# ATLAS Great Lakes Tier-2 (AGL-Tier2)

**Shawn McKee (for the AGL Tier2)**

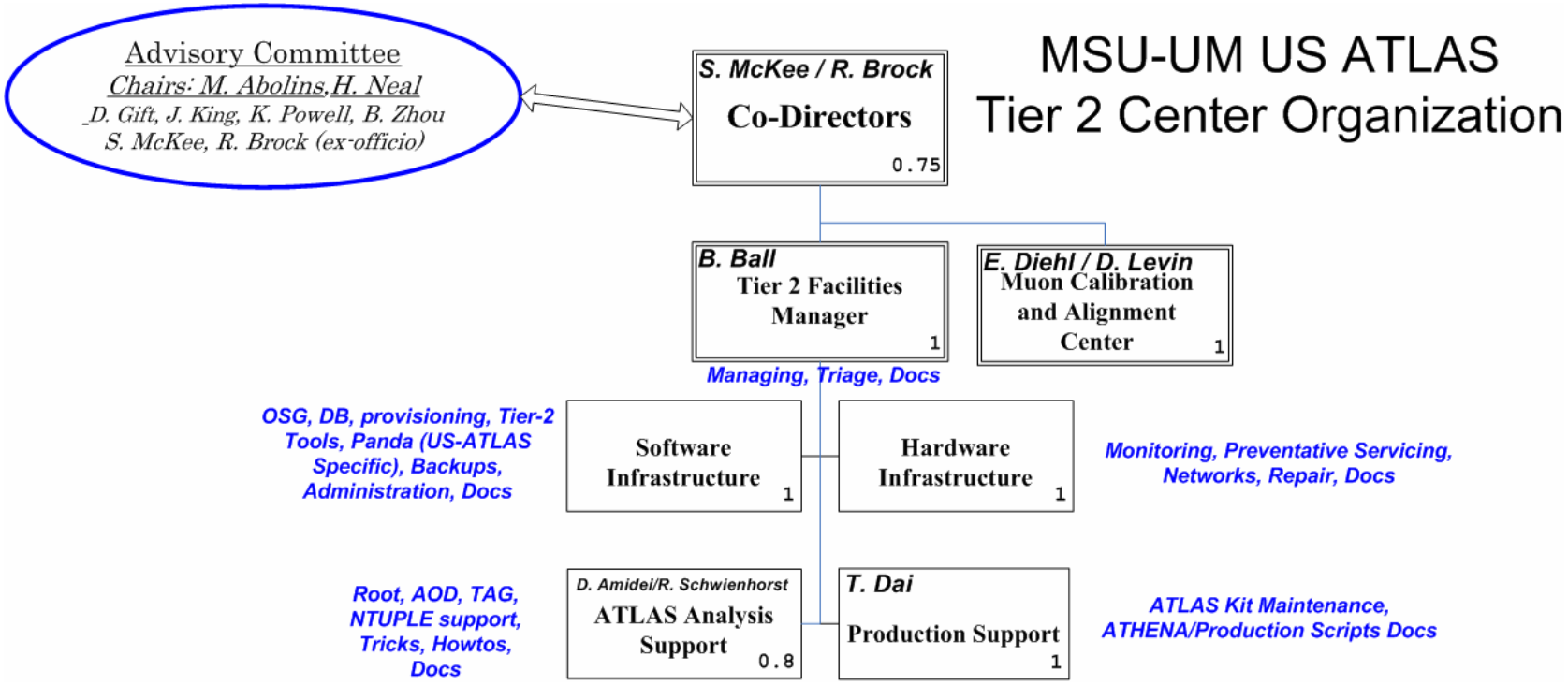
**University of Michigan**

**US ATLAS Tier-2 Meeting at Harvard**

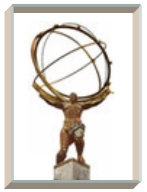
**Boston, MA, August 17th, 2006**

# The AGL Tier2

The ATLAS Great Lakes Tier-2 (AGL-Tier2) is a joint effort of Michigan and Michigan State Universities.



# AGL Personnel Involved



- We have a large group of faculty between MSU and UM who will be involved in related aspects of the AGL-Tier2
- Table at the right shows the expected contribution by name and year

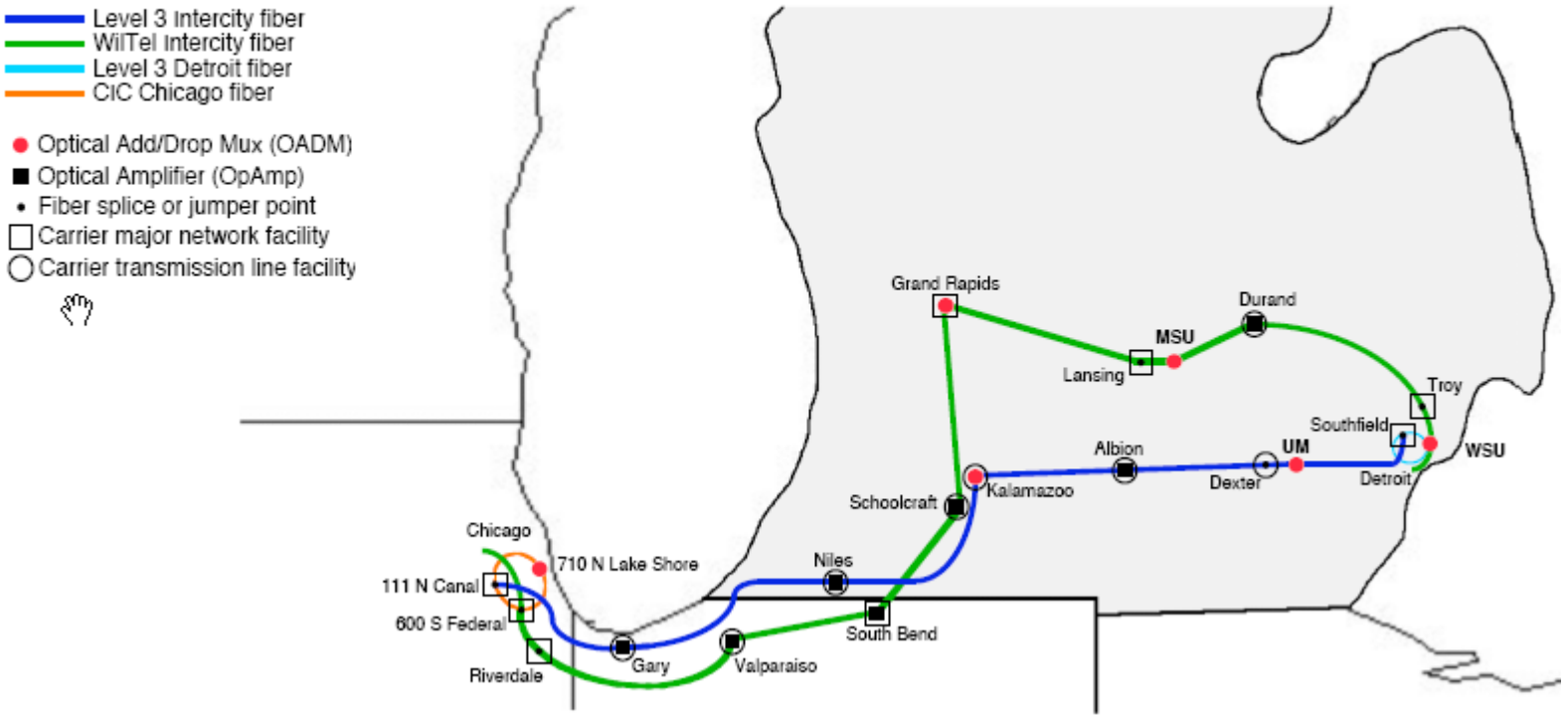
**Tier-2 Related Commitments**

| Name | UM TIER-2 ROLE | FY06 | FY07 | FY08 | FY09 | FY10 |
|---|---|---|---|---|---|---|
| D. Amidei | Software and Simulation | 0.1 | 0.15 | 0.2 | 0.25 | 0.25 |
| B. Ball | Facility Manager | *1* | *1* | *1* | *1* | *1* |
| J. Chapman | Muon Performance and Validation | 0.1 | 0.1 | 0.1 | 0 | 0 |
| T. Dai | Data Analysis and Services | 0.2 | 0.4 | 0.6 | 0.6 | 0.6 |
| E. Diehl | Muon Calibration and Alignment | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| S. Goldfarb | Muon Soft Coordination/Collab. Tools | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| D. Levin | Muon Calibration and Validation | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |
| S. McKee | Principal Investigator, Software | *0.5* | *0.5* | *0.5* | *0.5* | *0.5* |
| H. Neal | Simulation and Collaborative Tools | 0.25 | 0.25 | 0.25 | 0.25 | 0 |
| J. Qian | Software and Simulation | 0.1 | 0.15 | 0.2 | 0.25 | 0.25 |
| R. Thun | Muon Calibration and Validation | 0.1 | 0.1 | 0.1 | 0.1 | 0 |
| Z. Zhao | Data Validation and MC Simulation | 0.15 | 0.2 | 0.2 | 0.2 | 0.2 |
| B. Zhou | Muon Calibration and MC Simulation | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| **UM TOTAL** | *Leveraged* | **2.25** | **2.6** | **2.9** | **2.9** | **2.55** |
| Name | MSU TIER-2 ROLE | 2006 | 2007 | 2008 | 2009 | 2010 |
| M. Abolins | Software and Simulation | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| R. Brock | Center Co-Director, Software, Simul. | 0.3 | 0.3 | 0.4 | 0.4 | 0.4 |
| C. Bromberg | Software and Simulation | 0.1 | 0.1 | 0.15 | 0.15 | 0.15 |
| J. Huston | Software, calibration, and Simulation | 0.2 | 0.3 | 0.4 | 0.5 | 0.5 |
| J. Linnemann | Software and Simulation | 0 | 0.1 | 0.1 | 0.1 | 0.1 |
| B. Pope | Software and Simulation | 0.2 | 0.2 | 0.3 | 0.3 | 0.3 |
| R. Schwienho | Data Analysis, support and services | 0.2 | 0.2 | 0.3 | 0.3 | 0.3 |
| K. Tollefson | Software and Simulation | 0 | 0.05 | 0.1 | 0.15 | 0.15 |
| R. Hauser | Software, Data Analysis and services | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| **MSU TOTAL** | *Leveraged* | **1.7** | **1.95** | **2.45** | **2.6** | **2.6** |
| **TIER-2 LEVERAGED TOTAL** | | **3.95** | **4.55** | **5.35** | **5.5** | **5.15** |

*n* Means not counted in "leveraged" because they are dedicated to Tier-2

# 10GE Protected Network

## Michigan LambdaRail (MiLR)




- We will have a single "/23" network for the AGL-Tier2
  - Internally each site (UM/MSU) will have a /24
- Our network will have 3 10GE wavelengths on MiLR in a "triangle"
  - Loss of any of the 3 waves doesn't impact connectivity for both sites

# Existing MSU Cluster

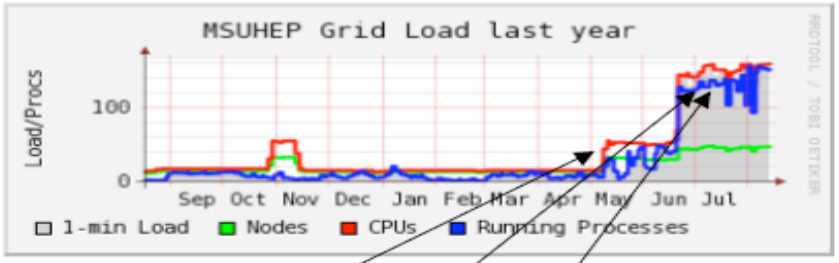
## MSU "WrigleyField" Cluster

*Devoted to DØ MC production since May*

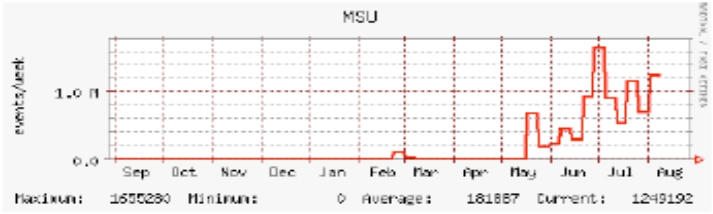
**MSUHEP Grid (2 sources)** (tree view)

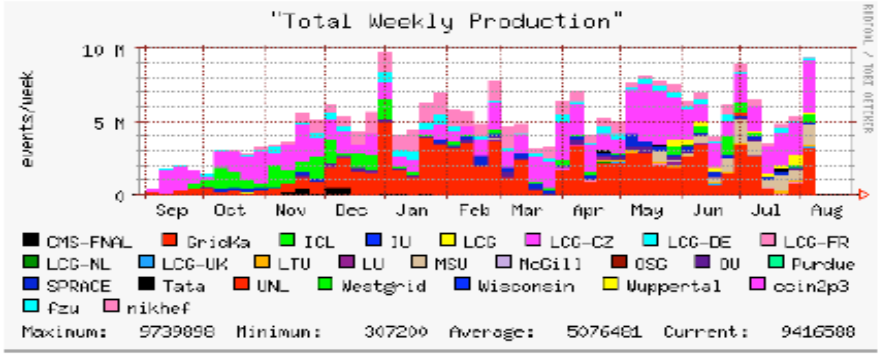
| | |
|---|---|
| CPUs Total: | 159 |
| Hosts up: | 47 |
| Hosts down: | 3 |

Avg Load (15, 5, 1m):




20 Western Scientific Opteron 246 2GHz, dual, single core; 2GB RAM, 250GB drives
(single head node with 250GB raid 1)
*upgraded to 20 Opteron 270 dual, dual core*
17 TeamHPC Opteron 270 2GHz, dual, dual core; 2GB RAM, 300GB drives



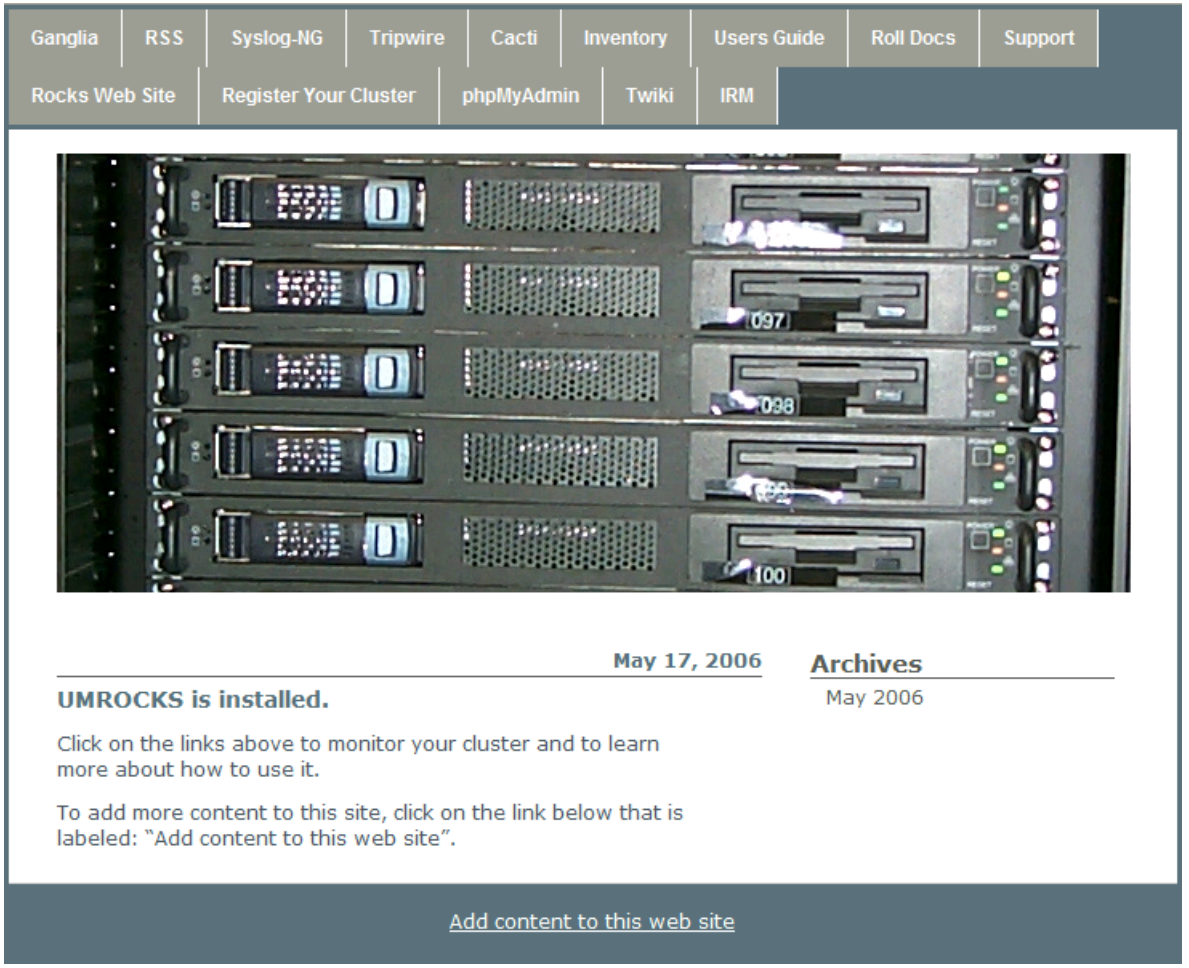



# Current UMROCKS Cluster

- We have a 5 rack AMD Athlon cluster with 70 operational nodes (2000/2400/2600 dual processor, 2 GB RAM)
- Two 100+GB disks
- Plan to have ~100 nodes operational
- ROCKS V4.1

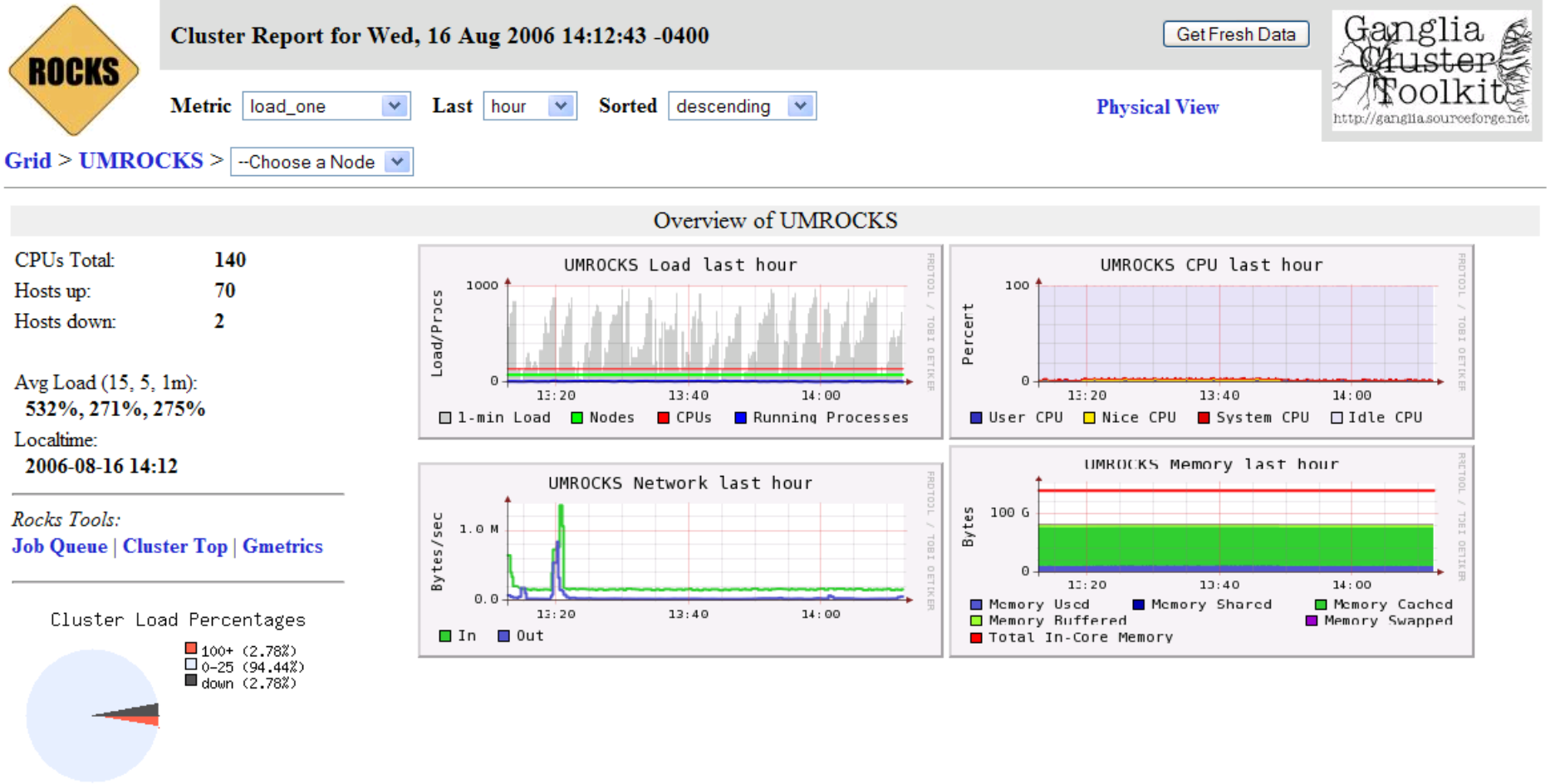
# UMROCKS Installation

- ROCKS V4.1 with SLC V4.3 (32 bit) OS
- Significant additional software: TWiki, Cacti, Inventory, Syslog, etc.
- Cluster running dCache as well

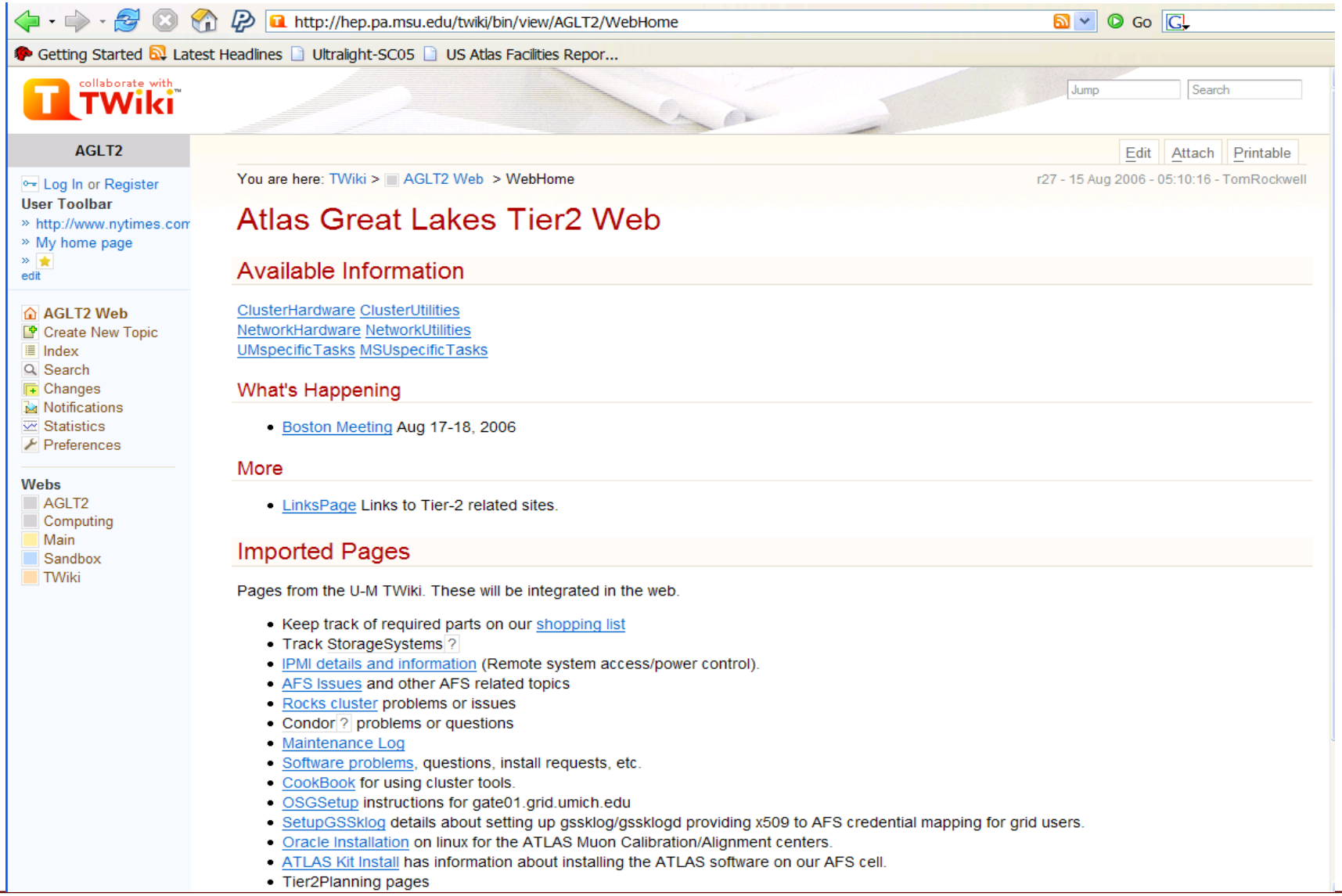


# Ganglia Info on Athlon Cluster

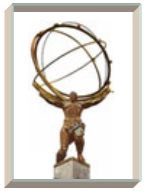
- Currently ~70 operational nodes – plan for ~100 from parts



# AGL-Tier2 TWiki



# Cacti Graphing/Monitoring



console | graphs | monitor | threshld | weathermap

Console -> Devices

Logged in as **admin** (Logout)

**Create**
New Graphs
**Management**
Graph Management
Graph Trees
Data Sources
**Devices**
Thresholds
Weathermaps
**Collection Methods**
Data Queries
Data Input Methods
**Templates**
Graph Templates
Host Templates
Data Templates
Threshold Templates
**Import/Export**
Import Templates
Export Templates
**Configuration**
Settings
**Utilities**
System Utilities
User Management
Network Tools
Logout User

**Devices** Add

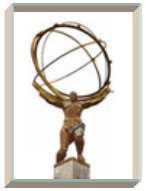
Type: Any | Status: Any | Search: | go | clear

<< Previous — Showing Rows 1 to 30 of 31 [1,2] — Next >>

| Description | Status | Hostname | Current (ms) | Average (ms) | Availability |
|---|---|---|---|---|---|
| Amazon (Cisco 6506) | Up | amazon.grid.umich.edu | 39.74 | 43.57 | 100% |
| atgrid | Down | atgrid.grid.umich.edu | 36.42 | 37.56 | 89.52% |
| attera | Up | attera.grid.umich.edu | 0.83 | 32.21 | 98.34% |
| gate01 | Up | gate01.grid.umich.edu | 36.56 | 37.98 | 85.98% |
| hypnos | Down | hypnos.grid.umich.edu | 37.83 | 35.91 | 76.85% |
| linat02 | Up | linat02.grid.umich.edu | 0.12 | 16.87 | 100% |
| linat03 | Up | linat03.grid.umich.edu | 0.14 | 22.31 | 99.98% |
| linat04 | Up | linat04.grid.umich.edu | 0.43 | 19.69 | 100% |
| linat05 | Up | linat05.grid.umich.edu | 0.78 | 16.87 | 99.81% |
| linat06 | Up | linat06.grid.umich.edu | 36.19 | 38.03 | 99.43% |
| linat07 | Up | linat07.grid.umich.edu | 37.02 | 36.55 | 100% |
| linat08 | Up | linat08.grid.umich.edu | 38.9 | 36.69 | 100% |
| linat09 | Up | linat09.grid.umich.edu | 37.28 | 38.05 | 99.78% |
| linat10 | Up | linat10.grid.umich.edu | 38.64 | 37.3 | 100% |
| linat11 | Up | linat11.grid.umich.edu | 37.31 | 37.91 | 99.98% |
| mars01 | Up | mars01.cern.ch | 170.81 | 167.36 | 99.6% |
| mars02 | Up | mars02.cern.ch | 169.52 | 167.1 | 99.55% |
| mars03 | Up | mars03.cern.ch | 171.71 | 167.41 | 99.57% |
| mars04 | Up | mars04.cern.ch | 171.9 | 167.18 | 99.56% |
| mars05 | Up | mars05.cern.ch | 170.83 | 167.02 | 99.54% |
| mercury | Up | 141.211.43.122 | 0.68 | 32.87 | 93.31% |
| ml-um | Up | ml-um.ultralight.org | 48.13 | 48.71 | 98.78% |
| Nile (Cisco 3512XL) | Up | nile.physics.lsa.umich.edu | 38.92 | 37.32 | 99.88% |
| Nile (Cisco 6509) | Up | nile.grid.umich.edu | 36.41 | 40.53 | 99.99% |
| Oracle server | Up | umors.grid.umich.edu | 0.69 | 8.9 | 99.95% |
| tera02 | Up | tera02.ultralight.org | 47.04 | 48.25 | 97.17% |
| UMFS01 | Up | umfs01.grid.umich.edu | 35.8 | 32.19 | 100% |
| umfs02 | Up | umfs02.grid.umich.edu | 35.74 | 37.38 | 99.93% |
| UMOPT1 | Up | 127.0.0.1 | 0.11 | 0.1 | 100% |
| UMROCKS | Up | umrocks.grid.umich.edu | 37.26 | 36.28 | 99.76% |

<< Previous — Showing Rows 1 to 30 of 31 [1,2] — Next >>

# AGL-Tier2 Inventory Software



OCS inventory — OCS next generation inventory — Ver. 4020

Logout Change pass

**All computers**

**112 Result(s)** (Download) Show: 15 Add column Reset

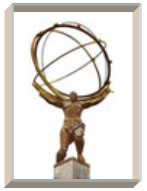
1 .... 8

| Tag | Last inventory | Computer | Operating system | Ram(MO) | CPU(MHz) |
|---|---|---|---|---|---|
| | 08/16/2006 13:12:04 | c-15-4 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 7984 | 1804 |
| | 08/16/2006 13:12:03 | c-15-2 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 7984 | 1804 |
| | 08/16/2006 13:12:03 | c-15-3 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 7984 | 1804 |
| | 08/16/2006 12:57:12 | ml-um | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 497 | 998 |
| | 08/16/2006 11:32:13 | umopt1 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 7984 | 1002 |
| | 08/16/2006 10:40:03 | c-10-116 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 2025 | 2000 |
| | 08/16/2006 10:39:05 | mars05 | Linux(Redhat / Scientific Linux CERN SLC release 3.3 (Beryllium)) | 3950 | 3600 |
| | 08/16/2006 10:38:04 | mars04 | Linux(Redhat / Scientific Linux CERN SLC release 3.3 (Beryllium)) | 3950 | 3600 |
| | 08/16/2006 10:36:04 | mars03 | Linux(Redhat / Scientific Linux CERN SLC release 3.3 (Beryllium)) | 2007 | 3600 |
| | 08/16/2006 10:21:12 | mars01 | Linux(Redhat / Scientific Linux CERN SLC release 3.3 (Beryllium)) | 3950 | 3600 |
| | 08/16/2006 10:21:09 | mars02 | Linux(Redhat / Scientific Linux CERN SLC release 3.3 (Beryllium)) | 3950 | 2800 |
| | 08/16/2006 10:13:03 | tera02 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 2026 | 1667 |
| | 08/16/2006 10:09:05 | atums2 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 3957 | 1994 |
| | 08/16/2006 10:05:20 | c-9-88 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 2025 | 1401 |
| | 08/16/2006 10:05:19 | c-12-49 | Linux(Redhat / Scientific Linux CERN SLC release 4.3 (Beryllium)) | 2025 | 1667 |

**Mass processing:** Frequency Deploy Delete

1 .... 8

# Existing Servers/Services

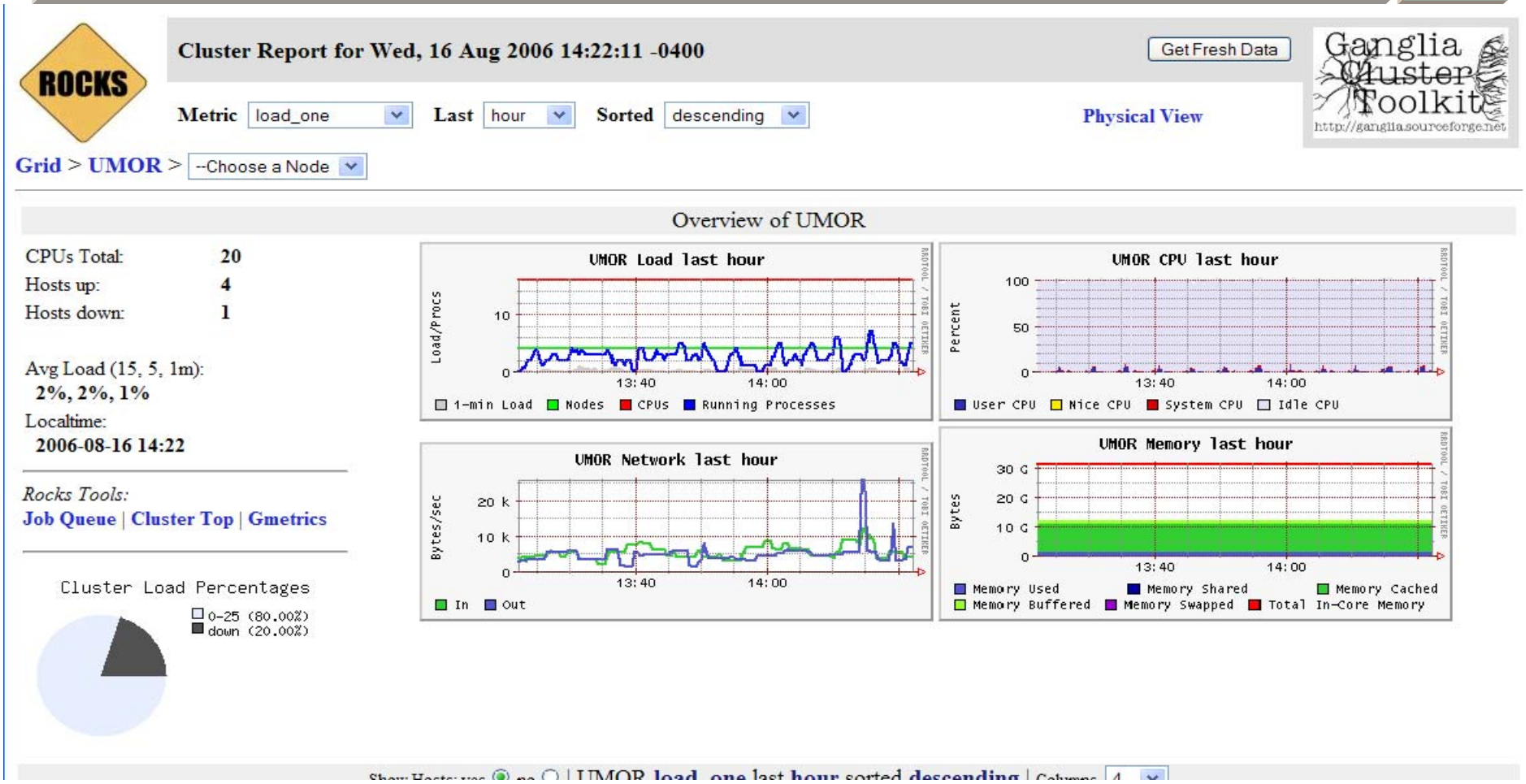


- In addition to the UMROCKS cluster we have a number of servers/services operational
- Two gatekeepers: Dual Xeon 3.6 (2MB cache), 4GB RAM, Intel SE7520AF2 motherboards, IMM card (IPMI) called gate01/gate02.grid.umich.edu
- AFS Cell atlas.umich.edu hosted on linat02/linat03/linat04.grid.umich.edu with file servers attera/linat06/linat07/linat08/atums1/atums2 (about 6TB)
- NFS data servers umfs01/umfs02/linat09/linat10/linat11 hosting about 22TB total
- Hypnos.grid.umich.edu is dCache headnode for UMROCKS
- Have MonALISA node at ml-um.ultralight.org and other monitoring services
- Oracle server on one of the “prototype” systems for Calibration/Alignment DB replication
- Planned servers: NDT node, GridFTP, DQ2

# OSG & ATLAS Software Status/Plans

- As shown we have both AFS and NFS storage at our Tier-2
- We plan to install software on AFS (good for readonly type data). OSG (0.4.1) and ATLAS software already in AFS (/afs/atlas.umich.edu)
- ATLAS software is mirrored via Pacman on our AFS cell at: http://gate01.grid.umich.edu/am-UM/ATLAS.mirror
- All users have their home space in AFS. Our system is setup to get Kereberos TGT (and AFS Tokens) at login via gssklog (instructions on TWiki)
- All OSG accounts created with “uniqname” IDs

# Prototype Opteron Cluster




- Testbed for dual dual-core systems (Opteron 280s, 4GB ram)

# Prototype Building Block Details

- We have purchased 5 dual dual-core Opteron 280 systems and an NFS storage server to test with.
- Worker nodes are using Supermicro H8DAR-T(1U) motherboards (AMD 8132 chipset), 4GB of RAM, dual dual-core Opteron 280, three 250GB SATA-II hot-swappable drives, CDROM (4 cores/1U)
- Disk server is a dual dual-core Opteron 280, 5U, 24 SATA-II (500GB) drives, dual 250GB system disks, 8GB of RAM, dual 1GE NICs, Areca 1170 RAID6 controller (11TB)
- Need to test I/O scaling using NFS with ~20 clients / storage server
- Possible to use 10GE NIC on storage server if network impacts the performance.

# Michigan State Space Planning

## MSU cluster space/upgrades


Bid in-hand from Clark Trombley Randers
(MSU-experienced, local engineering firm)
For 20ton precision, water-cooled, overhead, ducted system
Anticipate designing for ~12 10KW racks (4-5 anticipated for PA Department needs)
Read for upgrade by +10tons when required.

Work will go ahead, asap.
$250k in hand from internal, MSU sources
Since approval, accounts have been created and contact with the engineering firm has been re-established. MSU physical plant arranging first meeting in early September.
Brock will manage.

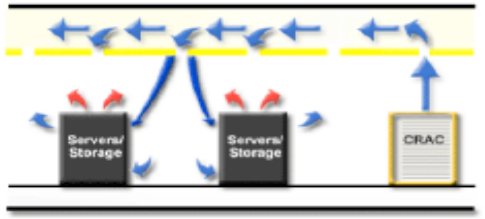



Figure 2. Overhead distribution system.

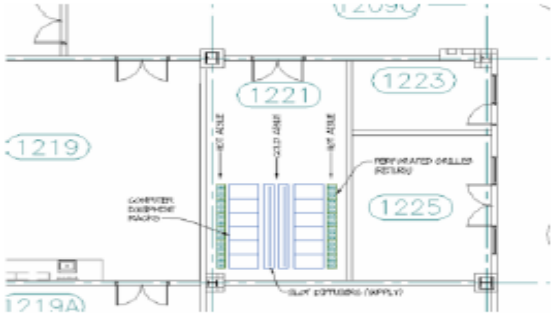
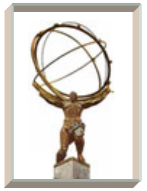



Figure 7. Computer Room 1221 with 12 racks and cooling distribution.

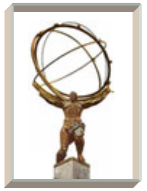
# Proposal Numbers for AGL-Tier2



| UM-MSU | Dedicated T2 Funded | | | Dedicated Other Funded | | | Leveraged | | Total | | Network |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Tier 2 (units) | CPU (SI2 k) | Disk (TB) | FTE | CPU (SI2 k) | Disk (TB) | FTE | CPU (SI2-yr k) | FTE | CPU (SI2 k) | Disk (TB) | WAN (Gbps) |
| **FY06** | 137 | 70 | 2.60 | 372 | 86 | 2.40 | 102 | 3.95 | 611 | 156 | 10+10 |
| **FY07** | 331 | 169 | 2.60 | 506 | 135 | 2.40 | 171 | 4.55 | 1008 | 304 | 10+10 |
| **FY08(HL)** | **605** | **309** | **2.60** | **645** | **162** | **2.40** | **261** | **5.35** | **1512** | **471** | **10+10** |
| **FY09** | 993 | 492 | 2.60 | 955 | 278 | 2.40 | 261 | 5.50 | 2210 | 770 | 10+10 |
| **FY10** | 1404 | 696 | 2.60 | 1339 | 359 | 2.40 | 206 | 5.15 | 2950 | 1054 | 10+10 |

- The table above is from our proposal. The FY06 numbers assumed full funding in FY06.
- We need to discuss the US ATLAS needs and timescales to deploy resources for the AGL-Tier2.
- Consideration is the availability of the long-term computer space at both MSU and UM
- We have 59 CPU-years of Opteron time from our Center for Advanced Computing which we will dedicate to the AGL-Tier2 (gate02/torque)

# Planning for AGL-Tier2 Profile



- MSU and UM are both working on high-quality server spaces.
- Michigan will share the Internet2/Merit space (MITC)
  - 4MW of power for 2MW of cooling and 2MW of equipment (Flywheels and generators)
  - ~280 racks of space
  - Lots of fiber access
- The MITC space is scheduled to be ready March 2007
- We have “interim” space in our colleges server room IF we need to use it (up to 6 racks worth), but this would require two moves.
- MSU space will be ready in the same timescale (spring 2007)

# Summary

- The AGL-Tier2 is very close to being able to provide cycles for US ATLAS.
- The AGL-Tier2 should truly be a single “site” from US ATLAS’s point of view, even though our equipment and services are distributed between two campuses 60 miles apart. This is because of MiLR (10GE) and a common network address block.
- We have some useful services running to help monitor and manage our equipment. If others are interested we are happy to share…
- The amount of cycles and storage required should be discussed as well as the time profile so we can optimize our plans.