysAdmins
- Full redundancy
  - 2 switches in SDX1 and 2 switches in SCX1.
  - Cards and Terminals have dual connections
- Less clutter
  - 2 optical fibers between SDX1 and SCX1, providing two independent connections, replaced ~100 copper cables
- Delayed updating existing systems as current available market hardware provides no redundancy and no major improvements

Satellite Control Room

- Experts base of operations
  - provides similar workspaces as in ACR
- Advanced debug tools allowed
- extensive use of IMACS and Mac MINI running SLC6
  - allowed for long lasting, system that can be replaced

CRD (KDE)

- Control Room Desktop
- Provides tools required per “seat” in the ACR
- No direct terminal access
  - controlled and authenticated access to terminal windows
- This implementation needed a version of KDE not available in SLC6