

US ATLAS SWT2 SITE REPORT

HORST SEVERINI, CHRIS WALKER, JOEL SNOW, PATRICK MCGUIGAN, MARK SOSEBEE

HEPIX MEETING FALL 2017

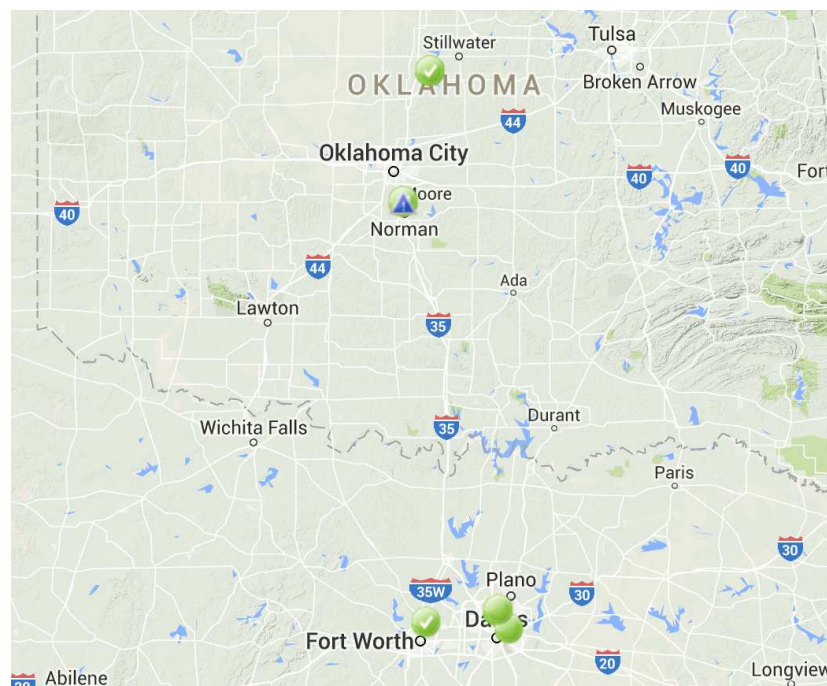
OCTOBER 2017

Outline

- Introduction
- Computing and Storage Hardware
- Network
- Upgrade Plans
- Opportunistic Cycles and Container Testing

US ATLAS SWT2 Center

- University of Oklahoma
 - Oklahoma Center for High Energy Physics (OCHEP)
 - OU Supercomputing Center for Education and Research (OSCER)
- University of Texas Arlington
 - Chemistry and Physics Building (CPB)
 - Arlington Regional Data Center (ARDC) in Fort Worth
- Langston University
 - Lucille



OU_OCHEP_SWT2 Tier 2 Hardware

- 72 Nodes (844 Slots) – 2 GB RAM per Slot
- 5 Support Nodes
- No local storage – Lustre storage (DDN 9900) died
- ROCKS 6 (RHEL 6), HTCondor, OSG 3
- Also, OUHEP_OSG Tier3 HTCondor Desktop Cluster
 - 275 cores, 400 TB usable XFS/NFS storage
 - Used for local Athena, root, and theory calculations
 - 5 node OSG Integration Testbed cluster

OU_OSCER_ATLAS Tier 2 Hardware

- 41 Nodes (1800 Slots) – 2-4 GB RAM per Slot
- 8 Support Nodes (2 head, 6 storage)
- 500 TB of usable xrootd storage (5 T630s with 16 8 TB drives)
- SALT (CentOS 7.3), SLURM 17.02, OSG 3.3
- Part of generic OSCER HPC cluster
- Rest of OSCER Schooner Hardware
 - 510 Nodes (about 10k Slots) – 2-4 GB RAM per Slot
 - Opportunistically available for ATLAS production

Lucille Hardware

- 30 Nodes (960 Slots) – 4 GB RAM per Slot
- 5 Support Nodes (2 head, 3 storage)
- 110 TB of usable XFS storage (MD1200/MD3200)
- ROCKS 6.1 (CentOS 6.9), HTCondor, OSG 3.3
- Cluster also has some GPU nodes that have been used for other OSG testing

UTA Tier 2 Hardware

- **SWT2_CPB**

- 439 Nodes (8736 Slots) – 2-3 GB RAM per Slot
- 42 Support Nodes (15 head, 27 storage)
- 5500 TB of usable xrootd storage (MD1000/MD1200/MD3X60)
- ROCKS 6.1 (SL 6.3), PBS/Torque, OSG 3.2

- **UTA_SWT2**

- 110 Nodes (2360 Slots) – 2 GB RAM per Slot
- 14 Support Nodes (9 head, 5 storage)
- 120 TB of usable xrootd storage (MD1000/MD1200)
- ROCKS 6.1 (SL 6.3), PBS/Torque, OSG 3.2

Network

- OU connected at 100 Gbps to I2 and ESnet via OneNet
- LU connected at 10 Gbps to OneNet
- UTA connected at 20 Gbps to LEARN via DMZ
- Working on 40 Gbps DMZ for OSCER cluster
- DMZ will run between OU, OSU, LU, TU, and UCO
- Hoping to get opportunistic cycles for ATLAS production on those HPC resources as well

Equipment Deployment

- Have 34 R430 and 7 R630 compute nodes deployed in Schooner as dedicated 'Condo' nodes, as well as 5 T630 storage nodes
- Have ordered and received 2 more T630s for a total of 700 TB usable xrootd storage
- Will keep running older OCHEP nodes until they die
- Space, Power, Cooling not an issue
- UTA used all available funds and deployed all hardware



Opportunistic Cycles

- 10k cores on Schooner
- Expect good usage as 'fair share OU user'
- Started running Event Service jobs on new hep_killable queue
- Got okay from OSCER Director to dynamically shorten requested maxwalltime of these jobs until they start, will not be preempted
- Constructed simple loop script which does that
- For now, only SCORE jobs; possible add MSCORE jobs if throughput is good enough

Container Deployment

- Singularity 2.2.1 (will upgrade to 2.3.2 soon)
- Ran almost successful test jobs on `OU_OSCER_ATLAS_TEST`
 - Solved initial bind-mount issues by enabling overlay
 - `singularity_options='-B /cvmfs,$workdir,/home --contain'`
 - Currently stuck on making HTCondor-Spool directory available in VM
 - Possibly still bind-mount related

Summary and Conclusions

- SWT2 active and successful ATLAS Tier2 Center
- Consistently in Top 3 in the US, Top 5 in the World; last month even Top 2 in both
- Very active in ATLAS/Panda Testing and Development
- Also OSG Integration Testing and Deployment
- Come visit our SWT2 Booth (989) at Supercomputing in Denver!

