

Canadian T1 & T2 sites News, Updates, and Issues

R. Tafirout, TRIUMF

ATLAS Computing Sites Jamboree,
January 18, 2017

Canadian Sites Evolution

- Tier-1: dedicated facility located at TRIUMF
 - Managed and operated by ATLAS-Canada, SFU and TRIUMF
- Tier-2's: shared facilities located at Compute Canada centres
 - National organization serving all research communities
 - Management structure and operations are more complex
 - Each ~year, ATLAS-Canada submits a proposal to the National Resource Allocation Committee (NRAC) to secure resources.
 - Two WLCG federations across 4 sites (was 5 prior to 2013).
- Same funding mechanism: Canada Foundation for Innovation (CFI) & provincial partners for matching
 - Tier-1: very successful in securing own funding since 2006
 - Compute Canada is refreshing all of its infrastructure and aging equipment (going from ~27 to 4-6 larger centres)
 - CFI would like a Tier-1 integration within Compute Canada, to minimize infrastructure and operating costs.

Tier-1 relocation plans

- TRIUMF infrastructure is aging and floor space limited: would require a new server room to expand further and to be able to provide the necessary tape capacity for 2017 (a critical item now).
 - New equipment will be installed at the new Compute Canada center at Simon Fraser U.; requires MoU and SLA between parties involved (to be finalized).
- New data centre at Simon Fraser U.:
 - 2 x 0.5 MW UPS capacity, backed up by a generator
 - Large floor space (new building being renovated)
 - Ensures proper expansion going into the future, overall power capacity not an issue.
- Distance between TRIUMF and SFU: ~28 km with ~1 ms RTT.
- Most of TRIUMF Tier-1 equipment reaching 5 years in 2017: we are extending warranties/support until early 2018. Full refresh needed.
- Want to minimize downtime at all cost during transition.

Tier-1 relocation plans

- TRIUMF infrastructure is aging and floor space limited: would require a new server room to expand further and to be able to provide the necessary tape capacity for 2017 (a critical item now).
 - New equipment will be installed at the new Compute Canada center at Simon Fraser U.; requires MoU and SLA between parties involved (to be finalized).
- New data centre at Simon Fraser U.:
 - 2 x 0.5 MW UPS capacity, backed up by a generator
 - Large floor space (new building being renovated)
 - Ensures proper expansion going into the future, overall power capacity not an issue.
- Distance between TRIUMF and SFU: ~28 km with ~1 ms RTT.
- Most of TRIUMF Tier-1 equipment reaching 5 years in 2017: we are extending warranties/support until early 2018. Full refresh needed.
- Want to minimize downtime at all cost during transition.

New Tier-1 deployment plan

- Implement a distributed Tier-1 during the transition phase
 - Only relevant and critical to storage services
 - achievable with dCache (NDGF Tier-1 is a prime example)
 - Other services should not be an issue (new CE, etc.)
- Phase 0: pre-production of initial services and testing (*now*)
 - Network configuration (new address space), bring LHCOPN to the new site (LHCONE already there). WAN limitation until Q4.
 - New tape library commissioning and distributed Tier-1 testing
 - Necessary equipment being installed: core switch, new HSM servers, SAN storage, basic infrastructure nodes.
- Phase 1: full production by April 1 2017 with the new tape capacity and related services fully online. Initiate tape migration.
 - Add additional pledged disk and cpu capacities:
need 1 PB disk, 1350 cores (not part of the refresh)
- Phase 2: end of 2017/early 2018, bring online new disk and cpu; complete data migration (mostly disk). Finalize tape migration.
- Current TRIUMF Tier-1 will remain online till ~ Q2 of 2018.

Tier-2 consolidations plan

- Our Tier-2 resources are secured from Compute Canada via a competitive process. For many years we have been using several sites. This will evolve going forward under a new operating model.
- Compute Canada (CC) is executing its consolidation plan and we were told ATLAS Tier-2 resources will be provided by 2 sites with very similar configurations (GP2 at SFU, and GP3 at Waterloo).
 - Various RFP's issued and some systems delivered already, commissioning has started
 - Infrastructure work still ongoing
- Working closely with CC subatomic physics national team (SPNT) via weekly meetings to ensure new sites are properly designed and configured for ATLAS use.
- Once sites are ready, need to work closely with ADC operations to shift production onto the new sites, including data migration.
- Until the new sites are fully commissioned and usable by ATLAS we continue to use our current sites at McGill, Simon Fraser, Toronto and Victoria.

Ops/ADC T1 issues (I)

1. low job efficiency, in particular for mcore jobs
 - single stuck athena process mobilizing all cores until killed by batch system (walltime limit). Should be solved now in pilot.
 - number of events / job: room for improvement ?
 - other causes
 - a) overloaded frontier servers
 - b) misconfigured access protocol during new site movers implementation.
 - c) misbehaving jobs, for example, some jobs spawned more than a hundred athena processes, making compute nodes unusable.

Ops/ADC T1 issues (II)

2. frontier servers were overloaded from the direct access
 - disable direct access to frontier servers from WNs ?
3. single stream and fixed time out parameters are used in data transfer
 - FTS provides the functionality for the dynamic stream number and time out parameters
 - we saw some transfers between CA sites and other sites timed out
4. some jobs used much more memory than it requested, in particular user analysis jobs
 - htcondor can monitor the PSS usage of jobs, we are testing monitoring and killing jobs based on PSS usage

cpu eff. (TRIUMF all)



Docker for worker nodes

- Last spring we began evaluating Docker as a solution for worker node deployment
- URL describing this work:
<https://twiki.atlas-canada.ca/bin/view/AtlasCanada/TRIUMFDocker>
- Goal: Run ATLAS SL-6 based Docker containers on SL-7 Docker engine nodes
- Current tests are done running containers with condor-master as the container application
- Good results comparing container to host performance with ATLAS software validation
- Good results comparing container to host performance for generic benchmarks:
 - HEPSPSPEC '06
 - iozone local I/O
 - data transfer protocol tests (gsiftp, dcap, root, http)
- Work to do: Distribution server and image signing, deployment strategies, container logging, image maintenance