T2_FI_HIP site report

Johan Guldmyr @ CSC.FI

- Hepix 2017 Spring
dCache pool server hardware Setup

HPE SL4510 Gen9 with 4x68TB disks – one server per chassis. 128GB RAM. For ALICE/NDGF-T1

2x P440 controllers, one spare disk per controller

2x RAID6 per controller: (16*2+1) + (17*2+1)

~240TB usable per server
201611 Hardware Failures 2/5

• 2016-11-13 Sunday
  - Disks in slots 34,52,54 failed ("aborted command" is the failure reason). XFS (sdb1): Log I/O Error Detected. Shutting down filesystem

• 2016-11-14 Monday
  - Wrote https://github.com/martbhell/python-hpadureport To easily compare HP ADU reports (XML)
  - NDGF Started draining the other pools. Open case with HPE – elevated.

• 2016-11-17 Thursday Pools drained ~100TB. Re-enabled logical drive and NDGF started to drain the pool

• 2017-11-18 Friday evening: 3 other disks “failed”. XFS_WANT_CORRUPTED_GOTO
201611 Hardware Failures 3/5

- 2016-11-21 Monday rebooted and reenabled logical drives. `xfs_repair`, mounted fs as read-only and rsync to an empty array on the same server.

- 2016-11-22 Tuesday io module assembly and cables from controllers to board in io-module replaced. Disks and backplane reseated. This meant taking all disks out of the server.

- 2016-11-25 Friday
  - rsync done. Could not start dCache pool as local metadata corrupted. Moved directory and let dCache fetch metadata from the central database instead.
  - P440 Controller in slot 2 replaced
201611 Hardware Failures 4/5

- 2016-11-28 Monday no more “Bus Faults” seen and dCache finished recovering files/metadata. NDGF moving files to other sites.
- 2016-11-30 Wednesday Pools empty – recreated array and filesystem.
TLDR; In total 8 files were lost for ALICE because of 3 disks in the same RAID6 lost contact with the RAID controller, on two separate occasions. That RAID6 had 1.2 million files. Fixed by replacing the RAID controller. The lost files were being written when the disks disappeared.

Nagios “check_hpadu” written to compare ADU reports
- https://github.com/martbhell/nagios-checks/tree/devel/plugins/check_hpadu
Physical Rack Move 1/2

- The one full rack of hardware CSC hosts & manages for HIP (Helsinki Institute of Physics) was moved on 20170403 a few hundred meters from one DC to another.

- All servers were kept in the rack.
  - 8 x HPE SL4510 Gen9 with 68 LFF disks each.
  - Each server weighs 100kg empty.

- Bought from HPE (a local moving company did the move) so the equipment was under warranty also during the move.

4/21/17
Physical Rack Move 2/2

- In October the Hack stopped working for some reason and the cluster started using Helsinki University's shared internet connection and nobody complained!
- No longer have a private OPN to Helsinki University for dCache traffic
  - No more dual interfaces on the servers!
  - No more DNS overrides!
  - No more static routes!
  - Better for CMS/HIP workloads in their new clusters from inside CSC's Openstack cpouta http://pouta.blog.csc.fi/ or hardware ones in other locations

4/21/17
dCache + CEPH

- Some basic functionality testing of the CEPH pool store backend in dCache
- Mostly works :)
- Not usable right now at CSC as dCache server needs to be able to access the-currently-very-integrated-with-openstack “CEPH network” to create Rados Block Devices
Some Ansible Roles Updates

- https://github.com/CSCfi/ansible-role-cvmfs
  - Now better support for Debian (thanks Ian Allison from pims.math.ca :)

- https://github.com/CSCfi/ansible-role-slurm
  - Many thanks to @jabl from Aalto University for pull requests
  - SLURM HA support
  - seff/smail e-mails on job completion show how much reserved and used resources the job used
  - Supports ansible “free” strategy to speedup ansible push

4/21/17
Grafana!

Average runtime across all sites: 24.5 min
Average runtime across site triton: 20.5 min
Max runtime across all sites: 4.56 hour
Max runtime across site triton: 1.354 hour
Grafana!

Tools used:
• Collectd
• Curl for sending annotations
• Grafana
• Grafana-heatmap Plugin
• Influxdb (opentsdb)
Fin