



The University of Alabama in Huntsville: Grid Integration with LEAD (Linked Environment for Atmospheric Discovery)



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 influences and outcomes 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 sNational Science Foundation Cooperative Agreement NSF 02-028, ANI-0123937.

Copyright © 2004 University of Alabama in Huntsville. The University of Alabama in Huntsville permits use of this content for noncommercial purposes with proper attribution. All rights reserved.



Executive Summary

It would be difficult to understate the demanding nature of meteorological applications on typical computer resources, and the degree to which the limitations of these resources have impacted the ability to predict severe weather events to date. Fortunately, computer scientists and meteorologists from the University of Alabama in Huntsville (UAH), and eight other universities, are working to address the computer science and meteorological challenges of mesoscale weather research through their participation in the Linked Environments for Atmospheric Discovery (LEAD) project. LEAD is a National Science Foundation (NSF) Information Technology Research (ITR) project that seeks to overcome a variety of the limitations that current computing technology places on accurate weather prediction.

In addition to their participation in the LEAD project, UAH is also a member of the NSF Middleware Initiative (NMI) Integration Testbed Program ¹. It is within this framework that the NMI Integration Testbed Program has furthered UAH's work in the LEAD program. The use of grid technology

is one of the key ways LEAD will provide researchers with access to large-scale computing resources, real-time weather data, models and analysis tools. The LEAD Grid and Web Services Testbed (GWSTB) team has been assembled with researchers from LEAD collaborating sites. Enabled by grid technology, the networked GWSTBs will allow researchers to develop on-demand hazardous weather detection systems.

Early in the project, researchers chose NMI software as the basis for middleware components to be used for LEAD. NMI was chosen because the software components are nicely packaged, well tested, documented clearly, and researchers have experience in working with NMI components. UAH researchers working on the LEAD project are currently assessing the use of specific NMI components. The GWSTBs plan to deploy [MyProxy](#) for credential management when the LEAD portal is ready for deployment. The ITSC-developed Algorithm Development and Mining (ADaM) Toolkit is one application that uses the GRIDS Center's [Globus](#). Using their experience in grid technology, ITSC researchers are using [Globus](#) to grid-enable ADaM.

The use of grid technologies supported by NMI components in LEAD project will help

¹ As part of its overall effort to develop and disseminate software that lets scientists and educators share resources across the Internet, NMI has begun a practical deployment and evaluation effort called the NMI Integration Testbed. Managed by Southeastern Universities Research Association (SURA) on behalf of the NMI-EDIT Consortium, the testbed consists of eight universities that participate in a closely coordinated effort to deploy and evaluate NMI technologies. <http://www.nsf-middleware.org/testbed/>.



researchers and regional forecasters conduct real-time analysis with data mining and detect severe weather events *as they are forming*, which is a major step in making accurate and timely weather predictions.

For more information about UAH Grid Integration with LEAD, contact Sandra Redman at sredman@itsc.uah.edu.



NMI Components Highlighted in this Case Study

The NMI components in this case study series encompass NMI Releases 1 through 4. Information about NMI Releases can be found at <http://www.nsf-middleware.org/>.

Globus

The GRIDS Center's Globus Toolkit is an open-source collection of modular technologies that simplifies collaboration across dynamic, multi-institutional virtual organizations. It includes tools for authentication, scheduling, file transfer and resource description.

Home site: <http://www-unix.globus.org/toolkit/>

MyProxy

The GRIDS Center's MyProxy is a credential repository for the Grid. MyProxy provides a set of flexible authorization mechanisms for controlling access to the repository.

Home site: <http://grid.ncsa.uiuc.edu/myproxy/>



The University of Alabama in Huntsville: Grid Integration with LEAD (Linked Environment for Atmospheric Discovery)

They say you can't change the weather and, granted, they may be right. However, this doesn't mean we cannot expect to lessen the damage caused by severe weather upon people and their environments. Every year, weather events, such as floods, tornadoes, and winter storms (termed "mesoscale" weather events) nationally cause hundreds of deaths and economic losses measured in the billions of dollars. Reducing these losses is dependent upon sufficient forewarnings of severe weather events. In turn, such prediction is dependent upon the information meteorologists have about prior and current weather events, along with their access to powerful computer technologies, new tools and meteorological applications. It would be difficult to understate the demanding nature of meteorological applications on typical computer resources, and the degree to which the limitations of these resources have impacted the ability to predict severe weather events to date.

Fortunately, computer scientists and meteorologists from the University of Alabama in Huntsville (UAH), and eight other universities², are working to address

the computer science and meteorological challenges of mesoscale weather research through their participation in the Linked Environments for Atmospheric Discovery (LEAD) project. LEAD is a National Science Foundation (NSF) Information Technology Research (ITR) project that seeks to overcome a variety of the limitations that current computing technology places on accurate weather prediction. Each participating LEAD university is contributing to the project through their unique area of expertise. UAH's contributions to LEAD are in the areas of developing semantic and interchange technologies for working with large, heterogeneous data sets, as well as developing advanced data mining technologies for phenomena detection and knowledge discovery.

In addition to their participation in the LEAD project, UAH is also a member of the NSF Middleware Initiative (NMI) Integration Testbed Program. Managed by the Southeastern Universities Research Association (SURA) on behalf of the NMI-Enterprise and Desktop Integration Technologies (EDIT) Consortium, the testbed consists of eight universities that participate in a closely coordinated effort to deploy and evaluate NMI technologies³. It is

² The eight other universities in the LEAD program are: Oklahoma University, Colorado State University, Millersville University, Indiana University, Howard University, The University Corporation for Academic Research (UCAR), and the University of Illinois Urbana Champaign, and the University of North Carolina.

³ <http://www1.sura.org/3000/NMI-Testbed.html>



within this framework that the NMI Integration Testbed Program has furthered UAH's work in the LEAD program. The use of grid technology is one of the key ways LEAD is addressing the computer resources needed for understanding weather systems through mesoscale weather research. This article will focus on how specific NMI grid components and lessons learned are being applied and integrated by UAH in the context of the LEAD project.

UAH Background in Middleware

"One of the primary missions of the Information Technology and Systems Center (ITSC) at the University of Alabama in Huntsville is to develop and transfer innovative technologies and knowledge into real-world areas. ITSC has been at the forefront of applying mining and other information technologies to scientific research domains, and is active in the areas of grid technology, collaborative computing, and high-performance networking." (1) As part of the process of bringing the benefits of grid technology to their research projects, UAH's ITSC became a participant in the NSF-funded NMI Integration Testbed Program. Through the NMI Testbed, the ITSC participates in "real life" evaluation and feedback on NMI middleware components, specifications, and services, such as the GRIDS Center's [Globus Toolkit](#)⁴ and the Open Grid Computing Environment (OGCE). At the same time, participation in the

Testbed allows UAH to further its goals in middleware research through collaboration with other institutions working within the Testbed. Already, the lessons learned from their participation as an NMI Testbed site have provided UAH's ITSC researchers with insights into some of the issues currently being worked within LEAD.

NMI Component Use in LEAD

The LEAD project, and weather research and prediction in general, present demanding computing challenges that grid technology is expected to help users overcome. LEAD will provide researchers with grid-based access to large-scale computing resources, real-time weather data, models and analysis tools. These grid resources will allow scientists to develop on-demand hazardous weather detection systems that can recognize and react to changing weather conditions in a manner that non-grid based computing technology simply cannot. For example, scientists will be able to direct Doppler radars to collect specific data or dynamically increase the resolution of a forecast model grid for regions with severe weather. Ultimately, these new grid-based resources and emerging meteorological applications will help researchers track weather systems, and make more accurate and timely severe weather forecasts.

A major requirement is to deploy a testbed environment for LEAD development and operations. UAH's Sandi Redman, a Research Scientist, is heading the LEAD

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



Grid and Web Services Testbed (GWSTB) team, which has been assembled with researchers from LEAD collaborating sites. UAH will host one of LEAD's five interconnected GWSTBs. LEAD GWSTBs are a customized, self-contained grid environment within which the LEAD system will be developed, tested, and used operationally during the lifetime of the project. Enabled by grid technology, the networked GWSTBs will allow researchers to use advanced applications to carry out mesoscale weather research. The LEAD GWSTB team meets biweekly via videoconference on the Access Grid to work on LEAD deployment issues, such as defining the GWSTB environment, developing procedures for account and configuration management, bug tracking, identifying system components, developing documentation, and verifying configurations.

Early in the project, researchers chose NMI software as the basis for middleware components to be used for LEAD. NMI was chosen because the software components are nicely packaged, well tested, documented clearly, and researchers have experience in working with NMI components. UAH researchers working on the LEAD project are currently assessing the use of specific NMI components. In addition, they are developing software applications for use within the project, and are defining LEAD's development and testing environments. These assessments have been informed by UAH's experience and insights gleaned during their participation as an NMI Testbed site. Currently, the GRIDS

Center's [MyProxy](#)⁵ and [Globus](#) are two of the NMI components being used in the LEAD project.

Users will connect to the LEAD portal for working within the LEAD environment. The LEAD portal is based on the Open Grid Computing Environment (OGCE). Development of the customized portal is underway. The GWSTBs plan to deploy [MyProxy](#) for credential management when the LEAD portal is ready for deployment. The LEAD portal will also include the integration of workflow orchestration and other services.

Developing grid-enabled applications for LEAD began with [Globus](#) Toolkit version 3.0.2. The LEAD GWSTB research sites standardized on this [Globus](#) version since it had been well tested in a number of research arenas, and some initial LEAD applications have a dependency on it. The ITSC-developed Algorithm Development and Mining (ADaM) Toolkit is one application that uses the GRIDS Center's [Globus](#). The ADaM toolkit consists of interoperable components that can be linked together in a variety of ways to aid researchers in defining and performing data mining operations on spatial data sets. UAH researchers are refining this software and developing new algorithms for LEAD. Using their experience in grid technology, ITSC researchers are using the GRIDS Center's [Globus](#) to grid-enable ADaM.

⁵ MyProxy information:
<http://grid.ncsa.uiuc.edu/myproxy/>



The UAH team is also studying the semantics used to define both data and applications in different systems and are developing a standard terminology to ensure interoperability with the many heterogeneous datasets of interest to LEAD. The Earth Science Markup Language (ESML) is a specialized markup language for Earth Science metadata based on XML, and provides the basis for core Interchange Technology that allows data/application interoperability in LEAD's grid environment.

Future NMI Component Use in LEAD

With their grid-enabled applications able to harness the combined computing resources available through the GWSTBs, researchers will be able to find relevant data at one remote location, then find a tool (such as UAH's data mining system) at another location, and use the computer resources at yet a third location to retrieve and analyze their data with new, customized data mining algorithms. The ability to conduct this type of

real-time analysis with data mining can help weather researchers and regional forecasters detect severe weather events as *they are forming*, which is a major step in making accurate and timely weather predictions. It is expected that NMI components will continue to help UAH researchers further enhance the applications being used in LEAD. Assessments of additional components are underway and will continue as the LEAD project evolves.

More Information

For more information about UAH Grid Integration with LEAD, contact Sandra Redman at sredman@itsc.uah.edu.

References

- (1) "University of Alabama in Huntsville NMI Testbed Program: A Testbed for Integrated Distributed Environments (TIDE)", Graves Dr. S.J, Redman, S, McEniry, M.



Links of Interest

The University of Alabama in Huntsville <http://www.uah.edu/>

Algorithm Development and Mining System (ADaM) <http://datamining.itsc.uah.edu/tools.html>

Earth Science Markup Language (ESML) <http://datamining.itsc.uah.edu/tools.html>

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

Information Technology and Systems Center (ITSC), UAH <http://www.itsc.uah.edu/>

Linked Environments for Atmospheric Discovery <http://lead.ou.edu/>

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

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

NSF Information Technology Research (ITR) <http://www.nsf.gov/home/crssprgm/itr/>

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

Open Grid Computing Environment (OGCE) <http://www.collab-ogce.org/nmi/index.jsp>