



# AGLT2 Site Report

Shawn McKee/University of Michigan  
Bob Ball, Chip Brock, Philippe Laurens, Mike Nila,  
*Forest Phillips*

HEPiX Fall 2017 / KEK Tsukuba, Japan



# AGLT2 Numbers

- 📄 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
  - 📄 9408 logical cores
  - 📄 MCORE slots 950 (dynamic) + 10 (static)
  - 📄 Additional 720 Tier-3 job slots usable by Tier-2
  - 📄 Average 10.21 HS06/slot
  - 📄 **6.85 Petabytes** of storage
  - 📄 Total of **96.1 kHS06**
  - 📄 Tier-2 services virtualized in VMware 5.5 (soon upgrading to 6.5)
- 📄 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 or 64x50G/25G ports
- 📄 High capacity storage systems have 2 x 50Gb or 2 x 100Gb bonded links
- 📄 40Gb link between Tier-2 and Tier-3 physical locations



# Personnel Updates

- ❏ Over the last year we have lost two tier-2 cluster administrators, one at UM and one at MSU
- ❏ As of September 1, 2017 we have added a new physics grad-student at MSU working on AGLT2 at 50% time: **Forest Phillips**
- ❏ I have the sad task of announcing that **Bob Ball** our Tier-2 Manager since we started in 2006, plans to retire next year 😞
  - ❏ Bob will likely start a phased retirement next April and be fully retired by about 1 year from now. He will be missed!
  - ❏ Hopefully Bob can make it to the Spring 2018 HEPiX!
  - ❏ For those interested, look for a job opening announcement around January/February of 2018

# Hardware Additions

- We are late on our FY17 hardware purchasing
  - Plan is to try to revive Dell ATLAS portal (at least for USATLAS sites)
  - Skylake processors look good but pricing is higher than before
- Focus will be on computing (rather than storage)
- Will buy some VMware storage at MSU
  - Need storage under warranty: will go with SAS attached MD3460, 20x1.2TB 10K SAS disks plus 40x6TB 7.2K NL-SAS disks
  - MSU will add a third VMware ESXi host (repurpose WN)
- Both MSU and UM sites may also purchase more 10G network ports
- Anticipate adding ~10 K-CPU-benchmarks of CPU
- Overall storage will stay around 6.8 PB till our next purchase

# Software Updates Since Last Mtg

- Tier-2 VMs rebuilt to use SL7 (old SL5)
- dCache updated to 3.0.28 (still Postgresql 9.5)
- HTCondor running version 8.4.11
- OSG CE updated to 3.3.25
- Various switch firmware updates applied
- Monitoring updates: OMD/check\_mk to 1.4.0p13
- Two “major” updates: Lustre and dCache/OVS

# Lustre at AGLT2

- Lustre at AGLT2 is used for “local” users (Tier-3)
  - About 1 PB of space running on SL6/Lustre 2.7/ZFS 0.6.5.4
- Upgrade targets: OSS and metadata servers
  - OS updated to SL7.4
  - Kernel 3.10.0-693.2.2
  - Lustre 2.10.1
  - ZFS 0.7.2 on the OSS, Idiskfs on MGS/MGT
- 20Gb/s bonded Ethernet on OSS
- Lustre Re-Exported to MSU WN via NFS from Lustre SL7.4 client
  - Same OS and kernel as on the OSS
- Data read rate shows significant increase (~x2; up to 2GB/s aggregate seen) over older SL6/2.7 combination using older SL6/2.7 clients

# dCache and Open vSwitch

- **AGLT2** has been planning to implement Open vSwitch (OVS) for almost a year
  - **Goal:** provide a mechanism to test Software Defined Networking for real LHC physics tasks
  - **OVS** gives us a means to get to the “source & sink” of data (visibility and control)
- Starting in September we began testing the OVS instructions worked out with help from **Ben and Galen Mack-Crane / CORSA**:
  - [https://www.aglt2.org/wiki/bin/view/Main/Open\\_vSwitch/InstallOVSonSL73](https://www.aglt2.org/wiki/bin/view/Main/Open_vSwitch/InstallOVSonSL73)
- As of the beginning of this month, all the **AGLT2** dCache storage servers are running SL7.4 and have their public IPs on OVS
  - Installing OVS and cut-over the IP to the virtual switch was done “hot” on production servers. Approximately 40 second downtime not a problem for transfers
  - [https://www.aglt2.org/wiki/bin/view/Main/Open\\_vSwitch/ConfigureOpenvSwitchSL6](https://www.aglt2.org/wiki/bin/view/Main/Open_vSwitch/ConfigureOpenvSwitchSL6)
- **Have been running for almost 2 weeks without any issues. We are now ready to start LHCONE P2P experiments and have interest from other ATLAS sites in trying this.**



# Future Plans

- Participating in **SC17**
  - Will demo use of **OSiRIS** storage service: sign up on the floor and use the storage during SC17 and possible use of object store for ATLAS
  - Demo with **MWT2** and **PRP**: ML on ATLAS calorimetry data AND on **perfSONAR** analytics. Data will be sourced from **AGLT2**
- Experimenting with SDN/OVS in our Tier-2 and as part of LHCONE point-to-point testbed.
- Update to **VMware** soon: new version (5.5->6.5), new configuration for HA, new ESXi host at MSU
- Working on IPv6 dual-stack for all nodes in our Tier-2
  - Have IPv6 address block for AGLT2 (spans UM/MSU)
  - Our perfSONAR nodes are all dual-stacked and working fine.
  - **Dual-stacking our dCache system is our next step. Planned for sometime in the next month or two.**
- Moving all worker nodes to **SL7** in the next two months

# Summary

- Monitoring stabilized, update procedures working well
- Tier-2 services and tools are evolving. Site continues to expand and operations are smooth.
- FUTURE: IPv6 for storage, SL7 WN, SDN testing, SC17

## Questions ?