AGLT2 Site Report

Wenjing Wu/University of Michigan
Dan Hayden, Philippe Laurens, Shawn Mckee

USATLAS WBS 2.3 Face to Face, SLAC

https://indico.cern.ch/event/1201515/ Nov 30, 2022

ATLAS Great Lakes Tier 2
UNIVERSITY OF MICHIGAN AGLT2 MICHIGAN STATE UNIVERSITY

AGLT2 Overview and Status (1)

- **HTCondor Cluster**:
  - 357 Worker Nodes (186/UM, 171/MSU), 18 kCores (59% @UM, 41% @MSU), 225 kHS06(60% @UM, 40% @MSU), Avg. 12.46 HS06/core
  - 2GB ~6.3 GB RAM/core, 1000 job slots for High Memory Queue (6 GB/core)
  - 14GB ~52 GB Disk/core, supports Merge Queue with higher disk requirement.
  - 2 x 10 or 2 x 25 Gbps bonded NICs for 90% work nodes (2 x 1 Gbps for the 10% oldest nodes)
- **dCache storage**:
  - 12 PB deployed (56% @UM,44%@MSU), **11 PB in space tokens**
  - 2 x 25 Gbps or 4 x 10 Gbps bonded NICs for almost all storage nodes (2x10 Gbps for 5x oldest)
- Continue to use and optimize the ATLAS@home backfilling on the HTCondor cluster.
  - optimize the cgroup and **BOINC** configurations to reduce the impacts on the CPU Efficiency of the Condor jobs.

AGLT2 Overview and Status (2)

- **Networking and site upgrades**
  - Established separate multipath 100G MSU-UM inter-site connectivity
  - We have completely (and very successfully) replaced our UM LAN and UM/MSU WAN network equipments starting June 2021 and finished WAN and routing changes in March 2022
  - MSU site migration to the MSU new data center completed in Fall 2021, which provided
    - 12x33kW racks with dual/true redundant power supplies, multipath 100G WAN and dual/redundant data switches with 25 Gbps ports in each rack, also with room for expansion.
    - including new rack network devices, optics, cabling, configuration, at no cost to AGLT2.
  - Upgraded UM networking gear,cabling and PDU
    - Separate new data (100 Gbps/port) and management switches (1 Gbps/port) for each Rack
    - multipath 100G WAN
    - Replaced all DAC cables with AOC, RJ45 with slim RJ45 to reduce cable spaces in Racks.
    - upgraded all racks with new Smart PDUs with individual socket meter and control

AGLT2 Equipment Purchase 2023

- Year 2023 Purchase Total Budget 415K (combine left over budget from 2022 and 2023)
  - $30K software, licenses and support infrastructure
  - $385K compute, storage and network (25% for computing, 75% for storage and network)
    - Estimated (20% up based on quote from 2022 purchase) to add 3.8 PB usable space (12 R740xd2 with 24*16TB disks) and 17.2 kHS06 (10 R6525)
    - network expenses are for cables and transceivers for the new storage and work nodes.($225/storage node, $100/work node)
  - Est. Total after 2023 retiring/purchase: **CPU: 225 - 6.628 - 10.8 + 17.2 = 224.77 kHS06 (save 5.4 kw power) / Disk: 12 - 1.5 + 3.8 = 14.3 PB**
  - To achieve the same ratio (64.8GB/HS06) for 2 sites: UM 9 storage, MSU 3 storage + 10 WN

Software and Technology Status (1)

- **AGLT2 runs a number of software packages required for an ATLAS site**:
  - OSG 3.6/HTCondor-CE 5.1.5/HTCondor 9.0.16
  - dCache 7.2.20 (most recent version: 7.2.27)
- **vmware cluster**
  - To host and manage critical services we rely upon VMware, which provide high availability and supports live migration of services to allow hardware, firmware and software updates
  - VMware 6.7U3 (plan to upgrade to 7.x as 6.7 already reached its EOL), new license issue blocks the progress. (allow only 32 core/node per license, our hardware is 64 core/node)
- **Storage**
  - Lustre(1.7PB, 2.12.8 on CentOS 7 and 2.15.1 on CentOS 8 stream)
  - NFS(0.5PB), AFS(1.8.7) and have collaborative access to Ceph.
- **Site Monitoring**
  - We have a combination of custom built monitoring tools, along with CheckMK(2.0.0-p6), Elasticsearch(7.17), Zeek(4.2.0), Elastiflow(5.3.4) and NetDisco to provide required management and operations visibility.
- **Tape Backup**: Amanda 3.5.1 on CentOS 7 (most recent version: 3.5.2 on CentOS 7)
- **Provision**: Cobbler 2.8.5 on CentOS 7 (most recent version: 3.3.3 on CentOS Stream 8)
- **Configuration**: CFEngine 3.7.2 on CentOS 7 and 3.12.4 on CentOS 8 Stream (most recent version : 3.20.0 on CentOS 8 stream)

Software and Technology Status (2)

Next generation of Operating System

- Currently our base operating system is CentOS 7.9 but we would like to migrate to either RHEL 9 or a RHEL 9 compatible OS (Rocky 9, Almalinux 9, CentOS Streams 9).
- We succeed in adding CentOS 8 Stream ir cobbler to build different flavors of nodes and make it work with CFEngine to configure the nodes.
- We are experimenting migrating different software systems to CentOS 8 Stream to prepare for the transition to 9.

software	Status	Notes
lustre server	finished	2.15.1 installed on CC8 Stream
lustre client	finished	2.15.1 installed on CC8 Stream
Amanda Server	In progress	No rpm, src tarball does not build on CC8 Stream
Amanda client	In progress	No rpm, src tarball does not build on CC8 Stream
openafs server	finished	1.8.8 built from src rpm, in umatlas repository
openafs client	finished	1.8.8 built from src rpm, in umatlas repository
CFEngine Server	Not started	Need to install 3.12.4 and test existing code
CFEngine client	finished	3.12.4 in umatlas repo, working with CFEngine 3.7.2 server and the majority of the existing code are fixed to work with the 3.12.4 client
OSG software	finished	OSG 3.6 installed on CC8 Stream
HTCondor client	In progress	9.0.17 installed on CC8 Stream, problem with starting jobs
HTCondor Server	Not started	Need to install 9.0.17 on CC8 Stream
Cobber server	In progress	Install cobbler 3.x on CC8 Stream
checkmk	In progress	Client rpm is ready for CC8 Stream

Research Areas (1)

- **VMware**
  - Continuing to use it for services (lots of resiliency and features). Both MSU and UM institutions have covered at least this year's licence costs (~$10K)
  - We added vSAN cluster (vSAN aggregates local or direct-attached data storage devices to create a single storage pool shared across all hosts in a vSAN cluster)
  - We deployed multi-100G TrueNAS iSCSI storage systems for VMware VMs, replacing unsupported Dell MDxxxx systems at higher performance and lower costs.
- **Network**
  - WLCG Security Operations Center: Hardware in place, needs Zeek deployment and MISP integration
  - WLCG Site monitoring in place, flow labeling on dCache hosts, deployed PTP

Research Areas (2)

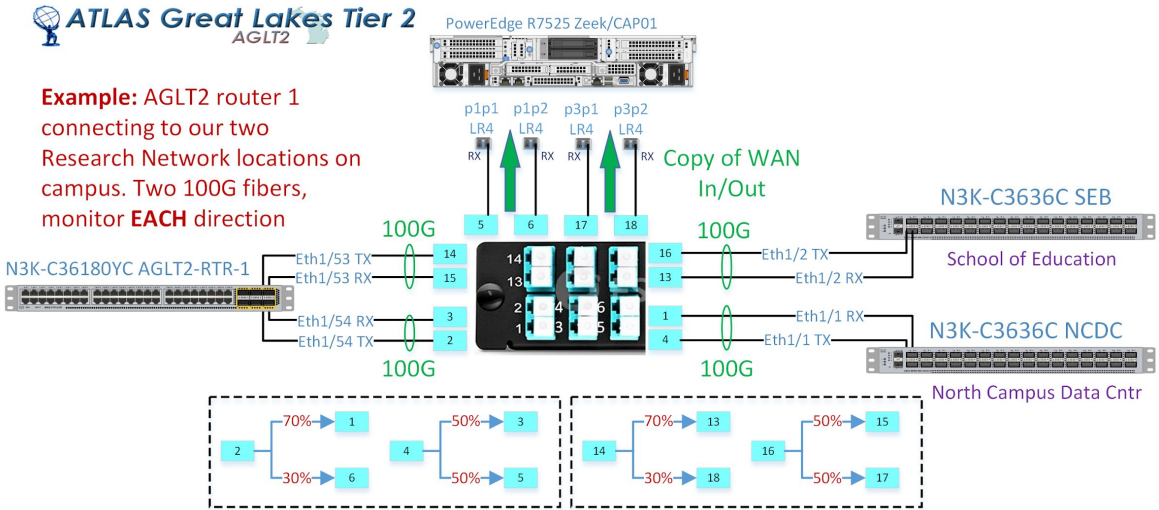
- **Monitoring**
  - Using CheckMK for service/host monitoring, packaged using containers with LE & NGINX
  - Continue to maintain/upgrade Elasticsearch for syslogging, security and monitoring
  - Elastiflow deployed monitoring UM in/out flows (utilizes our Elasticsearch)
- **BOINC**:
  - Improves overall cluster CPU utilization and fill the cluster during site downtime/cluster draining.
  - performance tuning/merge to gratia accounting.
- **Provision**:
  - Considering best options for future provisioning (Cobbler, Foreman, ?)
  - Using Ansible and Github for switch configuration and backup

Network Security

AGLT2 has been working with the WLCG SOC effort to help secure our networks while maintaining performance

Our original network had a Zeek+MISP+Elasticsearch setup for dual 40G. Cost to set up was about $2K plus repurposing an R630

Our new network is **4x100G**
We have purchased two "network capture" nodes (Dell R7525) each with two Bluefield-2 NICs (each 2x100G)
Have a milestone for April 2023 to get it into production…



CAP01/RTR-1 diagram (similar for CAP02/RTR-2)

Enabling PTP (1 / 2)

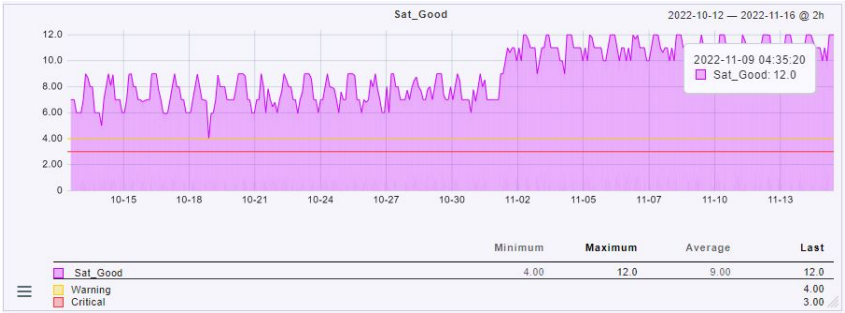
For about $1500, AGLT2 added dual GPS clocks to enable PTP

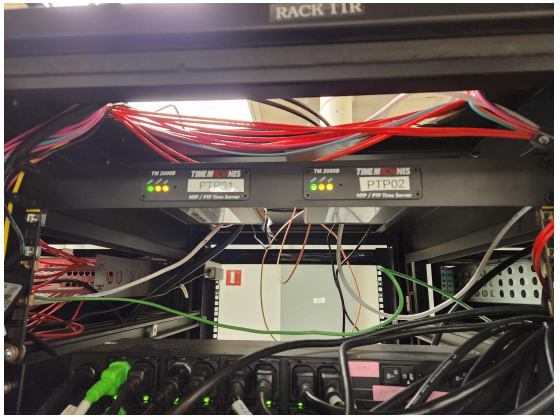
- Challenge is the antenna; ideally switch support
- PTP provides < 1 microsecond time accuracy
- Makes perfSONAR latency much more powerful **BUT** needs pS mods

  To do: PTP Clients (NTP ~20 u-secs)

Enabling PTP (2 / 2)

Start of Nov, we got a our antenna installed on the roof of Physics



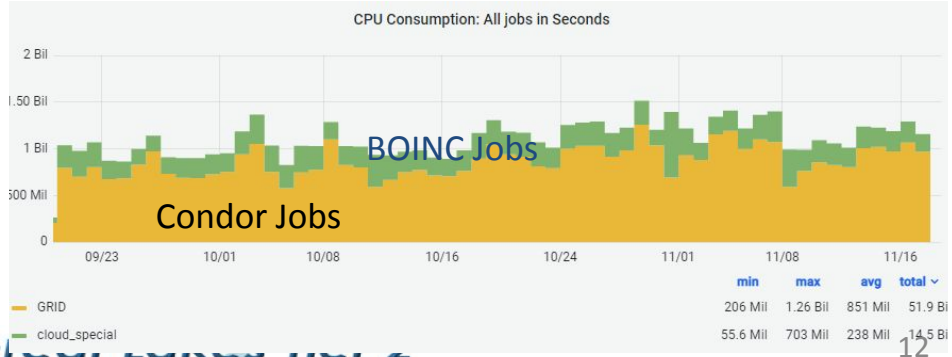


BOINC work

- **BOINC optimization**
  - configuring BOINC as a service and put it under the system.slice cgroup can reduce the CPU Efficiency loss for HTCondor jobs by 5%
  - having BOINC jobs use 50% of the cores (instead of 100%)can further reduce the CPU Efficiency loss for HTCondor jobs by another 5%.
  - we reconfigured the BOINC services and its cgroup configurations on all the work nodes.
- **Harvest from running BOINC jobs.**
  - It increases the CPU Utilization of the cluster, taking the past 2 months for example, the average cores for ATLAS is 16000, and the CPU Utilization reaches **98%** combining both Grid and BOINC jobs (80% for Grid, and 18% for BOINC jobs)
  - Fill the cluster during site downtime/cluster draining (HTCondor update)/grid service or network issues as BOINC jobs requires only the work node itself and intermittent network access

AGLT2 Site CPU Utilization(16000 CPU cores in 60 days)
Mon Sep 19 01:00:00 2022 to Fri Nov 18 00:00:00 2022

	cpu_eff	cpu_util	wall_util
BOINC	0.35	0.18	0.50
Grid	0.88	0.80	0.97
All	0.69	0.98	1.47



Site milestones/Concerns

- **Milestones:**
  - **See milestones spreadsheet, rows 99-103 for AGLT2 specific milestones and rows 82, 87, 91 and 95 for additional milestones.**
  - **AGLT2 Finish TrueNAS install (12/1/2022) - Not yet completed (waiting for network configuration)**
  - **Update VMware to 7.X (12/31/2022)**
  - **Enable UEFI Boot support in Cobbler (1/31/2023)**
  - **AGLT2 Implement SOC (4/15/2023)**
  - **Build new dCache nodes with JBOD /zfs (6/1/2023)**
- **Concerns**
  - **the transition from CentOS 7 to next generation OS (RHEL 9 flavor), software such as CFEngine, Amanda. We are trying to migrate software to CentOS 8 stream before their RHEL 9 releases are ready.**
  - **Because of the power limit(80kw UPS) in the UM Tier2 room, we might need to put more future work nodes to the MSU site.**

AGLT2 Contributions to the Facility

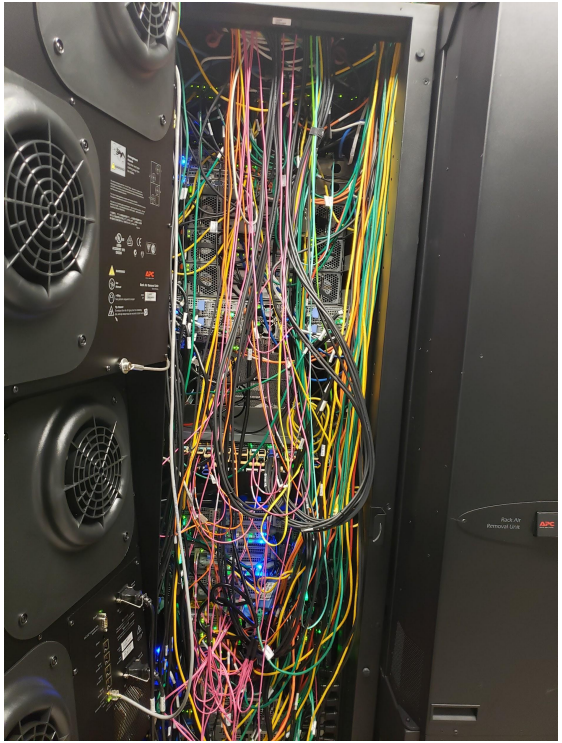
- **Documenting using BOINC backfilling jobs to improve the CPU Utilization of the cluster.**
- **Testing using PTP for precision time synchronization.**
- **Work in security area: WLCG SOC.**
- **Sharing expertise in deploying and using monitoring tools.**
- **Sharing knowledge about dCache operation and troubleshooting.**
- **Providing networking help in procurement, deployment and operations.**

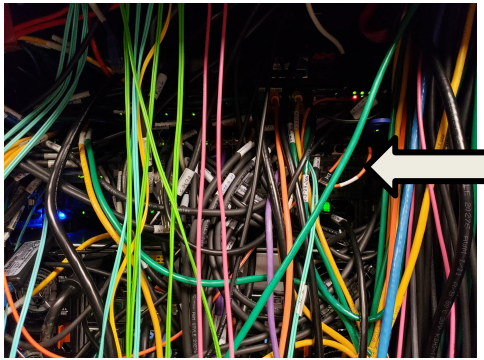
Questions or Comments?

Backup Slides

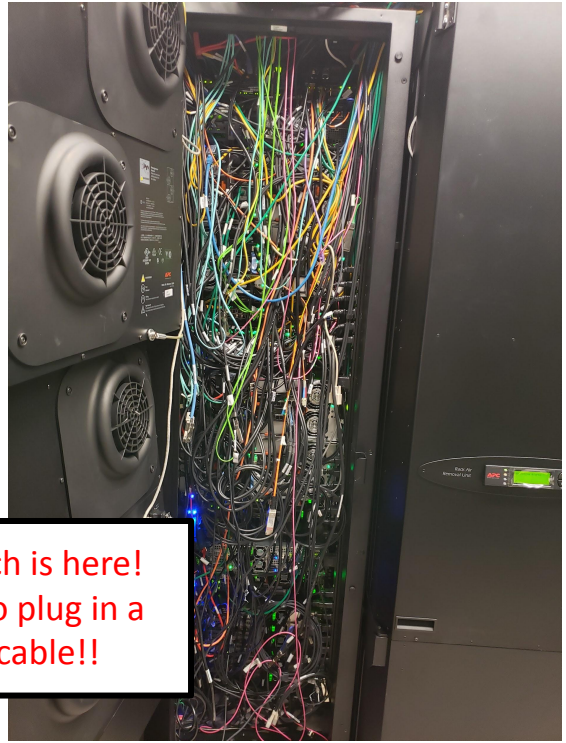
Part of Our Problem

Cabling: messy, wrong/missing labels, bad airflow, unworkable!





Switch is here!
Try to plug in a new cable!!



Network Monitoring and Measurement

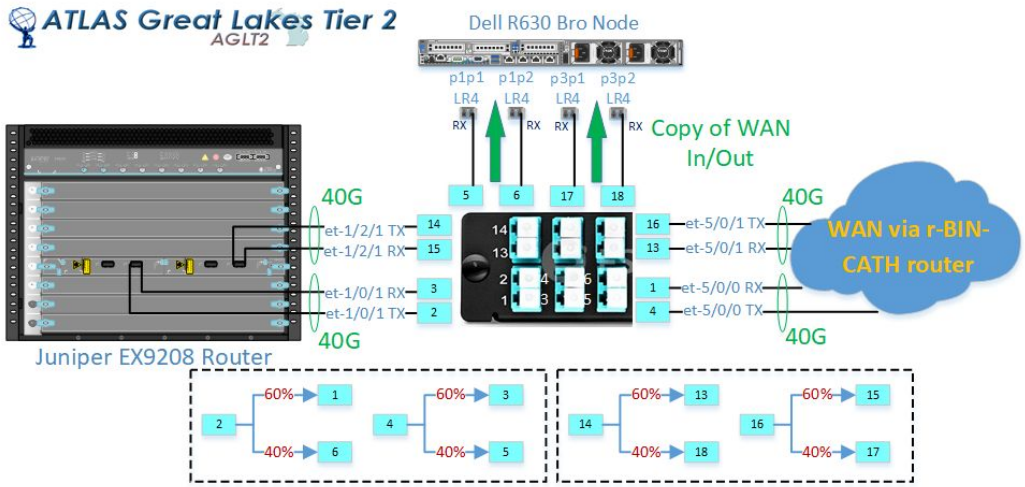
- **In order to see and understand how our network is performing.**
  - AGLT2 uses a combination of CheckMK, Elasticsearch and custom scripts to monitor our networks.
  - CheckMK provides port traffic/errors/discards (switch/server)
  - Elasticsearch tracks logging from devices and various metrics
  - For our new network we are deploying multiple 100G perfSONARs, strategically located to cover our border, storage location equivalent and inter-site.

Old Network Security

AGLT2 has been working with the WLCG SOC effort to help secure our networks while maintaining performance

Our original network had a Zeek+MISP+Elasticsearch setup for dual 40G. Cost to set up was about $2K plus repurposing an R630

Our new network is **4x100G**
We have purchased two “network capture” nodes (Dell R7525) each with two Bluefield-2 NICs (each 2x100G)
Have a milestone for April 2023 to get it into production…



Goals for our Network Rewire/Upgrade

For AGLT2, our goals is to fix a number of issues

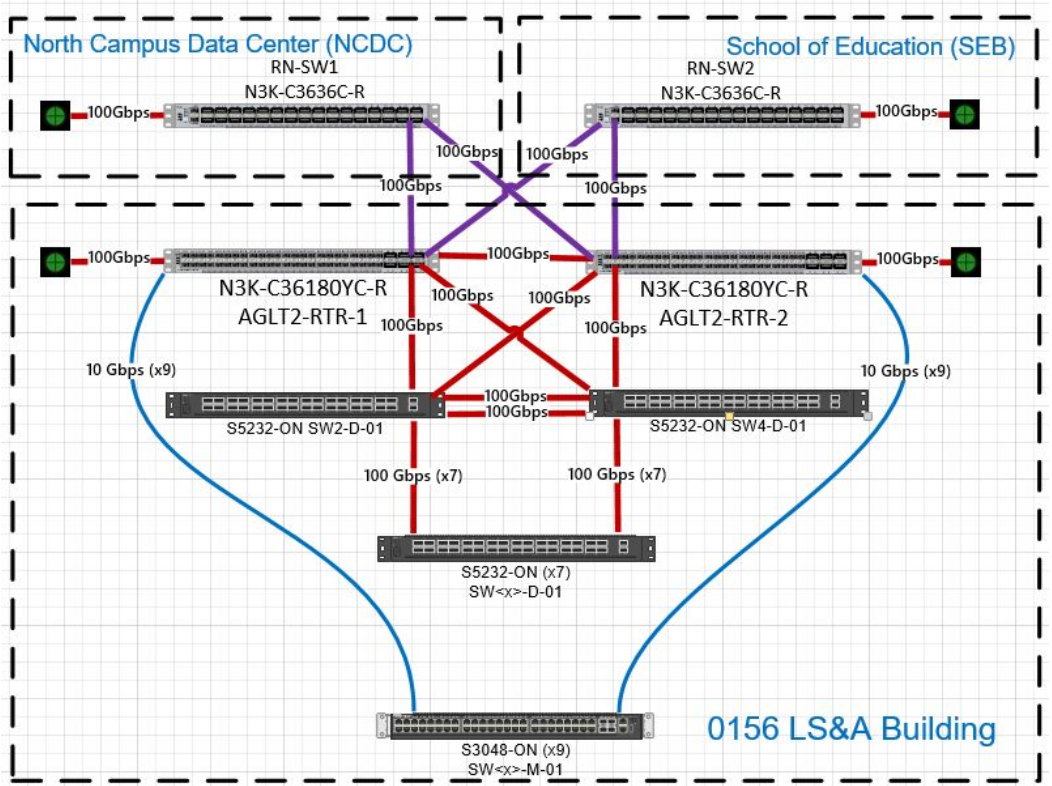
- Incrementally acquiring servers and network devices **created a mess** in terms of wiring and airflow
- Using multiple vendors switches (of different ages) has led to a **fragile network** that is unable to be optimized, managed and debugged
- Not have **central configuration** of our switches has led to mistakes and difficulty in implementing new features
- **Lack of resiliency makes** upgrades and config changes difficult, requiring downtimes to make major updates.

New AGLT2 UM LAN

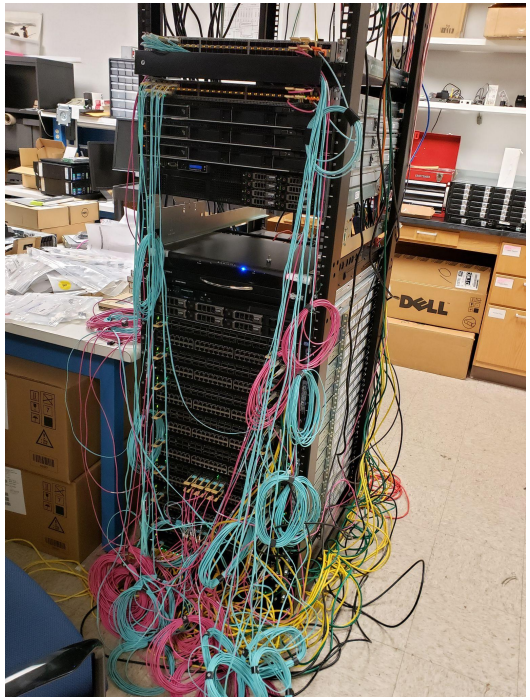
The new LAN/WAN design has a border, core and rack-level access on both **data** and **management** planes.

Unreached goal: VXLAN, EVPN to each rack (problem between Dell and Cisco)

Resiliency from LACP(VLT,MLAG) trunks between redundant switches



UM Switch Testing Rack



UM staged all the new LAN/WAN equipment in a testing rack
Allowed us to test and preconfigure switches using Ansible

UM Rewiring Details

Our plan is to completely recable our AGLT2 UM site, first removing all existing cabling (power, networking, etc) and then neatly installing new **pre-labeled** cables

- RJ45 will use “slim” cables
- DAC cables will be replaced by fiber optic cables + transceivers
- Use cable management (horizontal/vertical)
- Power cables are length optimized
- We will minimize inter-rack cabling

Have 1 week downtime planned (June 14-18)

Acknowledgements

We would like to thank MSU and UM networking personnel, especially Tim DeNike and Nick Grundler for contributing to AGLT2 networking and this presentation.

We want to especially acknowledge the National Science Foundation for their support on grant # 162473