5B.1

LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY (LEAD): WEB SERVICES FOR METEOROLOGICAL RESEARCH AND EDUCATION (http://portal.leadproject.org)

[*1,2]Kelvin K. Droegemeier, [9]Jay Alameda, [7]Legand Burge, [1]Keith Brewster, [3]V. Chandrasekar, [5]Marcus Christie, [4]Richard Clark, [8]Ben Domenico, [5]Dennis Gannon, [6]Sara Graves, [7]Everette Joseph, [5]Suresh Marru, [5]Beth Plale, [6]Rahul Ramachandran, [8]Mohan Ramamurthy, [10]Daniel Reed, [6]John Rushing, [9]Al Rossi, [6]Steve Tanner, [1]Kevin W. Thomas, [1]Daniel Weber, [9]Robert Wilhelmson, [8]Anne Wilson, [1,2]Ming Xue and [4]Sepideh Yalda

[1]Center for Analysis and Prediction of Storms and [2]School of Meteorology
University of Oklahoma
Norman, Oklahoma

[3]Colorado State University
Fort Collins, Colorado

[4]Millersville University
Millersville, Pennsylvania

[5]Indiana University
Bloomington, Indiana

[6]University of Alabama in Huntsville
Huntsville, Alabama

[7]Howard University
Washington, DC

[8]University Corporation for Atmospheric Research
Boulder, Colorado

[9]National Computational Science Alliance
Urbana, Illinois

[10]North Carolina State University
Chapel Hill, North Carolina

1. INTRODUCTION

Linked Environments for Atmospheric Discovery (LEAD; Droegemeier et al. 2005, 2007) is a National Science Foundation Large Information Technology Research (ITR) grant, now in its fifth of five years, that has pioneered new approaches for integrating complex weather data, assimilation, modeling, mining, and cyberinfrastructure systems in innovative ways to empower researchers and students with capabilities heretofore available at only a few major universities and research or operational centers around the world. LEAD brings these capabilities – using a service-oriented architecture and other relevant technologies as the underpinning – to users with the simplicity of familiar environments such as Amazon.com. By managing the complexity of inter-operative cyber tools and providing flexibility and ease in how they can be linked, LEAD allows students and researchers to focus their time on solving the science and engineering problems at hand, providing a means for more deeply understanding the tools and techniques being applied rather than the nuances of data formats, communication protocols, and job execution environments.

Foundational to LEAD is the notion that today's static environments for observing, predicting, and educating about mesoscale weather are fundamentally inconsistent with the manner in which such weather actually occurs, namely, with often unpredictably rapid onset and evolution, heterogeneity, and spatial and temporal intermittency. To mitigate this inconsistency, LEAD has created an integrated, scalable framework in which meteorological analysis tools, forecast models, and data repositories can operate as dynamically adaptive, on-demand, grid-enabled systems, thus allowing them to 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

[*] 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.

observing technologies to optimize data collection for the problem at hand.

The service-oriented architecture developed by LEAD now operates around the clock to support students and researchers nationwide. In a bold move to broaden the exposure of LEAD and assess its strengths and weaknesses as well as those of back-end services upon which it currently relies (e.g., the NSF TeraGrid), LEAD undertook several major projects in 2007 and implemented numerous additional capabilities. We describe them more fully in the sections that follow and discuss evolving plans to deploy LEAD as a community facility.

2. RECENT SYSTEM ENHANCEMENTS

Recent research and development in LEAD have focused heavily on improving system robustness via fault tolerance and greater service reliability, broadening applicability via enhancements to key services and the inclusion of new ones, and developing a better understanding of and expanding capabilities in dynamic adaptation.

More specifically for 2007, the LEAD metadata schema (Plale et al. 2004, 2007; Simmhan et al. 2006) was enhanced and fully integrated into LEAD, thereby allowing for fully automated metadata generation, including all products generated by workflows. The geospatial data query interface was enhanced and users now can search their own personal workspace (myLEAD) for any products created or imported. In order to facilitate understanding, particularly by students, of the complexities of meteorological terminology, the Noesis ontology tool (Ramachandran et al. 2006) has been fully integrated as well. Additional data format and processing capabilities were added and enhancements made to facilitate crosswalks among various repositories and catalogs. In addition to many real time data sets now available within LEAD, the data system now contains online archives of certain products dating back several hundred days (Baltzer et al., 2008; Wilson et al., 2007, 2008).

The LEAD portal interface (Gannon et el., 2007; Brewster et al. 2008; http://portal.leadproject.org) continues to be enhanced, and Java-based animation tutorials now are in place to lower the entry barrier, particularly for students. Additional enhancements have been added to facilitate greater flexibility in workflows, particularly with regard to parameter types. The workflow engine has been hardened and the NCSA Ensemble Broker (Alameda et al. 2007) has been linked to myLEAD, thus allowing users to generate hundreds of forecasts (ensembles or otherwise) and manage all of their results using LEAD data system tools. In fact, more than 1000 forecasts were so generated by NCSA during the spring, 2007 severe weather season (see §3.2 and Wilhelmson et al. 2008).

New monitoring and fault tolerance/recovery services have been added to LEAD and initial results are very encouraging. Efforts now are being directed toward approximating performance behavior to estimate required resources, and using an adaptive scheduler that can monitor the execution stage of each workflow component. Several experiments have been conducted in which workflows are launched automatically based upon events detected in streaming data (Plale et al. 2007), and this has been made possible in part by advanced data mining and clustering techniques developed for application to NEXRAD Doppler radar data (Li et al., 2008). Research in dynamic adaptation using simple meteorological models, and sophisticated data assimilation systems, is providing guidance that will be used to determine how models might best be configured to produce an "optimal" forecast. This includes the ability to estimate forecast sensitivity and request additional observations so as to reduce the impact of observation errors. These capabilities will be used to help achieve a fully closed-loop dynamically adaptive capability.

3. APPLICATION OF LEAD IN THE 2007 NOAA HAZARDOUS WEATHER TEST BED

3.1. Pre-Scheduled Deterministic and Ensemble Forecasts

For more than a decade, the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma has collaborated with the NOAA Storm Prediction Center (SPC) and National Severe Storms Laboratory (NSSL) to study fine-scale atmospheric predictability via real time forecasts performed during the U.S. spring severe weather season.[1] During spring 2005, this work involved using the WRF model to create 2 km grid spacing forecasts over 2/3rds of the continental US, with initial conditions specified by the NCEP operational model analysis (Kain et al. 2005). The forecasts provided dramatic evidence that the predictability of organized deep convection is, in some cases, an order of magnitude longer (one day) than suggested by prevailing theories of

[1] In 1997, CAPS and the Pittsburgh Supercomputing Center Shared the *Computerworld*-Smithsonian Award for this activity.

atmospheric predictability. No experiment was conducted in 2006 owing to the move of CAPS, SPC and the NSSL to a new building.

In 2007, with additional funding from NOAA, LEAD applied some of its technology to the NOAA Hazardous Weather Test Bed, which is a multi-institutional program designed to study future analysis and prediction technologies in the context of daily operations. The 2007 effort sought to go well beyond previous capabilities by addressing two important LEAD-related challenges: (a) The use of storm-resolving ensembles for specifying uncertainty in model initial conditions and quantifying uncertainty in model output; and (b) The application of dynamically adaptive, on-demand forecasts that are created automatically, or by humans, in response to existing or anticipated atmospheric conditions. Specific goals including providing an initial assessment of the following:

- Quantitative skill of storm-resolving ensemble forecasts compared to their deterministic counterparts at similar (experimental) and coarser (operational) grid spacings.
- Predictability of deep convection and organized mesoscale convective systems.
- The extent to which dynamically adaptive prediction leads to quantitative forecast improvements, possible negative consequences of adaptation, and an evaluation of strategies for making decisions regarding when, where and how to adapt.
- The ability of the TeraGrid to accommodate both scheduled and on-demand applications that have strict quality of service requirements and utilize a substantial portion of available resources for an extended period of time.

The 2007 Spring Experiment (see also Xue et al., 2007; Kong et al., 2007a; Weiss et al., 2007) extended from 15 April through 8 June with all forecasts run on dedicated NSF TeraGrid resources at National Center for Supercomputing Applications (NCSA) and Pittsburgh Supercomputing Center (PSC). The forecast suite included the following:

- A 33-hour, 10-member, 2/3rds continental US-scale (CONUS) (Figure 1) ensemble at 4 km grid spacing (run at PSC using a mixture of initial condition and physics perturbations);
- A 33-hour, single 2 km grid spacing deterministic forecast (same domain);
- One or more six- to nine-hour nested grid forecasts at 2 km spacing launched automatically over regions of expected severe weather, as determined by mesoscale discussions or tornado watches (run at NCSA); and
- One six- to nine-hour nested grid forecast, per day, at 2 km grid spacing launched manually when and where deemed most appropriate (run at NCSA).

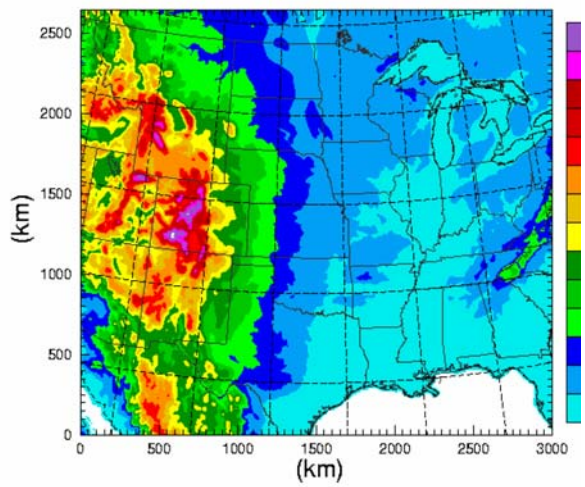

Figure 1. Domain of the 2 km grid spacing deterministic, and 4 km grid spacing ensemble forecasts of the 2007 Spring Experiment of the NOAA Hazardous Weather Test Bed. The domain covers approximately 2/3rds of the continental United States (CONUS) and color shading shows terrain height in meters.

Much of the technology used to run the ensemble forecasts already existed within CAPS and thus was utilized. CISCO Systems and OneNet, Oklahoma's state-wide ISP operated out of the Office of the State Regents for Higher Education, partnered to provide a dedicated single lambda from Norman, OK to PSC for the experiments. This link was used principally to handle the huge volumes of output produced by the 2 km, 2/3rds CONUS deterministic forecast.

An extremely important aspect of the Spring Experiment was that the daily forecasts were evaluated not only by operational forecasters in the NOAA SPC, but by dozens of faculty and researchers who visited the Hazardous Weather Test Bed in Norman, OK during the 7-week period. A formal procedure is employed by the SPC to evaluate the daily forecasts and additional details may be found in Kain et al. (2007).

To provide an example of the sorts of output generated by the WRF model in comparison to observations, Figure 2 shows NEXRAD radar composite reflectivity (maximum value across all altitudes within a given vertical column) at 1800 UTC on 24 May 2007 and Figure 3 shows the ensemble mean and spread of forecast reflectivity, the ensemble-derived probability of reflectivity exceeding 35 dBZ and a 'spaghetti' plot of 40 dBZ reflectivity contours, valid at 18 UTC on 24 May 2007. Because of the spatially discrete nature of the convection, the magnitude of the ensemble mean is not very meaningful (Kong et al. 2007b) – the ensemble mean will almost surely underestimate the intensity.

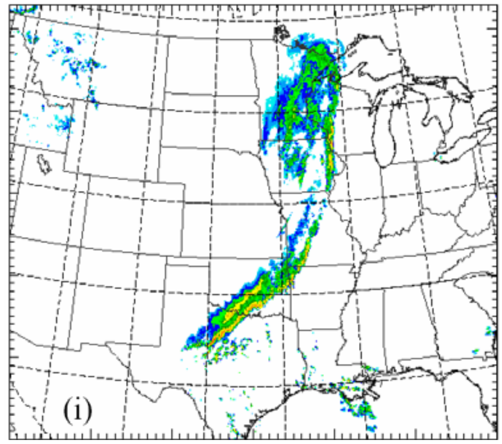

Figure 2. NEXRAD radar composite reflectivity (maximum value across all altitudes within a given vertical column) at 1800 UTC on 24 May 2007.

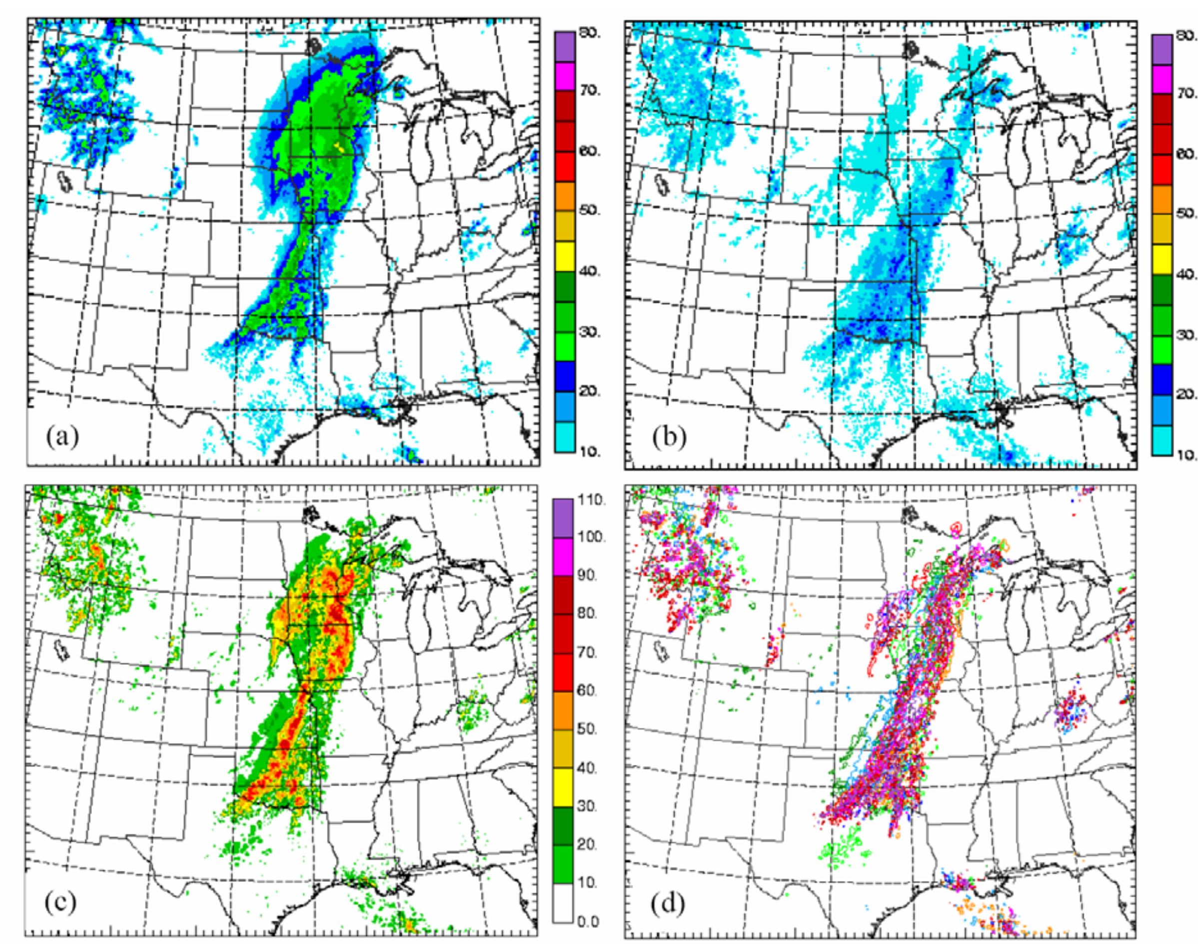

Figure 3. Forecast composite reflectivity ensemble mean (a) and spread (b), and ensemble-derived probability of composite reflectivity exceeding 35 dBZ (c) and the 'spaghetti' plot of 40 dBZ composite reflectivity contours, valid at 18 UTC, 24 May 2007, corresponding to a 21 hour forecast made with the WRF model as part of the 2007 Spring Experiment. From Xue et al. (2007).

The only region where the mean reflectivity exceeds 40 dBZ is southeast Minnesota (panel a), where the cold front consistently anchored the precipitation among most of the ensemble members (Xue et al., 2007). Maximum mean reflectivity remains in the 30-40 dBZ range along the line, and the positioning of the middle portion of this line is no better than that of the control or 2 km forecast.

Figure 3b shows that the spread of reflectivity has values between 20 and 25 dBZ along much of the line, indicating significant uncertainties at the convective scale. Figure 3c shows that the pattern of ≥50% probability of reflectivity exceeding 35 dBZ matches that of largest ensemble mean reflectivity very well, which therefore has a similar position error. The 'spaghetti' plot in Figure 3d suggests more spread in the position than the probability field and is therefore more indicative of position uncertainty.

3.2. Dynamically Adaptive Forecasts

As noted in the previous section, LEAD also conducted two-types of on-demand, dynamically adaptive forecasts: those launched at the discretion of forecasters and those launched automatically over regions of expected hazardous weather as determined by mesoscale discussions or severe weather watches. The former are described in Brewster et al. (2008), so we focus here on the latter.

Using its Ensemble Broker client (Alameda et al., 2007), the interface for which is known as Siege, NCSA developed a simple trigger capability which parsed Mesoscale Discussion and Severe Weather Watch information from the SPC, via an RSS feed, and instructed the ensemble broker to launch 6-hour WRF forecast workflows accordingly (Wilhelmson et al., 2008). Typically, 18-km singly nested and 2-km triply nested forecasts were triggered automatically using NAM data and the WRF processing package for initialization. This trigger service also generated identifiers which made deriving statistics from the end products relatively easy. The domain centers of all forecasts triggered automatically in this manner – over 1000 total – are shown in Figure 4, and quantitative evaluation of the results now is underway.

As noted previously, LEAD is an important research application (Science Gateway) for the NSF TeraGrid, and thus considerable effort was directed during the Hazardous Weather Test Bed to evaluate the stability and robustness of both the LEAD and TeraGrid infrastructures. TeraGrid reliability data for all WRF runs automatically triggered at NCSA are summarized in Table 1. The primary factors contributing to failures include machine hardware problems (Tungsten, in particular), gridFTP issues, GRAM issues, network issues, data interruption issues, as well as underlying issues with error detection in some of the initialization codes (e.g., failure to set an exit code, failure to check I/O status) as well as brittle paradigms for orchestration. Many of these issues have been addressed and recent tests have shown much greater degrees of reliability.

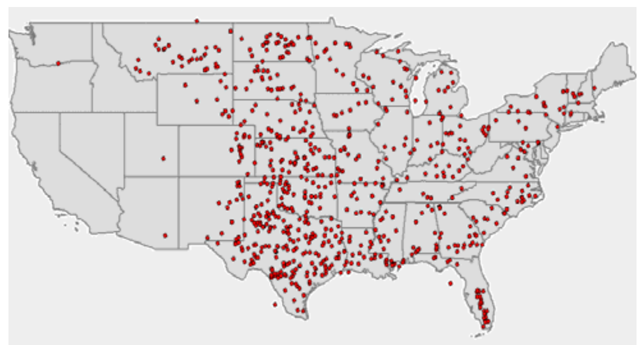

Figure 4. Domain centers of all nested grid WRF forecasts triggered automatically by NCSA based upon SPC Mesoscale Discussion and Severe Weather Watch information.

3.3. Plans for 2008

For the 2008 Hazardous Weather Test Bed, we will assimilate Level II data into the forecasts and attempt to bring the horizontal grid spacing of the 10-member ensembles down to 2 km and that of the single deterministic forecast down to 1 km. Additionally, we will emphasize study of the processes by which forecasters determine when and where to (manually) launch on-demand forecasts, and continue to study the tradeoffs of varying versus persistent model configurations.

4. APPLICATION OF LEAD IN WxCHALLENGE 2007

During the past few decades, the academic meteorology enterprise has supported a national collegiate weather forecast contest that seeks to engage both graduate and undergraduate students in practical forecasting under a variety of geographical and phenomenological circumstances. Known today as Weather Challenge (WxChallenge) and sponsored by the University of Oklahoma, each individual participant

Table 1. Summary statistics for WRF forecasts triggered automatically for the NOAA Spring 2007 Spring Experiment. MD indicates a trigger based upon the mesoscale discussion message and WW upon a weather watch. All forecasts were made on the Tungsten or Mercury systems at NCSA. For ADAS, ARPSPLT was not counted in the table as the multiple ARPSPLT runs were all part of a complete run. FAILED indicates a workflow was marked as failed. DONE indicates a workflow was successful. CANCELLED implies that the workflow was cancelled from the administrative interface (usually after either broker or event channel failure). Null refers to an event channel failure.

resolution	wf type	trigger type	by host:	FAILED		DONE		CANCELLED		null		total
2km	NAM-WPS-WRF	MD	tungsten	87	46.0%	91	48.1%	7	3.7%	4	2.1%	189
2km	NAM-WPS-WRF	MD	mercury	32	23.2%	94	68.1%	10	7.2%	2	1.4%	138
2km	NAM-WPS-WRF	WW	tungsten	34	49.3%	31	44.9%	2	2.9%	2	2.9%	69
2km	NAM-WPS-WRF	WW	mercury	12	23.1%	36	69.2%	3	5.8%	1	1.9%	52
18km	NAM-WPS-WRF	MD	tungsten	50	20.1%	194	77.9%	3	1.2%	2	0.8%	249
18km	NAM-WPS-WRF	MD	mercury	25	17.9%	111	79.3%	3	2.1%	1	0.7%	140
18km	NAM-WPS-WRF	WW	tungsten	19	18.4%	81	78.6%	3	2.9%	0	0.0%	103
18km	NAM-WPS-WRF	WW	mercury	9	17.6%	42	82.4%	0	0.0%	0	0.0%	51
20km	ADAS-WRF	MD	tungsten	26	68.4%	11	28.9%	1	2.6%	0	0.0%	38
20km	ADAS-WRF	WW	tungsten	14	73.7%	2	10.5%	3	15.8%	0	0.0%	19
2km	NAM-WPS-WRF	MD&WW	tungsten	121	46.9%	122	47.3%	9	3.5%	6	2.3%	258
2km	NAM-WPS-WRF	MD&WW	mercury	44	23.2%	130	68.4%	13	6.8%	3	1.6%	190
18km	NAM-WPS-WRF	MD&WW	tungsten	84	26.4%	225	70.8%	5	1.6%	4	1.3%	318
18km	NAM-WPS-WRF	MD&WW	mercury	37	19.3%	147	76.6%	6	3.1%	2	1.0%	192
20km	ADAS-WRF	MD&WW	tungsten	40	70.2%	13	22.8%	4	7.0%	0	0.0%	57
											total	1015

forecasts the maximum and minimum temperature, precipitation category, and maximum sustained wind speeds for selected U.S. cities. WxChallenge provides students an opportunity to compete against their peers and faculty mentors at other institutions (64 nationwide in 2006-2007), the prize being a trophy and bragging rights for a full year.

In spring 2007, 75 students and faculty from 10 institutions (7 non-LEAD and 3 LEAD; Table 2) that already were participating in WxChallenge, were invited to join a four week pilot program (two 2-week station forecasts plus a three week tournament extension) in which they were allowed to generate their own daily WRF-based forecasts from the LEAD portal and use them, in conjunction with standard products from the National Weather Service, to prepare their forecast (Clark et al. 2008). The pilot project was designed to make more broadly available the LEAD environments to a limited number of users for the purpose of evaluating system stability and reliability, ease of learning and use, the ability of the TeraGrid to accommodate dozens of on-demand forecasts, the potential benefits wrought by local models applied to local forecast problems, and the manner in which students chose to configure their forecast. Perhaps most importantly, WxChallenge placed very sophisticated technology under the control of students, thus providing a hands-on opportunity to learn about numerical modeling,

Participants were given the following: authorization to build experiments and compose workflows through the LEAD portal; sufficient computing resources to run WRF and save the WRF output in their myLEAD workspace; tools to visualize the WRF output; and users support. Computer time was provided by the TeraGrid via a community account. The users integrated WRF output products into their personal schema for preparing a forecast for stations previously selected by WxChallenge. Consequently, instead of students using traditional NWS model products or static models running locally, they could:

- Access the LEAD portal to establish a 5-km resolution nested WRF domain inside a prescribed 20 km CONUS
- Choose one of two options for model initialization: a) hourly gridded data from the ARPS Data Assimilation System (Univ. of Oklahoma) or, b) NAM gridded fields
- Choose one of two options for lateral boundaries (same as b)
- Launch a WRF run over the user-specified domain on non-local distributed computing resources
- Monitor the workflow
- Save the WRF output in the student's own myLEAD workspace
- If desirable, download the data onto their local computer
- Visualize the output using Unidata's IDV

- Use self-generated products to supplement the weather information currently used by students when preparing their WxChallenge forecast

During the seven week period of the pilot project, users from the 10 participating institutions launched a total of 279 workflows and generated 0.6 terabytes of output. Over 160 processors were reserved on the Tungsten system at NCSA five days each week from 10am to 8pm EDT. For the NAM-initialized WRF forecast, 78 percent of the workflows submitted completed successfully and 22 percent failed. The ADAS-initialized WRF forecast was less successful, with 36 percent completing as planned with 58 percent failing. The higher number of ADAS failures were associated with forecasting for Rapid City, SD and appear to be partly related to a missing terrain configuration and having ADAS components exit with the correct exit codes. We learned a great deal from this experience and the LEAD portal team has already begun to incorporate user suggestions (and bug discoveries) into subsequent releases.

Upon completion of WxChallenge, participants were asked to complete a survey, the results of which are being used to refine the production release of the portal and prepare for an expanded release to a much larger population of WxChallenge participants in fall 2008 (see Clark et al. 2008).

5. FUTURE SYSTEM ENHANCEMENTS

LEAD research and development during 2008 will focus principally on enhancements designed to increase the functionality and stability of the system and thus to enable broad, sustained use, maintainability and extensibility well beyond the NSF ITR grant, which ends in fall, 2008.

Foremost is the ability for users to edit namelist input files and incorporate them into workflows for key applications such as WRF and ADAS. This will greatly expand the ways in which key applications can be used, ranging, in the case of WRF, for example, from simple configurations (e.g., cloud model, dry boundary-layer model) to the very sophisticated (parametric studies involving multiple types of observations). Closely related is the capability for users to edit, compile and manage their own versions of application source codes (e.g., WRF) – a capability which now exists but has not yet been exposed in a general fashion. Also being developed is the capability for users to run "parametric workflows," in which selected parameters of the WRF (for example) can be varied systematically over a specified range within a single workflow (e.g., testing various values of computational mixing or user-specified parameters in physics parameterizations). These capabilities are essential for LEAD be of practical value to its primary audience of graduate and undergraduate student researchers.

Many capabilities tested experimentally during 2007 will be made available to users in 2008, including the service factory, which is used to convert applications into web services (such services, or those already available, can be used outside of LEAD as stand-alone applications or as services in other service-oriented architectures). Likewise, the ability for users to compose their own workflows will be exposed within the portal and additional, commonly used workflows will be added to the repository. Options for repeating workflows at specified times or time intervals, and pre-scheduling of workflows, will be made available, as will the important capability of terminating or restarting an executing workflow. At the present time, workflows execute on pre-defined resources; however, users will need the ability to specify a particular resource, or ask the system to broker the necessary resources based upon specific quality of service requirements. Some of this capability exists today but will be enhanced during 2008 to broaden the options made available. Fault tolerance, scheduling and monitoring will continue to be enhanced, and existing capabilities within the development system will be hardened and made available in the production version, particularly for ensembles.

One of the most important research components in 2008 involves developing and studying the behavior of workflows that are

- <u>Event-driven</u> (responding to weather or other triggers, especially real time CASA observations)
- <u>Priority-based</u> (execution urgency based upon user-assigned priority)
- Dynamically adaptive (change in response to the availability of resources, data)
- <u>Fault-tolerant</u> (ability to recover from application or resource failures while meeting the specified quality of service constraints)
- <u>Data-driven</u> (observations selected by the user and processed through ADAS or 3DVAR for analysis, visualization, or subsequent prediction using WRF)

Some of these (e.g., event-driven, dynamically-adaptive, fault-tolerant) already are being explored and special emphasis will be given to those driven by data. Toward that end, we plan to bring the

WRF 3DVAR package into LEAD as an option to ADAS for creating three-dimensional atmospheric analyses. Indeed, ADAS and the WRF 3DVAR system can be used sequentially (e.g., for single Doppler wind analysis), though such capability is beyond the scope of what can be accomplished during 2008. By adding the WRF 3DVAR system, we not only will enhance LEAD but make it much more attractive to the WRF community.

Finally in 2008, we will conduct a fully closed-loop, dynamically adaptive, on-demand, event-triggered WRF forecast workflow on the TeraGrid using real time dynamically adaptive radar data from the Center for Collaborative, Adaptive Sensing of the Atmosphere (CASA); Brotzge et al., 2006; Plale et al., 2006).

6. DEPLOYMENT OF LEAD AS A COMMUNITY RESOURCE

From the beginning, the LEAD vision has been to not only conduct excellent research and develop exciting and powerful technologies, but to do so in a practicable way that transforms meteorological research and education and has value to other disciplines. However, the transformation can occur only when LEAD is transitioned from a research project and made available as a persistent, stable facility upon which the community can rely. This notion was expressly stated in the original LEAD proposal, with Unidata as the envisioned home for community deployment:

> *"As a virtual extension of the user's desktop, and via deployment through Unidata – which involves approximately 150 organizations encompassing 21,000 university students, 1800 faculty, and hundreds of operational practitioners – LEAD will enable researchers, users, educators, and students to use atmospheric models and other tools in more realistic, real time settings than is now possible."*

The outcomes and impacts realized by LEAD, and those forthcoming in the final year of the ITR grant, 2008, provide a strong foundation upon which to build a persistent LEAD cyberinfrastructure. Based upon considerable interest in LEAD by the atmospheric science community, as measured, for example, by the strong positive feedback following the 2006 Unidata Workshop (Baltzer et al. 2007), LEAD proposes to seek funding for a formal deployment at Unidata. Our vision of the future is a primarily service-oriented environment of data streams, historical case study-type data sets, assimilation and modeling tools, mining and analysis engines, and visualization capabilities that are as pervasive in atmospheric, ocean and Earth science research and education as are desktop computers. Indeed, LEAD's extensibility makes it is ideally suited to application in regional climate, air quality, hydrology and oceanography.

The LEAD vision also involves building upon growing education and outreach programs (Clark et al., 2007) to extend the many LEAD resources into progressively lower grade levels and into communities for which even basic capabilities are unavailable. Because weather is experienced by every human and is an excellent motivating factor for studying science, our vision includes using LEAD to stimulate interest and broaden participation in STEM (science, technology, education, mathematics) education at the grade levels where most students choose, sometimes unwittingly, to avoid science as a career.

7. 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. Support for the NOAA Hazardous Weather Test Bed was provided by NOAA C-STAR Grant NA17RJ1227/NAGO000070029 to the University of Oklahoma.

8. 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.

Baltzer, T., and Co-Authors, 2008: The LEAD test bed system at the Unidata Program Center: a medium term archive of meteorological data. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, Amer. Meteor. Soc.

Brewster, K. and Co-Authors, 2008: Use of the LEAD portal for on-demand severe weather prediction. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, Amer. Meteor. Soc.

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.

Clark, R.D., S. Yalda, D. Gannon, B. Plale, T. Baltzer and E.C. Meyers, 2007: Integrating LEAD research in undergraduate education. Preprints, *23rd Int. Conf. on Interactive Information Processing Systems for Meteorology*, 14-18 January, San Antonio, TX, Amer. Meteor. Soc.

Clark, R. and Co-Authors, 2008: The LEAD-WxChallenge pilot project: Enabling the community. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, 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.

Droegemeier, K.K., and Co-Authors, 2007: A new paradigm for mesoscale meteorology: Grid and web service-oriented research and education in LEAD. Preprints, *23rd Conf. On Integrated Information and Processing*, 15-18 January, San Antonio, TX, Amer. Meteor. Soc.

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.

Kain, J.S., S.J. Weiss, M.E. Baldwin, K.K. Droegemeier, D. Bright, J.J. Levit, D. Weber and K.W. Thomas, 2005: How much resolution is enough? Comparing daily WRF ARW forecasts at 2 and 4 km grid spacing in severe convective weather environments during the 2005 SPC/NSSL Spring Program. Preprints, *11th Conf. on Mesoscale Processes*, Amer. Meteor. Soc., Albuquerque, NM.

Kain, J.S. and co-authors, 2007: Some practical considerations for the first generation of operational convection-allowing NWP: How much resolution is enough? Preprints, *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred*., Salt Lake City, Utah, Amer. Meteor. Soc.

Kong, F., M. Xue, Kelvin K. Droegemeier, D. Bright, M. C. Coniglio, K. W. Thomas, Y. Wang, D. Weber, J. S. Kain, S. J. Weiss, and J. Du, 2007a: Preliminary analysis on the real-time storm-scale ensemble forecasts produced as a part of the NOAA hazardous weather testbed 2007 spring experiment. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred.,* Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 3B.2.

Kong, F., K.K. Droegemeier and N.L. Hickmon, 2007b: Multiresolution ensemble forecasts of an observed tornadic thunderstorm system, Part II. *Mon. Wea. Rev.,* **135***, 759-782.*

Li, X., and Co-Authors, 2008: Storm clustering for data-driven weather forecasting. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, 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., S. Movva, P. Cherukuri, and S. Graves, 2006: Noesis: An Ontology-based Semantic Search Tool and Resource Aggregator. Geoinformatics Conference 2006, Reston, Virginia.

Simmhan, Y.L., Sangmi Lee Pallickara, Nithya N. Vijayakumar, and Beth Plale, 2006: Data Management in Dynamic Environment-driven Computational Science IFIP Working Conference on Grid-Based Problem Solving Environments (WoCo9) August 2006, to appear as Springer-Verlag Lecture Notes in Computer Science (LNCS).

Weiss, S. J., J. S. Kain, D. R. Bright, J. J. Levit, G. W. Carbin, M. E. Pyle, Z. I. Janjic, B. S. Ferrier, J. Du, M. L. Weisman, and M. Xue, 2007: The NOAA Hazardous Weather Testbed: Collaborative testing of ensemble and convection-allowing WRF models and subsequent transfer to operations at the Storm Prediction Center. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred*., Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 6B.4.

Wilhelmson, R.B.. and Co-Authors, 2008: Automatic triggering of high-resolution forecasts in response to severe weather indicators from the NOAA Storm Prediction Center. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, 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.

Wilson, A., and Co-Authors, 2008: Create an archive with the THREDDS data repository. Preprints, *24th Conf. On Integrated Information and Processing*, 21-24 January, New Orleans, LA, Amer. Meteor. Soc.

Xue, M., F. Kong, D. Weber, K. W. Thomas, Y. Wang, K. Brewster, K. K. Droegemeier, J. S. K. S. J. Weiss, D. R. Bright, M. S. Wandishin, M. C. Coniglio, and J. Du, 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed 2007 spring experiment. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred*., Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 3B.1.