

# Developing Grid based infrastructure for climate modeling

John Taylor,<sup>1,2</sup> Mike Dvorak,<sup>1</sup> and Sheri Mickelson<sup>1</sup>

<sup>1</sup>Mathematics and Computer Science, Argonne National Laboratory, Argonne, Illinois 60439  
and Computation Institute, University of Chicago, IL, 60637

<sup>2</sup>Environmental Research Division, Argonne National Laboratory, Argonne, Illinois 60439  
{jtaylor dvorak mickelso}@mcs.anl.gov  
<http://www-climate.mcs.anl.gov/>

**Abstract.** In this paper we discuss the development of a high performance climate modeling system as an example of the application of Grid based technology to climate modeling. The climate simulation system at Argonne currently includes a scientific modeling interface (Espresso) written in Java which incorporates Globus middleware to facilitate climate simulations on the Grid. The climate modeling system also includes a high performance version of MM5v3.4 modified for long climate simulations on our 512 processor Linux cluster (Chiba City), an interactive web based tool to facilitate analysis and collaboration via the web, and an enhanced version of the Cave5D software capable of visualizing large climate data sets. We plan to incorporate other climate modeling systems such as the Fast Ocean Atmosphere Model (FOAM) and the National Center for Atmospheric Research's (NCAR) Community Climate Systems Model (CCSM) within Espresso to facilitate their application on computational grids.

## 1 Introduction

The Grid is currently considered an emerging technology with the potential to dramatically change the way we use computing resources such as commodity and high performance computers, high speed networks, experimental apparatus and security protocols [1]. In a recent article on the emergence of computational grids, Larry Smarr speculated that computational grids could rapidly change the world, in particular Smarr noted that:

“Our nation’s computer and communication infrastructure is driven by exponentials: it’s on steroids. Every aspect – the microprocessor, bandwidth and fiber, disk storage – are all on exponentials, a situation that has never happened before in the history of infrastructure” [1]

The Grid primarily refers to the underlying hardware and software infrastructure needed to undertake computational tasks using distributed resources. Under ideal conditions the Grid would appear to the user as a single system. The details of the Grid infrastructure are ultimately intended to be invisible to the user just as the

infrastructure of the electric grid are largely hidden from the electricity consumer[1]. As advanced computational grids are developed, e.g.: the US National Science Foundation TerraGRID project, they offer the potential to address far more challenging problems in climate science than have been addressed in the past using a single high performance computer system.

## 2 Grid based infrastructure for climate modeling

In order to develop software to exploit the grid infrastructure we need to be able to:

- access information about the grid components
- locate and schedule resources
- communicate between nodes
- access programs and data sets within data archives
- measure and analyze performance
- authenticate users and resources

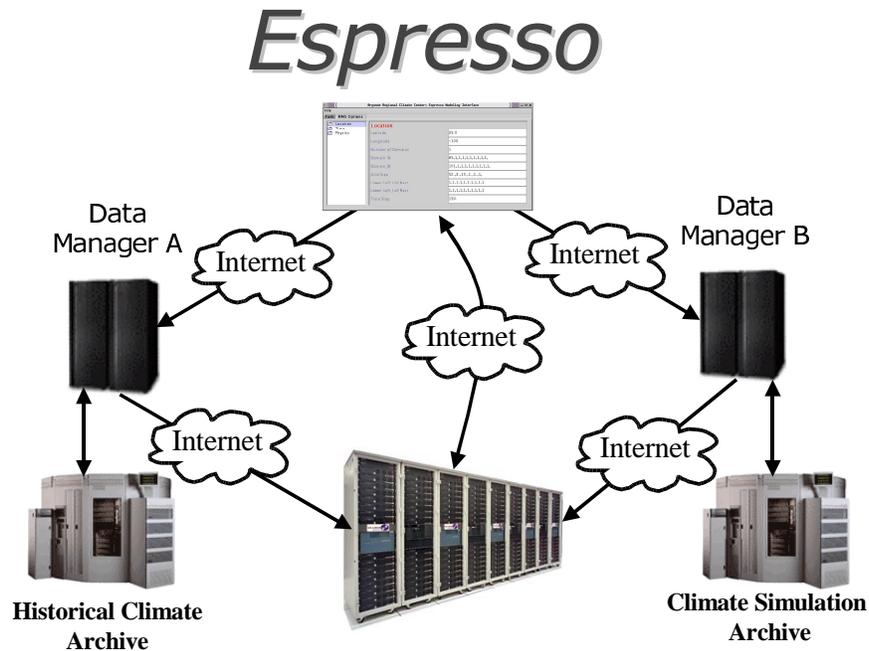
All these activities must be undertaken in a secure environment. The Globus project has developed an integrated suite of basic grid services collectively known as the Globus Toolkit [2]. These services and their Application Programming Interface (API) effectively define a Grid architecture on which we can develop Grid applications tools. We adopted the Globus Toolkit API as our underlying architecture for the development of our Grid based tools for climate modeling applications. We selected the Java programming language and the Globus Java Commodity Grid Kit, known as the Java CoG Kit [3], for our software development.

Building the software infrastructure to fully exploit computational grids remains a major challenge. In this paper we discuss current efforts at Argonne National Laboratory (ANL) to build Grid based infrastructure to facilitate climate modeling. We provide an overview of the Grid based software tools that we are currently developing and practical examples of the climate modeling results that we have obtained using this infrastructure.

The emerging Grid technology will allow far more challenging climate simulations to be performed. These simulations may ultimately be used to assess the impacts of global climate change at the regional scale. In a recent IPCC report on *The Regional Impacts of Climate Change* it was concluded that:

The technological capacity to adapt to climate change is likely to be readily available in North America, but its application will be realized only if the necessary information is available (sufficiently far in advance in relation to the planning horizons and lifetimes of investments) and the institutional and financial capacity to manage change exists [4] (IPCC, 1998).

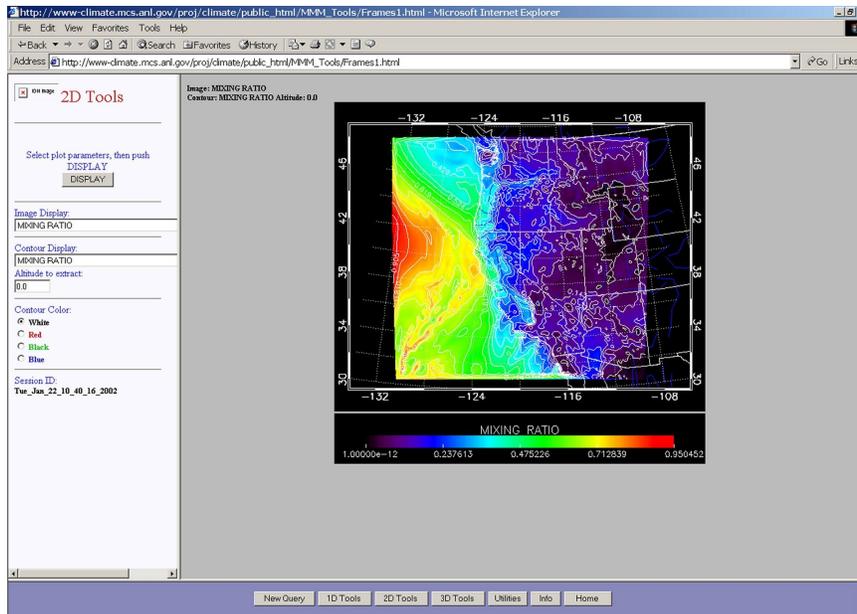
It was also acknowledged by IPCC that one of the key uncertainties that limit our ability to understand the vulnerability of sub-regions of North America to climate



**Fig. 1.** Espresso allows the user to access remote computational resources and data sets on a climate modeling grid via an easy to use scientific modeling interface. Espresso uses the Globus Java CoG Kit [3] to access Grid services. In the above example we use Espresso to perform a climate model simulation on a remote machine while accessing input data sets from remote data archives.

change, and to develop and implement adaptive strategies to reduce vulnerability, was the need to develop accurate regional projections of climate change, including extreme events [4]. In particular we need to account for the physical-geographic characteristics that play a significant role in the North American climate, e.g. the Great Lakes, coastlines and mountain ranges, and also properly account for the feedbacks between the biosphere and atmosphere [4].

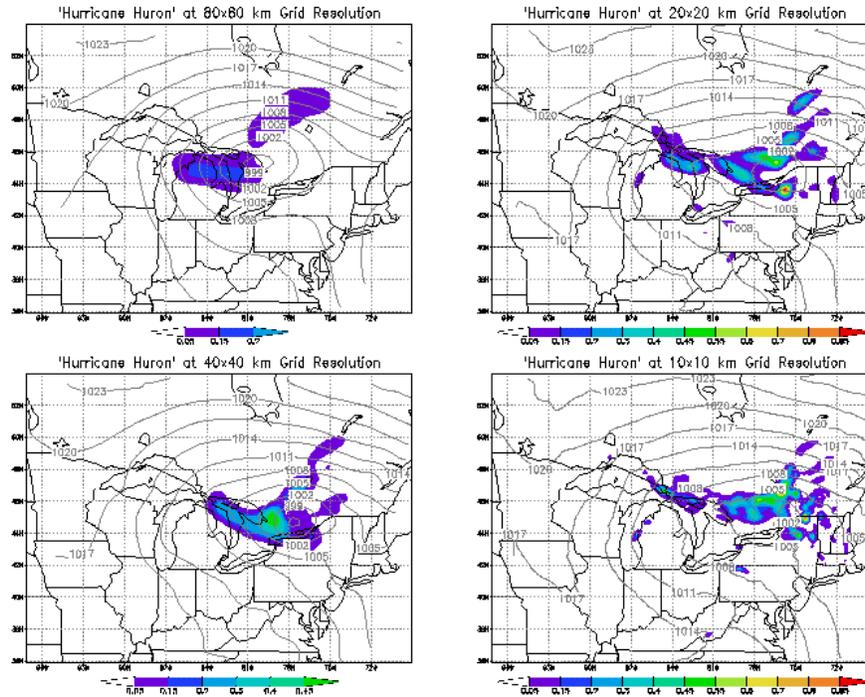
The potential impacts of global climate change have long been investigated based on the results of climate simulations using global climate models with typical model resolutions on the order of hundred's of kilometers [5,6]. However, assessment of impacts from climate change at regional and local scales requires predictions of climate change at the 1-10 kilometer scale. Model predictions from global climate models with such high resolutions are not likely to become widely available in the near future. Accordingly, at ANL we have begun developing a climate simulation capability, at both the global and regional scales, for application on computational grids.



**Fig. 2.** Web interface displaying the water vapour mixing ratio data from a high resolution (10km) climate simulation of the western United States at 07Z 3 December 1996. Much detail of the complex topography of the western United States is revealed at this model resolution. The data set from this simulation consists of 10 GB of data, all of which can be viewed via the web analysis tool. The model simulation is for the period 1-10 December 1996 and shows an advancing cold front producing significant precipitation over northern California. All plots are generated and displayed online based on user requests. The next generation of this web tool will allow interactive data analysis in addition to data display.

Our long term goal is to link the predictive global climate modeling capability with the impact assessment and policymaking communities. The regional climate simulation system at Argonne currently includes:

- A Java based Scientific Modeling Interface (Espresso) to facilitate task management on the Grid. See [7,8] for further details on the development of Espresso.
- A high performance version of MM5v3.4 modified to enable long climate simulations on the Argonne 'Chiba City' 512 processor (500 Mhz Pentium III) Linux cluster.
- An interactive web based visualization tool to facilitate analysis and collaboration via the web.
- An enhanced version of the Cave5d software capable of working with large climate data sets.



**Fig. 3.** Precipitation and surface pressure for hour ending 12Z 11 September 1996 at 80, 40, 20 and 10km grid resolution. Precipitation intensifies by nearly an order of magnitude as we go to higher model resolutions and occupies a smaller more sharply defined rainbands. Increasing precipitation intensities will alter the fraction of rainfall allocated to storage in the soil ie soil moisture and runoff which in turn will alter rates of decomposition and photosynthesis particularly under water limited conditions.

The regional climate model used in this study is based on the Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) fifth generation mesoscale model (MM5). In brief, MM5 is a three-dimensional, non-hydrostatic, and elastic mesoscale model. It uses finite differences and a time-splitting scheme to solve prognostic equations on an Arakawa type-B staggered grid. Its vertical coordinate, though defined as a function of the reference-state pressure, is similar to a terrain-following coordinate. For case studies, MM5 employs observed wind, temperature, and humidity as the initial and boundary conditions. MM5 incorporates realistic topography and sophisticated physical processes to represent the appropriate forcing for the development of the observed weather system. These physical processes include clouds, long- and shortwave radiation, and the surface fluxes of heat, moisture, and momentum. A more detailed description of MM5 is provided by, Chen and Dudhia [9], Chen et al. [10], Dudhia [11] and Grell et al. [12].

The NCAR MM5 modeling system consists of six programs: TERRAIN, REGRID, RAWINS, INTERP, MM5, and GRAPH [9]. Each program is executed in

the above order interdependently. The programs are composed of a series of scripts that traditionally have been time consuming to modify and execute. Recently, this process has been simplified by using the Espresso Scientific Modeling Interface. Espresso was developed at ANL by the Mathematics and Computer Science Division (MCS) staff, including John Taylor, Mike Dvorak and Sheri Mickelson. The design of Espresso [7,8] has benefited from an earlier climate workbench[13].

In order to facilitate Internet based collaboration we have developed a web browser application that enables access to the output of regional climate model runs generated using the MM5 regional climate modeling system. Fig. 2 above illustrates a typical session. The web browser uses the native MM5 data format, thus avoiding the need to store duplicate copies of model output, and works efficiently with gigabytes of data. The web tool was developed using IDL/ION software. An enhanced version of this web tool is currently awaiting installation at MCS.

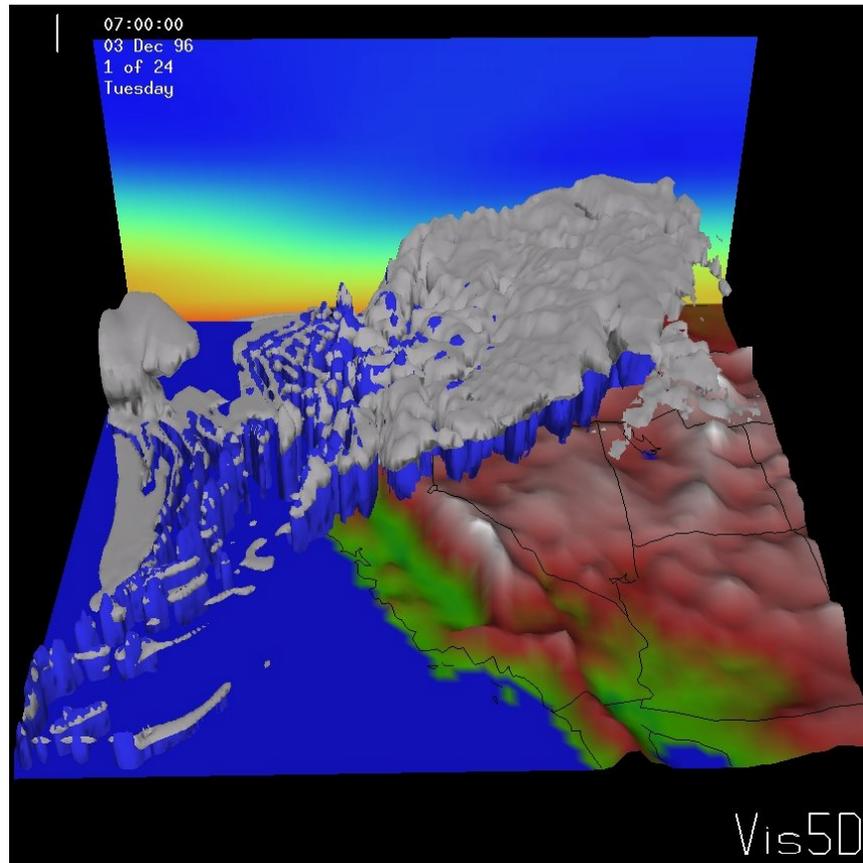
### **3 Example Model Results**

We have performed preliminary model runs in climate mode using a recent release of the MM5 modeling system (MM5v3.4) looking at extreme events using the forerunner to the Espresso scientific modeling interface. Mesoscale resolution climate models provide a consistent framework for us to investigate the link between incoming solar radiation, climate and extreme weather. The following series of experiments were undertaken to illustrate the importance of enhanced model resolution to simulating the weather, climate, and atmospheric transport processes that impact on extreme weather events.

In Figure 3, above, we illustrate the results of our model simulations of 'Hurricane Huron'. We have performed four simulations for the period 6-15 September 1996, at a range of model grid resolutions, 80, 40, 20 and 10km over an identical 2000x2000 km region centered over Lake Huron. By performing our model simulation over an identical spatio-temporal domain we can study the effect of grid resolution on model physics.

Figure 3 also illustrates that hourly rainfall intensity increases dramatically, by nearly an order of magnitude, as we go to higher model resolutions. The pattern of rainfall also changes from broad scale low intensity rainfall at 80km grid resolution, to high intensity rainfall with significant spatial structure associated with formation of well defined rain bands. We use NCAR/NCEP Reanalysis Project wind fields to provide boundary and initial conditions.

We have also performed a high resolution (10km) climate simulation using MM5v3.4 over the western United States for the period 1-10 December 1996 using the Climate Workbench [13]. Such high resolution simulations have not previously been undertaken for the western United States. The purpose of this study was to investigate the ability of the MM5v3.4 model to predict intense precipitation over the complex terrain of the Western United States. During this time La Nina conditions persisted in the tropical Pacific. La Nina conditions are usually associated with strong rainfall events over Northern California. Rainfall events associated with La Nina conditions play an important role in increasing water storage levels in Northern California.



**Fig. 4.** This figure shows isosurfaces of cloud water and rain water mixing ratios from a high resolution (10km) climate simulation of the western United States at 0700 on 3 December 1996. The figure shows an advancing cold front producing significant precipitation over the Northwest and Northern California. Much of the detail of the complex topography of the Western United States is revealed at this model resolution. The model results above are derived from a MM5V3.4 simulation performed for the period 1-10 December.

Figure 4, above, shows an advancing cold front producing significant precipitation over the Northwest and Northern California for 3 December 1996. The model results look very promising as they are in qualitative agreement with observed precipitation. Further work is needed to fully assess the ability of the MM5 model to correctly predict intense precipitation over the complex terrain of the Western United States.

## 4 Conclusions and Future Research

We will continue to address the key scientific and computational issues in climate modeling [14] and their importance to simulating global climate and the climate of the United States by:

- Developing Grid based tools such as the Espresso Scientific Modeling Interface in order to be able to efficiently manage climate simulations using the full capabilities offered by today's distributed computing environments.
- Contributing to the performance improvement of long term global and regional climate simulations by developing new physical parameterizations suitable for incorporation into climate models.
- Developing computer code capable of running efficiently on the new generation of high performance computers such as commodity clusters.
- Performing challenging climate simulations which test the limits of the available technology.
- Building analysis and visualization tools capable of working with the large data sets (order terrabyte to petabyte) generated by global and regional climate model simulations.
- Defining and delivering quality data products via the web through the development of web based interfaces for analysis and visualization.

## Acknowledgments

This work was supported in part by the Laboratory Director Research and Development funding subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy, under Contract W-31-109-Eng-38, and through a NSF Information Technology Research Grant ATM-0121028.

## References

1. Foster, I and Kesselman, C. (1999) *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers.
2. The Globus Home Page. <http://www.globus.org/>
3. von Laszewski, G., Foster, I., Gawor, J., and Lane, P. (2001) A Java commodity grid kit. *Concurrency and Computation: Practice and Experience*, 13:645-662.
4. IPCC (1998) *The Regional Impacts of Climate Change*, Cambridge Univ. Press, Cambridge.
5. IPCC WGI (1990) *Climate Change: The IPCC Scientific Assessment* R.A. Houghton *et al.*(eds),Cambridge Univ. Press, Cambridge, UK.
6. IPCC WGI (1996) *Climate Change 1995: The Science of Climate Change* R.A. Houghton *et al.*(eds),Cambridge Univ. Press, Cambridge, UK.
7. Mickelson, S., Taylor, J. and Dvorak, M. (2002) Simplifying the task of generating climate simulations and visualizations. To appear in the *Proceedings of the 2002*

*International Conference on Computational Science*, Springer-Verlag, Berlin, Germany.

8. Dvorak, M., Taylor, J. and Mickelson, S. (2002) Designing a flexible grid enabled scientific modeling interface. To appear in the *Proceedings of the 2002 International Conference on Computational Science*. Springer-Verlag, Berlin, Germany.
9. Chen, F. and Dudhia, J. 2001: Coupling an Advanced Land-Surface/Hydrology Model with the Penn State/NCAR MM5 Modeling System. Part I: Model implementation and Sensitivity, *Monthly Weather Review*, in press. ( See also Pennsylvania State University / National Center for Atmospheric Research, MM5 Home Page <http://www.mmm.ucar.edu/mm5/mm5-home.html>).
10. Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.L. Pan, V. Koren, Q.Y. Duan, K. Ek, and A. Betts, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.*, 101, 7251-7268.
11. Dudhia, J., 1993: A nonhydrostatic version of the Penn State / NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, 121, 1493-1513.
12. Grell, G.A., J. Dudhia and D.R. Stauffer, 1994: The Penn State/NCAR Mesoscale Model (MM5). NCAR Technical Note, NCAR/TN-398+STR, 138 pp.
13. Taylor, J. Argonne National Laboratory, Climate Workbench [http://wwwclimate.mcs.anl.gov/proj/climate/public\\_html/climate-workbench.html](http://wwwclimate.mcs.anl.gov/proj/climate/public_html/climate-workbench.html), 2000.
14. Giorgi, F. and L.O. Mearns (1999) Introduction to special section: Regional climate modeling revisited, *Journal of Geophysical Research*, **104**, 6335-6352.