



Georgia State University: Shifting the Grid Paradigm with NMI

NSF Middleware Initiative (NMI) Integration Testbed Case Study Series

Series contact: Mary Fran Yafchak, Southeastern Universities Research Association,
maryfran@sura.org.

The NMI Integration Testbed Program provided practical evaluation of NMI components within the context of real projects and application scenarios from June 2002 through November 2004. During that time, NMI Testbed sites collectively submitted over 220 evaluation reports to middleware component developers as direct feedback into the NMI development cycle. Site representatives also actively inspired, promoted and facilitated the integration of middleware throughout their institutions.

The NMI Integration Testbed Case Study Series documents the most significant outcomes and influences of NMI Testbed sites' middleware integration efforts, highlighting intersections with established projects, application contexts and influences, drivers for innovation, decision points and challenges. Through this documentation, the work of these pioneering institutions is captured to provide a breadth of insight and approaches for others to use towards successful middleware development and deployment.

This NMI Integration Testbed Case Study Series is sponsored by the National Science Foundation Middleware Initiative-Enterprise and Desktop Integration Technologies (NMI-EDIT) Consortium of EDUCAUSE, Internet2, and SURA. Additional support was provided by the National Science Foundation Cooperative Agreement NSF 02-028, ANI-0123937.

Copyright © 2005 Board of Regents of the University System of Georgia by and on behalf of Georgia State University. All rights reserved. Content may be used for non-commercial purposes with attribution.



Executive Summary

In an older paradigm of grid computing, grids were seen as an “essentially solved” distributed computing problem. For the past several years, the perception of grid computing throughout Georgia State has been evolving from the older view of grids to the new paradigm in which grid computing provides a middleware infrastructure to coordinate distributed resources.

When Georgia State set about deploying a central identity management infrastructure several years ago, their deployment path provided them the opportunity to work with grids - an opportunity they seized and that helped them come to embrace the new grid paradigm.

There were a number of key elements that facilitated Georgia State’s grid paradigm shift, including participation in the NMI Integration Testbed, visionary campus leaders, and an inspired, effective team of grid champions on campus.

Tasked with spearheading their identity management infrastructure deployment and integration of middleware for the campus, Georgia State’s Advanced Campus Services (ACS) unit built the campus grid initiative on

a foundation of collaboration. These collaborations included a wide range of constituents from the campus community, as well as individuals from peer institutions and higher education IT organizations such as Internet 2.

The ACS employed several methods to engage the campus community in a grid exploration, including the creative use of limited resources and multiple communication forums that focused on understanding Georgia State’s potential grid applications and engaging application users in an ongoing exploration of grids. The ACS methodology has been successful in slowly, but surely, shifting the campus community’s grid paradigm from old to new.

For more information about Georgia State’s NMI Grid development, contact:
Art Vandenberg at avandenberg@gsu.edu.



Georgia State University: Shifting the Grid Paradigm with NMI

In an older paradigm of grid computing, grids were seen as an “essentially solved” distributed computing problem. In this view, grids were part of the fabric of the distributed computing model that personal computing technology had long ago entrenched across the computing landscape of so many institutions. This older view, however, does not fully consider more recent developments and trends in the computing science and networking fields. In the new paradigm, grid computing provides a middleware infrastructure to coordinate distributed resources by integrating enterprise identities, security policies, web services, portals, and technology management tools. Rather than a static definition around groupings of computer clusters, “grid computing is increasingly being viewed as the next phase of distributed computing” (1).

For the past several years, the perception of grid computing throughout Georgia State has been evolving from the older view of grids to the definition provided by the new paradigm. This article will examine the drivers and circumstances at Georgia State that engendered a change in both the perception and use of grids on campus. The intent of this examination is two-fold; the reader should gain a basic understanding of how “new” grids are used, as well as some non-technical methods for developing a

campus grid initiative. As we examine the “recipe” of drivers and processes that contributed to the change in the perception and use of grids at Georgia State, the reader should keep in mind that another institution could (in fact probably should) adapt the recipe to meet their specific needs and environment if they intend to shift the grid paradigm on their own campus.

A Foundation for Grid Exploration

Changes in Institutional Needs

Georgia State University is a public, doctoral/research extensive university. Georgia State’s Information Systems & Technology (IS&T) group supports around 27,270 students (Fall 2004) with 250 fields of study offered through six colleges, and generated \$60 million in research income in FY2004. In the late 1990’s, Georgia State began making plans to address some significant changes in their users’ computing needs. The nature of Georgia State’s research projects increasingly require project teams comprised of researchers at multiple institutions and universities who must share information and other resources across their institutional boundaries. In addition, Georgia State wanted to increase the availability of courseware for both on-campus and remotely located students – especially since a significant portion of its students are not in traditional on-campus



housing. To meet these needs, Georgia State would need to deploy a central identity management infrastructure. Though not apparent at the time, the path Georgia State would follow over the next several years to implement this infrastructure would provide them the opportunity to work with grids - an opportunity they seized and that helped them come to embrace the new grid paradigm.

Collaboration and Strong Leadership

As a member of Internet 2 (I2), Georgia State engages in collaborative activities that help them identify important trends, on both local and national levels, in their research and education environment. Georgia State had been aware of I2 middleware initiatives since 1999, and this awareness provided them with a clear vision of how a central identity management infrastructure could support their users' research and education activities. Due in large part to the leadership of then CIO, J. Reid Christenberry, when Georgia State set about implementing their new identity management infrastructure in 1999, they incorporated Internet2 and EDUCAUSE middleware software and standard practices at the identity management core. In order to ensure proper focus on their middleware initiative, Georgia State created the Advanced Campus Services (ACS) unit in February 2000, and charged it with coordinating the new infrastructure. While Georgia State's grid paradigm did not include middleware at the time the ACS was created, their definitions

of both grids and middleware were about to undergo significant change.

When One Door Opens, So May Another With Art Vandenberg as ACS Director, the university began the investigation, design and deployment of their new middleware infrastructure¹. Vandenberg reached out to I2 as a source of guidance and inspiration. Vandenberg's investigative and collaborative nature served not only the university's identity management infrastructure well, it would also be a key driver of the university's changing view of grids. Among the I2 members Vandenberg met through I2 working group activities were representatives from the Southeastern Universities Research Association (SURA). After SURA was selected to manage the National Science Foundation (NSF) Middleware Initiative (NMI) Integration Testbed² on behalf of the NMI-EDIT Consortium³, they released a Call for Participation (CFP) for institutions to participate as NMI Testbed sites.

The NMI Integration Testbed program was to include eight universities collaborating in a closely coordinated effort to deploy and evaluate NMI middleware technologies in "real life" scenarios. Realizing this type of hands-on, collaborative work could be a key

¹ NMI-EDIT's Georgia State case study: http://www.nmi-edit.org/pdf/GSU_nmiCaseIdMgtFinal16Oct2004.pdf

² The NMI Integration Testbed: <http://www1.sura.org/3000/NMI-Testbed.html>

³ The NSF Middleware Initiative - Enterprise and Desktop Integration Technologies (NMI-EDIT) Consortium of Internet2, EDUCAUSE, and SURA. <http://www.nmi-edit.org>



part of the overall strategy to bring middleware to their campus, Georgia State responded to the NMI Testbed Site CFP. Though not selected as a funded Testbed site, an undaunted Vandenberg, with the support of the university and its CIO, accepted SURA's offer to participate in the Testbed as an unfunded site. This was a significant opportunity for Georgia State, founded on a commitment to developing middleware infrastructure and a belief that collaborative participation could help Georgia State advance their new central identity management infrastructure

In For A Penny, In For A Pound

The middleware software that was to be evaluated within the NMI Integration Testbed included components that were primarily directory and identity management focused (NMI-EDIT), as well as components typically used in grids (GRIDS Center components). Since the mission of ACS initially was the deployment of Georgia State's identity management infrastructure, ACS viewed working with NMI-EDIT components as the central focus of their Testbed participation. However, since these new grid technologies were associated with directories, identity, authentication, authorization, and PKI, Vandenberg and his team were also open to testing and evaluating grid-focused components.

Seeing The Potential in Grids

Early in the Testbed, the Georgia State team familiarized themselves further with the modern grid paradigm. In addition to the infrastructure issue of linking grids to identity management infrastructure issues, Vandenberg was familiar with the potential for grids to harness the power of distributed computing resources. He hoped their work with grids in the Testbed might help Georgia State pool and share their disparate computing resources across the campus (and beyond), thus providing them with an opportunity to increase the available computing resources to their researchers, students and faculty without significant cost increases - and perhaps even at lower cost. Additional potential benefits of grids that warranted further investigation are the ability to leverage grid middleware for accessing campus computing resources more securely and effectively, and the potential to help solve computing research problems that, without grids, are difficult or infeasible to solve (3).

It's one thing to see the potential in the new grid paradigm, and another to see the potential realized. ACS's first order of business, then, in their grid initiative was to understand the nuts and bolts of creating a campus grid with middleware. After evaluating the first NMI software package within the NMI Integration Testbed in the early spring of 2002, Vandenberg and his



teammate, Victor Bolet, were impressed, but realized that to fully evaluate the software components, they needed to put them through their paces within the context of a real application. To ferret out those Georgia State applications that could benefit from grids, Vandenberg and Bolet began what has become ACS's continuing discussion with the campus user community about the new grid paradigm.

Multifaceted Strategy – The Recipe for Change

To determine which applications were “grid-appropriate”, Vandenberg and his team would need to learn more about application requirements, while application users would need to develop an understanding of how grids could benefit their work. ACS would employ a multifaceted strategy to advance the new grid paradigm. The strategy would, first and foremost, need to facilitate the learning process between ACS and potential grid users and make use of various communication tools to reach the broadest possible audience. All of this would also need to be done as cost effectively as possible.

The components of Vandenberg's strategy listed below comprise Georgia State's recipe for changing the perception and use of grids on the Georgia State campus:

- Collaborate
- Make Creative Use of Limited Resources
- Communicate

- Understand the Applications
- Engage the Researcher

Collaborate

Shifting the Georgia State grid paradigm was a journey the campus community would take together. It wouldn't be enough to build a campus grid in a vacuum. To ensure their new campus grid would meet their needs, the community had to build a mutual understanding of grids through a continuous information exchange. Georgia State's ongoing collaboration held to the following principles:

- Strive to meet mutual objectives of faculty, researchers and IT staff. Common objectives often center on project funding, publication of work, and a desire to expand student learning opportunities by including them in projects whenever possible.
- Engage peers. Both ACS and potential grid users should reach beyond the campus to find peers to increase their opportunities for, and exposure to, new ideas.
- Participate in collaboration. Collaborators need to be active in their group - volunteer, be enthusiastic and work to develop cooperative approaches to common problems.
- Grow relationships through personal contact.

The GRID Group @ GSU

During Georgia State's first year in the NMI Integration Testbed, Vandenberg engaged a number of researchers and faculty one-on-



one in discussions about grid technology and their research projects and applications, determining that there were indeed applications on campus that could take advantage of grids. While this smaller discussion format was an effective approach initially, there was a need for a broader discussion forum where potential grid users could come together to share ideas and learn about grids together with Vandenberg's team. To address this need, in the spring of 2003, the GRID Group @ GSU began meeting. ACS issued an open invitation for group membership, with previously identified researchers and faculty invited as core researchers.

The importance of gathering potential grid users into a working group forum should not be underestimated. Dr. Xiaochun He, a physicist who worked very closely with ACS in the past several years to grid-enable his Cosmic Ray Lab (see section "Experimental High Energy Nuclear Physics" below), believes that, in the absence of such a forum, these individuals would likely have been slower to come together, given the tendency for researchers to work sometimes independently (and, perhaps, with relative isolation) in their own labs and offices.

The GRID Group @ GSU developed the following list of grid objectives:

- Build effective grid infrastructure on the Georgia State campus
- Identify Georgia State applications that can benefit from grids

- Attract and retain faculty and students
- Foster research activities and opportunities
- Encourage funding

Vandenberg's philosophy on structured groups like this is that mutual group goals should determine the nature of the group's activities, including how often they meet and for how long the group continues. During 2003, the group met regularly, working to absorb new information about grids and middleware as fast as the ACS team was gathering it in the NMI Integration Testbed. Since that time, the group has established various other avenues of communication (e.g., side groups for proposal development, cross-membership in other groups) and has shifted to an occasional meeting schedule.

Creative Use of Limited Resources

Employ Students as Staff

Though the ACS exploration and development of a Georgia State campus grid is supported by campus leadership, the reality is that funding for ACS's grid-based activities has been tight. Yet Vandenberg and his team have used creative approaches to address staffing needs and technical resources. Hiring students is a most effective means that ACS has used to stretch its smaller-than-needed grid development budget. ACS has leveraged funds and coordinated student positions funded under various grant programs, such as the NMI Research Experiences for Undergraduates (REU) program, to consistently employ graduate and



undergraduate students in their grid initiative. Since October 2000, Vandenberg estimates that middleware-related ACS programs have generated work opportunities for more than twenty students.

After losing a full-time employee just eight months after the ACS was created, Vandenberg was able to redeploy that salary line item to hire graduate students instead. In one case, a student employee became so valuable to the ACS research grid effort that, upon graduation, the student was hired as a fulltime employee. When they hire former student employees, ACS gains employees who appreciate the learning opportunity they were provided as students and are already trained and able to contribute to the organization.

Students are valued employees for ACS. There have been several instances in collaborative projects where student employees have provided important technical (computer and grid-based) support for Georgia State researchers that decreased the technical learning curve for the researcher and, more importantly, allowed them to focus on the research science itself. In addition, Vandenberg notes that often times students don't see as many obstacles in their paths – maybe they have an *advantage*: they don't know the saying, "things cannot be done."

Providing students with meaningful, real-world experiences at Georgia State means

including student positions into collaborative grid projects with other departments. Where researcher's application work might benefit from "getting on the grid", ACS has shared student employees with other departments. These collaborations include projects with, for instance:

- Professor Vijay K. Vaishnav, Computer Information Systems – Vandenberg and Bolet presented topics on grid computing for Georgia State's *CIS 9220 Topics in Information Technology Systems* course.
- Professor Yi Pan, Computer Science – CS Graduate Student Nova Ahmed worked with Vandenberg and Pan to deploy a new multiple genome alignment algorithm to the grid (see below section "Applications: Grid-Enabling Them").
- Professor Xiaochun He, Physics – Nicole Geiger, Physics, and Anish Shindore, CIS, worked as REU students for the NMI Integration Testbed, building and demonstrating a muon-detector grid for Dr. He's Cosmic Ray Lab, and compiling the "Catalog of Grid Applications (see section below).
- Professor Ying Zhu, Computer Science – working with ACS and his CS graduate student James Reid to create a grid for graphics rendering to support his graphics class and their projects.

Grid middleware is still considered by many to be cutting edge and this can be intimidating (even for technical staff). The fact that ACS students have been able to make such significant contributions to



Georgia State's grid initiative is indicative of the students' skills and also illustrates that grids, when approached with a can-do attitude, can be implemented. Student employees have been key players in many of the ACS's key grid activities including enabling grid applications, portal development, grid node installation and configuration, and installing Ganglia to monitor node activity. Students have demonstrated and discussed grid capabilities during GRID Group meetings - something Vandenberg has found can have a real impact on engaging the whole Grid Group.

Turn "Trash into Treasure"

ACS has also begun to salvage surplus computer hardware and turn this almost discarded equipment into productive grid nodes. In one example, two REU students developed a grid demonstration project for the Cosmic Ray Lab (see "Experimental High Energy Nuclear Physics" section below) by building ten PCs from the parts of forty surplus units, then putting the PCs "on a grid" by installing and configuring the GRIDS Center's *Globus*⁴ middleware components on them.

Communicate

There is no single or best means of communication for reaching a campus community as large and varied as Georgia State's. The ACS grid team would therefore

need to employ multiple communication techniques. Existing campus communication tools would be used, and where these didn't meet the goals of the ACS grid initiative, they would develop new tools.

- Use both one-on-one and group discussion formats.
- Leverage your people network. Vandenberg's prior work on the campus Y2K preparedness project brought him into contact with many departments and users on campus. Vandenberg continued to build these relationships and was able to create new contacts in turn.
- Coordinate within the IS&T organizational structure. The CIO and six Directors (ACS is one of the six units) coordinate technical strategy and objectives for the enterprise – grids are a key component of research computing infrastructure for the campus.
- Use the Focus-IT Newsletter. As part of its commitment to participatory, consensus-driven operations, IS&T publishes this newsletter on a monthly basis, circulating it to faculty and staff. Vandenberg has used the newsletter for ongoing outreach and education to the broader campus community about grid and other ACS initiatives.
- Use prototyping and application demonstrations to show the benefits and use of grids.

Visual Communication - A Grid Demonstration

No matter who your audience is, there are few better ways to champion a new concept than letting your users see and, ideally,

⁴ Globus Toolkit information: <http://www-unix.globus.org/toolkit/>



interact with, an example of the concept. Realizing that the grid landscape is full of new terms and concepts, Vandenberg's team began evolving a grid demonstration tool in mid-2004, beginning with the question, "What functions do we need to show our users so they can appreciate the grid?" Georgia State was evolving their grid demonstration in tandem with their participation in the Testbed, and Testbed sites had common concerns when it came to reaching out to potential grid users on their campuses. As Testbed site representatives' skills with middleware and campus grid resources grew in the Testbed, so did their understanding of the elements and process of creating an effective grid demonstration.

The NMI Integration Testbed sites worked together to prepare a joint set of grid demonstrations for the Internet2 Fall 2004 Member Meeting. For this meeting, Georgia State created a 7-step demonstration that featured their Cosmic Ray Lab muon detector grid (see "Experimental High Energy Nuclear Physics" section below), modeling grid nodes that store data collected by remotely located muon particle detectors. The functionality of the grid was shown as a "user" entered the OGCE portal, generated a grid credential, located a set of collected particle data, set up and ran a simulation, transferred the output data using gridFTP, and visualized the results with a 3D wire frame viewer

In December 2004, the GRID Group @ GSU presented the Cosmic Ray Lab muon detector grid demonstration to a small but targeted group of Georgia State researchers. Their conclusion that, as a general set of steps, many scientists (besides physicists) can use the 7-step process in their research, confirmed the benefit of further development of a Georgia State campus grid. To better convey the value of grids to more researchers from additional fields, ACS is making plans to expand the research applications used within the grid demonstration since it should preferably feature "the researcher's science" (e.g., molecular manipulations for chemists or biologists). Ideally, the grid demonstration project will be expanded with several such applications, providing a growing library of demonstration cases.

After the Demonstration – Persistent Grid Resources

The ACS team discovered over time that they would need more than a demonstration to help potential and new grid users come to understand and utilize grids effectively. One issue with trying to convey the potential in the new grid paradigm is the relatively brief period in which the demonstration is typically run. People often need to see new concepts more than once, and certainly might want to do so with a demonstration of grids as they contemplate how their application might best make use of this technology. Vandenberg's team has been working on a grid demonstration that would be more available and accessible (preferably on-demand).



While the ACS's 7-step demonstration is a straightforward process, grid users (both potential and certainly actual) need access to persistent grid resources to investigate, program and run applications on. In addition to enhancing the current grid demonstration, ACS is also building a grid lab at Georgia State to support exploration by campus users.

Understand the Applications

Applications: First, Find Them

The Catalog of Grid Applications

Early in the ACS grid development initiative, Vandenberg created the Catalog of Grid Applications⁵ project as a means to discover the state of grid research at Georgia State (who is implementing grids and for what applications), and as a subsequent means to track the use and progress of the applications. Two REU students initially populated the catalog with Georgia State's applications, then their inventory work expanded to data from all NMI Integration Testbed sites. Currently, the catalog includes application data that represents the work of over 475 researchers from twenty universities and national labs. As part of ongoing efforts to reach potential grid users on their campus, the cataloged data is shared with potential users to help them see how others are using the grid and how that might apply to their own work.

The Catalog may also facilitate inter-institutional research and cost-sharing collaborations. For example, Georgia State physicist Dr. Xiaochun He uses data generated and stored at Brookhaven National Laboratory. The optical networks like National Lambda Rail that Dr. He could use to access the data can be quite costly. By using the Catalog of Grid Applications to find similar particle physics researchers at other institutions, Dr. He may be able to foster a collaborative solution with these other institutions. Such a team could share the costs of optical network access or develop a research project where such access is required.

Applications: Grid-Enabling Them

If there is any facet of changing the perspective and use of grids on a campus that requires collaboration and creativity, it's grid-enabling an application. If an application is already running, rewriting it for the grid needs to have some motivating factor. In determining whether there is sufficient motive, one should consider the following:

Will the application run faster? Consider a genome alignment algorithm, where two gene sequences are being compared to find similarities. A single CPU processor can complete the job, but using a grid of processors might enable the work to be parallelized – the gene comparison could be split among, for instance, 10 processors that each do a smaller piece of gene-to-gene comparison and communicate their results

⁵ Catalog of Grid Applications
http://art12.gsu.edu:8080/grid_cat/index5.jsp



to one another to produce the final result. While the inter-processor communication adds overhead beyond that involving only one processor, the speed and efficiency of having 10 processors working together may offset this.

If not faster, then can more jobs be run in the same time? It may not be a question of speeding up a particular process. For instance, Dr. Zhu's graphic students have projects that render an animation file by rendering each frame. If there are 1000 frames, each may take only several minutes, but a single processor must do each one in turn; 1000 frames at just one minute a frame ends up taking 16 hours. Having access to 10 processors in a grid means the overall time can be reduced maybe 10 times over.

Will the infrastructure be better? Using grid infrastructure can enable a more secure environment, the ability to share distributed resources among many users, or the aggregation of separate resources into a pool of resources where the "whole is greater than the sum of the parts."

Rewriting applications requires time and effort, and may be non-trivial, so grid-enabling an application is helped if there are examples or precedents. It does present challenges, but these are not insurmountable, as the ACS team has discovered. A case in point was Professor Yi Pan, Computer Science, and Art

Vandenberg's collaborative project to grid-enable a genome alignment application.

Nova Ahmed, a Computer Science graduate student of Professor Pan's, was developing a memory efficient genome alignment algorithm for running on a shared memory machine. The shared memory architecture was specifically designed to be a tightly coupled set of processors, sharing memory, and with fast interconnection of the multiple processors. In contrast, a grid is a loosely coupled set of processors, each with its own memory and interconnects that are typically much slower, on the order of 10/100mbit switched Internet speeds. There was some skepticism that the algorithm could be adapted to run well in a loosely coupled grid environment, or that it would achieve significant speedup if it were adapted. The intent of the collaborative research, then, was to see if Ahmed could successfully grid-enable the genome application. Ahmed was successful in her work - the algorithm performed well in a grid cluster environment - and it has provided an illustrative example of how applications can be grid-enabled, and the benefits that can bring. ACS is working within SURAgid⁶ (the post-NMI continuation of NMI Testbed Grid) on grid deployment issues where the genome alignment application continues to provide a reference example.

⁶ Work begun in the NMI Integration Testbed is being continued in the SURAgid. See: <http://www1.sura.org/3000/SURAgid.html>



Engage the Researcher

Researchers and the Grid Paradigm

Early conversations the ACS team had with campus members about grid development did not always produce immediate results. Some researchers, faculty or management were not particularly convinced of the “new grid” ideas Vandenberg broached them with, they believed that independent clusters of workstations they were using were “all the grid” they needed. In some cases, their applications had not yet moved to a stage (larger application problems, or more complex problems) where the benefit of the new grid paradigm would be apparent. In other cases, there was an acknowledgement of the “new grids” but a lack of clarity about how they worked (e.g., concerns about the security, who owns or controls the resources on a grid, how are usage policies set, etc.). In situations like these, while the concept of the new grid was common to those in the conversation, the technical details were not, and campus policies didn't exist to deal with the security and access issues the “new grid” presented.

What the ACS team and the campus community had to do was to come to a common understanding of both the definition and usage details (technical and policy-based) involved in Georgia State's grid deployment. The ACS team used a collaborative approach to reach a consensus about the new grid paradigm. Vandenberg's group focused on finding those things ACS and researchers have in

common – a desire to collaborate and to include students in the collaborative work, a need to seek independent funding and to publish their work, and an interest in the research topics themselves (e.g., computer science, bioinformatics). With these common interests in mind, Vandenberg, Bolet and the GRID Group @ GSU began looking for those researchers whose view of grids was open to the new paradigm, or at least researchers who were receptive to exploring the potential of new grid technologies.

Researchers as Grid Partners

Of the researchers ACS has collaborated with during the past three years of their grid initiative, two stepped-up to the grid-challenge early on and partnered with Vandenberg and his team. Their work has helped advance the acceptance of the new grid paradigm on campus and provided an example of the type of collaborations and applications that the new paradigm can foster.

Semantic Facilitator™ SM

Vandenberg and Professor Vaishnav have built a Semantic Facilitator™ SM to display directory objects in a clustered and meaningful way. An initial prototype has been implemented and its clustering capability demonstrated. The two are now investigating how grid-based resources may permit the Semantic Facilitator™ SM to operate more efficiently. One aspect of the clustering algorithm requires exploring many parameter sets to find those parameters that



more closely match a human expert's clustering. Using a grid would enable these many parameter sets to be explored in parallel. Another idea is to use grid-distributed resources to cluster many large sets of data from multiple directories in parallel to speed up results.

Vandenberg and Vaishnav not only share a common definition of the new grid, they share a vision for its future. Both see a potential ubiquity for grids rivaling that of the Internet, though whether it will progress with the rapidity of the Internet remains to be seen. In the meantime, Vandenberg and Vaishnav agree that to make grids as practical and useful as the Internet, further development, both in the lab and real world, is required.

Experimental High Energy Nuclear Physics

During one of Vandenberg's initial discussions with Georgia State faculty about the ACS's grid initiative, he was referred to Professor Xiaochun He in the Physics Department as a potential grid-user. Dr. He's works in particle physics, an area of physics that depends on the analysis of huge data sets resulting from collider experiments and often makes the job of the researcher like looking for a needle in a haystack. Vandenberg and Dr. He's early conversations centered on the experience Dr. He already had with the PHENIX grid. Their conversations make an excellent example of the mutual learning that deploying a campus grid, and grid-enabling

applications for it, depend upon. The PHENIX grid provided a talking point for Vandenberg and Dr. He regarding the intersection of NMI-EDIT and GRID's Center middleware components for sharing resources across inter-institutional grids

One of the most significant outcomes of Dr. He's collaboration with the ACS team is the grid-enabling of the Georgia State's Cosmic Ray Muon Project⁷. The muon particle detector is planned as a distributed network of detectors installed at selected middle and high schools in Georgia, and two Georgia State REU students have done much of the deployment work. The project provides a means of outreach to enrich the education experience for students and their teachers by using grid technology to manage and access the detector's distributed resources including sensors to detect and an application to simulate the trajectory of muon particles. Without being grid-enabled, the logistics of coordinating muon measurements at the nearly 100 schools targeted for inclusion in the project would likely be overwhelming.

The project, one of the first undertaken by ACS, provided them with a real world environment in which they could evaluate NMI middleware components and spurred discussion between ACS, the GRID Group @ GSU, Dr. He and his physics coworkers.

⁷ *NMI GRID Middleware Enhances Georgia State's Cosmic Ray Muon Project*
<http://www1.sura.org/3000/NMI-Testbed/GSU-GridMuonDetector.pdf>



The NMI Integration Testbed program provided Vandenberg's team with real life experience in using NMI components, while the discussions between GRID Group @ GSU and Georgia State's physicists helped illuminate the potential benefits the grid could have for the muon detector project.

The nature of the high energy and particle physicist's research requires the researcher to work on the development of the tools that enable their research – including computing resources like grids - since the physicist is always looking for faster, on-demand access to the huge volumes of data they must work with. The field also requires patience to reach long-term research goals. Likewise, the deployment of a grid is a long-term undertaking, especially when it is intertwined with a requisite shifting of the campus community's grid paradigm. Having a partner such as Dr. He, able to think in the long term and that has the skill to think beyond his field of study, has been an important advantage for Vandenberg and his team. It should be noted that while physics departments might be one of the first places one might look, grid application deployment partners can be found in unexpected fields as well, such as the arts and humanities⁸.

A Shifted Grid Paradigm Evidence of Progress

The definition and view of grids on the Georgia State campus has most certainly changed since the ACS was created. One of the strongest indicators of this was the reaffirmation this past fall by J.L. Albert, Georgia State's CIO since March 2004, of the institution's commitment to support ACS's research computing and grid initiatives. Additional confirmation comes in the form of new computing hardware. Georgia State is installing a 1,000-node United Devices Grid MP – providing a grid that will harvest the idle cycles of workstations in labs and classrooms across the campus. ACS is also hoping to grid-enable a 24-node, dual CPU Beowulf cluster. Georgia State's growing bioinformatics and research computing initiatives were bolstered by the recent acquisition of an ATIPA cluster and Georgia State's VP for Research, as well as a number of research and faculty members, increasingly interested in grids are including grid solutions in research proposals. ACS has received requests from Computer Science faculty to grid-enable groups of their classroom and research workstations to facilitate their work, for example, in bioinformatics, visualization, and for Computer Graphics students' use in rendering activities.

⁸ See the technical supplement *Building a Campus Grid: Concepts and Technologies* for additional information on this subject at: <http://www1.sura.org/3000/NMI-Testbed/SURA-CampusGrid.pdf>



Conclusion

Georgia State's grid paradigm shift was not a matter of happenstance, rather it came about through visionary campus leadership, and an inspired, effective team of grid champions on campus. The University's participation in the NMI Integration Testbed helped facilitate the identity management infrastructure they were implementing and their exploration of grids. The ACS employed several methods to engage the campus community in a grid exploration, including the creative use of limited resources and multiple communication forums that focused on understanding Georgia State's potential grid applications and engaging application users in an ongoing exploration of grids. The ACS methodology has been successful in slowly, but surely, shifting the campus community's grid paradigm from old to new.

More Information

For more information about Georgia State's NMI Grid development, contact:

Art Vandenberg at avandenberg@gsu.edu.

References

- 1) OGSA
http://www.gridforum.org/L_WG/News/OGSA%20Flyer_v31.pdf
- 2) I. Foster, C. Kesselman, J. Nick and S. Tuecke, *The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration*, 2002.
- 3)
http://www.gridforum.org/ggf_grid_understand.htm



Links of Interest

Advanced Campus Services <http://www2.gsu.edu/~wwwacs/>

Catalog of Grid Applications http://art12.gsu.edu:8080/grid_cat/index5.jsp

Cosmic Ray Measurement, Department of Physics and Astronomy, Georgia State University
<http://phynp6.phy-astr.gsu.edu/~cosmic/>

Georgia State University <http://www.gsu.edu>

Global Grid Forum <http://www.gridforum.org/>

GRID Group @ GSU http://www.gsu.edu/~wwwacs/GRID_Group/Index.htm

GRID technology <http://www.globus.org/> and <http://www.globus.org/research/papers.html>

GRIDS Center <http://www.grids-center.org/>

Memory Efficient Pair-wise Genome Alignment Algorithm -- A Small-Scale Application with Grid Potential, in Grid and Cooperative Computing. Ahmed, N., Pan, Y., and Vandenberg, A.
<http://www.springerlink.com/index/6QKX7FVA226FEGQ6>

Middleware infrastructure <http://middleware.internet2.edu/>

National Lambda Rail <http://www.nlr.net>

NMI-EDIT <http://www.nmi-edit.org/>

NMI Integration Testbed Program <http://www1.sura.org/3000/NMI-Testbed.html>

NSF Middleware Initiative <http://www.nsf-middleware.org/>

NSF Research Experiences for undergraduates (REU) Program
<http://www.nsf.gov/home/crssprgm/reu/start.htm>