AGL-T2: ATLAS Great Lakes Tier 2 Computing Center

University of Michigan

and

Michigan State University

University of Michigan: Shawn McKee (PI), Dante Amidei, Robert Ball (Tier-2 Manager), J.W. Chapman, TieSheng Dai, Edward Diehl, Thomas Finholt, Steven Goldfarb, Daniel Levin, Homer Neal, Jianming Qian, Rudolf Thun, Zhengguo Zhao, Bing Zhou

Michigan State University: Raymond Brock (co-PI), Maris Abolins, Carl Bromberg, Reiner Hauser, Joey Huston, James Linnemann, Bernard Pope, Reinhard Schwienhorst, Kirsten Tollefson

Shawn McKee 83467

Digitally signed by Shawn McKee 83467
DN: cn=Shawn McKee 83467, ou=People, dc=org, doegrids
Reason: I am the author of this document
Date: 2006.05.15 14:27:29 -04'00'

Introduction

The University of Michigan and Michigan State University hereby submit a proposal for a powerful Tier-2 Computing Center in support of the ATLAS Experiment. Our goal: to reach the high luminosity processing and storage milestones by the second fiscal year of the project. With combined commitments from both institutions totaling over $4M, we will reach that goal and will eventually provide resources to U.S. ATLAS of nearly 3M SI2k in processing power and 1 PB of disk storage.

The **ATLAS Great Lakes T2 (AGL-T2)** is a collaboration of physicists and computer professionals from the University of Michigan (UM) and Michigan State University (MSU) who share common interests in LHC physics and a motivation to enable the success of the ATLAS experimental program. The physical proximity of our two institutions is enhanced by the unique statewide infrastructure of high-speed connectivity allowing us to leverage separate university infrastructures into a tightly integrated system which will present itself to the user community as a single, logical facility.

Collectively, in the proposed Center, UM and MSU would comprise the largest confederated entity in ATLAS in terms of the number of active faculty. We offer a very significant leveraging of U.S. ATLAS computing resources as a result of the strong contributions made to the project by our two institutions. Our researchers have already made enormous contributions to the detector design and development through our work on the muon detector, hadron calorimeter, and trigger systems. Our group contains the U.S. ATLAS Networking Coordinator. We have demonstrated our ability to contribute to ATLAS computing by out-performing even proto-type Tier 2 centers during Data Challenge-2. UM was responsible for providing more simulated events for the Rome Physics Workshop than any other single institute in ATLAS.

The High Energy Physics faculties at Michigan State University and the University of Michigan have a long history of fruitful physics and technical collaboration. For example, both groups are collaborators in the DØ and CDF experiments. The two campuses are located only 60 miles apart, and intracampus visits have been frequent. Knowledgeable computing staff exists on both campuses, providing a level of expertise backup that would be beneficial for the smooth operation of a Tier-2 Computing Center.

In our proposal we have concentrated on the level of service we can provide to the ATLAS community. We realize that service is the primary function of a Tier-2 center and that a key evaluative factor is how much computing U.S. ATLAS can obtain for the budget it has available. Our proposal offers a higher computing capacity to the US ATLAS collaboration than most of the existing Tier-2 Computing Centers. We also note that in the long run the effectiveness of Tier-2 centers will depend on the expertise and commitment of the researchers at the host sites. With our detailed expertise of the ATLAS trigger and the muon systems– including the Muon Calibration Center responsibilities–with our cutting edge research in UltraLight, and our leadership role in collaborative tools, we can not imagine a more suitable operating environment into which a Tier-2 Computing Center could be located.

Below we describe in detail a system designed to reliably serve the analysis and simulation requirements of physicists within the U.S. ATLAS collaboration.

Roles of the AGL-T2 Center

U.S. ATLAS Tier 2 Computing Centers are to support the simulation and analysis activities of U.S. ATLAS as well as specific detector calibration efforts. This obligates us to a series of necessary functions and capabilities.

Physics Analysis and Simulation Support

The prospects for a national analysis effort are daunting and the Tier-2 centers should play a role in enabling efficient access to data and expertise. Both of our institutions intend to be actively involved in analysis and will maintain Tier-3 facilities. So, the physicists involved in the Tier-2 center will be users of AGL-T2 and therefore "in the know," and able to be helpful to the virtual communities which will rely on us.

Simulation will be a focus of AGL-T2 and, as will be demonstrated below, significant resources are planned in order to make this a major center for Monte Carlo generation of important processes. The University of Michigan has been deeply involved in data challenges and highly successful in simulating large amounts of ATLAS physics events.

Data Access

The primary job of the Tier-2 centers is to provide seamless access to ATLAS data and simulation samples of interest to the community. With our planned storage capacities and the fully-protected, fast networks afforded by the Michigan LambdaRail (MiLR) and our 10 Gbits/sec campus networks we intend that AGL-T2 users will have efficient reliable access to ATLAS data. Our ability to present a single, logical identity to the community should lead to a seamless data access capability for users.

Unique Responsibilities to the Muon Community

The UM group has significant responsibilities with the production and performance evaluation of the US endcap muon chambers. Through special agreement with the muon community, any Tier-2 center approval will include a separate focus on this important U.S. ATLAS responsibility.

The University of Michigan has committed, along with sites in Rome and Munich, to serve as a calibration center for the ATLAS muon system and signed an MOU with the US ATLAS Research Program Office to formalize this role. We will require computing resources to fulfill this obligation. Of course our center's activities will be determined by the US ATLAS Resource Allocation Committee (RAC) which oversees all US ATLAS resource allocations.

Personnel and Expertise

The personnel of AGL-T2 are experimental high energy physicists and computer professionals with experience in grid computing, the design and support of large-scale hardware for high energy physics experiments, and track records in large project management.

Personnel

The personnel on this proposal include all of the ATLAS physicists from both institutions. Details of their commitments to AGL-T2 are in the Appendix. Management of the Center includes the PI and co-director, co-PI and co-director, and the Tier-2 Manager.

SHAWN MCKEE, PI. McKee has a strong record of leadership in the development of grid, computational and network resources for physics experiments, and is the Network Project Manager for U.S. ATLAS and the Co-Chair of the Internet2 working group on high energy and nuclear physics. He is also Co-Chair of the Open Science Grid Networking Technical Group and helped to found the MGRID initiative at the University of Michigan, continuing to serve on its technical leadership team. In addition he has led Michigan's involvement in the NMI testbed (2002-2005) and continues to participate in SURAgrid, a regional grid initiative. McKee also has a leadership role beyond ATLAS in the LHC-Optical Private Network group where he heads the monitoring effort. Of special importance is his management of the existing Tier-3 cluster at the University of Michigan which has been extremely successful in the ATLAS data challenges and Rome Physics meeting data production (described below).

RAYMOND BROCK, CO-PI. Brock has been a member of the DØ collaboration since its inception. In addition to various responsibilities in physics analysis and detector construction, he was the co-head of the DØ Database, Data-handling, and Datagrid group. During that time, the SAMGrid concept matured and was converted to a Computing Division project. Subsequently, MSU acquired significant computing capability and has recently joined the SAMGrid network as a processing site. Brock has extensive management experience for technical projects within DØ and served as chairperson of the Department of Physics and Astronomy for seven years.

ROBERT BALL, TIER-2 MANAGER. Ball has been involved with physics and computing for over twenty years, including establishment of a parallel computing farm used by the CERN L3 experiment. More recently, he managed UNIX clusters and high-performance networking for the UM physics department and currently manages our UMROCKS cluster.

AGL-T2 has significant local physics, simulation, computing and management expertise. We summarize our faculty and senior personnel Tier-2 related commitments in Table 3 in the Appendix. Shown there are the expected commitments and roles for the five years of the proposal.

Institutional Expertise and Experience

The University of Michigan has been a very active participant in U.S. ATLAS grid and simulation efforts. The first U.S. ATLAS grid meeting was hosted in Ann Arbor in 2001. During the last two years we have successfully utilized a large fraction of the UMROCKS cluster to run ATLAS event generation and simulation jobs for DC2. The UMROCKs cluster allowed us to sustain 172 jobs at once, exceeding the running job capacity of any other single site in Grid3.

Even though we were not a Tier-2 prototype center and entered the process later than most sites we were successful. We were able to run large numbers of jobs simultaneously and were frequently one of the top three sites in terms of successfully completed jobs, even though our processors were not as powerful as many of the other sites.

Our contributions during DC2 were not limited to CPU cycles and storage. We produced detailed job monitoring scripts which interfaced job schedulers like Condor and PBS with the Capone job submission system. The resulting web pages allowed job submitters and site managers access to the relevant information from one location. We were also successful in locating problems in the job submission and management systems as well as in versions of the event generation and simulation code, helping to resolve problems occurring at other sites as well.

UM has extensive experience in operating large computer clusters for physics purposes and we list the following highlights. In the last two years:

- We successfully participated in Data Challenge 2 during its 2004-2005 duration generating 6% of the total US contribution by leveraging local University resources, including the engineering cluster called "Hypnos" (with 200 CPUs) and local MGRID computing and network expertise.
- UM has extensive experience with computer clusters, networking and grids. For example we have successfully deployed and maintained dCache (since spring 2005), the Grid3/OSG software stack (since their inception) and ATLAS software to support our analysis and simulation work and provide resources to the U.S. ATLAS grid testbed.
- ROCKS, Grid3, and ATLAS software were customized to provision our nodes for Rome physics data production. In about 10 weeks time period, we were able to produce over **3 million** full simulation MC events for the Rome physics meeting. Those events intensively used for ATLAS Muon Spectrometer performance studies; for diboson physics studies and for Λb polarization studies. Our presentations of the studies were well received in the Rome Workshop.
- We continued MC event production after the Rome workshop with the goal of producing sufficient background events from Standard Model processes to understand the ATLAS detector sensitivity for new physics discoveries with early LHC data. In the past several months, we have enlarged our MC data sample size by a factor of 10. This has allowed us to understand the fake event rate, and develop analysis methods to increase the background rejection power. As an example, for diboson physics, we have improved signal-to-background ratio by at least a factor of three.
- The UM DØ group has played a leading role in the DØ experiment. Qian served in leadership positions for many years. Among them, he was a co-leader of DØ software and computing project between 2002-2003 and the Physics Coordinator between 2003 - 2005. Much of the current DØ Run II physics analysis infrastructure and procedure was established under his leadership.
- The UM CDF group has extensive and current experience with large scale data handling. Dan Amidei was Deputy Manager of the CDF Run II Computing Upgrade 1998-2000. UM staffs and manages the high level dataset production and validation for the whole CDF top group. Amidei chaired a recent task force on CDF Disk Usage, and a UM-Fermilab collaboration developed and commissioned (Dec. 2005) a 50 TB dCache distributed disk pool system now used by the entire collaboration.
- UM constructed 80 of the large muon chambers for the ATLAS MDT endcap, has provided the readout multiplexer for all MDT chambers, has led the certification team for the front-end electronics, has assembled and supplied all the data acquisition stations for chamber certification and commissioning, was the leading institution in the certification of all 240 US supplied MDT chambers at CERN, and has and continues to lead the muon endcap commissioning effort at CERN.

The Michigan State University group has been involved in technical projects on DØ and ATLAS for many years,

- MSU has played a leadership role in the design, construction, analysis, and management of the DØ and CDF experiments.

- MSU (Brock) has just recently begun large scale simulation with a new 37 node system of mixed-capability Opteron processors within the DØ grid environment.
- We (Abolins, Brock, Linnemann, Schwienhorst, Tollefson) have made major contributions to the DØ and CDF experiments' Trigger systems. This includes design, construction, commissioning, simulation, and upgrading of the DØ L1 and L2 trigger electronics for Runs I and II as well as design and implementation of the L3 system in CDF. In addition, Tollefson has been the head of the CDF Trigger Dataset Working Group and has extensive experience in trigger database design and implementation.
- The CluED0 cluster at Fermilab was designed and built by MSU postdocs (including Hauser) for local desktop machines and grew to become a significant resource (consisting of 365 machines from 40 institutions including Fermilab) serving the entire collaboration.
- We (Abolins, Hauser) have participated in the architectural design and testing of the High Level Trigger (HLT) system of ATLAS. In particular we are active in the Data Collection and Dataflow groups, working on the development of common libraries and various error handling and fault tolerance programs.
- MSU ATLAS hardware activity has involved the design, construction and testing of the trigger Supervisor (L2SV) and the Region of Interest Builder (RoIB) (which organizes and transmits region of interest information from the Level-1 trigger to the Level-2 system) as well as participation in the R&D and design for the ATLAS hadron calorimeter (Tilecal). MSU led one of the two instrumentation sites in the US, and during 1999-2003, we instrumented and tested 32 10-ton Tilecal modules.
- At MSU we designed and built ATLAS cryostat scintillators. We also designed and built scintillation counters (MBTS) to be used for triggering in the commissioning run. MSU is partially responsible for developing the trigger using these counters.

Physical Infrastructure

The physical infrastructure of the AGL-T2 center is significant in both the available networking commitments and the resources devoted to this proposal from both institutions.

Networking Resources

UM and MSU are well positioned to provide excellent network connectivity for our proposed Tier-2 Computing Center, both in the Local and Wide-Area network. Equally important will be our ability to share a common subnet so that external access to our distributed Tier-2 center is seamless. This is made possible by the MiLR (Michigan Lambda Rail) infrastructure which was developed and is owned by a State of Michigan university consortium[1].

The MiLR infrastructure is shown schematically in Figure 1. It consists of a diverse fiber ring from Chicago through UM to Detroit (4 fiber pairs) with 3 pairs continuing on through MSU to Grand Rapids and back to Chicago. Currently both UM and MSU have three 10-Gbits/sec wavelengths "lit" on MiLR. For our Tier-2 center we will have access to three University provided 10-Gbits/sec wavelengths: UM to StarLight, MSU to StarLight and MSU to UM. They

[1] University of Michigan, Michigan State University, and Wayne State University.

are within a single loop in the State of Michigan and may be shared with other campus activities. AGL-T2 will be the primary user of these links.

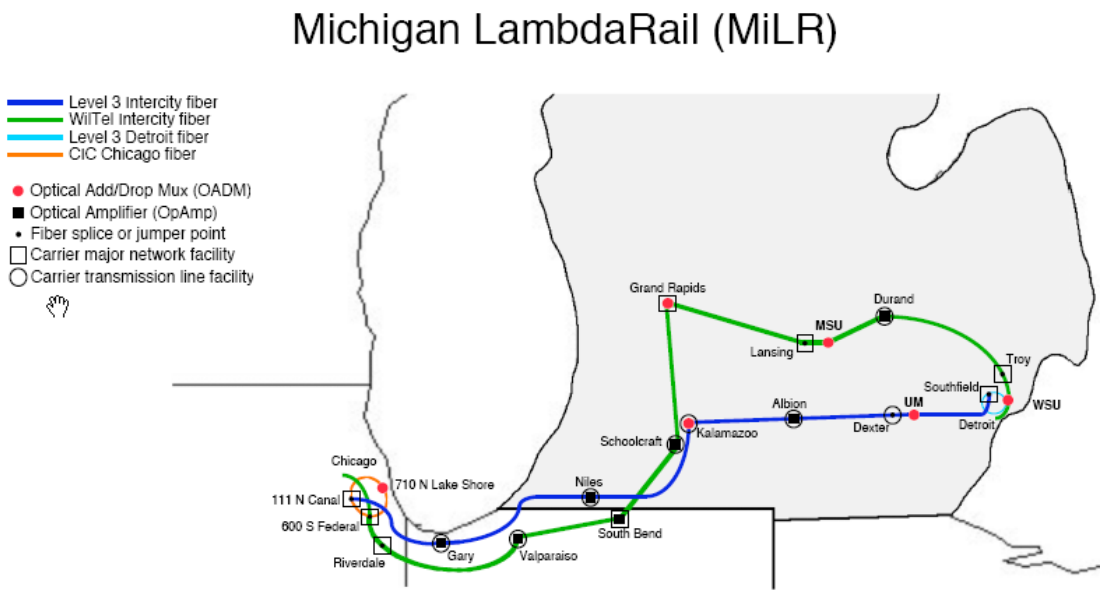

Figure 1: The MiLR (Michigan Lambda Rail) fiber map.

This loop construction is very powerful and results in a "protected" 10 Gbits/sec link for the AGL-T2 Center. If any link in the loop fails both sites remain connected, both to each other and the outside world, at 10 Gbits/sec. We plan to initially provision a /23 subnet (~510 host addresses) for the Tier-2 center using addresses from Merit Network, Inc.[2] Thus from the outside we are a single site/subnet. Internally we will each have a /24 subnet (~254 hosts) and use standard network peering to route, providing fail-over and load-balancing for our center.

As discussed in the Related Tier-2 Research Efforts section in the Appendix, UM is one of the main ATLAS collaborators on UltraLight and has an experimental "lambda" (10 Gbits/sec) connection to StarLight/NLR (National Light Rail) over MiLR. This link could be used to provide additional failover capability if required. Because of this work we already have extensive experience with 10 Gbits/sec Ethernet networks.

We plan to utilize the Michigan UltraLight network chassis which has a total of 12 instrumented 10 Gbits/sec interfaces to connect the AGL-T2 at UM to MiLR. Our plan is to share this chassis, using VLANs to isolate production and research activities. Since this chassis lists for over $250K this is a significant contribution to the center. We expect the UltraLight connection to migrate toward "production" use starting in the second year of the center. Figure 2 illustrates the basic 10 Gbits/sec architecture for UM, while Figure 3 shows the layout for MSU.

Michigan State will contribute a dedicated on-campus fiber loop and install both Layer 2 and Layer 3 switches to connect the MSU portion of AGL-T2 to MiLR through the existing MiLR RAYexpress OADM chassis. The existing fiber loop on campus will be connected to the Biological and Physical Sciences building (BPS) which houses the Physics and Astronomy

[2] Merit is a non-profit regional research and education network governed by the public universities in the State of Michigan.

Department server room and computer personnel. Fiber will be pulled within BPS to complete the circuit to the servers.

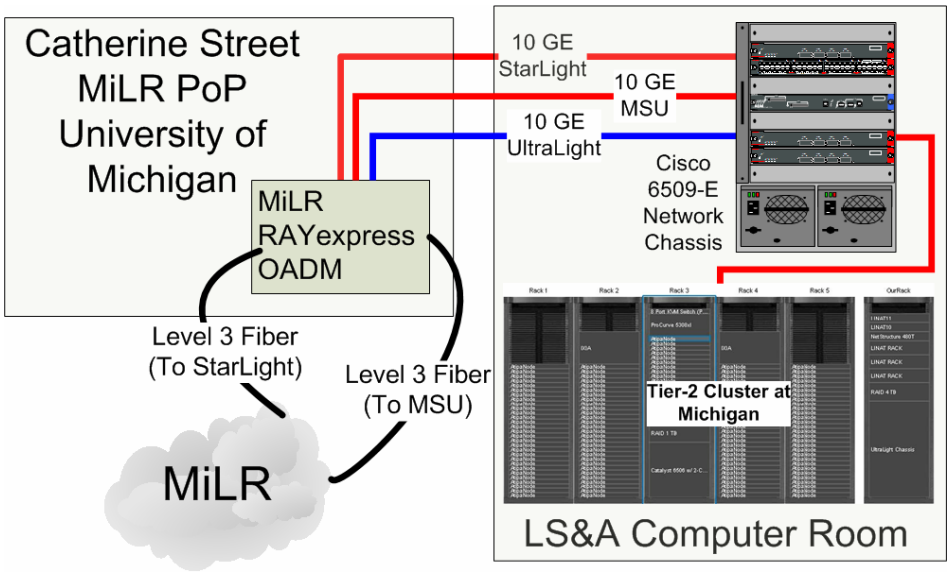

Figure 2: University of Michigan campus connection to MiLR.

Because of MiLR we have an exceptional ability to inexpensively access multiple lambdas to support the needs of the center. In the future we could easily add additional 10 Gbits/sec wavelengths, as required.

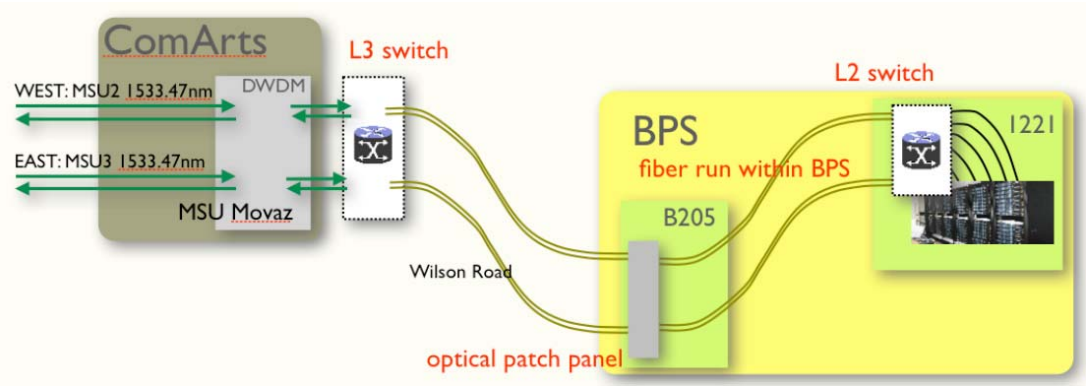

Figure 3: Michigan State University 10 Gbits/sec connectivity between MiLR and Tier-2 site. Shown are planned new Layer 2 and 3 switches.

Computing Resources

We have designed a five year deployment plan for a robust, manageable Tier-2 computing center, distributed between MSU and the UM. This distribution is feasible because of the unique capabilities described in the previous section. Both institutions will provide high-quality computing space (described under "Operational Plans" below) to house the Tier-2 equipment.

For our hardware planning, we have targeted all equipment to have a nominal 4 year lifetime. Funds spent in the 5th year will be replacing equipment acquired in the 1st year, and so on. ***For planning purposes we have used specific equipment and quotes, but our actual purchases will be done in consultation with the Tier-1 and other Tier-2 sites***. We will continue to run older equipment as long as it remains viable and we have the space, power and cooling to support it.

For the following, in order to estimate computing and storage capacity in future years we assume a doubling time of 2 years for storage and 2 years for processing power. That is, for a fixed amount of dollars (accounting for inflation) the amount of storage or computing power will double in the assumed time frame. Our calculations include 3% inflation per year for all costs.

Finally, we intend that our future equipment will be identical in all respects between the two institutions and expect to coordinate procurement and bidding between the two sister State of Michigan institutions. This should leverage quantity discounts and also allow for mutual monitoring and even sharing of hardware maintenance responsibilities, adding to our flexibility and responsiveness.

Dedicated, Tier-2 Funded

Available funding, detailed in Table 2 in the Budget section, allows the acquisition of CPU and storage capacity well beyond that required of a Tier-2 Computing Center. The Infrastructure section of this proposal details our assumptions regarding the evolution of CPU speed and storage unit size with time. In both cases, purchases will be made late in the budget year and in consultation with the Tier-1 and other Tier-2 centers to take advantage of hardware improvements to the extent possible.

Anticipated purchase coordination between UM and MSU will produce a consistent AGL-T2 infrastructure. To aid us in defining the best hardware we are prototyping (with Tier-3 funds) the systems described below and will take delivery of 4 compute nodes and 1 storage node in mid-May 2006. This prototype will be extensively tested by using ATLAS software, running multiple simulation and analysis jobs and monitoring I/O capabilities and performance.

Our compute node consists of a dual-CPU, dual-core Opteron 285 (2.6GHz) 1U rack-mount chassis. Each system has dual Gbits/sec NICs and is packaged with 8GB of RAM (2GB/core following the ATLAS requirement), 3 250GB SATA2 hard drives (for use with dCache), and an IPMI 2.0 card for remote power control. The Opteron architecture was chosen following discussions with LHC collaborators (see talks at HEPIX meeting in Rome, April 2-5, 2006) and others for multiple reasons: scalable use of multiple processor cores for program execution, reduced power (and heat output) when compared to Intel processors, 64-bit capability and compatibility with recent BNL purchases.

Our disk storage unit utilizes the same processor and RAM configuration. It comes with 24 500GB SATA2 hot-swappable hard drives, and an Areca ARC-1170, 24-port SATA2 RAID controller (perhaps we will use ARC-1280 when available). Configured with hardware RAID6 we retain 11TB of robust storage. We have the possibility to add a 10 Gbits/sec network interfaces if this is useful for overall performance of the storage nodes.

We note that the storage capacity shown for AGL-T2 includes storage from the above described storage nodes and dCache space (from the compute nodes). The fraction of “dCache” storage is initially 45% of the total space, decreasing to ~20% by the 5th year.

Dedicated, Non-Tier-2 Funded

Because of generous support from our institutions we will be able to acquire over $800K of new compute and storage equipment dedicated to the AGL Tier-2 (see budget section). This allows us to provide more computing capacity to the Tier-2 in the first three years

We are dedicating our existing UMROCKS cluster (see Figure 4) to the AGL-Tier-2. This facility was successfully utilized during Data Challenge 2 and UM's Rome data production. This system provides 130,000 SpecInt 2000 (SI2K) and 16 TB of local disk. Each node is composed of a dual Athlon 2000/2400/2600 MP+, 2 GB of RAM, two 100 GB IDE disks and a dual Fast Ethernet NICs. We plan to maintain and use UMROCKS as long as it is viable.

Figure 4: Views of the UMROCKS cluster showing racks, typical rack contents and water cooling

Our relationship with the UM Center for Advance Computing (CAC) has also resulted in the dedication of 59 CPU years of the Nyx Opteron cluster for U.S. ATLAS Tier-2 work. This is an aggregate of 66,000 SI2K.

Non-dedicated

Each institution will have computing Tier-3 capabilities which will serve the physics analysis needs of our local and CERN-based physicists. Fractions of these resources will be available for non-dedicated Tier-2 usage.

Funding from UM resources has allowed us to expand our current computational capability through the purchase of dual Opteron 280 systems. The first "prototype" as described below will shortly be delivered to UM, providing 25,000 SI2K of processing in addition to an 11TB (reliable space) disk server. Upon successful validation of this system we plan to expand it this summer to 110,000 SI2K (going to Opteron 285s) of computing power and 44TB of space. We have funding to add to this system each of the following two years and we expect to provide 50% of the CPU cycles as a "leveraged" contribution to AGL-T2.

Significant other resources will be available on the UM campus for the duration of our proposal. We are already working closely via our MGRID initiative with many other groups who have significant computing resources. We will continue to work closely with many other colleges and disciplines as part of our Tier-2 outreach effort and because of this close association we anticipate having some low priority access to unused cycles from the Center for Information Technology Integration, and the School of Information.

MSU currently maintains 37 nodes of mixed class computing. These are recently being used for DØ Monte Carlo production, but half will be available for Tier-3 computing as needed by new personnel already hired come aboard. This equipment consists of 20 Opteron 246, dual single

core processors and 17 Opteron 270, dual, dual core processors for a total of approximately 90,000 SI2K capability. In addition, 2 more racks worth of Opteron 270-class processors are funded and out for bidding. Of this total, 50% will be for Tier-2 and 50% for Tier-3 use. MSU's existing nodes are housed in an APC NetShelter enclosure, with APC power distribution units (PDU), and an APC UPS. The quality rack securely mounts the nodes, eases maintenance and provides for free air-flow. The PDUs are integrated with the cluster control systems providing monitoring of the electrical current drawn (essential to avoid circuit breaker trips) and remote control of the individual outlets. The UPS provides battery backup for the cluster head-node and network switches.

Total AGL Tier-2 Resources

The results of summing the components described above are summarized in Table 1. We note that because of strong support from our Universities the housing of the Tier-2 equipment will be fully covered and will require no funds from the Tier-2. This support includes all power, cooling and campus networking required for the Tier-2.

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

Table 1: The AGL Tier-2 breakdown of resources from all sources by category and year.

We point out that the equipment contributions alone from UM and MSU amounts to more than $1M and that this makes possible a ***total computing capability of over 1MSI2k and 300TB of storage by only the second year of the center***. We believe that this represents a significant demonstration of commitment by our institutions.

Infrastructure and Operational Plans

Our plans for operating and organizing an AGL-T2 are designed for close coupling of the personnel and resources on the two campuses

Infrastructure

Equipment for the ATLAS Great Lakes Tier-2 Computing Center will be housed in localized sites, one at the University of Michigan and one at Michigan State University.

The UM space will be in the College of LS&A's server room, designed specifically to handle large scale computing clusters. All rack power is provided by UPS units, and is battery backup protected so that in the event of a major power failure a 5 minute grace period will allow generators to come on line that can power the entire room for a few hours, or half of the capacity-filled room indefinitely. Rack capacity is estimated to be ~8kW which will easily allow racks to be filled to capacity. Each rack is cooled by chilled forced air from beneath the raised floor. Temperature and humidity sensors built into each rack supplement the external sensors that monitor the ambient room conditions.

The MSU site will be housed in the Physics Department server room in the Biomedical & Physical Sciences Building. Completed in 2001, the BPS building provides a convenient and secure location for the MSU Tier-2 site. The server room will be upgraded to 20 tons of specialized computer room air conditioning. The upgraded room will also feature a 400 Amp electrical distribution panel, key card security access, and smoke detection system. It is located adjacent to the Physics and Astronomy Department electrical engineer and electronics shop, as well as the department computer technicians. Room temperature, humidity and A/C performance will be monitored in real-time. Tier-2 storage systems, head nodes and network equipment will be powered through UPS units. The BPS building is serviced by 2 750 KW backup generators. Generator power will be allocated to support the Tier-2 head nodes and network equipment in the event of a power failure lasting longer than the UPS units' run time. This room also houses computer systems for the Physics Department and other College of Natural Science users but the AGL-T2 will be the major user and drives the renovations.

The network backbone for node to node communications within each center will run at 1 Gbits/sec. Taking advantage of MiLR (see Networks and Physical Infrastructure) we will have a 10 Gbits/sec interconnect in place between the MSU and UM clusters. This fast interconnect will allow us to run the entire distributed center as if it were housed under one roof. Disk servers could interface at 10Gbits/sec as well.

The environmental sensors in the racks and the facility allow for quick notification in cases where manual intervention could be required. Further, because identical equipment and IPMI management controllers will be deployed to all nodes, both groups can monitor operations at the whole center. The proximity of the two universities allows personnel from either to assist the other in cases where such help is needed for an intervention, providing us with coupling as tight as if we were housed together.

Keeping touch with the sensors is just one aspect of monitoring the status and health of the AGL-T2 Center. It is our intention to provide web-based information for users of the center, providing pages of information about the center and directions for use, along with links to other centers and software information. Links for email questions about issues concerning the center itself will be in place to assist in this process. It is our goal that even a novice user can quickly and easily learn to use the AGL-T2 center.

Operations

AGL-T2 will operate fully attended from 9am-6pm Eastern Time, Monday through Friday, excluding holidays. Personnel will be accessible during this time both by desktop and cell phone. "Fast access" email will also receive attention as appropriate, with simple requests receiving immediate attention, and more difficult tasks assigned to the appropriate manager. It is our intention to give the highest quality assistance possible to such requests so that on-going work by the facility users will be minimally impacted.

During off hours, calls for help will be queued until the next fully attended time period. However, a staff person will be assigned to weekend check-in service, looking to handle problems which do not require the full attention of the weekday hours. Automatic monitoring of both hardware and software conditions will also take place during this time, using systems like Nagios, MonaLisa and Ganglia, with a call-in via phone page for any serious problem.

Major upgrades of either hardware or software will take place during off hours. Every attempt will be made to stage such occurrences, and to maximize notice, such that the center will continue to operate whenever possible, albeit at a reduced capacity. These outages will be scheduled in consultation with the system users, avoiding conflicts with major analysis work deadlines to the extent possible.

Tier-2 Organization and Staffing

AGL-T2 emphasizes the central importance of Tier-2 services and available technical personnel resources will be deployed to efficiently serve that role (see Figure 5 below). They will fall into three broad categories: local cluster maintenance at each campus and overall Center operations. Actual makeup of the latter will be somewhat dependent on the ultimate U.S. ATLAS computing model and will be determined as this becomes clearer. In line with this, specific personnel in the third category will have full-Center responsibilities regardless of their physical home. We outline our known and potential personnel allocation in this section.

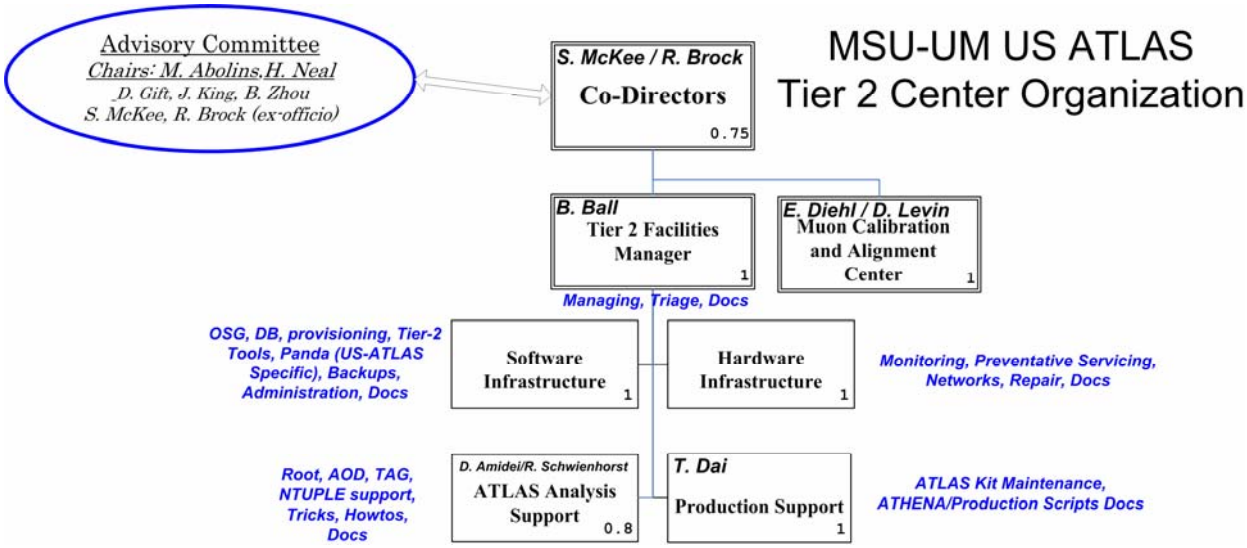

Figure 5: Organization of the AGL Tier-2

Local Operations, UM. The cluster operation at the Ann Arbor campus consists of student personnel for day-to-day hardware maintenance and a 0.4 FTE professional systems manager and 0.5 FTE for the Co-Director. These are UM dedicated positions.

Local Operations, MSU. The cluster operation at the East Lansing campus consists of student personnel for day-to-day hardware, backup maintenance, 0.5 FTE for hardware and backup support, and 0.5 FTE of professional systems management. These are MSU dedicated positions.

Overall Center Management and Operations. The Center Manager (Ball) will have responsibility for day-to-day operations, ensuring maximum availability and smooth functioning of the combined center. He will coordinate equipment upgrades and replacement, interface to the Universities on issues of facility support, establish policy in consultation with the Directors, and supervise a staff of support personnel, including students, and help in the day-to-day operation of the center. We expect one-half of the manager's time will be devoted to Software Support, with the help of other software professionals, insuring applications needed for the Tier-2 center are deployed, configured and updated as needed. The Center Manager is a UM dedicated position.

In addition to these support resources, an additional FTE is provided by MSU for technical software/grid support and 1.6 FTE's are planned from the Tier-2 budget.

AGL-T2 will benefit from an Advisory Committee for support and advice which will include the Associate Provost for Academic, Information and Instructional Technology Affairs from the UM (J. King) and the Vice Provost for Libraries, Computing & Technology from MSU (D. Gift). The Advisory committee will be chaired by M. Abolins (MSU) and H. Neal (UM).

Budget

We have been able to build our proposal based upon significant contributions from both the University of Michigan and Michigan State University. Specifically, we bring the following resources to our proposal[3].

University of Michigan:

- A contribution of **$400K** from the UM Office of the Vice-President for Research (OVPR). These resources are allocated during the 5 years to purchase additional dedicated computing and storage nodes for the Tier-2 center.
- A contribution of **$300K** from the UM Provost's Office. These resources provide 0.4 of a dedicated Tier-2 FTE as well as 50% of the UM MiLR connection and racks and power distribution for the Tier-2.
- In kind contributions totaling over **$800K** contributed by UM Physics and the UM College of LS&A to support Tier-2 activities include:
    - Dedicated use of 59 CPU years of the Nyx Opteron cluster (66k SI2K) in the Center for Advanced Computing for U.S. ATLAS.
    - Dedicated use of the UMROCKS (AMD Athlon) cluster (~130k SI2K) consisting of 100 nodes of dual Athlons (MP 2000/2400/2600)
    - Leveraged use of 50% of the new UM Tier-3 computing power (Tier-3 is $487K total, 391k SI2K by year 3)
    - Housing (Space, power, cooling and campus networking) for the Tier-2
    - Provisioning 50% of a MiLR wavelength for 5 years (other 50% from Provost's support)
    - Shared use of a Cisco 6509-E (UltraLight) and 6506 network chassis with 12 10 Gbits/sec Ethernet ports.

Michigan State University:

- **$863K** for equipment and personnel from the MSU Provost, Vice President for Research and Graduate Studies, College of Natural Science, and Department of Physics and Astronomy to include:
    - **$485K** for new equipment, including 551k SI2K of computing (by the end of 5 years), 159 TB of storage (by the end of 5 years), archival/backup system for the two-campus Center, protected 10 Gbits/sec connectivity on the MSU campus.
- Leveraged use of the MSU Tier-3 computing power.

[3] Prior to any MOU our grants office requested the following text: Due to the difficulty in tracking the referenced cost sharing, documentation will not be maintained.

- Housing (Space, power, and cooling) in renovated computer server rooms within the Biological and Physical Sciences building (~$200K site upgrade).

Each institution:

- Significant computing and networking research activities in support of the AGL-T2 (See Appendix for details).
- Use of a UM-MSU 10 Gbits/sec Ethernet campus link, in addition to our Tier-2 link to StarLight.
- Low cost additional 10-Gbits/sec connections as required for future capability.

UM-MSU	Equipment					Personnel			Indirects	Total
Tier 2 (units)	CPU $(K)	Storage $(K)	Network $(K)	Mainten. $(K)	Equipment Total $(K)	Salary $(K)	Benefits $(K)	Travel $(K)	(52%) $(K)	Tier-2 $(K)
FY06	$ 109.4	$ 83.1	$ 20.6	$ 10.7	$ 223.8	$ 172.5	$ 51.8	$ 10.0	$ 121.8	$ 579.9
FY07	$ 112.7	$ 85.6	$ 3.2	$ 10.1	$ 211.6	$ 177.7	$ 53.3	$ 10.3	$ 125.5	$ 578.5
FY08	$ 116.1	$ 88.2	$ 3.3	$ 10.4	$ 218.0	$ 183.1	$ 54.9	$ 10.6	$ 129.3	$ 595.8
FY09	$ 119.5	$ 90.8	$ 3.4	$ 10.7	$ 224.5	$ 188.5	$ 56.6	$ 10.9	$ 133.1	$ 613.7
FY10	$ 123.1	$ 93.6	$ 3.5	$ 11.0	$ 231.2	$ 194.2	$ 58.3	$ 11.3	$ 137.1	$ 632.1
TOTAL	$ 580.8	$ 441.3	$ 34.2	$ 52.8	$1,109.1	$ 916.1	$ 274.8	$ 53.1	$ 646.9	$3,000.0

Table 2: AGL Tier-2 budget (in $K).

The Tier-2 budget ($600K/year) allocation is shown in Table 2. The amount budgeted by year for compute nodes (CPU), storage, maintenance, and network is subtotaled as "Equipment". The CPU and Storage column funds will be used to purchase the corresponding capacity in Table 1. The Network column will be used to pay for connecting the AGL-T2 over the MiLR network to its peers (ESNet, Abilene, NLR, etc) via Merit and represents paying the initial cost for three 10GE ports and an ongoing "peering" charge. The Salary column is the "normalized" salary available assuming carry-over between years and a 3% inflation rate. Benefits are calculated at 30% and indirects at 52%. The total for five years is three million dollars.

Summary

To summarize our proposal and highlight its strengths and unique attributes we note:

- Our combined UM / MSU group, with 21 faculty, would be the largest active ATLAS research group.
- Our collective groups have a broad range of expertise and demonstrated capability in large scale computing tasks in HEP at the Tevatron experiments and in ATLAS.
- With over $4 Million of leveraged and in-kind resources ($2.5M resources and $1.5M in personnel) our proposal demonstrates extraordinary institutional support.
- Our unique networking capabilities, built on a dark fiber infrastructure, allows AGL-T2 to have a fully protected 10 Gbits/sec Ethernet circuit interconnecting our distributed site to the rest of the world.
- We have a significant, unique and important R&D program of relevance to Tier-2s.
- We are firmly committed to delivering the computing resources (as shown in Table 1) which will exceed the call for proposal specifications by more than 150% in FY08 (assumed high-luminosity year).

APPENDIX

Personnel Details for AGL-Tier2

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

Table 3: Senior personnel/faculty commitment to the AGL-Tier-2.

Related Tier-2 Research Efforts

Ongoing research of importance to an AGL Tier 2 Computing Center include:

UltraLight, a NSF ITR program, started September 15, 2004 thru 2008. The focus of UltraLight is to explore future possibilities for e-Science by integrating agile, configurable, high-performance networking in our grid systems. LHC physics is the focus. (http://ultralight.caltech.edu)

GridNFS, an NMI development project (http://www.citi.umich.edu) was funded and started work on September 30, 2004. Its focus is on delivering a "grid-aware" wide area network version of NFS (Network File-System), starting from the work on NFS V4 done at CITI. Open Science

Grid (OSG) is one of the targeted deployment and testing communities for this work. We are also collaborating with FNAL and DESY on integrating dCache and GridNFS/NFSv4.

TeraPaths, Shawn McKee is collaborating with BNL on the "Prototype QoS/MPLS Enhanced Network Services at the ATLAS Tier 1 at BNL" project. This project's goal is to develop and deploy network/bandwidth management for the Tier1 site and eventually the Tier2 US sites. This effort is synergistic with the UltraLight program and will involve U.S. ATLAS sites and CERN.

Our perspective is that all of the above funded efforts are very important to our future LHC infrastructure. Having a Tier-2 co-located with these efforts is critical to ensuring that the work being done will be useable, effective and integrated with the ATLAS computing model. An AGL Tier-2 assures the applicability and effectiveness of this work for U.S. ATLAS.

Additionally we have a pending proposal to NSF called **PLaNetS** (Physics Lambda-based Network System) targeted at providing vital network tools and services for US involvement in the LHC, as well as other major NSF- and DOE-funded data intensive science programs by developing a core suite of high performance end-to-end data transfer tools and applications. Both Tier-1 centers are collaborators on this proposal.

These activities should be viewed as augmenting our Tier-2 proposal while requiring few enabling resources from the Tier-2. The only load on the system comes from integrating and testing these capabilities as part of the Tier-2 operations. As an example, these activities and the infrastructure which come with them make it much easier (in $ and expertise) to deploy and access cutting-edge network infrastructure in support of ATLAS (note the use of the UltraLight high-end switch to enable 10GE access at Michigan). The simplest example is enabling effective use of 10GE circuits which would allow a Tier-2 center to move 1TB of data from another site in approximately 15 minutes.