

South Africa ALICE Computing and WLCG and grid

Sean Murray
ALICE
CHPC,CSIR
University of Cape Town
May 4 2017

Happy Starwars Day.



Outline

Current Status (last year)

Going Forward

CPU

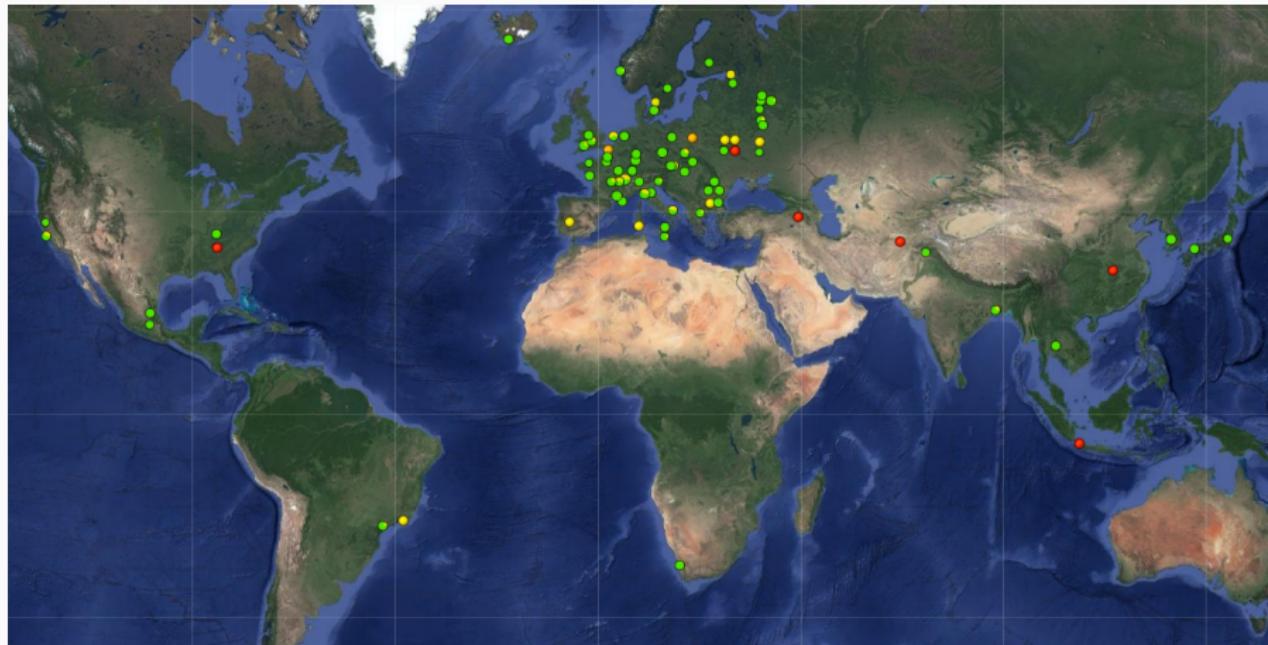
Storage

Network

Alternate users, the cpu slush fund.

Summary

Location



Commitments

According to Tender :

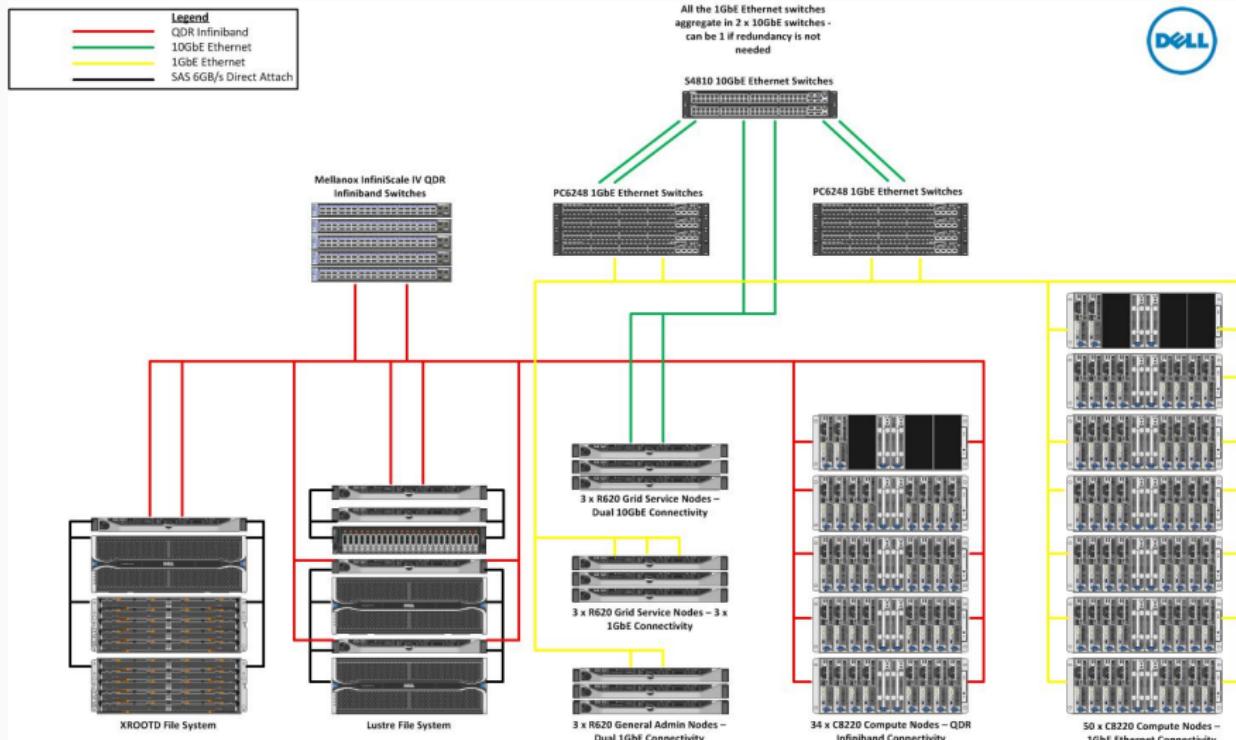
- ALICE 600 cores
- ATLAS 600 cores
- ALICE 400TB [383TB]
- ATLAS 400TB [248TB]

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, 2x bonded 1G ethernet
- 34 nodes of 48 cores 96GB RAM and 1TB, QDR infiniband, 6 lost to another project
- 100TB of Lustre on the 34 nodes with QDR infiniband. now dead.
- 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. dead
- 104 TB EMC for ATLAS. dead and now revived.

Reduction in data sizes is due to reorganisation for reliability.

Improvements

Since my epoch . . .

- Fixed ALICE error rates, mostly, or know the cause with solutions pending.
- ALICE concurrent jobs from 700 to 2100 (briefly).
- max out bandwidth now regularly in both directions, and now storage is getting hit.
- ATLAS running pilot jobs.
- Storage cleaned up.
- New storage quoting, now changed to intergrated RFP.
- Plan in place to reinstall entire site, and being tested.
- xcat replaced by foreman.

Transparency

Historically this has not been great, so . . .

- Federated logins to zabbix and grafana, and public.
- All code on github in line with AAROC, after getting burnt at last employer.
- All issues PUBLIC on github, redmine is more admin than I want.
- A few privacy problems need to be ironed out still, like network diagrams, and passwords.
- AAROC slack channels for reporting everything.

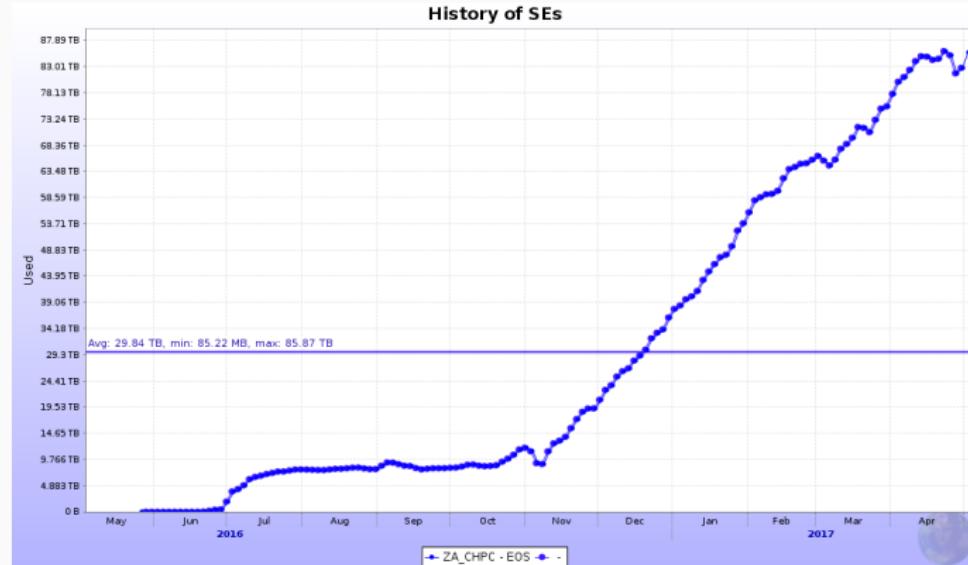
Current Performance last year



- 890k ALICE jobs in last year, 465 the previous year.
- 980 Avg concurrent jobs up from 704.
- 91TB of 383TB
- 235.1TB(139) in 404.5TB(97) out.
- Average 8.9MB/s in and 14.3MB/s

South Africa — Total number of jobs by Resource Centre and Month (Official VOs)											
Resource Centre	Aug 2016	Sep 2016	Oct 2016	Nov 2016	Dec 2016	Jan 2017	Feb 2017	Mar 2017	Apr 2017	May 2017	Total
ZA-CHPC	236,378	162,598	96,835	70,310	54,516	95,524	125,329	78,680	62,587	1,747	1,328,984
ZA-IUCT-ICTS	468	788	603	4,121	0	3,269	1,180	0	0	0	23,920
ZA-UJ	290	428	0	5,021	0	2,283	5,113	13,528	4,420	311	35,775
ZA-WITS-CORE	2,638	6,827	10,980	49,102	0	30,678	12,266	38,463	118,427	0	348,795
Total	239,774	170,641	108,418	137,154	54,516	131,694	143,888	130,671	185,434	2,056	1,704,415
Percent	13.82%	9.84%	6.25%	7.91%	3.14%	7.59%	8.30%	7.59%	10.69%	0.12%	

Graph of Storage Use



Availability / Reliability

Function	Last 365 days
Availability	98
Reliability	99
Storage	94.5

Storage down time is primarily due to storage test failures due to network starvation.

- bouncy running jobs is due to chasing the queue down and up.
- error_e caused by errors of things taking too long to get data.
- random increases in jobs due to no jobs matching (graphs) and
- can not start more than about 50 jobs at a time, to stay stable.
- This is all due to network problems, a standing joke, the network is for all on site.

Downtimes

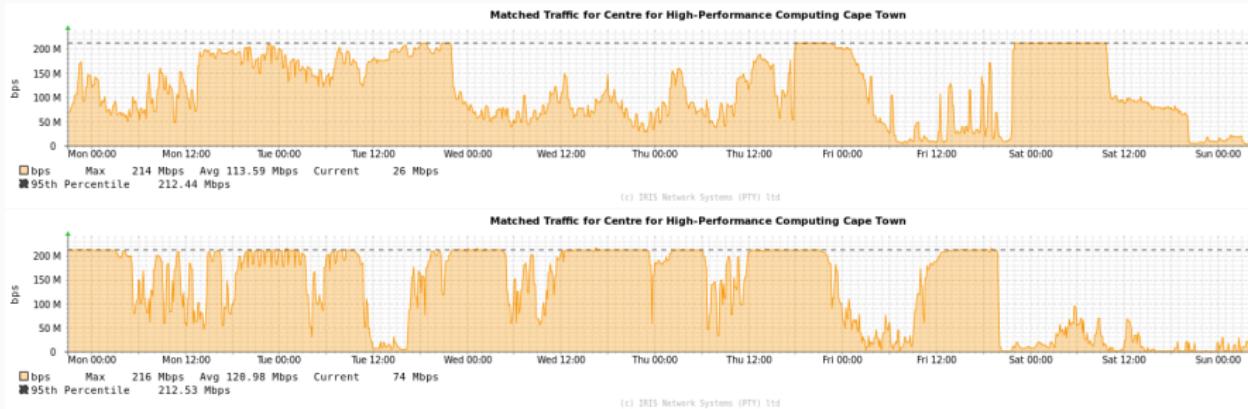
The big ones are, ignoring network starvation :

21 May Site Power upgrade whole weekend.

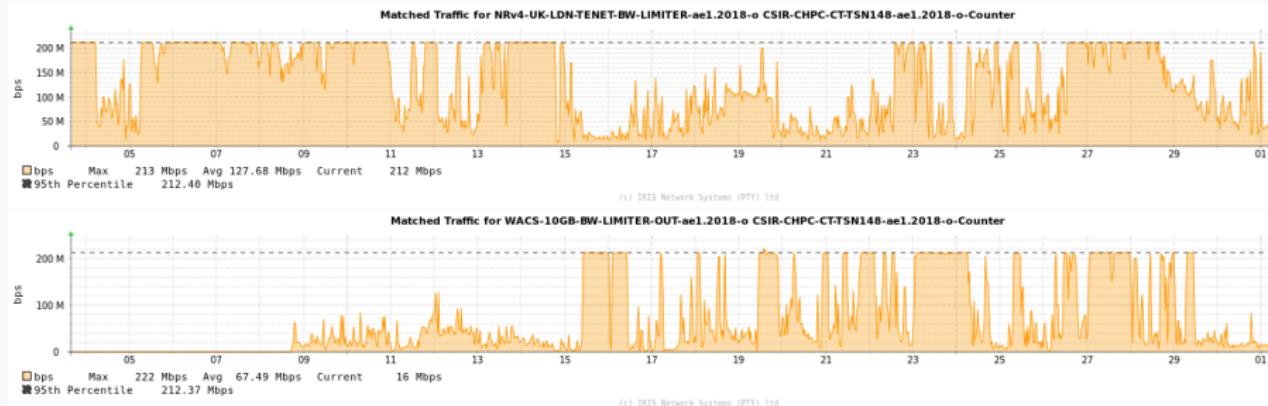
22 Jun General Power failure, storage out for 2 days, SAM tests in
unknown for 1 week.

10 Oct power test down for 1 day.

Data Traffic Mid June



TENET perimeter graphing



Network maximisation

It was always a concern for when our storage becomes sufficiently full to be useful to others.

- March 2, problems in Mexico and our storage got hit outgoing, 2 days cpu down
- March 4 another person liked our data, 2 days cpu down.
- Then, because our storage is under utilised, it became a replication site.
- When in and out on storage is maxed out the compute blows up....



Status this morning

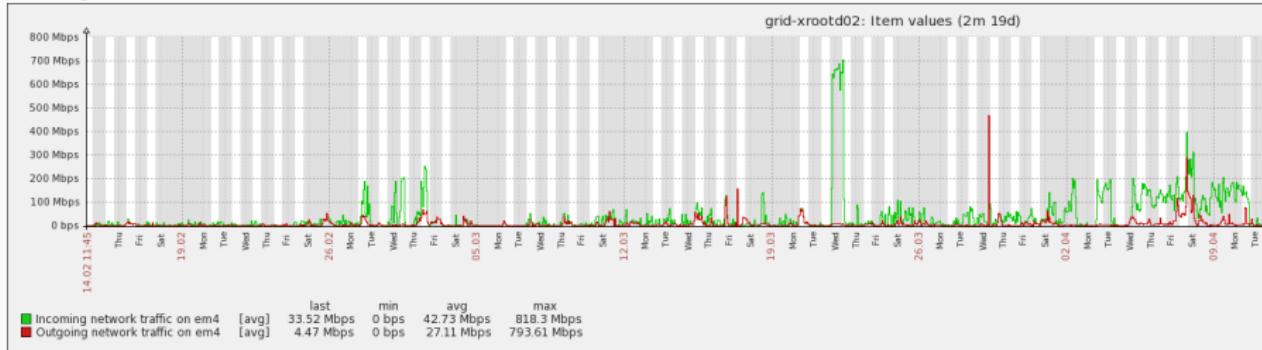


Another example

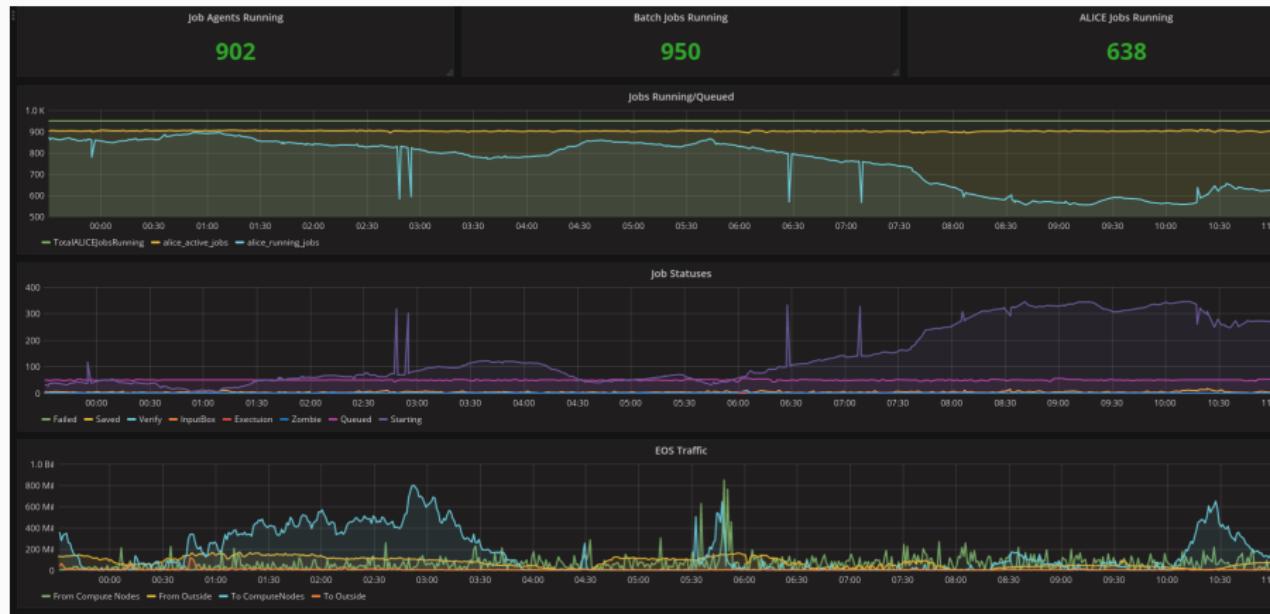


Storage Swamping network

Network was bad and then since March the storage has made things worse.



Processing Today

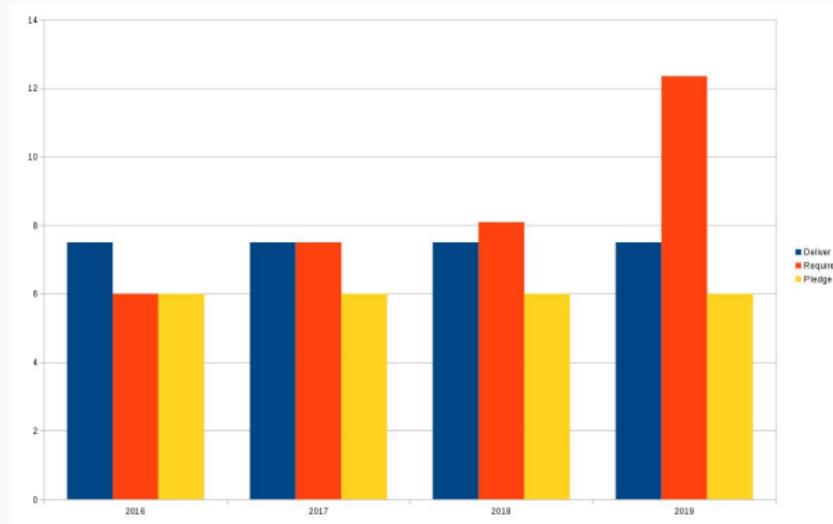


provided and (required)

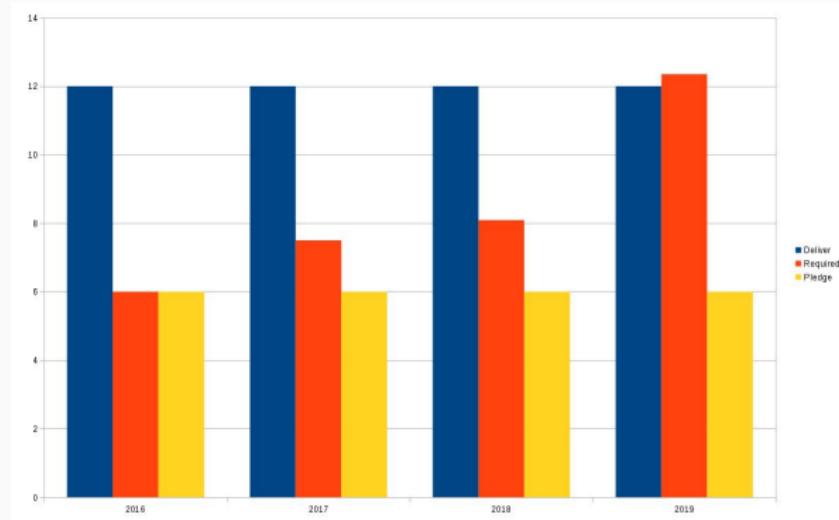
Resource	2016	2017	2018	2019
CPU kHS06	7.5 (6)	7.5 (7.5)	7.5(8.09)	7.5(12.35)
Storage TB	384 (550)	384 (682)	384(990)	384(1100)

- If our HS06 value is incorrect and is in fact 10 then we are cpu sorted till 2019.
- Pledges are wrong, but still under.
- We could double our pledges but with our network it is meaningless.

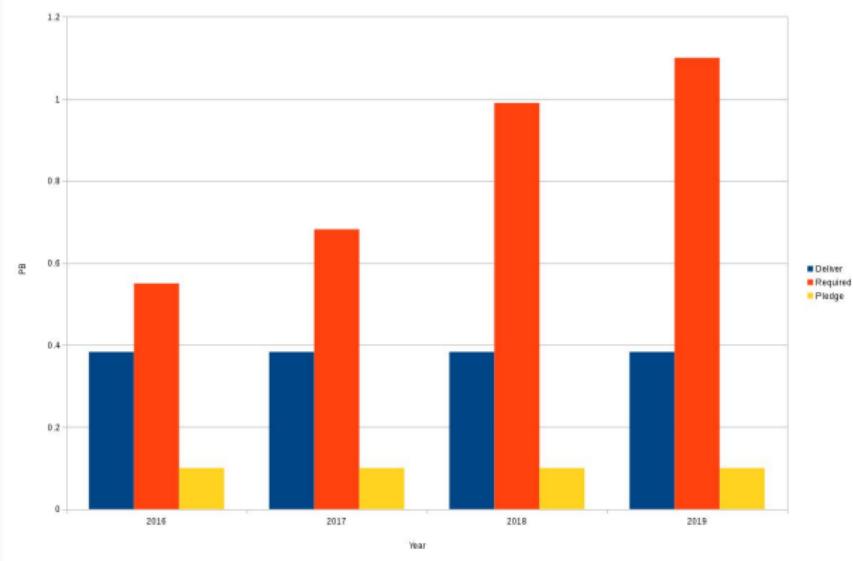
Cpu graphically hepspec06=8



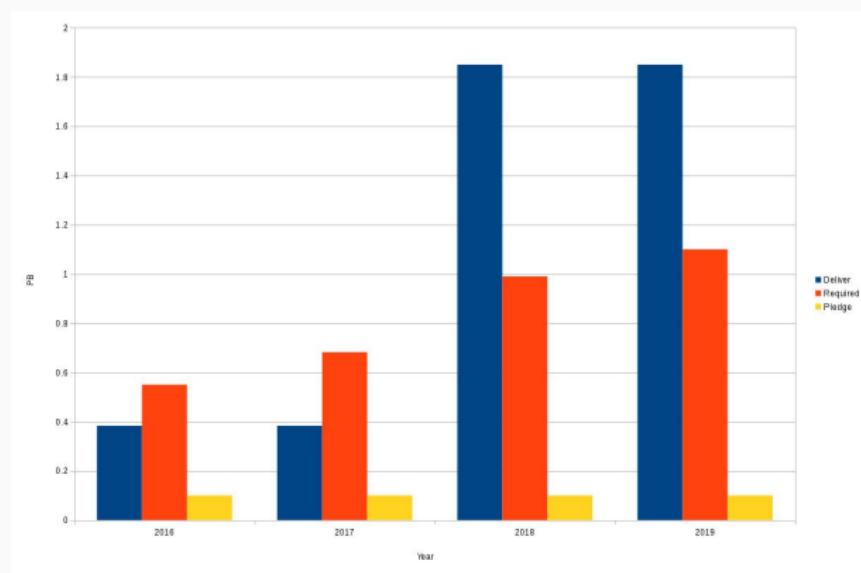
Cpu graphically hepspec06-10



Storage graphically



Storage graphically with DIRISA



Outline

Current Status (last year)

Going Forward

CPU

Storage

Network

Alternate users, the cpu slush fund.

Summary

cpu growth

No real plans here(money), migrate computing into the lower spec 28 nodes, pending tier1 growth.

Project for openstack cloud.

By the end of this funding projections, plan is to be a tier1.

Alternate funding.

Enter DIRISA



science
& technology

Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA



our future through science

CSIR Operating Unit ...1

CSIR Operating Unit ...2

CSIR Operating Unit ...x

CSIR Operating Unit: Meraka Institute

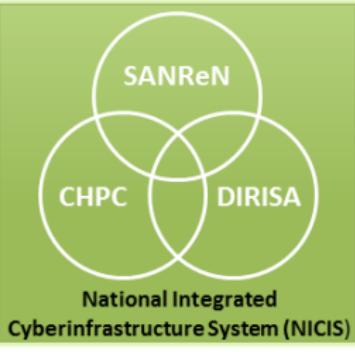
ICT R&D Core

Integrative Systems CA

Telecommunications
and Media CA

Smart Technologies CA

Earth Observation CA



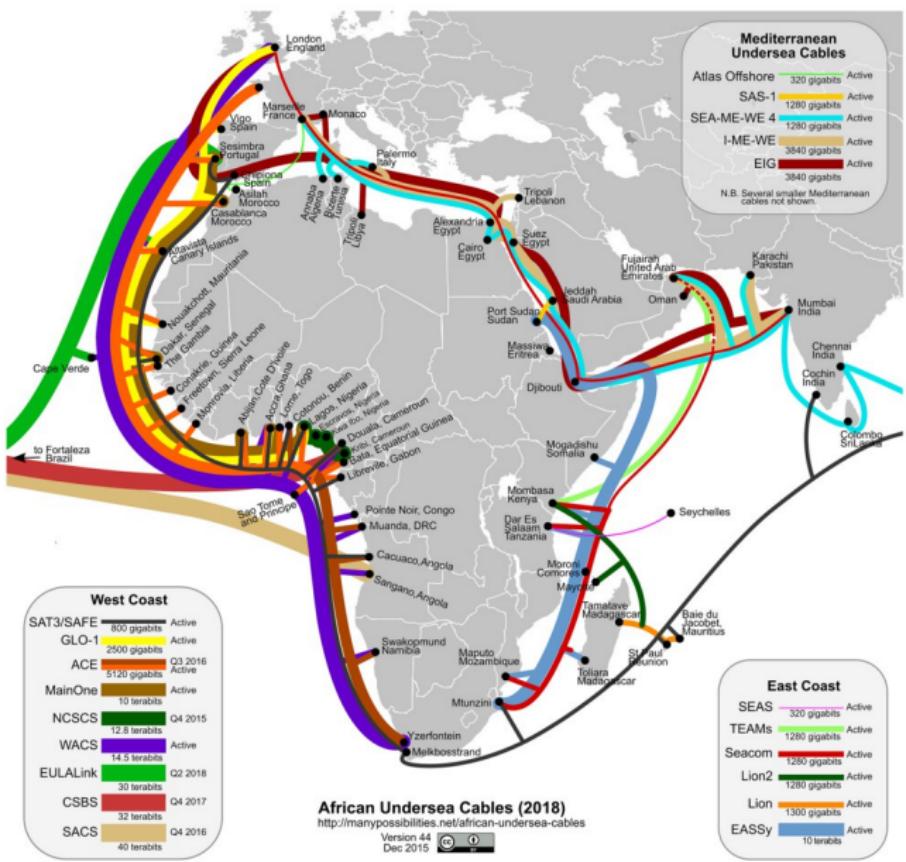
temporary fix, some outstanding issues.

- After much frustration with the 100TB of Lustre
- Tried to restorage EMC, unsupported infiniband, but currently cvmfs stratum1 test.
- old ibm 50TB. yes desperate.
- We have 1PB of lustre, connected to 28 grid nodes.
- 480MB/s write and 2GB/s read. May need some tuning.
- infiniband to 10G bridge
- "donate" some grid nodes as fst's for eos.
- a means of swamping our network test till August, explanation coming.

Storage Growth

- DIRISA 3PB EOS RFP in the works for, we are glued onto the end of a 20-40 PB RFP.
- 1.5PB eos for ALICE
- How to stagger ? or get all on day 1.

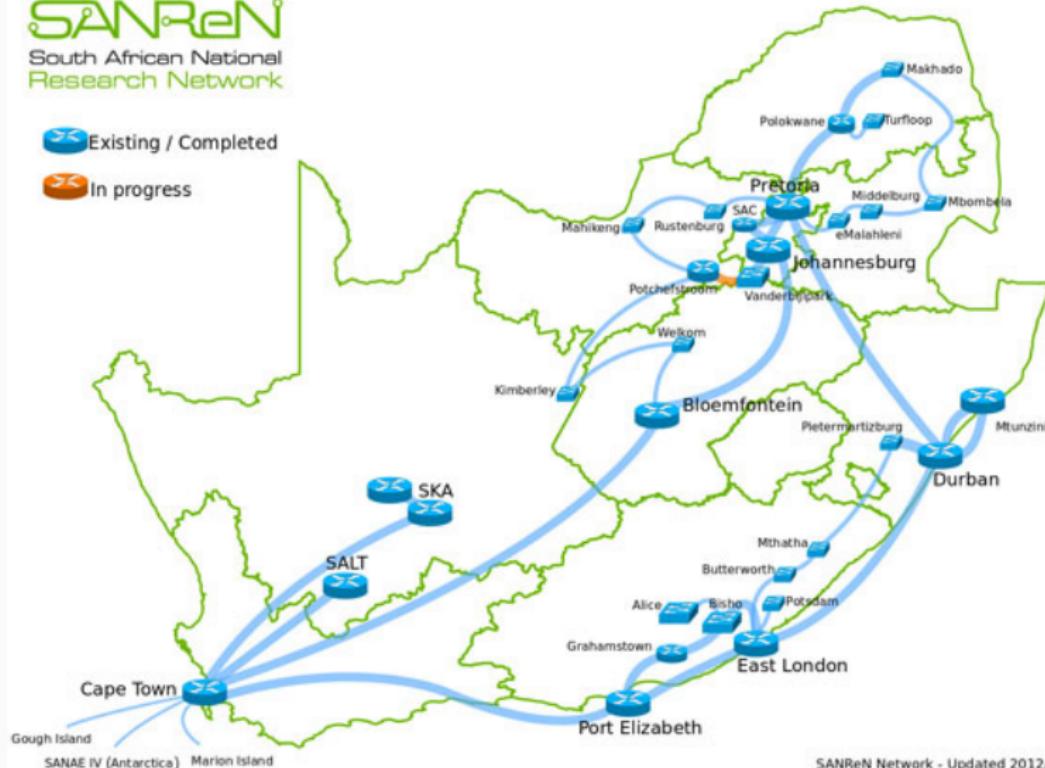
Network Connection



Domestic Network



- Existing / Completed
- In progress



SANReN Network - Updated 2012/11/26

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.
- SANREN joins GNA-RE agreement

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.
- SANREN joins GNA-RE agreement
- nren-nren 10Gbps test system.

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.
- SANREN joins GNA-RE agreement
- nren-nren 10Gbps test system.
- Come August we pay M&O nren-nren
- on LHCONE or not.

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.
- SANREN joins GNA-RE agreement
- nren-nren 10Gbps test system.
- Come August we pay M&O nren-nren
- on LHCONe or not.
- have the quote to lay the fibre for completely new ip4/6 network separate from current network.

The plan

- Currently we have 212Mbps and pay 180kZAR (12k EUR) shared.
- Various avenues investigated.
- SANREN joins GNA-RE agreement
- nren-nren 10Gbps test system.
- Come August we pay M&O nren-nren
- on LHCONe or not.
- have the quote to lay the fibre for completely new ip4/6 network separate from current network.
- Meeting next week with SANREN to iron out details, possibly stay on current network and bump up to 10G international.

Outline

Current Status (last year)

Going Forward

CPU

Storage

Network

Alternate users, the cpu slush fund.

Summary

28 nodes

I got the go ahead early Julys 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 RSA federated identities, and perrun. No user account admin.
- code based on CODE-RADE, or LHC experiments from CVMFS.
- Local Storage for users, eos and lustre.

Lengau 32k cores

- Our main flagship machine is called Lengau (Cheetah).
- 1.1 PFlops, no gpu, 24 cores per node, 128GB RAM (64 on 1/4), 4PB Lustre, FDR interconnect.
- cvmfs via nfs. 3/4 have local disk.

There is a project for a openstack based science cloud across institutions driven primarily by SKA. Status unknown.

Outline

Current Status (last year)

Going Forward

CPU

Storage

Network

Alternate users, the cpu slush fund.

Summary

Summary

Things are improving.

- The big impediment of network is on the mend.
- Storage is shortly to increase greatly and be accessible thanks to network
- all cpu's will run thanks to network.

Freebsd vs linux