5B.5

LEAD: AUTOMATIC TRIGGERING OF HIGH RESOLUTION FORECASTS IN RESPONSE TO SEVERE WEATHER INDICATIONS FROM THE NOAA STORM PREDICTION CENTER

Robert Wilhelmson, Jay Alameda, Albert Rossi, Shawn Hampton
National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Brian Jewett
Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Daniel Weber, Kevin Thomas, Yunheng Wang, Kelvin Droegemeier
Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK, 73019

1. INTRODUCTION

In Spring, 2007, LEAD, using a trigger developed by NCSA/CORE, launched WRF forecasts in support of NOAA's Hazardous Weather Testbed (hwt.nssl.noaa.gov/Spring_2007/) (HWT); this was one of three aspects of LEAD's collaborations with the HWT which are described elsewhere in this session. The trigger determined when and where forecasts would take place by continuously monitoring and parsing Mesoscale Discussion and Severe Weather Watch products from the NOAA Storm Prediction Center via an RSS feed. 6-hour WRF forecast workflows were then launched, monitored, post-processed and archived by the workflow broker (http://broker.ncsa.uiuc.edu) . Typically, 18-km, singly-nested and 2-km triply nested forecasts were triggered automatically for each SPC bulletin, using NAM data and the WPS package for initialization. 20km ARPS Data Analysis System (ADAS) initialized WRF forecasts were also triggered. The domain centers of all WRF forecasts triggered automatically in this manner are shown in the figure below. Overall, more than **1000 forecasts were initiated in this manner**.

Corresponding author address: Jay Alameda, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL, 61801: jalameda@ncsa.uiuc.edu

stepping stone to building a cyberenvironment which aims to provides flexibility and ease of use far beyond the current range of typical problem solving environments.

This work, besides providing an opportunity to have ready access to high resolution forecasts over regions of interest based on SPC observations, provided an excellent opportunity to stress-test all aspects of the workflow and computational systems involved. This support of the SPC Spring Experiment also served as an important mechanism for further integration of the workflow broker into LEAD, where it soon will be added as a service. The triggered runs, carried out on production rather than dedicated resources, also pointed out the critical need for sophisticated quality of service improvements needed so that all of the important runs on a busy severe weather day complete within the needed window of time; static mechanisms to provide reserved portions of a machine are inadequate in this sense to deliver forecasts in a timely fashion on a busy day, while not waste resources on a calm weather day

2. PROBLEM DESCRIPTION

Historically, the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma has collaborated with the NOAA Storm Prediction Center (SPC) to study atmospheric predictability through real time forecasts during the United States spring severe weather season. This work has demonstrated that organized deep convection is much more predictable than previously thought. In the Spring of 2007, as part of the NOAA Hazardous Weather Testbed, LEAD sought to apply its technology to the numerical prediction of deep convection expressed via two LEAD challenges: a) the use of storm-resolving ensembles for specifying uncertainty in initial conditions, as well as quantifying uncertainty in model output, and b) the application of dynamically adaptive, on-demand forecasts created automatically or manually in response to atmospheric conditions. This paper focuses on the latter challenge, namely, forecasting in response to predicted conditions as annotated by the Storm Prediction Center's mesoscale discussions or severe weather watches.

3.DYNAMIC ADAPTIVITY THROUGH THE LEAD WORKFLOW BROKER.

The plan for the NOAA Hazardous Weather Testbed collaboration aimed to support a number of experiments, namely:

1. a 10-member ensemble of forecasts run on the Cray XT3 at Pittsburgh Supercomputing Center (PSC), 2/3 CONUS, and 4km resolution
2. a single 2km forecast, 2/3 CONUS, also run on PSC's Cray XT3
3. a single 2km forecast, domain determined at daily meeting with SPC staff, run at NCSA,
4. triggered 2km 6 hour forecasts, smaller domain centered at the centroid of a region corresponding to a Mesoscale Discussion or Severe Weather Watch, also run at NCSA.

Forecasts 1 and 2 were to be run in a production sense using ARPS control (as documented in [Xue], [Kong] and [Weiss]) ; forecast 3 was to be run "on-demand" through the LEAD Portal, and forecast 4 was to be run using the workflow broker; note that there was a need to obtain the results from all of these forecasts in a sufficiently timely fashion as to be useful to forecasters at the Storm Prediction Center. The workflow broker, formerly known as the workflow broker, was described in [Siege], and is schematically depicted in Figure 1:

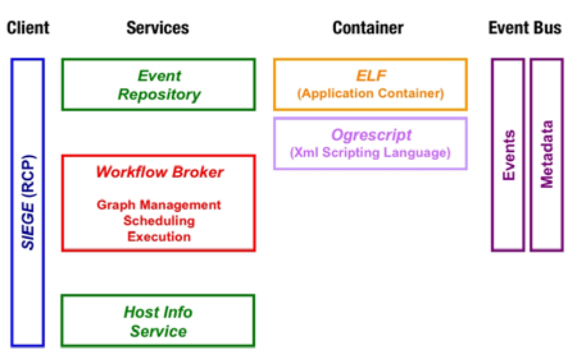

Figure 1: The LEAD Workflow Broker and associated services. Siege is the graphical user interface that the end-user would use; the workflow broker handles parametric expansion; while ELF and Ogrescript manage host-local orchestration.

Additionally, we were to explore use of the workflow broker for managing forecasts 1 and 2, especially in light of the parametric variations for forecast 1. In preparation for the spring forecast experiment, the workflow broker team laid out the following development plan:

- develop a ADAS-WRF workflow following the workflow conventions desired by the Oklahoma team.
- Develop a trigger service, which parses the RSS feed from the SPC for information on mesoscale discussions and weather watches, derives the center latitude and longitude of the region corresponding to the event, and then triggers a sub-domain 6 hour forecast of our choosing.
- develop an experimental job-quality-of-service improvement capability in the broker by the use of the MOAB scheduler used at NCSA on its production resources.
- Enhance the broker's parametric expressivity through a refactoring of key broker components, as driven by needs of the 10 member ensemble for the spring experiment..

With our initial requirements for the triggered runs and 10 member ensemble forecasts in hand, we collaboratively developed a 20km ADAS-initialized WRF forecast, including postprocessing triggered by the appearance of each output file from WRF. We also developed an 18km 1-nest forecast, using NAM data and the WRF WPS package for initialization, and subsequently developed a 3-nest 2km high resolution version of the same NAM-WPS-WRF workflow. The low-resolution versions of the workflows were useful for debugging and operational testing; the high resolution versions were of course the desired product from the simulations.

In order to support runtime modifications to WRF execution to support triggered runs at various locations, we developed a general namelist modification capability, which we are currently making use of this through a series of Ogrescript tasks; for example, one can use the namelist template "namelist.wrf" to generate a namelist instance, here denoted as "namelist.input", with the variable "run_hours" being replaced by a value determined at runtime, as follows:

```
<namelist-substitution
namelistTemplate="${runtime.dir}/namelist.wr
f"
targetFile="${runtime.dir}/namelist.input">
        <substitutions>
                <map-entry
        key="run_hours">${WPS_fcst_len}</map-
        entry>
        </substitutions>
</namelist-substitution>
```

Parametric variations, then, in things such as namelists, are accomplished simply by wrapping a local workflow template script with a parameter variation description; the workflow broker then expands the template into the full set of instances based on that description at the point that the script is processed for submission to the hosts.

4. RESULTS

In the course of the Spring 2007 NOAA Hazardous Weather Testbed, we were able to trigger more than 1000 runs based on RSS feed content from the Storm Prediction Center for mesoscale discussions and severe weather watches. The runs were carried out on NCSA's Intel Itanium cluster, "Mercury" and Intel Xeon cluster, "Tungsten". Details of the success rates are shown in Figure 2.

resolution	workflow 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
												1015

A few notes of explanation:
The data for these statistics are in the multiple tabs in this worksheet.
For ADAS, we did not count ARPSPLT in the scoring, as the multiple ARPSPLT runs were all part of a complete run.
The original.summary.dat tab is the raw data from the broker
the last tab, extracted.summary.tab, is the "left over" jobs - which are test or development jobs, not truly triggered runs

FAILED: a workflow was marked as Failed
DONE: a workflow successfully completed
CANCELLED: a workflow was cancelled from the administrative interface (usually after either broker or event channel failure)
null: one symptom of (typically) an event channel failure -

Figure 2. Triggered run statistics for the 2007 Spring Experiment.

Note that we encountered reliability issues with the ADAS-WRF workflow, due in part to the multi-step sequence of the workflow (and combined with reliability issues on Tungsten) which contributed to a relatively low success rate for the workflow. Also, note that the reliability issues we experienced on Tungsten also impacted adversely the larger 2km NAM-WPS-WRF runs.

The majority of the runs – 6 hour WRF forecasts, with both 18 km singly nested and 2km triply nested forecasts were triggered using NAM data and the WPS package for initialization. Note that the trigger capability developed also generated identifiers, which made deriving the statistics portrayed in Figure 2 straightforward. Figure 3 shows the domain centers of the automatically triggered forecasts.

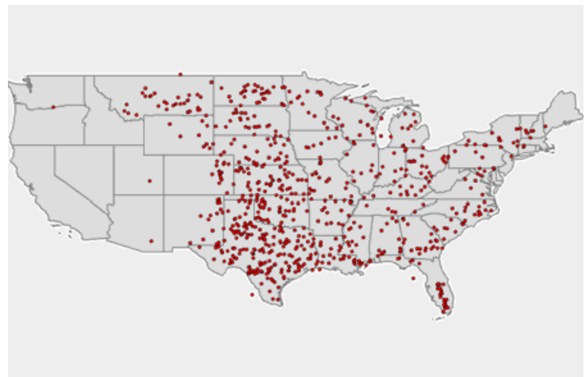

Figure 3. Domain centers of all nested grid forecasts triggered automatically by NCSA based upon SPC Mesoscale Discussion and Severe Weather Watch information

We have also hosted the results from the triggered runs at Brian Jewett's web site [RT], which provided easy access to automatically generated images from all of the runs.

5. DISCUSSION

In the course of the Spring 2007 NOAA Hazardous Weather Testbed, we demonstrated the feasibility of dynamically launching and managing large numbers of runs on multiple compute platforms at NCSA. For this experiment, we were using the reservation capabilities of the resources to provide for priority execution of the jobs. Unfortunately, since the reservation was finite in size, on busy weather days, we could easily saturate the reservation and then provide less-than-timely results on later triggers. Also, as one could see from the web display of the results, it can be a formidable challenge to know what results you may have – even displaying the thumbnails of one image of each of >1000 runs can tax a modern web browser!

A number of improvements in our experiment plan, then would be in order to have a more scalable and more effective Spring 2008 experiment. First, quite often a weather watch will closely follow a mesoscale discussion. In this case, a modification of the run (either submission of a run based on the weather watch combined with a cancellation of the mesoscale discussion run, or suppression of the secondary weather watch run) would be in order. Presumably, this would require development of logic to determine if the domain of the mesoscale discussion and watch were substantially the same.

Additionally, motivation for improvements in quality of service by direct negotiations with the cluster-local scheduler is critical for improving forecast quality-of-service. Work had started on this aspect of the project; more work including some modifications of the cluster-local scheduler, are required in order to flexibly and optimally provision computational resources to allow for forecasts to be generated in a sufficiently timely fashion.

Automated evaluation of the computational results, perhaps using the newly released MET package [MET] from the NCAR Developmental Test Center. This likely would require a compound triggering, for instance, triggering of the evaluation once both the model run is complete and the required observational data is available.

Perhaps out of scope of the LEAD project, but definitely important to obtaining reliable forecast results – would be the ability to reliably handle hardware exceptions on the computational clusters would be incredibly valuable. Currently, events such as filesystem issues, or network connectivity issues are not cleanly handled, which implies not possible to propagate back up from the remote computer, and consequently, very difficult to diagnose the fault when it occurs. As computational systems become larger and more complex, this issue will become even more critical, especially for usage modes which couple the computational resource with software capabilities such as the workflow broker.

One thing that we have made progress on, that we believe will positively impact our ability to manage the set of results, was the development of a capability to publish metadata into the myLEAD data workspace [RelayAgent]. This will likely also require additional work to couple easy to understand web views of the data with the additional metadata attributes, but represents a substantial step towards managing the large numbers of runs that would be a result of another successful spring experiment.

Finally, strides in reliability and scalability enhancement have been made through the addition of support for gsissh-enabled job submissions, and de-coupling a striped gridftp server set by running instances of non-striped gridftp that can be used independently to scalably transfer large numbers of files of of various sizes. We are also investigating submission of multiple jobs into a broker glide-in that would provide for efficient management of a large number of jobs within a single batch submission.

6. ACKNOWLEDGEMENTS

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). This work has also been supported by SCI03-30554, SCI04-38712, and SCI96-19019. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

7. REFERENCES

[Siege] Siege – MRD-Public – Confluence, http://torcida.ncsa.uiuc.edu:8080/confluence/display/MRDPUB/Siege, also, Alameda, J., R. Wilhelmson, A. Rossi, S. Hampton, B. Jewett, T. Baltzer, A. Wilson, 2007: Siege: A Graphical User Interface to Enable Management of Large Numbers of Weather Simulations, 23rd Conference on IIPS, San Antonio, TX, American Meteorological Society Annual Meeting, 8B.3

[Xue] 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 experi-ment. 22nd Conf. Wea. Anal. Forecasting/18th Conf.Num. Wea. Pred., Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 3B.1.

[Kong] 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, 2007: 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., Park City, Utah, Amer. Meteor. Soc., CDROM 3B.2

[Weiss] 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., Park City, Utah, Amer. Meteor. Soc., CDROM 6B.4.

[RT] Triggered WRF forecast tests at NCSA, http://rt.atmos.uiuc.edu/trigger/.

[MET] MET Users Home Page, http://www.dtcenter.org/met/users/.

[RelayAgent], MyLEAD Relay Agent, http://torcida.ncsa.uiuc.edu:8080/confluence/display/MRDPUB/MyLEAD+Relay+Agent.