5B.6

THE LEAD-WxCHALLENGE PILOT PROJECT: ENABLING THE COMMUNITY

Richard D. Clark
Millersville University, Millersville, PA

Suresh Marru, Marcus Christie, and Dennis Gannon
University of Indiana, Bloomington, IN

Brad Illston
WxChallenge, Norman, OK

Kelvin K. Droegemeier
University of Oklahoma, Norman, OK

Thomas Baltzer
Unidata Program Center, Boulder, CO

1. INTRODUCTION

During the past few decades, the academic meteorology enterprise has supported a national collegiate forecast contest that seeks to engage mostly undergraduate students with some graduate students and faculty in practical forecasting under a variety of geographical and phenomenological circumstances. Known today as The Weather Challenge (WxChallenge) and sponsored by the University of Oklahoma, each participant forecasts the maximum and minimum temperature, precipitation amount, and maximum sustained wind speed for select North American cities. WxChallenge provides students an opportunity to compete against their peers and faculty mentors at other institutions (64 nationwide in 2006-2007) for honors as the top weather forecaster in the nation.

2. PILOT PROJECT

In spring 2007 the NSF-funded ITR project, Linked Environments for Atmospheric Discovery (LEAD) engaged a manageable subset of the WxChallenge community and provided access to the LEAD Gateway portal and its underlying services. Since its inception LEAD has espoused the motivation to empower a community of users, and to provide the necessary infrastructure and tools to facilitate research and education.

The goal of the so-called "LEAD-WxChallenge Pilot Project" was to begin ushering in a fundamental paradigm shift in how experiments are conceived and performed, in the structure of user application tools and middleware, and in methodologies used to observe and model the atmosphere.

Students were provided with the ability to generate, run, analyze, and visualize their own WRF forecasts using the web-enabled service oriented architecture developed by LEAD researchers. For seven weeks, the 75 students and faculty from 10 institutions were given an unprecedented opportunity for enhancing their understanding of numerical modeling, high performance computing, physical parameterization schemes, data assimilation, and workflow orchestration. It also demonstrated one of LEAD's primary goals; to democratize and empower students by lowering the barrier for access to complex, integrated services, thereby allowing users the freedom to select inputs such as initialization fields, set model domains, and run WRF at a time and location determined by the user, and not constrained as in static, fixed and prescribed operational environments. In return, LEAD developers benefited from user feedback that exposed strengths and weaknesses, and provided a better sense of

the challenges and resource requirements associated with maintaining a reliable and persistent system aimed at enabling a larger community.

3. TECHNICAL NOTES

Participants in the pilot project were given authorization to the LEAD Gateway portal to access data, build experiments, and compose workflows with sufficient computing resources to run WRF and save the WRF output in myLEAD workspace. LEAD also provided tools for visualizing the output, as well as user support. Students were able to configure and run a 42-hour forecast using the WRF system at a horizontal resolution of 5 km. Students integrated output products generated by the WRF run into their personal decision-making schema for preparing a forecast for stations previously selected by WxChallenge.

Over the seven week pilot project participants launched a total of 279 forecast workflows and generated 0.6 TB of data. Over 160 processors were reserved on a multi-processor server five days each week from 10am to 8pm EDT. For the NAM initialized WRF forecast, 78 percent of the workflows submitted were successful, 22 percent failed. The ADAS initialed WRF forecast was less successful with 36 percent successful and 64 percent failing. When the WxChallenge ended and the pilot project concluded, the participants were asked to complete a survey, the results of which are being used to refine the portal and functionality and prepare for an expanded release to all institutions participating in the spring 2008 WxChallenge forecast competition.

Results of the survey were largely positive with participants finding the portal “easy and intuitive,” and the workflow and experiment builder a “great concept” and “highly intuitive.” Particularly noteworthy were comments on the workflow monitor, a feature that allows the user to follow the progress and status of the experiment. One participant reported that [the workflow monitor] was a “very powerful feature that allows the user to interact with the forecast he/she is creating.”

4. SPRING 2008 FORECAST COMPETITION

The spring 2007 pilot project gave LEAD developers the confidence that a stable, persistent, and reliable infrastructure could be sustained when scaled to a larger community of users. Thus, beginning in spring 2008, resources are being allocated to engage all 67 institutions participating in the WxChallenge. As part of the spring project to test the capabilities of an on-demand real-time numerical weather forecasting system, LEAD is offering accounts, one per institution (with a backup); to WxChallenge participants to submit customized WRF forecasts for the contest. Up to two forecasts per day may be submitted by each institution during the spring semester 2008. Details about account set-up, and instructions on the interactive process to set-up and run the forecasts will be provided separately to the local data managers. The goal of the spring 2008 project is to introduce the LEAD technology to a broader and more diverse student population participating in the WxChallenge, allow them to generate their own daily WRF-based forecasts from the LEAD portal, and integrate self-generated WRF model products with other models and observations into their forecast decision-making process. It is intended that these forecasts benefit all the students and that the students will have some input in the daily model set-up decisions.

Each day students will be running 2 to 3 forecasts between 16 UTC (11am EST) and 21 UTC (4 pm EST). All workflows have to finish execution including post processing and results have to be submitted to the community no later than 00 UTC (7pm EST) to allow time for the integration of WRF products into the student forecasters decision-making process.

LEAD developers anticipate the need for 150 concurrent workflow runs typically for 20 to 180 minutes each, plus post processing, resulting in 1200 Jobs within span for three hours. For each workflow, 1.6 GB of data is ingested with 1.2 TB/day of output data that must be stored in MyLEAD workspace. Over the full spring project period, 110 TB of output will be generated.

5. ACKNOWLEGEMENTS

LEAD is a Large Information Technology Research (ITR) Grant 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), ATM03-31586 (Howard University), ATM-0331587 (University Corporation for Atmospheric Research), and ATM-0331578 (University of Illinois at Urbana-Champaign, with a sub-contract to the University of North Carolina). The LEAD team wishes to thank the WxChallenge and the students and faculty who were part of the spring 2007 pilot for participating and providing valuable feedback.