

Automating Climate Science: Large Ensemble Simulations on the TeraGrid with the GriPhyN Virtual Data System

Veronika Nefedova¹, Robert Jacob¹, Ian Foster^{1,2}, Zhengyu Liu³, Yun Liu³,
Ewa Deelman⁴, Gaurang Mehta⁴, Mei-Hui Su⁴, Karan Vahi⁴

¹*Mathematics & Computer Science Division, Argonne National Laboratory*
²*Computation Institute, Argonne National Laboratory & University of Chicago*

³*Center for Climatic Research, University of Wisconsin-Madison*

⁴*Information Sciences Institute, University of Southern California*
nefedova@mcs.anl.gov

Abstract

Ensemble simulations are a promising technique for identifying the signal of atmospheric response to extra-tropical sea surface temperature variability with high statistical significance. The basic idea is to perform multiple simulations from slightly different initial conditions and then to study the average signal of the ensemble. A significant obstacle to performing such ensemble simulations is the bookkeeping required to prepare, execute, and track the progress of hundreds of different computations. We describe an ensemble simulation experiment in which the Fast Ocean Atmosphere Model was run on the U.S. TeraGrid. In this experiment, we used the GriPhyN Virtual Data System to manage our ensemble simulations and their execution on distributed resources, achieving dramatic (order-of-magnitude) reductions in turnaround time relative to previous manual experiments.

1. Introduction

In spite of intensive studies over five decades, the mechanism by which the mid-latitude (roughly 30–50N) atmosphere responds to an anomaly in the extra-tropical sea surface temperature (SST) remains one of the most perplexing and controversial issues in current climate research. Poor understanding of this atmospheric response has severely hindered the efforts of the climate community to understand climate variability and climate change from seasonal to multi-decadal and multi-century time scales.

Ensemble simulations of a coupled climate model are a promising technique for identifying the signal of atmospheric response to extra-tropical SST variability with high statistical significance. In an ensemble climate simulation, the same climate model is integrated forward from several different initial conditions. For climate prediction simulations of the next 100 years, ensemble sizes are typically on the order of 10 members. Recent results suggest that, for mid-latitude variability, we need to examine the daily and seasonal evolution of the coupled atmosphere-ocean response for a few months to a few years in order to fully understand the mechanism of the atmospheric response. The large atmospheric weather noise at daily time scales means that the ensemble size needs to be on the order of 1000 to obtain statistically significant results. While 10 integrations can be tracked by an individual investigator, tracking 1000 requires automation and thus a software solution.

To address this need for automation, we have applied a Grid-based workflow management system, the Virtual Data System (VDS) [1], to control the execution of a large ensemble integration on the U.S. TeraGrid cyberinfrastructure (www.teragrid.org) and have demonstrated order-of-magnitude improvements in user productivity relative to manual approaches.

2. Ensemble system technologies

We describe the FOAM application, the VDS software, and the experimental setup.

2.1. Fast Ocean Atmospheric Model

FOAM is a fully coupled, mixed-resolution, general circulation model of the Earth's climate. "Mixed-resolution" means it uses a combination of a low-resolution atmosphere model and a medium-resolution ocean model. Each model solves the full three-dimensional "primitive equations" for a fluid on a rotating sphere and includes parameterization packages for subgrid scale and nonfluid processes (referred to as "physics" in atmosphere modeling). The atmosphere component of FOAM is PCCM3, a combination of the parallel atmosphere model PCCM2 [2] and the column physics of CCM3 [3]. The fluid dynamics are solved with the spectral transform method [2] with a truncation of R15 corresponding to 4.5 degrees latitude and 7.5 degrees longitude. PCCM3 uses a two-dimensional decomposition that allows it to use more processors at a given resolution than does CCM3. The ocean component of FOAM is OM3, a finite-difference, hydrostatic, incompressible z-coordinate model that uses a fully explicit time-split integration scheme. OM3 also has a two-dimensional decomposition. The resolution is 128x128 in the horizontal (1.4 degrees latitude and 2.8 degrees longitude) with 24 unevenly spaced vertical levels with maximum thickness of 500 m and maximum depth of 5500 m. The land model uses a simple four-level diffusion model for soil temperature with prescribed vegetation types and a prognostic soil hydrology and river routing scheme [3]. The sea ice model contains only thermodynamics from CSIM 2.2.6 [5].

The simulated climate of FOAM compares well with observations and models with similar and higher resolution [4]. FOAM's mixed-resolution and efficient parallel numerics allow it to achieve simulation speeds of up to 150 years/day on modest numbers of processors (16-32). FOAM is well suited to its intended applications in long-term climate variability and paleoclimate and has been shown to have good interannual and interdecadal variability [4,6].

2.2. Virtual Data System

The Virtual Data System (VDS) provides mechanisms that allow a user first to describe (in a declarative manner) the steps required to produce a desired data product and to produce that data product in various execution environments, including distributed Grids. A Virtual Data Catalog can be used by application environments to describe a set of application programs ("transformations") and then to track all data files produced by executing those applications ("derivations"). This catalog contains the "recipe" to produce a given logical file and, on request,

produces a directed acyclic graph of program execution steps that will produce the file.

Workflows consist of graphs of application invocations and can be expressed in a location-independent, high-level Virtual Data Language (VDL) [7] that frees the workflow from specifying details of the location of files and programs in a distributed environment.

VDL workflows can be executed in a variety of environments ranging from the desktop to Grids such as the Open Science Grid (www.opensciencegrid.org) and TeraGrid.

VDL definitions are stored in the Virtual Data Catalog that tracks the provenance of all files derived by the workflow.

Both VDL definitions and the workflows that it generates can be expressed in XML documents. The workflow is expressed as a "DAG in XML" or "DAX" document, which VDS planners transform to generate executable forms of the workflow for different environments.

VDS planners include those contained within the Pegasus package [8] and other experimental research planners used to prototype new planning features and executable code generation approaches. We used the Pegasus planner in this work. This planner queries various catalogs to collect information about the various grid sites, the location of required input data sets and any already generated output data, and the location of application programs on various Grid sites. Using this global view of the Grid, Pegasus removes from the workflow jobs that have already been executed, schedules the application jobs on various Grid sites, and adds data movement jobs to move the data consumed and generated by the application. Newly derive data is also registered in data catalogs for future discovery and reuse.

This planned workflow is rendered as a set of DAG (directed acyclic graph) descriptions that are then passed to the Condor DAGman system [9], which manages their execution (e.g., by handling restarts in the event of failure), dispatching tasks to remote sites via the Globus-based Condor-G library [10].

2.3. Experimental design

Mid-latitude variability has been studied extensively by using atmosphere-only models [11,12] and atmosphere models coupled to a "slab" representation of the ocean [13,14]. Conflicting results from these approaches suggested that mid-latitude response to SST should be studied with a fully coupled model [15].

A typical experiment begins by introducing a temperature anomaly in the northwest Pacific before the beginning of winter. To generate the initial conditions for our ensemble, we performed a single 250-year integration of FOAM with a checkpoint, or restart, dataset saved each November 1. An ensemble size of 250 was chosen because it is a significant increase over a typical manual ensemble size and clearly requires automation. Because the climate system and any model (such as FOAM) are chaotic, each of the 250 checkpoint datasets will differ slightly from the others. By integrating over four years starting with initial conditions in which the same temperature anomaly is imposed, we can determine the average response of the model to the anomaly by averaging the four-year period from each member.

The software system to run the integration involved three major tasks. The first was to set up each individual ensemble, which included retrieving a specific restart data set, moving it to the run location on the target computer, and creating a small file FOAM uses to determine how long to run and what file to start from. The second task was to run each simulation, monitor its execution, handle restarts on failure, and collect output data. The third task was to postprocess the output, converting it to the NetCDF format and moving it to archival storage.

3. Building the ensemble system

We made several modifications to both FOAM and VDS in order to run our ensemble simulations on TeraGrid. The FOAM modifications were minor, while modifications to VDS introduced several new features that are beneficial to both our project and others.

3.1. Modifications to FOAM

FOAM required only minor modifications to work within the Ensemble System. Like many large applications, FOAM outputs several diagnostics and other information each time step. Typically this data is output to standard output, which is then redirected to a file so that the user can view this information as the model runs. To detect when the model has finished running, VDS typically monitors the standard output. A standard FOAM run does not produce any failure code in the case of unexpected problems during the run (e.g., requested CPU time exceeded, disk space exhausted). We therefore introduced a simple marker into the standard output of VDS. The redirection of standard output to a file was removed, and a simple

“Done” message was added to the end of the program so VDS would know when an individual member had finished successfully. Since a human was not likely to watch all the ensemble members, we also reduced the amount of output sent to standard output.

We were also concerned about the volume of data that would be produced if the normal number of variables were output for each ensemble member. Thus, we modified FOAM to reduce the output to only those fields of interest in mid-latitude variability.

Other modifications concerned the scientific experiment we wished to perform. In particular, we added code to the ocean model to introduce a temperature anomaly into the north Pacific at the start of the integration.

3.2. Modifications to VDS

```
TR FOAM(input RunParamsName, output DirAtmos,
        output DirOcean, output DirCoupl,
        input WorkDirName, none vds_workdir,
        none vds_group)
{
  argument = ${RunParamsName};
  profile globus.maxWallTime = "5";
  profile vds.workdir=${vds_workdir};
  profile vds.group=${vds_group};
}

TR DirGen(input Finished, input WorkDir,
          input Anomstart, output WorkDirName)
{
  argument = ${Finished};
  argument = ${WorkDir};
  argument = ${Anomstart};
  profile globus.maxWallTime = "3";
}

DV FOAM_RUN1->FOAM(
  RunParamsName=@{input:"run_params":"tmp"},
  WorkDirName=@{input:"foam-mem0001":"tmp"},
  DirAtmos=@{output:"foam01/atmos":"tmp"},
  DirOcean=@{output:"foam01/ocean":"tmp"},
  DirCoupl=@{output:"foam01/coupl":"tmp"},
  vds_workdir="foam01",
  vds_group="Foam1"
);
```

Figure 1. VDL fragments for FOAM runs

The VDL fragments in Figure 1 illustrate the VDL notation. The first two statements define the transformations FOAM and DirGen. These transformation definitions act as function definitions and specify the formal arguments to an application and the details of how those arguments are passed to and from the application represented by the TR definition. In FOAM, the arguments include an input file (run_params) as well as the working directory path and output directory paths. The running environment for

each FOAM run (i.e., for each ensemble member) is set up by DirGen. The third statement specifies a FOAM derivation. Derivations, defined by “DV” statements, specify the actual arguments to be passed to a transformation. File names used as arguments in DV statements are “logical names,” mapped to physical file names at run time [16].

The VDL specifications for each ensemble run (250 members) are generated automatically by a script. The workflow for each ensemble member consists of five steps: working directory generation (as each ensemble member has to be ran from the unique directory), FOAM run itself, and three postprocessing jobs (Atmospheric, Coupler, and Ocean). The results are transferred back to the user by GridFTP once the computations are completed. Figure 2 below shows the whole ensemble workflow.

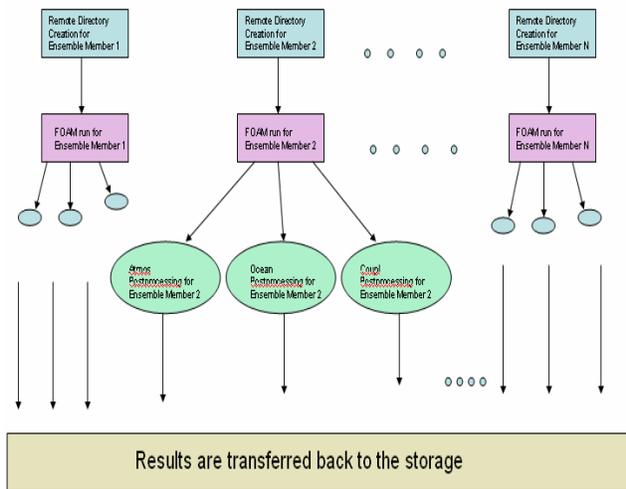


Figure 2: FOAM ensemble

VDS had to be modified in several ways to accommodate our specific workflow needs. First, we added the option to specify various working directories for various workflow components, in order to allow for simultaneous execution of 250 independent workflows. Second, we implemented changes that allowed us to run different workflow components on different remote schedulers. More time-consuming and parallel FOAM jobs were run on the PBS scheduler, while the less time-consuming single-threaded jobs (e.g., the DirGen task used to prepare the initial directories for each ensemble member) were run on the fork scheduler.

4. Running on TeraGrid

The execution of the workflows is performed in automated fashion, as follows:

1. A script generates the description of the computation as a set of VDL statements. This specification provides a high-level description of the tasks associated with the ensemble. Each ensemble member comprises a FOAM run and three postprocessing jobs for data obtained from the atmosphere model, the ocean model, and the coupler.
2. Pegasus generates a mapping of the VDL workflow onto TeraGrid resources, represented as a directed acyclic graph and associated job submission files for the data transfers and job submissions required to execute these tasks.
3. The resulting DAG is executed on the TeraGrid. For each ensemble member, a remote directory is created and populated with the initial and boundary conditions. The FOAM job is run (which takes approximately three hours on eight processors), the postprocessing jobs are run, and the results are transferred back to storage.

Running on SDSC’s and NCSA’s TeraGrid resources, we typically complete an ensemble of 250 members in two to two and a half days, using roughly 7,000 CPU-hours for an average parallelism of approximately 140. This time can be improved significantly by several factors, such as making reservations ahead of calculations, changing TeraGrid settings to allow more simultaneous processes per user, and running at multiple sites.

5. Results

In Figure 3 we show results for both a 250-member ensemble integration performed with VDS on the TeraGrid (bottom) and a second 160-member ensemble integration performed at the National Center for Atmospheric Research (NCAR: top).

Each panel shows a time series over four years of simulation. The green (light) dots are for individual ensemble members while the red (dark) line is the ensemble mean. The dashed lines show the 95% confidence level. The first column shows the averaged sea surface temperature from 30-50N and 140-180E. The second column shows the atmosphere’s geopotential height (GPH) at 500 mb averaged over 30-60N and 180-220E and the third column shows the atmosphere-ocean heat flux averaged between 30-50N and 140-180E.

The scattering in all three fields from the ensemble members (green) shows how strong the variability is in

this region. A clear seasonal cycle can be seen in the envelope formed by the ensemble members. The scattering is strongest in the winter for the atmospheric response (GPH) and the heat flux, while the SST is strongest in late summer. The atmospheric response is barely above the noise level, suggesting that the ensemble size should be increased. The mean response (red line) for the SST differs for the two ensembles. The reason is probably that the manual integration began on September 1 compared to the November 1 start for the VDS ensemble. We will explore this issue in future simulations.

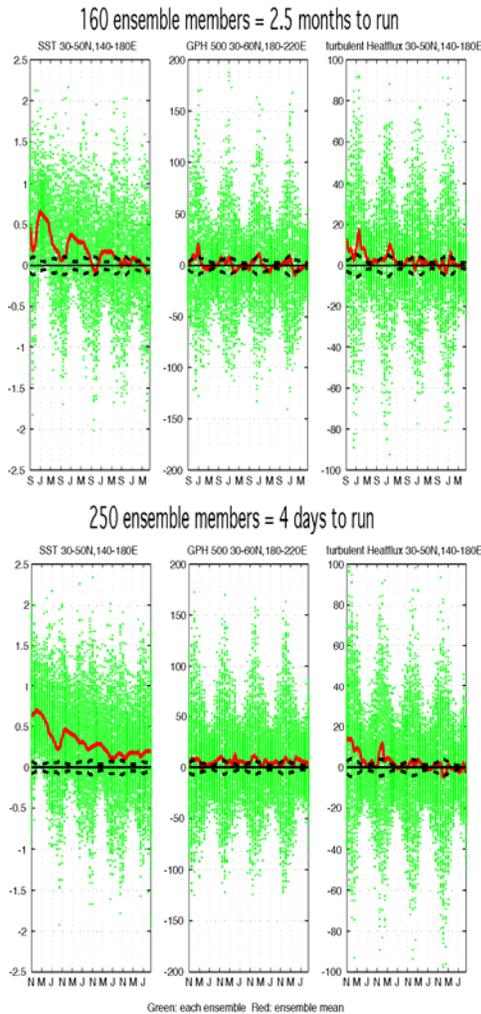


Figure 3. Average SST, Geopotential height and heat flux from ensembles conducted manually at NCAR (top) and with VDS on TeraGrid (bottom).

The figure also notes the time taken for the two studies: 2.5 months (75 days) for 160 integrations in the case of the manual ensemble, vs. 2.5 days for 250

integrations in the VDS case. Thus we achieved an improvement of a factor of 45 in turnaround time when using VDS relative to the manual case. Interpreting this result, we note that both ensembles could have been executed faster in different circumstances: the VDS ensemble could have been completed in less time if we had used multiple sites, while the manual ensemble was hindered by a limit of at most four jobs at once in the queue on the platform used, as well as by competition for resources on the single platform, which created long wait times between integrations. However, we do feel that the automation achieved with VDS played a significant role in accelerating the ensemble generation process.

6. Related work

We have referred in the text to other ensemble studies of climate variability. We note briefly some other approaches to managing ensemble studies.

Many systems have been developed to manage the execution of large number of systems tasks, e.g., SETI@home [17] and Condor DAGman (which we use in this work).

Nimrod [18,19] simplifies parameter studies by allowing users to use a declarative language to specify which parameters should be varied and how, in order to explore a parameter space. VDL’s declarative syntax achieves a similar goal but permits the specification of a richer set of workflows. APST is another system specialized for parameter studies [20].

Also relevant is work on coordination languages and workflow systems, although we are not aware of projects that have used those systems to manage large ensembles of parallel program executions, as here.

7. Conclusions

We used the GriPhyN Virtual Data System to construct a workflow system for performing ensemble climate simulations. The automated results are similar to those done manually but can be performed far more quickly due to the automation of many time-consuming management tasks.

FOAM’s speed means that the integration performed here could be done quickly. We have focused so far on the interannual response to SST anomalies, but daily response is also of interest. The high daily variability in the mid-latitude regions means that even larger ensembles will be required. In order to further simplify the use of ensemble integrations, it

will be necessary to bring more parts of the workflow, such as the generation of analysis plots such as those in Figure 3, into the VDS-based workflow system. We may also need to further develop our planners to allow for more flexible site selection, choice of schedulers for various parts of the workflow, and so forth.

Acknowledgments

We thank the entire VDS support group, and in particular Mike Wilde and Jens Vöckler, for their help with the use of VDS and with the VDS installation at the University of Chicago. This work was supported in part by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under contract W-31-109-Eng-38 and by the NSF Middleware Initiative's GRIDS Center project.

References

- [1] I. Foster, I., J. Voeckler, M. Wilde, Y. Zhao, "Chimera: A Virtual Data System for Representing, Querying, and Automating Data Derivation", *14th Intl. Conf. on Scientific and Statistical Database Management*, 2002.
- [2] Drake, J., Foster, I., Michalakes, J., Toonen, B., and P. Worley, P., "Design and Performance of a Scalable Parallel Community Climate Model", *Parallel Computing*, 21(10), 1995, pp. 1571-1591.
- [3] Kiehl, J., Hack, J., Bonan, G., Boville, B., Briegleb, B., Williamson, D., Rasch, P., "Description of the NCAR Community Climate Model (CCM3)", *National Center for Atmospheric Research Tech Note*, NCAR/TN-420+STR, Boulder, CO, 1996
- [4] Jacob, R. "Low Frequency Variability in a simulated atmosphere ocean system", Ph.D. Thesis, *University of Wisconsin-Madison*, 1997.
- [5] Bettge, T., J. W. Weatherly, W. M. Washington D. Pollard, B. P. Briegleb, W. G. Strand Jr., "The NCAR CSM Sea Ice Model", National Center for Atmospheric Research Tech Note., NCAR/TN-425+STR, Boulder, CO, 1996
- [6] Liu, Z, Wu, L., Gallimore, R., Jacob, R., "Search for the origins of Pacific decadal climate variability", *Geophys. Res. Lett.*, 29, doi10.1029/2001GL013735, 2002, p.1401.
- [7] Zhao, Y., Dobson, J., Foster, I., Moreau, L., Wilde, M., "A Notation and System for Expressing and Executing Cleanly Typed Workflows on Messy Scientific Data", *SIGMOD Record*, 34 (3). 2005, pp. 37-43.
- [8] Deelman, E., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Blackburn, K., Lazzarini, A., Arbree, A., Cavanaugh, R., Koranda, S., "Mapping Abstract Complex Workflows onto Grid Environments", *Journal of Grid Computing*, Vol.1, no. 1, 2003, pp. 25-39.
- [9] Thain, D., Livny, M., "Building Reliable Clients and Services", *The Grid: Blueprint for a New Computing Infrastructure* (2nd Edition), Morgan Kaufmann, 2004.
- [10] Frey, J., Tannenbaum, T., Foster, I., Livny, M. and Tuecke, S., "Condor-G: A Computation Management Agent for Multi-Institutional Grids", *Cluster Computing*, 5 (3). 2002, pp.237-246.
- [11] Kushnir, Y., Robinson, W. A., Blade, I., Hall, N. M. J., Peng, S., Sutton, R. "Atmospheric GCM response to extratropical SST anomalies: Synthesis and evaluation." *J. Climate*, 15, 2002, pp. 2233-2256.
- [12] Peng, S., Robinson, W. A., Hoerling, M. P., "The modeled atmospheric response to midlatitude SST anomalies and its dependence on background circulation states", *J. Climate*, 10, 1997 pp. 971-987.
- [13] Yulaeva, E., Schneider, N., Pierce, D., Barnett, T., "Modeling of North Pacific climate variability forced by oceanic heat flux anomalies", *J. Climate*, 14, 2001, p.4027.
- [14] Sutton, R., Mathieu, P. P., "Response of the atmosphere-ocean mixed layer system to anomalous ocean heat flux convergence", *Quart.J.Roy.Meteor. Soc.*, 128, pp.1259-1275.
- [15] Liu, Z., Wu, L., "Atmospheric Response to North Pacific SST; The Role of Ocean-Atmosphere Coupling", *J. Climate*, 17, 2004, pp.1859-1882.
- [16] Chervenak, A., Foster, I., Kesselman, C., Salisbury, C., Tuecke, S., "The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets". *Journal of Network and Computer Applications*, 23. 2000, pp. 187-200.
- [17] Anderson, D.P., Cobb, J., Korpella, E., Lebofsky, M. Werthimer, D., "SETI@home: An Experiment in Public-Resource Computing." *Comm. ACM*, 45 (11), 2002, pp.56-61.
- [18] Abramson, D., Sosic, R., Giddy, J., Hall, B., "Nimrod: A Tool for Performing Parameterised Simulations Using Distributed Workstations". *5th IEEE Intl. Symp. on High Performance Distributed Computing*, 1995.
- [19] Abramson, D., Giddy, J., Kotler, L., "High Performance Parametric Modeling with Nimrod/G: Killer Application for the Global Grid?", *Intl. Parallel and Distributed Processing Symp.*, 2000, pp. 520-528.
- [20] Casanova, H., Berman, F. "Parameter Sweeps on the Grid with APST". *Grid Computing: Making the Global*

Infrastructure a Reality, John Wiley & Sons Publisher, Inc.,
2003.