An Automated Decision and Control System for Numerical Weather Prediction Research and Education: Results from a Recent Experiment

P14b

D. Weber[1], K. W. Thomas[1], K. Droegemeier[1], Y. Wang[1], and D. O'Neal[2]
[1]Center for Analysis and Prediction of Storms, University of Oklahoma
[2]Pittsburgh Supercomputing Center

Spring Experiment Computing and Manpower Requirements

Since the 2005 Spring experiment forecasts were conducted at higher resolution (2-km horizontal grid spacing) than previously attempted, the amount of effort to upgrade the WRF and ARPS software to enable greater than 2Gb files and the parallelization of several pre- and post-processing applications, was significant. Figure 1 presents the manpower requirement for the 2005 Spring Experiment.

Computing:
•System testing, optimization and troubleshooting prior to the experiment
•1156 dedicated processors, 1100 for the forecast and 56 for post-processing for 8 hours from 10pm-6am local time. Total time requested - 400,000 nodes hours on LEMIEUX at the Pittsburgh Supercomputing Center.
•A minimum of 380Gb of disk space per forecast
•Remote job submission permission

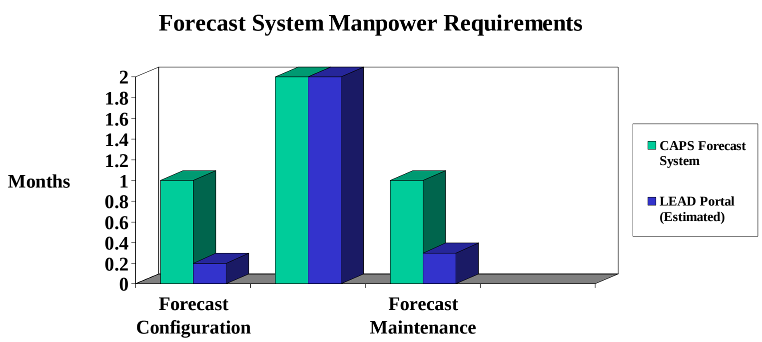

Figure 1. Manpower Requirement for the 2005 Spring Experiment in Months

Forecast Fault Assessment - Preliminary Findings

- **Network considerations:**
  •Combined amount of data transferred to PSC from OU was on the order of 100-300 Mb.
  •Average transfer rates between OU and PSC were on the order of 3 Mb/sec
  •Optimized network data movement is needed for future experiments
  •Several transfers were restarted due to failed transmission/time outs
- **Computing considerations:**
  •*Did not experience a compute node failure during the experiment!*
  •Per processor performance is on the order or 10% efficient, due to cache thrashing and intensive memory references
  •Optimized code and/or an increased number of processor is needed for planned future Spring Experiments
- **Disk System Performance**
  •Sufficient disk space and file archiving capabilities existed during the duration of the experiment
  •Model simulations wrote data in a split file format, allowing each processor to produce a data file, or 1100 files per forecast write
  •Disk performance is diminished when accessing/handling 30,000+ files in a single directory
  •Several disk problems occurred which halted the computing and file I/O intensive tasks (files writes from the model and post-processing during the graphics generate, may have been a result of larger than physical memory usage on the post processing nodes)

WRF Forecast and Observed Radar Reflectivity

The WRF forecasts were initiated at 00Z and extended to 30 hours in order to allow forecasters to review and use data that would cover the current day's convective cycle (typically 21Z to 06z). The forecast from the last day of the experiment, June 4, 2005 is presented in Figure 2. The graphs show the 24-hour simulated radar and observed ADAS radar analysis valid for 00Z June 5, 2004, each depicting a squall line developing in east central Oklahoma and extending northward into Kansas and Missouri. The placement of the forecast storms at hour 24 is within a few counties of the observed convection. Note the forecasts were initialized from the 40km ETA model forecast data without radar or other observations.

(A)

(B)

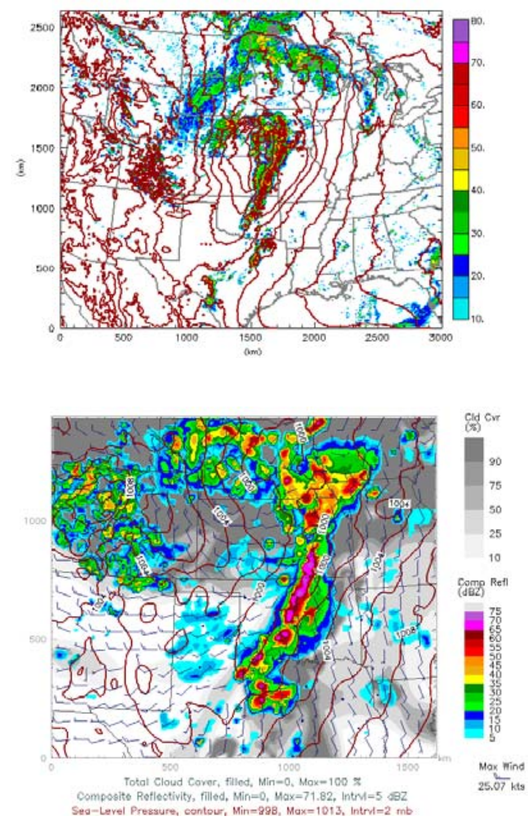

Figure 2. WRF predicted radar reflectivity (24-hour) (a) and ADAS radar analysis (b) valid 00Z June 5, 2005. The WRF forecasts used a horizontal grid spacing of 2-km.

Summary

The goal of the LEAD portal is to reduce significantly the amount of preparation associated with creating, running, and analysis of complex workflows associated with mesoscale meteorology research and education. The current CAPS forecasts system is very capable of automating a real time forecast cycle, including pre- and post—processing and verification processing, but is not easily upgraded, modified and maintained. A fault assessment of the Spring 2005 forecast experiment operations revealed weak points in the handling of files and significantly reduced internet network performance. In addition, significant fault tolerance research and development is needed to improve the error handling aspects of a workflow, especially dynamically data driven experiments, that are minimally addressed within the current CAPS system.

This work is supported primarily by the National Science Foundation under the following cooperative agreements: ATM03-31574, 31578, 31579, 31480, 31586, 31587, 31591, and 31594. 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.