Site Report HPC2N

NDGF AHM Fall 2022

2022-10-10 — 2022-10-11

Center storage, Lustre

- Versions
  - SFA400NVXE (FW 12.0.0.1 & 04050206) plus 4 enclosures (F03090224)
  - Exascaler 5.2.5
  - Insight 4.1
  - CentOS 7.9
- A couple of new LBUGs triggered
- The previous hotpool problem with shared libs (and other mmaped files) vanished after the last upgrade when we also converted the last subtrees with FPL stripes to the default policy.
- We are currently testing some Lnet tuning that we hope will reduce the risk of nodes going offline during if too many have problem. It might also reduce the risk of our over-time reboots of the servers. See the last two sessions at LAD'22.
- Upgrade of Exascaler planned for early next year.
- DDN Insight still have some issues (ex. size of storage and list of active high load jobs)

Kebnekaise/Cluster

- Ubuntu 20.04.5 LTS
- Slurm 21.08.7
- HW works really well
  - The cluster no longer has an active HW support, the plan is to retire problem nodes and have as spared during the remaining lifetime of the cluster.
- Kebnekaise no longer part of the national resources after 2022-12-31
  - The number of active SNIC compute projects on cluster is therefore declining
  - The remaining projects have been taking advantage of this by being able to run more jobs in their own projects well above their allocation.
- A few extra maintenance windows due to modifications of data centre
  - mainly the ventilation and power supply structure, see later slide

Abisko dismantled, central IT borrows space

- University IT (ITS, formerly UMDAC) is refurbishing/redesigning their data centre
- ITS now temporarily housed in the old Abisko racks
  - Only approx. 60-80kW, so not noticeable wrt power/cooling
- Temporary kludges for UPS power of ITS equipment
  - Including emergency cooling, which also needs to cater for HPC2N equipment if we get cooling loss but still have power.

Feed from central IT UPS being discontinued

- ITS has decided (without talking to any stakeholders) that we will no longer get any power feed from the ITS UPS (which is backed by a diesel generator) in the storage data center shared between HPC2N, Computing Science and Informatics.

- Seems like our suggestion to let the HPC2N UPS take the role will get approved.
  - At least there are now plans on how to get diesel generator feed to it etc.

- Had hoped for the old UPS feed to be converted to real redundancy from other source but looks like it will be from the same in-house source.

Feed from central IT UPS being discontinued(2)

- Status
  - Cabling done to feed Computing Science/Informatics from HPC2N UPS
  - Will hopefully get a decision on moving forward with getting diesel generator feed to it now that there is a plan.
  - All management/economy-issues unsolved
    - HPC2N is billed for energy consumption by that UPS
    - Computing Science and Informatics pay rent that includes power
    - How will the formal agreement on responsibility/maintenance/etc look like?

Tape/backup

- Hardware unchanged
  - IBM TS4500 library (2000 slot capacity)
  - 6x TS1155 tape drives
  - Dell R740xd with approx. 200T spinning disk, 4x16G FC, 2x25GigE
- About to procure new TSM server
- Added even more SNIC tapes to handle increased copy pool volume
- Most hardware issues related to tapes being broken
  - Not that many, but more than you expect of tapes with "lifetime" warranty
- TSM server software still on version 8.1.14.100
  - Support TOTP MFA for admin accounts, seems to work

Tape/backup (2)

- WLCG specific tape status
  - No ALICE tape data (migrated to UIB)
  - ATLAS part mostly full
    - There are some free space after reclamations, works well with HPC2N as overflow when load gets too high for the UiO tape write pool.
  - Handling 2023 pledges
    - Want to buy new tape drives if IBM ships next generation, but given that the world has become even more broken, unclear if the roadmap still holds.
    - Fallback is to buy more tapes for the TS1155 drives we have. It hurts to buy old stuff, but it should be manageable.

LHC OPN

- Finally got the attention of the networking people
- Multiple challenges wrt delivery times of equipment etc.
- Three options
  - Steal the 100G uplink UmU wants to use, connect directly to our top-of-rack switch and let SUNET do all routing.
  - Let UmU networking have the 100G and have them aggregate a gaggle of 10G links into a dedicated LHC OPN switch/router and connect a 100G from there to our top-of-rack switch.
  - New dedicated 100G link from HPC2N to Stockholm

Conclusion of the HPE support saga ?

- 4 node setup for ARC cache. Delayed delivery. 1 broken @ delivery
- Took HPE well over 6 month to fix it
  - Been stable since then … except for a once in a blue moon raid hiccup
- Asked for small compensation, got BIG from HPE
  - credits for a complete machine (77.500 SEK)
- Took 4 months get the credits and a quote for support extension
- Support extension ordered
- 44 kSEK left to buy more stuff! What to buy?