8B.1

A NEW PARADIGM FOR MESOSCALE METEOROLOGY: GRID AND WEB SERVICE-ORIENTED RESEARCH AND EDUCATION IN LEAD

*[1]Kelvin K. Droegemeier, [2]Tom Baltzer, [2]Anne Wilson, [2]Mohan Ramamurthy, and [3]Katherine Lawrence

[1]School of Meteorology
University of Oklahoma
Norman, Oklahoma

[2]Unidata Program Center
University Corporation for Atmospheric Research
Boulder, Colorado

[3]School of Information
University of Michigan
Ann Arbor, Michigan

1. INTRODUCTION

On 1 October 2003, the National Science Foundation began funding a Large Information Technology Research (ITR) grant known as Linked Environments for Atmospheric Discovery (LEAD). A multi-disciplinary effort involving nine institutions and more than 100 scientists, students and technical staff, LEAD is creating an integrated, scalable framework in which meteorological analysis tools, forecast models, and data repositories can operate as dynamically adaptive, on-demand, grid-enabled systems that a) change configuration rapidly and automatically in response to weather; b) respond to decision-driven inputs from users; c) initiate other processes automatically; and d) steer remote observing technologies to optimize data collection for the problem at hand. Although mesoscale meteorology is the particular science domain to which these concepts are being applied, the methodologies and infrastructures being developed are extensible to others including medicine, ecology, hydrology, geology, oceanography and biology.

LEAD is targeted principally toward the meteorological higher education and research communities, though the project also is developing learning communities, centered around teacher-partners and alliances with educational institutions, to bring the benefits of LEAD technologies to grades 6-12 (Clark et al., 2007).

2. PROJECT OBJECTIVES

LEAD has two major objectives. The first is *to lower the entry barrier for using, and increase the sophistication of problems that can be addressed by, complex end-to-end weather analysis and forecasting/simulation tools.* Existing weather tools such as data ingest, quality control, and analysis/assimilation systems, as well as simulation/forecast models and post-processing environments, are enormously complex even if used individually. They consist of highly sophisticated software developed over long periods of time, contain numerous adjustable parameters and inputs, require one to deal with complex formats across a broad array of data types and sources, and often have limited transportability across computing architectures. When linked together and used with real data, the complexity increases dramatically. Indeed, the control infrastructures that orchestrate interoperability among multiple tools – which notably are available only at a few institutions in highly customized settings – can be as complex as the tools themselves, involving thousands of lines of code and requiring months to understand, apply and modify.

Although many universities now run experimental forecasts on a daily basis using public-domain software such as the Weather Research and Forecast (WRF) model (Michalakes et al. 2000), they do so in very simple configurations using mostly local computing facilities and pre-generated analyses to which no new data have been added. LEAD seeks to democratize the availability of advanced weather technologies for research and education, lowering the barrier to

---

* Corresponding Author Address: Prof. Kelvin K. Droegemeier, University of Oklahoma, 120 David L. Boren Boulevard, Suite 5900, Norman, OK 73072. Email: kkd@ou.edu. Visit LEAD on the web at http://leadproject.org.

entry, empowering application in a grid context, increasing the realism of how technologies are applied, and facilitating rapid understanding, experiment design and execution.

The second objective involves *improving our understanding of and ability to detect, analyze and predict mesoscale atmospheric phenomena by interacting with weather in a dynamically adaptive manner.* Most technologies used to observe the atmosphere, predict its evolution, and compute, transmit and store information about it operate not in a manner that accommodates the dynamic behavior of mesoscale weather, but rather as static, disconnected elements. Radars do not adaptively scan specific regions of storms, numerical models mostly are run on fixed time schedules in fixed configurations, and cyberinfrastructure does not allow meteorological tools to operate on-demand, change their mode in response to weather, or provide the fault tolerance needed for rapid reconfiguration. As a result, today's weather technology, and its use in research, operations and education, are far from optimal when applied to any particular situation (Droegemeier et al. 2005). To address these severe limitations, LEAD is

- Developing capabilities to allow models and other atmospheric tools to respond dynamically to their own output, to observations, and to user inputs so as to operate as effectively as possible in any given situation;

- Developing, in collaboration with the NSF Engineering Research Center for Collaborative Adaptive Sensing of the Atmosphere (CASA; Brotzge et al. 2006; Plale et al. 2006), capabilities to allow models and other atmospheric tools to dynamically task adaptive observing systems, with an emphasis on Doppler radars, to provide data when and where needed based upon the application, user or situation at hand;

- Developing appropriate adaptive capabilities within supporting IT infrastructure.

3. SYSTEM CAPABILITIES

LEAD comprises a complex array of services, applications, interfaces, and local and remote computing, networking and storage resources – so-called *environments* – that can be used in a stand-alone fashion or linked together in workflows to study mesoscale weather; thus the name "Linked Environments for Atmospheric Discovery." This framework provides users with an almost endless set of capabilities ranging from simply accessing data and perhaps visualizing it to running highly complex and linked data ingest, assimilation and forecast processes in real time and in a manner that adjusts dynamically to inputs as well as outputs. A brief overview of the LEAD service-oriented architecture (SOA) is presented in §4 and additional detail can be found in Droegemeier et al. (2005).

Figure 3.1 shows the logical structure of the LEAD environments. At the fundamental level of functionality, as shown by the top horizontal gray box, LEAD enables users to accomplish the following:

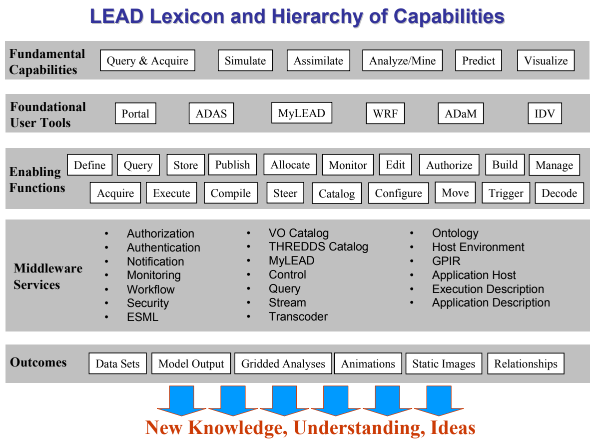


Figure 3.1. The fundamental capabilities (top gray box) and tangible outcomes (bottom gray box) of LEAD are enabled by a rich fabric of tools, functions and middleware services that represent the LEAD research domain.

- *Query for and Acquire* a wide variety of information including but not limited to observational data (including real time streams) and gridded model output stored on local and remote servers, definitions of and interrelationships among meteorological quantities, the status of an IT resource or workflow, and education modules at a variety of grade levels that are designed specifically for LEAD.

- *Simulate and Predict* using numerical atmospheric models, particularly the WRF model system now being developed by a number of organizations. The WRF can be run in a variety of modes ranging from basic (e.g., single vertical profiles of temperature, wind and humidity in a horizontally homogeneous domain) to very complex (full physics, terrain, and inhomogeneous initial conditions in single forecast or ensemble mode). Other models (e.g., ocean) can be included but are not fundamentally part of the LEAD system now being created.

- *Assimilate* data by combining observations, under imposed dynamical constraints, with background information to create a 3D atmospheric gridded analysis. As noted in the tools description below, LEAD supports the ARPS Data Assimilation System (ADAS) and soon will incorporate the WRF 3D Variational (3DVAR) Data Assimilation System (Barker et al. 2005).

- *Analyze and Mine* observational data and model output to obtain quantitative information about spatio-temporal relationships among fields, processes, and features.

- *Visualize and Quantitatively Evaluate* observational data and model output in 1D, 2D and 3D frameworks using batch and interactive tools.

LEAD comprises a large number of tools ranging from simple services to highly sophisticated meteorological, data mining and visualization packages. Within this array we define a sub-set of foundational application or productivity tools that include:

- LEAD Portal (http://portal.leadproject.org), which serves as the primary though not exclusive user entry point into the LEAD environments;

- ARPS Data Assimilation System (ADAS; Brewster 1996), a sophisticated tool for data quality control and assimilation including preparation of model initial conditions;

- myLEAD (Plale et al. 2004), a flexible personalized data management tool that at its core is a metadata catalog. myLEAD stores metadata associated with data products generated and used in the course of scientific investigations and education activities.

- Weather Research and Forecast model (WRF; Michalakes et al. 2000), a next-generation atmospheric prediction and simulation model that runs on single or multiple processors at grid spacings ranging from meters to hundreds of kilometers;

- Algorithm Development and Mining (ADaM; Rushing et al. 2005), a powerful suite of tools for mining observational data, assimilated data sets and model output; and

- Integrated Data Viewer (IDV; Murray et al. 2003), a widely used desktop application for visualizing, in an integrated manner, a broad array of multi-dimensional geophysical data.

The power of LEAD lies not only in the capabilities of its various tools but more importantly in the manner in which they can be linked together to solve a broad array of problems, as shown schematically in Figure 3.2. The tangible outcomes (bottom bar in Figure 3.1) include data sets, model output, gridded analyses, animations, static images, and a wide variety of relationships and other information that leads to new knowledge, understanding and ideas. The fabric in Figure 3.1 that links the top set of requirements with the bottom set of outcomes – namely, the extensive middleware, tool and service capabilities – is the research domain of LEAD.

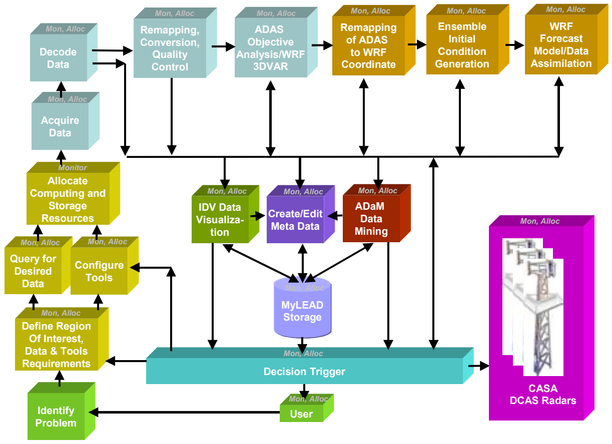


Figure 3.2. Conceptual/functional linkages among components of LEAD. Virtually any mesoscale research or educational problem can be mapped onto this figure.

4. SYSTEM ARCHITECTURE

In the language of computer science, a service is an entity that carries out a specific operation, or a set of operations, based upon requests from clients, e.g., booking airline flights or looking up the address of a friend. Web services are networked services that conform to a family of standards that specify most aspects of a service's behavior and have been developed by a number of organizations. The LEAD architecture is a "Service Oriented Architecture" (SOA), which refers to a design pattern based upon organizing all of the key functions of an enterprise or system as a set of services. The work of the enterprise or system is carried out by *workflows* that orchestrate collections of service invocations and responses to accomplish a specific task. SOAs are being deployed widely in the commercial sector and form the foundation of many scientific “grid” technologies.

As shown in Figure 4.1, the LEAD SOA is realized as five distinct yet highly interconnected layers (detailed descriptions of each service are provided in Droegemeier et al. 2005). The bottom layer represents raw resources consisting of computation as well as application and data resources, e.g., on the TeraGrid. At the next level up are web services that provide access to “raw/basic” capabilities and services for accessing weather data. LEAD is leveraging these resources from other projects and modifying them as appropriate. A wide variety of configuration and execution services compose the next layer and represent services invoked by LEAD workflows. They are divided into four principal groups, the first being the application and configuration service that manages the deployment and execution of fundamental user applications such as the WRF model, ADAS data assimilation system, and ADaM data mining tools.

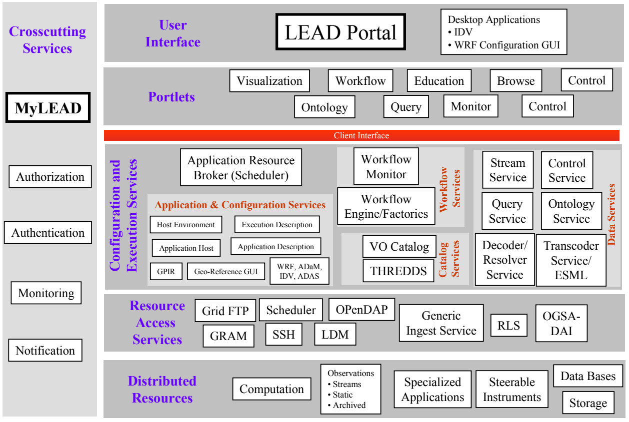


Figure 4.1. The LEAD service-oriented architecture.

For each of these, additional services are needed to track deployment and execution environment requirements to enable dynamic staging and execution on any of the available host systems. A closely related service is the application resource broker, which is responsible for matching the appropriate host for execution to each application task based upon time constraints of the execution and other factors. Both of these services are invoked by workflow services, which drive experimental workflow instances. Catalog services control the manner in which a user discovers data for use in experiments via a virtual organization (VO) catalog. Finally a host of data services are used to search for and apply transformations to data products. An ontology service resolves higher-level atmospheric concepts to specific naming schemes used in the various data services, and decoder and interchange services transform data from one form to another. Stream services manage live data streams such as those generated by the NEXRAD Doppler radar network.

Several services are used within all layers of the SOA and are referred to as crosscutting services, indicated in the left column of Figure 4.1. One such service is the notification service, which lies at the heart of both static and dynamic workflow orchestration. Each service is able to publish notifications and any service or client can subscribe to receive them. Another critical component is the monitoring service, which provides, among other things, mechanisms to ensure that desired tasks are completed by the specified deadline – an especially important issue in weather research and education.

A vital crosscutting service that ties multiple components together is the user metadata catalog known as myLEAD. As an experiment runs, it generates data that are stored on the LEAD Grid or elsewhere (e.g., TeraGrid) and cataloged to the user’s myLEAD catalog. Notification messages generated during the course of workflow execution also are written to metadata and stored on behalf of a user. A user accesses metadata about the products used during or generated by an investigation through a set of metadata catalog-specific user interfaces built into the LEAD Portal. Note that users can edit metadata, and that LEAD has developed a specific schema based upon existing standards. Through these interfaces the user can browse holdings, search for products based on rich meteorological search criteria, publish products to broader groups or to the public, snapshot an experiment for archiving, or upload text or notes to augment the experiment holdings. Authentication and authorization are handled by specialized services based upon grid standards.

Finally, at the top level of the architecture in Figure 4.1 is the user interface, which consists of the LEAD portal (see §7) and a collection of “service-aware” desktop tools . The portal is a container for user interfaces, called portlets, which provide access to individual services. When a user logs into the portal, his or her grid authentication and authorization credentials are loaded automatically. Each portlet can use these certificates to access individual services on behalf of the user, thus allowing users to command the portal to serve as his or her proxy for composing and executing workflows on back-end resources.

Alternatively, users may access services by means of desktop tools. For example, the Integrated Data Viewer (IDV) can access and visualize data, and provide domain sub-setting capability, using a variety of sources including OPeNDAP servers. Similarly, the workflow composer tool can be used to design a workflow on the desktop that can be uploaded to the user's myLEAD space for later execution, as at the Unidata Users Workshop where LEAD recently was demonstrated as a tool for modeling in education (Clark et al. 2007).

The SOA architecture is especially attractive from the user point of view because it allows for loose coupling of capabilities represented by the services. Within this framework, other types of services may be invoked such as machine-specific desktop applications (e.g., high-end visualization tools) that typically require software installation.

5. INTEGRATIVE TEST BEDS AND STRATEGIC TIMELINE

From project inception, the LEAD team faced the challenge of striking an appropriate balance between software developed to explore fundamental concepts in meteorology and computer sicence and the instantiation of this software as stable, persistent capabilities in well engineered systems deployed for use by the broader community. Although resource constraints prevent full attention from being given to both research and deployment, LEAD has created conceptual and practical frameworks for achieving what is believed to be an appropriate balance.

As shown in Figure 5.1, which was inspired by D. McLaughlin, the director of CASA, LEAD research begins with fundamental ideas (oval in the lower center of the diagram) that lead to basic research and an overarching system architecture. The resulting software components, most of which are developed piecemeal, then are integrated to provide increasingly greater functionality as part of so-called integrative test beds (ITBs). A vitally important part of LEAD, ITBs are experimental frameworks for instantiating LEAD system components in a coordinated manner to evaluate fundamental concepts in an end-to-end fashion using selected end user testers. The former ensures an integrated systems-orientation to testing while the latter allows LEAD to obtain very specific user feedback on system design, capabilities, performance, etc. The integrative test beds spawn new ideas, from which new capabilities are developed and tested in a cyclic manner (Figure 5.1).

Note that most of the end user testers have been identified and educated about LEAD by the Education and Outreach Thrust. They include teachers in grades 6-12, students in grades 6 through graduate school, researchers in meteorology and computer science, non-research faculty, attendees at workshops (e.g., WRF Community Workshop, Unidata Users Workshop), and other special groups (e.g., participants in the NOAA/NCAR Developmental Test Bed Center). Efforts now are being extended to expose LEAD to other communities (e.g., oceanography), though judiciously owing to limited resources.

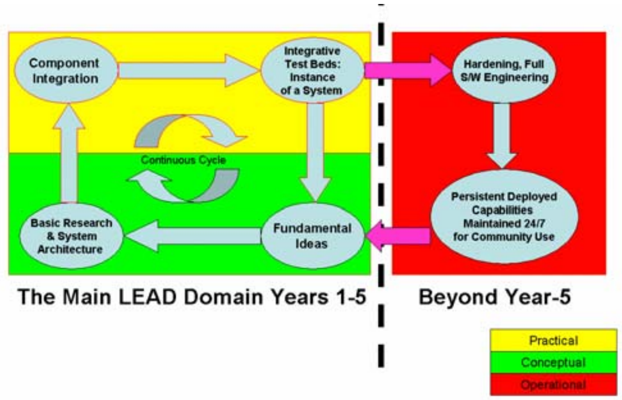


Figure 5.1. The LEAD research process and role of integrative test beds as an instantiation of end-to-end systems concepts evaluated by end user testers. Concept from D. McLaughlin, Director of CASA.

Figure 5.2 shows the LEAD strategic timeline, which is founded upon three ITBs that sequentially build upon one another in a manner similar to radar test beds in CASA (Brotzge et al. 2006, 2007). Each is designed around a specific set of capabilities and each has a principal goal that addresses key research and education issues. As each ITB proceeds, early capabilities within them will become more stable and will be transitioned to the Unidata deployment, (see §9), while other capabilities are added.

6. DYNAMIC ADAPTATION

As noted in §1, the second goal of LEAD, in addition to democratization, is dynamic adaptivity to weather of meteorological tools, cyberinfrastructure and observing systems, particularly Doppler radars via CASA. This is the transformative research challenge upon which LEAD was founded and to meet it, LEAD has created a research agenda centered around

Figure 5.2. The LEAD strategic timeline built around three integrative test beds. Concept from D. McLaughlin, Director of CASA

several use case scenarios. It is important to recognize that adaptation can take many forms but in all cases *the objective of adaptive systems is to improve upon their static counterparts in some manner, ideally one that formally optimizes or at least quantitatively improves upon certain aspect(s) of performance*. Systems or components may adapt in time, space or modality and the adaptation can be automated, manual, objective, heuristic, etc. Further, adaptation can occur in a variety of locations within the system (i.e., within the LEAD environments), at multiple levels and in highly connected, nonlinear ways.

Given this complexity, LEAD has framed its associated research agenda by key issues and questions in adaptivity that implicitly include concepts of streaming and on-demand functionality. The list below is not intended to be exhaustive but rather representative of the most compelling issues relevant to and being addressed by LEAD:

- When is adaptation useful?
- Can the costs and benefits of adaptation be quantified?
- What types of adaptation are possible and most effective and how can they be chosen and combined?
- How is adaptivity triggered/controlled?
- What elements of the system can or should adapt (application, workflow, cyberinfrastructure, observing systems, combinations of these)?
- How can one deal with loss of resources or less than ideal availability to achieve the required adaptation?
- What metrics can be used to measure the effectiveness of adaptation?
- What negative consequences exist to adaptation?
- Can "optimal" adaptivity be defined?
- What are the time scales of adaptivity and what controls them?
- Do adaptivity and on-demand functionality need to be pre-scheduled to any extent?
- What triggers the decision to adapt and how is the decision communicated across the system?
- How can applications most effectively be maintained in "stand-by" mode, ready to be invoked by an adaptive trigger?
- What does quality of service mean in an adaptive system?

Initial evidence for the value of adaptation was presented by Droegemeier et al. (2005). LEAD is now studying dynamic adaptivity in a variety of ways, ranging from simple one-dimensional models to real-data simulations using ensemble Kalman filtering ina full-physics prediction system.

A canonical example of adaptation in the latter is shown in Figure 6.1, where streaming observations, say from NEXRAD (Plale et al. 2007) or CASA radars (yellow left-pointing arrow), are mined for specific features by a persistent data mining agent (Graves et al. 2007). If an event is detected, the system automatically launches a WRF forecast (far right), with a brokering service seeking to obtain the required resources on the TeraGrid. In some cases, the forecast would need an entire machine and would have to begin immediately in order to retain value. As the forecast output is produced, it is directed to the same data mining engine (dark green arrows) to identify specific features that, if found, might be used to retask the CASA radars (aqua arrows). The process then repeats in a closed-loop fashion. LEAD is experimenting with individual components of this capabilitiy and plans to have a fully dynamic system *for experimentation* in late 2007.

It is interesting that dynamic adaptation may in some cases be viewed as less effective than its static counterpart. For example, in operational forecasting, the evaluation of a model operated daily in the same configruation allows forecasters to identify idosyncracies that can be factored into analysis of the output. If, as is the case in a dynamically adaptive system, the forecast configuration changes frequently in response to changing weather, the skill obtained through repeated examination of a static framework would be diminished or perhaps lost. This issue will be addressed in upcoming experiments with NOAA (§9).

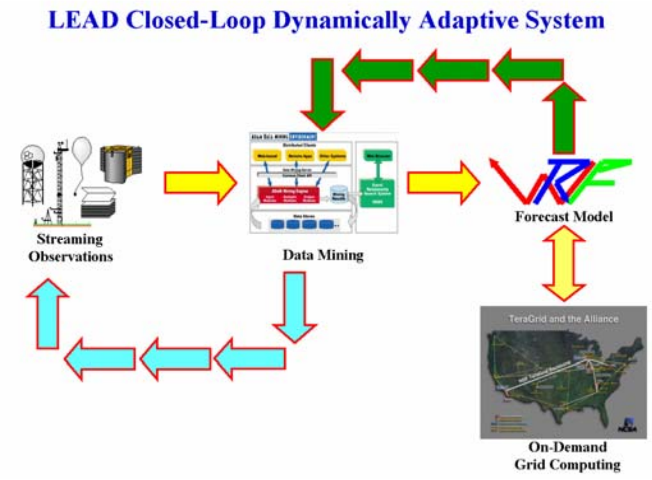


Figure 6.1. The closed-loop dynamically adaptive LEAD system, in which streaming data, workflows, tools and cyberinfrastructure mutually interact with mesoscale weather.

7. USER INTERFACES

As described more fully in Gannon et al. (2007), the principal user interface to LEAD is a portal – a customizable and dynamic web framework that supports grid and web services in a manner similar to commerical environments with which most are familiar, e.g., Amazon.com. As shown in Figure 7.1, the LEAD portal (http://portal.leadproject.org) contains customized pages for specific users (e.g., researchers, educators, students, visitors) and many of its capabilities can be accessed without obtaining an account. The portal provides links to the data query interface (Figure 7.2), education modules (Clark et al. 2007), visualization system and other resources as well as easy access to the most popular services, e.g., visualization, running a forecast or simulation, and accessing weather data. The interface is built upon Gridsphere (www.gridsphere.org), with custom graphics designed by the University of Michigan using human-computer interface principles to ensure that all concepts are intuitive and easy to implement.

For example, the data query system (Figure 7.2) allows users to specify temporal and spatial boundaries of data, the data sets or fields desired, and other attributes. Glossaries and ontology catalogs are available to maximize the richness with which queries can be executed. Upon submission, the request proceeds through the data system (Plale et al., 2004; Wilson et al. 2007) and returns the location of the desired information – which then can be placed in the user's myLEAD space or downloaded to the desktop. At the present time, approximately eight of the most widely used data sets are being incorporated (NEXRAD Level II and III, METAR, RAOBS, GOES Visible and Infrared, NCEP NAM forecasts, and persistent ADAS analyses generated by CAPS). LEAD envisions expanding the data system to more than 20 data sets, with the latest 6 months of data available online.

8. ORCHESTRATION AND MONITORING

As described more fully in Ramrkrishnan et al. (2007) and Ramachandran (2007), LEAD uses graphical tools for composing, compiling, and monitoring workflows. Figure 8.1 shows the workflow composer in event-monitoring mode for an experiment in which the WRF model is initialzed using a NAM analysis and run on the TeraGrid. Each box represents a service in the workflow and the lines connecting them show communication links. In general, a service ingests information as well as delivers it (unless it's the first or last segment in the chain). In the example shown, the WRF model service is depicted in the upper center of the figure. The colors of the boxes indicate the execution status of each service: those colored black have completed, those colored green are in progress, those colored yellow have not yet begun, and services suffering a fault are shown in red. Details of the execution process are provided in the dialog box at the bottom of the figure.

Although users can compose workflows, LEAD maintains a repository that eventually will contain hundreds or perhaps thousands of workflows. For example, workflows already exist for running an ensemble of forecasts, creating an analysis over a specific region of the country, and ingesting and displaying radar data. Further, virtually any application can be cast as a service via the Service Wrapper Toolkit provided within the LEAD environments. One of the first such applications so modified was the WRF model. Efforts now are underway to integrate the WRF graphical configuration tool (Smith et al. 2007) in LEAD, though the capability to edit WRF input files in LEAD already has been developed (Ramachandran et al. 2007).

LEAD also has been experimenting with a desktop client service for orchestrating large numbers of numerical simulations or forecasts, particularly for ensembles. Known as Siege (Alameda et al. 2007), it can be used with its own interface or as a service within the LEAD portal framework.

Linked Environments for Atmospheric Discovery (LEAD) makes meteorological data, forecast models, and analysis and visualization tools available to anyone who wants to interactively explore the weather as it evolves. The LEAD Portal brings together all the necessary resources at one convenient access point ... read more

Researchers	With university, government, or industry affiliations
Educators	At college and university level, high school, or middle schools
Students	At graduate, undergraduate, middle and high school levels
Visitors	Newcomers and the curious

COPYRIGHT © 2006 LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY. ALL RIGHTS RESERVED.

Figure 7.1. The LEAD Portal home page at http://portal.leadproject.org.

Data Products:

☑ All
☐ Surface METAR Obs
☐ UpperAir Balloon Soundings (Rawinsondes)
☐ NEXRAD Level II
☐ NEXRAD Level III
☐ GOES Imagery
☐ NCEP NAM Forecast
☐ ADAS
☐ WRF

Data Categories: ☑ Any

☐ Temperature
☐ Wind
☐ Pressure
☐ Moisture

eg: -37.68

	Layer
☑	Search Area
☐	Metar (labeled with site name)
☐	Metar (labeled with temperature)
☐	Radar Sites
☑	Pop. Places

☑ Any temporal range

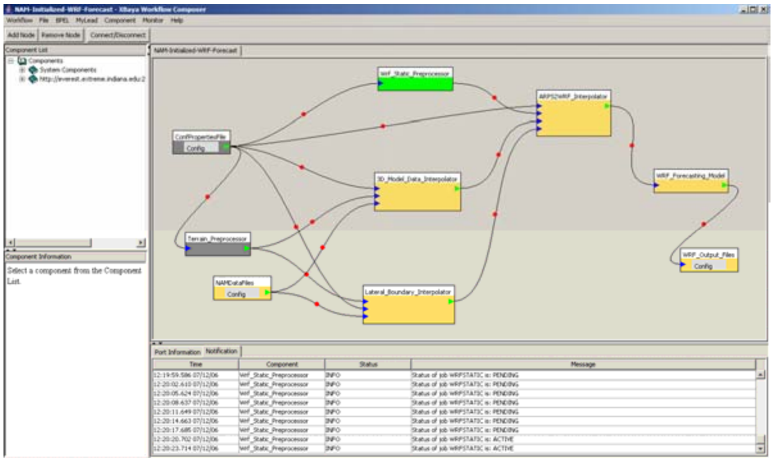
yyyy-mm-ddThh:mm:ss
eg: 2005-11-15T08:30:56

Begin Time: 2006-11-4T0:0:0

End Time: 2006-11-4T24:0:0

COPYRIGHT © 2006 LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY. ALL RIGHTS RESERVED.

Figure 7.2. The LEAD geographical data query tool.



Select a component from the Component List.

Time	Component	Status	Message
12:19:59.586 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:02.610 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:05.624 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:08.637 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:11.649 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:14.663 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:17.685 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: PENDING
12:20:20.702 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: ACTIVE
12:20:23.714 07/12/06	Wrf_Static_Preprocessor	INFO	Status of job WRFSTATIC is: ACTIVE

Figure 8.1. The LEAD Workflow Composer in event monitoring mode, shown here for the workflow used at the Unidata Users Workshop.

9. BRINGING LEAD TO THE COMMUNITY

As noted in §5, LEAD utilizes integrative test beds to develop and evaluate capablities. Once certain of these have become sufficiently stable, they are moved into the UCAR Unidata Program for broad deployment to the community. Unidata reaches approximately 150 organizations encompassing 21,000 university students, 1800 faculty, and hundreds of operational practitioners.

The initial step in exposing LEAD to a broad group of potential users occurred during the July 2006, Unidata tri-annual User's Workshop in Boulder, Colorado. The theme was "Expanding the Use of Models as Educational Tools in the Atmospheric & Related Sciences" (see http://www.unidata.ucar.edu/community/2006workshop/). The workshop afforded a unique opportunity to unveil the rapidly maturing capabilities of LEAD to a select group of university researchers and faculty who are particularly interested in running forecast models and applying related tools in the classroom (Baltzer et al. 2007).

Thursday, 13 July was devoted entirely to LEAD with the goals of a) introducing LEAD to a broad array of users via hands-on experimentation; b) obtaining both quantitative and qualitative feedback to inform future LEAD research and education programs; c) conducting the first scalability and stability tests of the entire LEAD system, including in particular the ability of the TeraGrid to handle the simultaneous submission of dozens of WRF jobs; and d) expanding the LEAD user tester community and establishing new collaborations.

Eighty-two faculty and staff from the Unidata community attended the workshop, and the more than 50 individuals participating in the LEAD laboratory session (split into two groups of approximately 25) were provided with the capability to launch WRF forecasts on TeraGrid resources at the National Center for Supercomputing Applications (NCSA), University of Illinois (Figure 9.1).

A second potential venue for bringing LEAD to the community is the Developmental Test Bed Center (DTC; http://www.dtcenter.org) at the National Center for Atmospheric Research. Sponsored by the NSF and NOAA, the DTC provides a national collaborative framework in which numerical weather analysis and prediction communities can interact to accelerate testing and development of new technologies as well as techniques for research applications and operational implementation – all in a way that mimics, but does not interfere with, actual forecast operations. It is anticipated that the DTC will become the focal point for mesoscale model experimentation and the transfer of new concepts and technologies into operational practice. LEAD is exploring ways in which it can assist the DTC in achieving its mission.

Figure 9.1. LEAD PI Anne Wilson of Unidata answers a user question during the LEAD laboratory session of the Unidata Users Workshop while Tom Baltzer of Unidata (extreme right) looks on.

Third, LEAD will be used in the spring, 2007 Spring Program conducted jointly by the NOAA Storm Prediction Center and National Severe Storms Laboratory (e.g., Kain et al. 2005). Plans consist of running a daily 30-hour, 5- or 10-member CONUS WRF ensemble at 4 km grid spacing, automatically launching 2 km grid spacing forecasts on-demand over tornado watch regions, and supporting one forecaster-initiated, on-demand 2 km grid spacing forecast per day.

The fourth venue of bringing LEAD to the community exploits a strategy for empowering new generations of scientists: engaging them during their student carrers in the process of fundamental change. Toward this end, LEAD plans to make available to 10 U.S. institutions certain WRF capabilities for use in the national collegiate weather forecast contest, now known as Weather Challenge (http://wxchallenge.com). These institutions will form a pilot program during the spring 2007 forecasting period, the ultimate goals of which are to a) use LEAD as a hands-on mechanism for educating students about numerical weather analysis and prediction, b) understand the cognitive processes by which students will choose when, where and how to make their own numerical forecasts; c) evaluate the impact of WRF output on student forecast skill to determine how and whether such self-initiated, customized forecasts add value; and d) evaluate the LEAD framework, SOA methodology, and ability of the TeraGrid to accommodate on-demand forecasts. Funding for extending this pilot program to larger numbers of students and institutions over multiple years will be sought.

10. ACKNOWLEDGMENTS

LEAD is funded by the National Science Foundation under the following Cooperative Agreements: ATM-0331594 (University of Oklahoma), ATM-0331591 (Colorado State University), ATM-0331574 (Millersville University), ATM-0331480 (Indiana University), ATM-0331579 (University of Alabama in Huntsville), ATM-0331586 (Howard University), ATM-0331587 (UCAR), and ATM-0331578 (University of Illinois at Urbana-Champaign with a sub-contract to the University of North Carolina). CASA is funded in part by the Engineering Research Centers Program of the National Science Foundation under NSF Cooperative Agreement EEC-0313747 to the University of Massachusetts-Amherst.

11. REFERENCES

Alameda, J. and Co-Authors, 2007: Siege: A graphical user interface to enable management of large numbers of weather simulations. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Baltzer, T. and Co-Authors, 2007: LEAD at the Unidata workshop: Demonstrating the democratization of NWP capabilities. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Barker D. M., Y.-R. Guo, W. Huang, H. Huang, S. Rizvi, Q. Xiao, and M.-S. Lee, 2005: WRF-VARA Unified 4/4D-Var Data Assimilation system for WRF. Preprints. 6th WRF/MM5 Users Workshop. NCAR. Boulder, CO.

Brewster K., 1996: Application of a Bratseth analysis scheme including Doppler radar data. Preprints, *15th Conf. Wea. Anal. Forecasting*., Amer. Meteor. Soc., Norfolk, VA, 92–95.

Brotzge, J., K.K. Droegemeier, and D.J. McLaughlin, 2006: Collaborative Adaptive Sensing of the Atmosphere (CASA): New radar system for improving analysis and forecasting of surface weather conditions. *J. Transport. Res. Board*, No. 1948, 145-151.

Brotzge, J. and Co-Authors, 2007: CASA IP1: Network operations and initial data. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Clark, R.D., S. Yalda, D. Gannon, B. Plale and T. Baltzer, 2007: Integrating LEAD research in undergraduate education. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Droegemeier, K.K. and Co-Authors, 2005: Service-oriented environments in research and education for dynamically interacting with mesoscale weather. *Computing in Science and Engineering*, **7**, 12-29.

Gannon, D., and Co-Authors, 2007: The LEAD science portal problem solving environment. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Graves, S. and Co-Authors, 2007: Deployable suite of data mining web services for online science data repositories. Preprints, 23rd *Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Kain, J. S., S. J. Weiss, M. E. Baldwin, G. W. Carbin, D. R. Bright, J. J. Levit, and J. A. Hart, 2005: Evaluating high-resolution configurations of the WRF model that are used to forecast severe convective weather: The 2005 SPC/NSSL Spring Program. Preprints*, 21th Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction*, Washington, D. C., Amer. Meteor. Soc., CD-ROM, 2A.5.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2000: Development of a next-generation regional weather research and forecast model. Proceedings, Ninth ECMWF Workshop on the use of Parallel Processors in Meteorology, Reading, U.K., November 13-16, 2000. Argonne National Laboratory preprint ANL/MCS-P868-0101.

Murray, D., J. McWhirter, S. Wier, and S. Emmerson, 2003: The Integrated Data Viewer—a web-enabled application for scientific analysis and visualization. Preprints, *19th Conf. On Integrated Information and Processing*, 8-13 February, Long Beach, CA, Amer. Meteor. Soc.

Plale, B., J. Alameda, R. Wilhelmson, D. Gannon, S. Hampton, A. Rossi, and K.K. Droegemeier, 2004: User-oriented active management of scientific data with my LEAD. *IEEE Internet Computing*, **9**, 27-34.

Plale, B., D. Gannon, J. Brotzge, K.K. Droegemeier and Co-Authors, 2006: CASA and LEAD: Adaptive cyberinfrastructure for real-time multiscale weather forecasting. *IEEE Computer*, **39**, 66-74.

Plale, B. and Co-Authors, 2007: Real time filtering and mining of NEXRAD streams for mesoscale forecast and prediction. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Ramachandran, R. and Co-Authors, 2007: A Java-based graphical user interface for numerical simulation and analysis. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Ramakrishnan, L. and D. Reed, 2007: Monitoring and orchestrating an mesoscale meteorological cyberinfrastructure. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Rushing, J., R. Ramachandran, U. Nair, S. Graves, R. Welch, and A. Lin, 2005: ADaM: A Data Mining Toolkit for Scientists and Engineers. *Computers & Geosciences,* **31,** 607-618.

Smith, J.S. and Co-Authors, 2007: WRF portal and domain wizard: Graphical tools for WRF. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

Wilson, A., T. Baltzer and J. Caron, 2007: The THREDDS data repository (TDR) for long term data storage and access. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.