RSA WLCG Tier2 Facility plus other

Sean Murray
ALICE
CHPC
CSIR

May 4 2017

Outline

Commitments

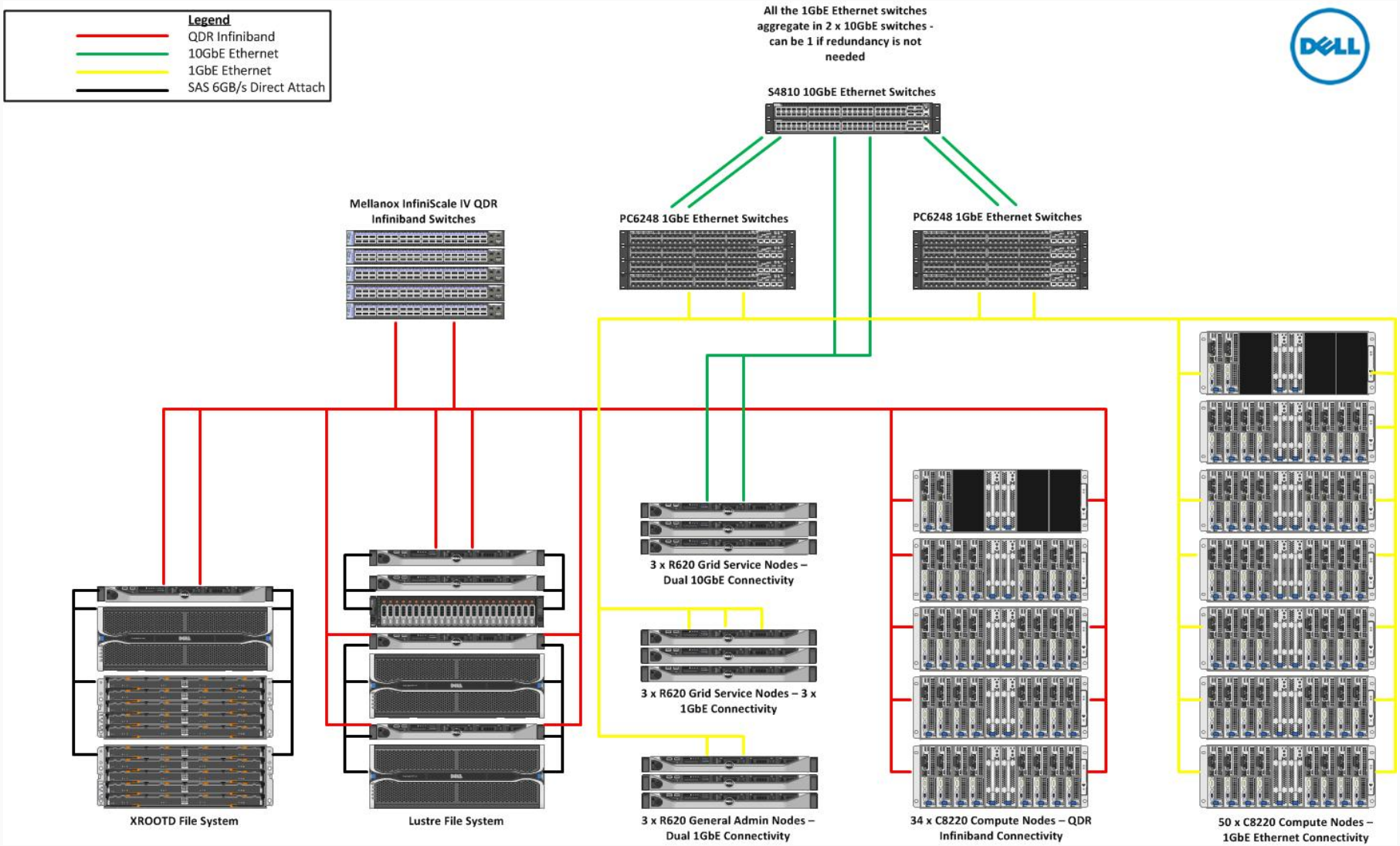
According to Tender :

- ALICE 600 cores
- ATLAS 600 cores
- ALICE 400TB
- ATLAS 400TB

According to :
https://wlcg-rebus.cern.ch/apps/pledges/resources/

- 6000 HEPSPEC06 cores (560 of our cores)
- 100TB storage
- All ALICE.

Computing Infrastructure



Current hardware

- 50 nodes of 48 cores 192GB RAM and 1.6TB of SSD, 1G ethernet
- 34 nodes of 48 cores 96GB RAM and 1TB, FDR infiniband, 6 “stolen”
- 100TB of Lustre on the 34 nodes with FDR infiniband.
- 9 management servers, lower spec
  - compute element (head node,ce),
  - storage element 2 redirectors, 2 storage nodes with direct attached multipath storage
  - authentication, user interface (gone), monitoring, provisioning.

Current Storage

- 383TB EOS for ALICE, down from 440TB
- 252 TB EOS for ATLAS, down from 400TB
- 107 TB lustre for 34 nodes.
- 104 TB EMC for ATLAS.

Reduction in data sizes is due to reorganisation for reliability.

Current Performance

- ———-465k ALICE jobs in last year
- ———- Avg concurrent jobs 704.
- ———ALICE 358TB storage properly live 24 June, Consumed 6.4TB
- ———Data traffic consumed 70TB in and 60TB out in the last 3 months
- ———11MB/s in and 7.5MB/s

pic from monalisa TODO ngi_africa accounting screen shot for alice

Resource Centre	Aug 2016	Sep 2
ZA-CHPC	236,378	162,

Availability / Reliability

CHECK NEW DATA

Function	May	April	March
Availability	90	97	91
Reliability	100	90	90

Problems in monitoring infrastructure in June, pending recalculation.

explain the decrease of running jobs. impacts of network ....
error_e caused by errors of things taking too long to get data.
random increases in jobs due to no jobs matching (graphs) and

Grafana ALICE

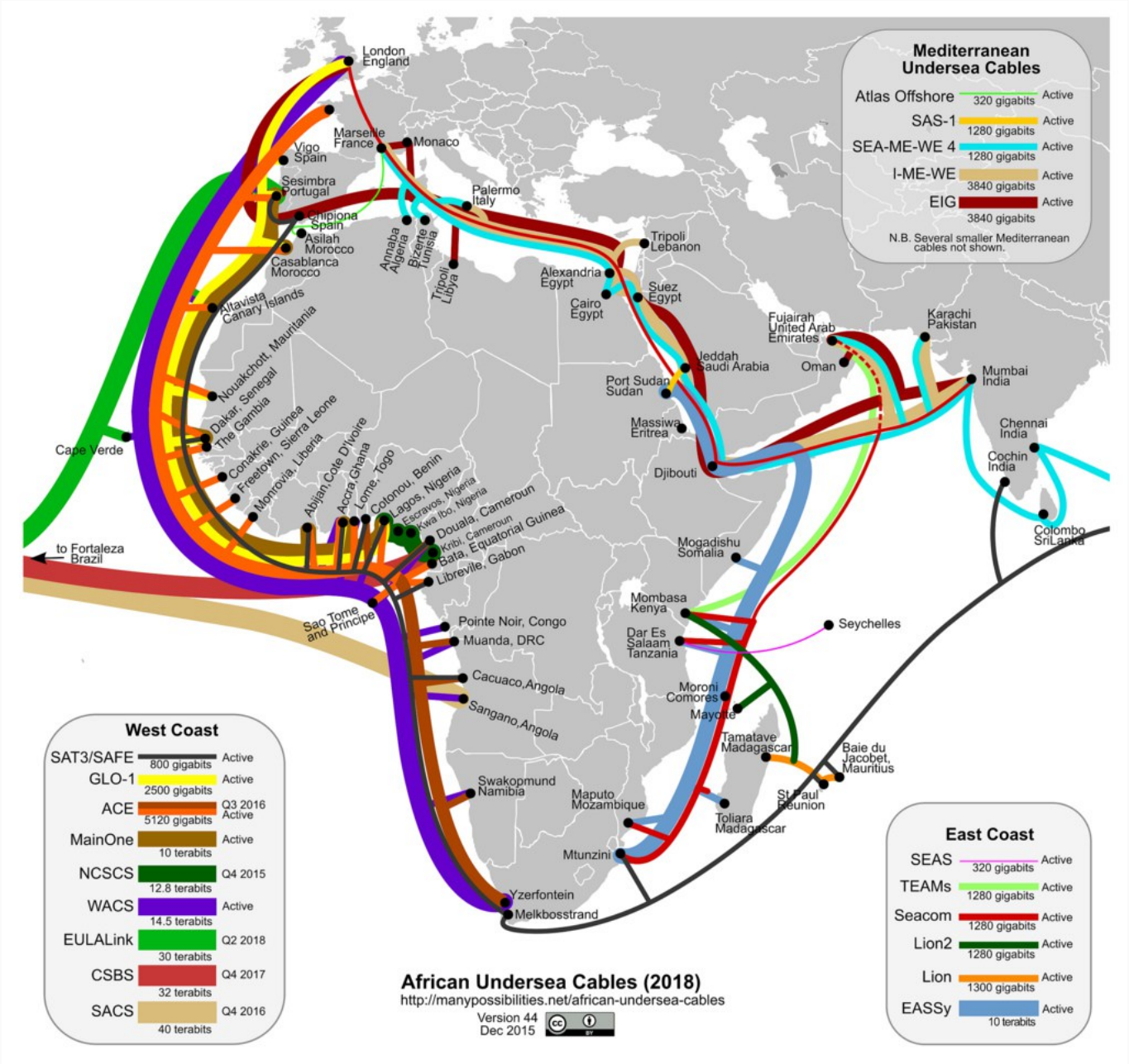
This is currently being expanded to pull in more from MonaLisa, and more experiment and storage specific metrics to help to trivialy diagnose and forewarn problems.

Problems

It has not been an easy 4 months, the big ones are :

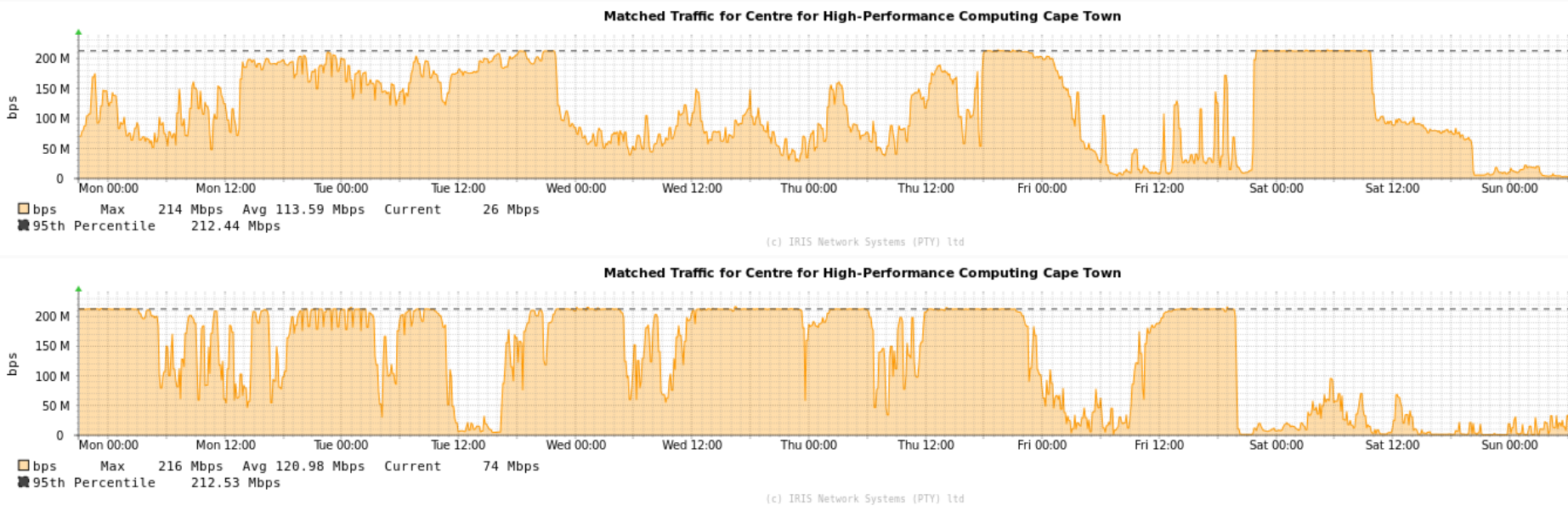
21May	Site Power upgrade whole weekend.
22 Jun	General Power failure, storage out for 2 days, SAM tests in unknown for 1 week.
0 October	power test down for 1 day.

Network Connection



Domestic Network

Data Traffic Mid June



Improvements

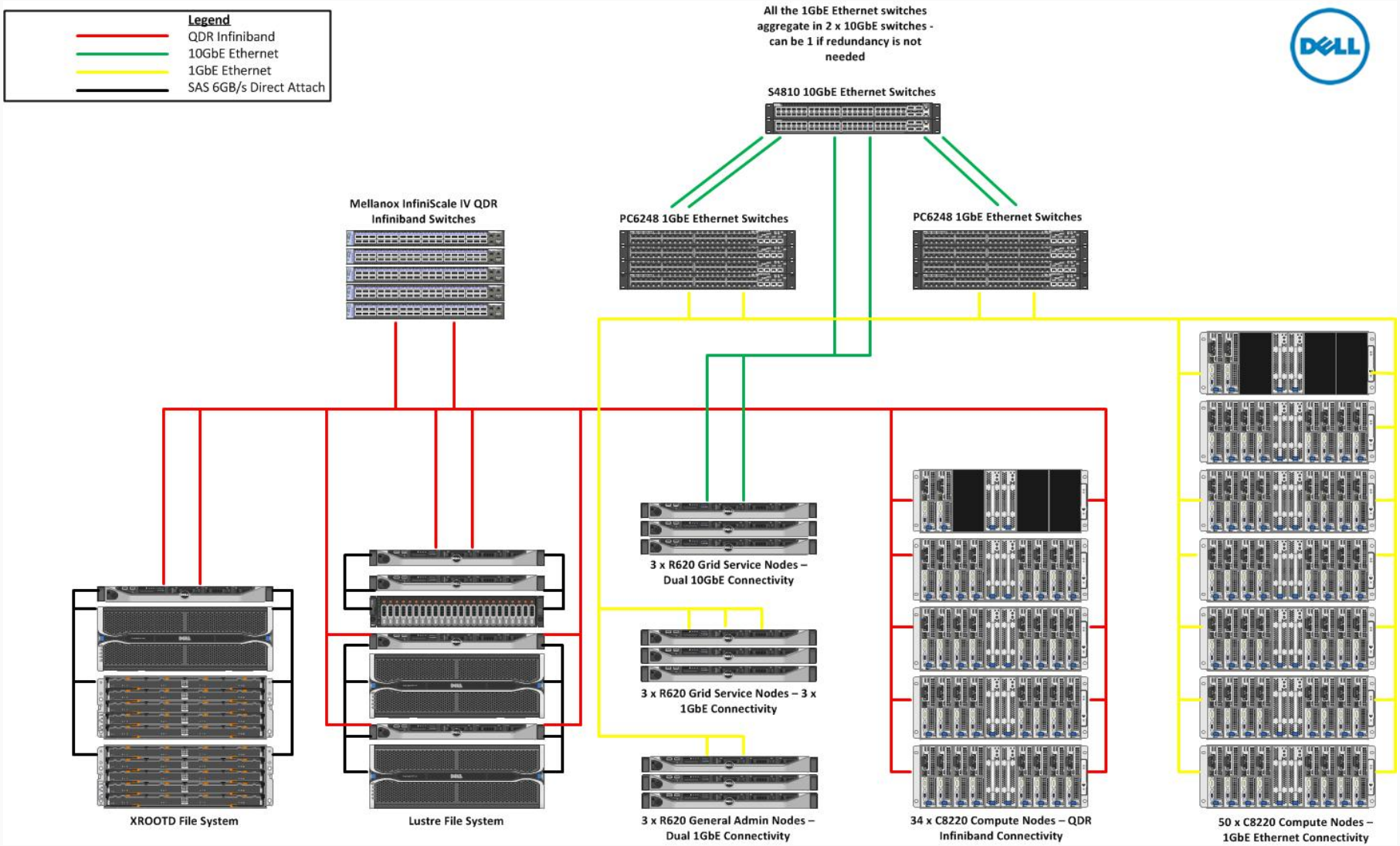
Since my epoch . . .

- Fixed ALICE error rates, mostly.
- ALICE concurrent jobs from 700 to 1500.
- max out bandwidth now regularly.
- ATLAS running pilot jobs.
- Storage cleaned up.
- New storage quoting
- Plan in place to overhaul, and being tested.

upgrades

- attempts to delay till CC7 validation have failed.
- Puppify, to auto site deployment, r10k an issue.
- Transition to foreman from xcat.
- Fix ATLAS Storage (reinstall)
- upgrade monitoring server to zabbix3 and new grafana to technical reasons.
- Rewire whole network, re-power to monitored pdu, and monitor all.
- Add inherent redundancy into 10G interfaces on ce,se,se2.
- Reinstall while TRYING to keep A/R. problems are vobox and ce.
- Storage, we need an additional 750TB for ALICE and ATLAS.

Computing Infrastructure



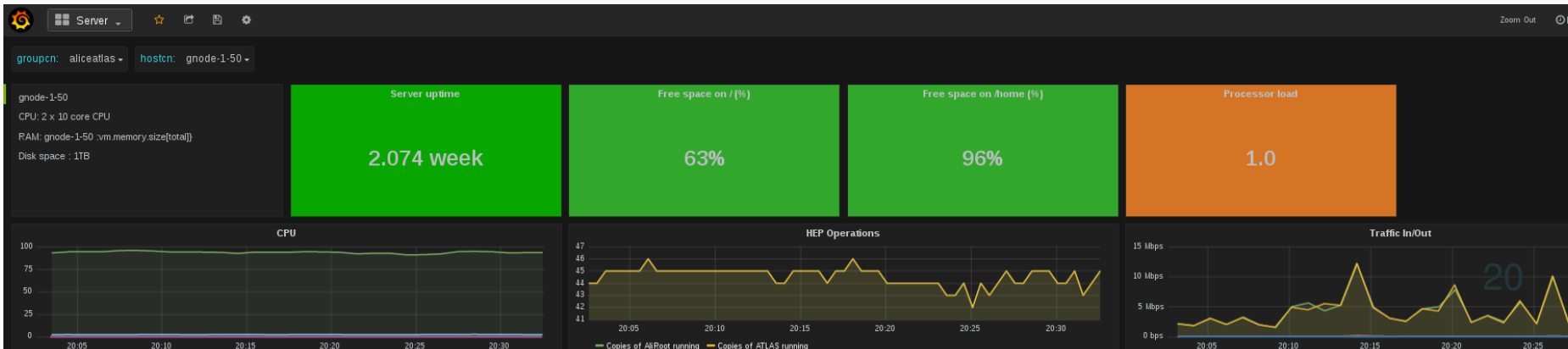
Transparency

Historically this has not been great, so . . .

- Federated logins to zabbix and grafana, i.e. You
- All code on github in line with AAROC
- All issues PUBLIC on github
- Still have GGUS for normal tickets.
- Some training on the user analysis facility (hopefully online)
- A couple of things remain private like network diagrams, obviously, and passwords.
- AAROC slac channels.


ALICE
Zoom Out
Job Agents Running
1800
Batch Jobs Running
1800
ALICE Jobs Running
1797
Jobs Running/Queued
Batch Jobs Running
Batch Jobs Queued
Running Jobs
Jobs Agents
Job Statuses
alice_failed_jobs
alice_saved_jobs
alice_error_v_jobs
alice_error_jb_jobs
alice_error_e_jobs
alice_zombie_jobs
ALICE EOS
No datapoints
ALICE Zombies Running
ALICE Running


ATLAS
Zoom Out
Total ATLAS Jobs Running
3
Jobs Running/Queued
Total ATLAS Jobs Running
Total ATLAS Jobs Queued


Server
Zoom Out
groupcn: aliceatlas
hostcn: gnode-1-50
gnode-1-50
CPU: 2 x 10 core CPU
RAM: gnode-1-50 :vm.memory.size[total]}
Disk space : 1TB
Server uptime
2.074 week
Free space on / (%)
63%
Free space on /home (%)
96%
Processor load
1.0
CPU
HEP Operations
Copies of AliRoot running
Copies of ATLAS running
Traffic In/Out

Outline

28 nodes

I got the go ahead 3 weeks ago to claim 28 of the 34 nodes back for SAGrid and HEP user analysis.

- Go back to SAGrid to support anybody on SAGrid VO.
- hep user analysis, based on federated identities, no user account admin.
- code based on CODE-RADE, or LHC experiments.
- Local Storage for users, eos and lustre.

Outline

Tier 1 Technical Requirements

Its a long list, I wont bore you or me with the intricate details of the MOU.

First and foremost :

Tier 1 Technical Requirements

Its a long list, I wont bore you or me with the intricate details of the MOU.

First and foremost :

A STABLE, RELIABLE PROVEN Tier2

The criticality of that can not be under estimated.

Tier 1 Technical Requirements

- Custodial storage of raw data.
- O(10k) cores
- single digit PB disk
- single digit PB online Tape library, custodial raw data.
- redundant links on LHCOPN (light paths to cern) 10Gbps.

This is the easy part its just a question of money.
The human and process requirements are more onerous.

Tier 1 SLA

- 99% uptime when beam on.
- 4 hour response to failures or degraded service (20%).
- A long term commitment to be a tier1.
- 12 h max delay to responding to operational problems.
- 12 h for network degradation (20%)

People Requirements

- 24x7 operation.
- storage specific experts
- experiment specific experts
- tiered support.
- operators on call.
- wlcg membership, meetings, meetings, meetings.

Outline