



AGLT2 Site Report

Shawn McKee/University of Michigan

Bob Ball, Chip Brock, Philippe Laurens, ~~Ryan Sylvester,~~
Richard Drake

HEPiX FALL 2016 / LBNL



Site Summary

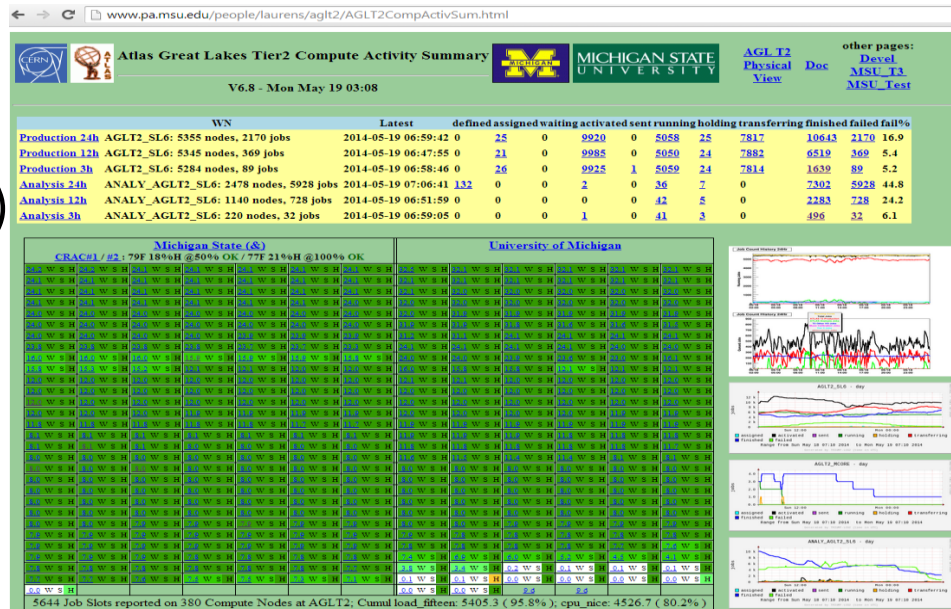
- 📄 The ATLAS Great Lake Tier-2 (AGLT2) is a distributed LHC Tier-2 for ATLAS spanning between UM/Ann Arbor and MSU/East Lansing. Roughly 50% of storage and compute at each site
 - 📄 8200 single core job slots
 - 📄 MCORE slots 720 (dynamic) + 10 (static)
 - 📄 720 Tier-3 job slots usable by Tier-2
 - 📄 Average 10.08 HS06/slot
 - 📄 4.7 Petabytes of storage (soon ~7 PB)
 - 📄 Total of 83.5 kHS06
 - 📄 Tier-2 services virtualized in Vmware 5.5
- 📄 2x40 Gb inter-site connectivity, UM has 100G to WAN, MSU has 10G to WAN, lots of 10Gb internal ports and 20 x 40Gb ports, 32x100G/40G ports
- 📄 High capacity storage systems have 2 x 50Gb bonded links
- 📄 40Gb link between Tier-2 and Tier-3 physical locations

AGLT2 Monitoring

AGLT2 has a number of monitoring components in use

As shown in before we have:

- Customized “summary” page->
- **OMD (Open Monitoring Distribution)** at both UM/MSU
- **Ganglia**
- Central syslog'ing via ELK: **Elasticsearch, Logstash, Kibana**
- **SRMwatch** to track dCache SRM status
- **GLPI** to track tickets (with FusionInventory)



Personnel Changes

- 📄 Our new hire **Ryan Sylvester** (announced at the last HEPiX) has moved on to a new job working for the Veterans Administration.
 - 📄 As Ryan was a air-force it was a logical career move
- 📄 For the near-term, we are not planning on hiring a replacement
 - 📄 We have been very successful in getting good undergraduates to work on an hourly basis for us
 - 📄 Looking to augment our current student with 1-2 more

Hardware Changes Since Last Mtg

- ☐ This summer we had a one-time infusion of NSF funds to ensure our ATLAS MOU was based upon reliable, warrantied equipment
 - ☐ AGLT2 received **\$398K additional funds** for hardware
- ☐ We negotiated good pricing with Dell and purchased:
 - ☐ 55 R630s; 2xE2640v4 processors, 128GB ram, 2x800 SSDs
 - ☐ 4 Storage nodes (R730xd+MD3460+MD3060e; 120x8TB)
 - ☐ 2 MD3060e shelves (each 8x60TB disk)
 - ☐ All 5-year warranties
- ☐ Total **addition** was **24 kHS06**, 4.8 PB(raw) / **3.84 PB(r6)**
 - ☐ Still have 3 storage units to commission (at SC16)
 - ☐ 24 R630s not yet in production

Software Updates Since Last Mtg

- Tier-2 VMs rebuilt to use SL7 (old SL5)
- dCache updated to 2.13.42 and Postgresql 9.5
- HTCondor now running version 8.4.7
- OSG CE updated to 3.3.11
- Various switch firmware updates applied in summer; bios/firmware on Dell systems in August
- Configuration of Mellanox SN2700 (32x100G) completed.
 - Using it as “core” storage/cluster switch at UM
 - Supports 100G, 40G, 2x50G, 4x25G and 4x10G on ports

Lustre at AGLT2

- We have updated our Lustre storage, using new hardware and incorporating old servers
- The new Lustre server and storage shelves were racked in the Tier-2 center last year
 - Lustre version 2.7.0 was installed and new file system created (Using ZFS)
 - Old files from /lustre/umt3 were copied to the new system
 - Old Lustre servers were then recycled to increase total storage in new file system
 - New Lustre file system is (re)mounted at the old location, /lustre/umt3
- Current AGLT2 lustre size is now 1.05 PB (retired some *old* storage)
- **Next up:** go to the almost released 2.9+ZFS (v0.7) version

Possible Relocation at UM

- 📄 I reported in spring that **AGLT2** may need to move to a new physical location at the University of Michigan
- 📄 **Good news:** After exploring our requirements the University is instead planning to have the architects involved with the building renovation work around our Tier-2 center
- 📄 Baseline plan is the we are **NOT planning on moving**
 - 📄 **However, if there are snags in the design this could again change**

Future Plans

- Participating in SC16 (simple infrastructure for 100G+)
- Still exploring OpenStack as an option for our site. Testing Ceph for a back-end.
 - Our Tier-2 will be first OSiRIS Client (see Thursday's talk)
- Working on integration of OVS (Open vSwitch) on production storage and Software Defined Networking (OpenFlow). We are experimenting with SDN in our Tier-2 and as part of LHCONE point-to-point testbed.
- Working on IPv6 dual-stack for all nodes in our Tier-2
 - Have IPv6 address block
 - Ticket reopened with IT. Hung up on engineering discussion ☹️

Summary

- Tier-2 services and tools are evolving. Site continues to expand and operations are smooth.
- Monitoring stabilized, update procedures working well
- FUTURE: OpenStack, IPv6, SDN

Questions ?

AGLT2 100G Network Details

100G to WAN works

Last Fall: "normal" FTS

transfer hit 2.73 Gbytes/sec

