Lancaster: Free Range and Organic.
HEPSYSMAN SITE REPORT

Matt Doidge, Robin Long, Peter Love
14th June 2017

What is there to be said...

Lancaster University

- We’re home to a Tier 2 Gridsite, which is entwined with the University’s "HEC" cluster.
  - We have root powers and root responsibilites, many perks but a few restrictions.
- EPP group primarily involved with ATLAS and T2K, but diversifying.
- Matt and Robin man the Tier 2, Peter keeps the local users in check. Mike Pacey herds the local-local users as HEC admin.

...that hasn't been said before?

Lancaster University

- The local cluster runs a small HTCondor batch system.
- The HEC runs S(on of)GridEngine, shared areas are NFS for the grid and panasus for the local-local users).
- The Tier 2 uses DPM (1.9) - not using DOME yet.
- Still using CREAM for our CE, monitoring is a mixture of ganglia, icinga and a dashing dashboard.
- Most Tier 2 (and many Tier 3) services are hosted on the University VMWare infrastructure.
- The Tier 2 site connects to the world over the University's 10G "Back Up Link", which we have exclusive use of when it's not needed.

Okay, it's not all the same...

Lancaster University

- ZFS has landed.
  - Peter has been deploying this for local storage, using the native nfs export options - about 1/2 is now ZFS.
  - The latest batch of Tier 2 Storage ditch raid cards for HBAs and ZFS.
- Making IPv6 progress.
- Tentative CentOs7 prodding - some services already there.
- Slowly getting install builds under control with Ansible.

Mildy Interesting Miscellanea.

Lancaster University

- Almost all workers and storage are now 10Gbit connected, most with Mellanox cards and the mellanox ofed drivers (a few chelsio cards and legacy 1Gb).
- Our latest batch of storage are "38-bay", the usual 36-bay nodes with mirrored 2.5 inch SSDs for the OS and 36 6TB drives on an HBA.
  - Some fun and games getting the install on the right drive!
  - Had to disable the HBA "on boot" to stop the bios being overwhelmed with drive options!
- Installs are done using cobbler and a lot of doidgey bash scripts over pdsh.
  - I'm a chronic Yaimophobe, which has extended to a fear of Puppet.
  - Slowly moving to ansible as the lightweight solution.
- WN software is being served up through cvmfs via grid.cern.ch.

Intel Outside.

Lancaster University

- For the first time in nearly a decade, we didn't buy our usual *Twin*^{2} chassis - we got the Intel equivalent instead.
- Other then IPMIView not working for them anymore we noticed no real gotchas. It's early days yet but we like the build quality.
- Of note that we played with the performance settings, switching to "Balanced Performance" from "Balanced Power" increased various benchmarks by 10% whilst only increasing power consumption by a few %.

C7 Samurai.

Lancaster University

- After inroads by Robin CentOs 7 takeup has slowed.
  - Squids, BDII, "ancillery" servers done.
  - Which leaves the hard stuff.
  - Low motivation to roll out C7 on new disk servers when we have a tried and true SL6 install.
- The HEC has pressure from the local-local user's needs to upgrade.
- Have a working C7 compute install, cvmfs has C7 WN software in it - next step is to plug a few nodes into an atlas queue.

Paving the Road to Heck.

Lancaster University

- C7 isn't the only thing on our "to-do list of shame", we have other good intentions.
  - Meant to deploy an ARC CE ages ago (still on CREAM, because it works).
  - Need to move our DPM headnode to new (or virtual) hardware. But our SE is holding on after a bit more turning.
  - Ganglia is showing its age but it's working for us.
  - Have long standing load issues with our NFS server which mounts the grid home areas and sandbox (a BLAH/TMPDIR issue).
  - Been meaning to deploy BRO for ages after promising David I'd look into deploying it- sorry bro!
  - Have a plan to look at XROOT caching on the Worker Node, as part of our C7 install.

The Odd Couple.

Lancaster University

- Our partnership with the HEC has a few caveats worth noting, because sharing always means some compromise.
  - The HEC schedules on VMEM as well as CPU slots. This causes a number of issues for atlas. But there is a plan...
  - The multiuser environment prevents using hyperthreading and requires 4GB/core RAM - no gazillion-core boxes.
  - All these job types don't always tessellate well (we've had to balkanise a bit).
  - Purchasing takes into account more benchmarks then HEPSPEC06 (not that I mind this).
  - We can't be quite as agile as we'd like to be due to long running local-local users jobs (although security patches trump user jobs).
  - We have to be good cluster citizens - no running extra services on the compute (bar cvmfs), no swapping!
- But the benefits (an extra, expert sysadmin; shared resources; a nice machine room; direct link to IT services) far outweigh the costs of having to play with others!

IPv666

Lancaster University

- Our IPv6 deployment is going okayish - but we never made an Address Plan!
- Perfsonar Boxen and RIPE probe v6ed.
- Our Networking admins were nervous about v6ing our whole SE though.
  - We’re the heaviest single user, and we have special routing rules in place so our traffic goes over the Carlisle "back up link".
  - Fear that a sudden deluge of grid traffic over IPv6 would break everything.
- After some iperf tests with Brian the fears might have been half right - outbound IPv6 traffic might not be respecting the routing rules.
  - That’s about it, we’re still investigating. Sadly there was no Thunder for Brian to steal!

Tarball Plug.

Lancaster University

- For his sins Matt is the keeper of the WN and UI tarball, now in it's UMD4 incarnation (and C7 versions too).
- Among other places this software can be found in /cvmfs/grid.cern.ch/
- It's a good way of getting an instant UI.
- VO information is kept here in etc/grid-security/ - this is as up to date as possible but any mistakes or ommissions please poke me.

GridPP39 is in Lancaster!

Lancaster University

14th-15th of September 2017.

Matt Doidge, Robin Long, Peter Love
14th June 2017