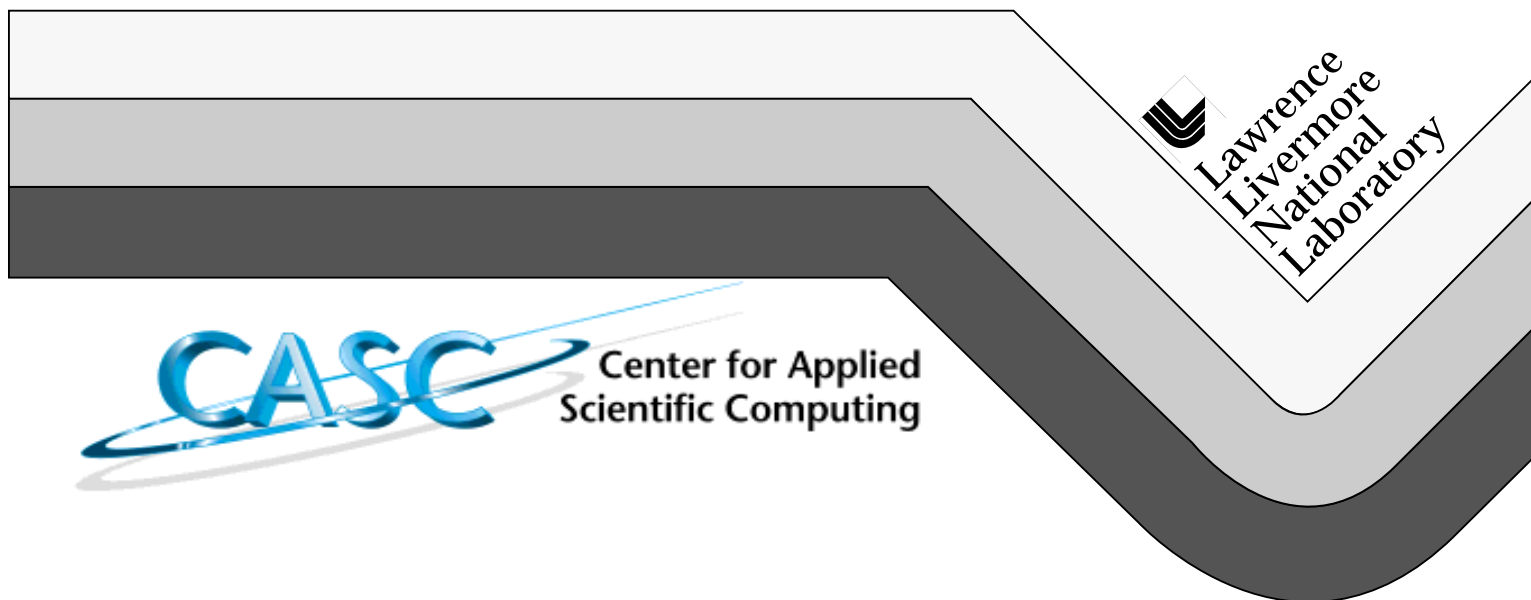


Climate System Modeling on Massively Parallel Systems: LDRD Project 95-ERP-47 Final Report

A.A. Mirin
W.P. Dannevik
B. Chan
P.B. Duffy
P.G. Eltgroth
M.F. Wehner

Lawrence Livermore National Laboratory



DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This report has been reproduced
directly from the best available copy.

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (615) 576-8401, FTS 626-8401

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161

Climate System Modeling on Massively Parallel Systems: LDRD Project 95-ERP-47 Final Report*

A. A. Mirin[†] W. P. Dannevik[†] B. Chan[†] P. B. Duffy[†] P. G. Eltgroth[†]
M. F. Wehner[†]

1 Purpose

Global warming, acid rain, ozone depletion, and biodiversity loss are some of the major climate-related issues presently being addressed by climate and environmental scientists. Because unexpected changes in the climate could have significant effect on our economy, it is vitally important to improve the scientific basis for understanding and predicting the earth's climate. The impracticality of modeling the earth experimentally in the laboratory together with the fact that the model equations are highly nonlinear has created a unique and vital role for computer-based climate experiments. However, today's computer models, when run at desired spatial and temporal resolution and physical complexity, severely overtax the capabilities of our most powerful computers. Parallel processing offers significant potential for attaining increased performance and making tractable simulations that cannot be performed today.

The principal goals of this project have been to develop and demonstrate the capability to perform large-scale climate simulations on high-performance computing systems (using methodology that scales to the systems of tomorrow), and to carry out leading-edge scientific calculations using parallelized models. The demonstration platform for these studies has been the 256-processor Cray-T3D located at Lawrence Livermore National Laboratory.

2 Technical Progress

The climate models with which we are working are our enhanced versions of the atmospheric general circulation model developed at UCLA [1] and the oceanic general circulation model developed at the NOAA Geophysical Fluid Dynamics Laboratory [2]. We had previously converted both models to run on massively parallel computing systems using a domain-decomposition, message-passing approach, with portability between computer architectures being attained through preprocessing techniques [3].

2.1 Code Development and Optimization

The first step was to port both codes to the Cray-T3D. We then focused on improving interprocessor communication performance. To accomplish this we implemented a message-passing layer on top of the T3D's Shared Memory Communications Library (SHMEM); that is, we disguised the SHMEM calls as message-passing calls and installed them in

*Work performed under the auspices of the U.S.D.O.E. by LLNL under Contract No. W-7405-ENG-48.

[†]Lawrence Livermore National Laboratory, Livermore, CA.

our Macro Interface for Communication and Allocation (MICA) [4], which is the principal vehicle for attaining message-passing in both the oceanic and atmospheric models. Use of the existing message-passing programming model helped keep source code changes to a minimum. We also instituted SHMEM collective calls in the ocean code (none were needed in the atmospheric code). We found that the use of SHMEM improves communication throughput by a factor of 3 to 10 over the default PVM, and that for certain parameter regimes with larger processor configurations, the ocean code could execute several times faster with SHMEM. Speedup of the atmospheric code was more modest because of much better communication efficiency a priori.

We next undertook a comprehensive evaluation of oceanic and atmospheric code nodal performance on the T3D using the *Apprentice* analysis tool. This enabled us to identify those areas of code that were taking the most time. We removed bottlenecks where possible, but oftentimes performance degradation was due to the small cache on the T3D. As part of this evaluation process, we investigated Cray’s internal benchmarking library *benchlib*, which contains fast, but possibly less accurate versions of transcendental functions. We found that the atmospheric code in particular could execute substantially faster ($\sim 20\%$) using this library.

The focus then shifted to removing the major sources of load imbalance, most notably the longitudinal Fourier filters, which are applied only at the higher latitudes. Using a complicated communication stencil, we redistributed the filtering workload among all the processors. The algorithm for doing this is general and flexible, working for all processor decompositions and extending easily to other variables to be filtered [5]. This, together with other improvements, led to a factor of two-to-thirty improvement in the filtering. Because the scaling of filtering with mesh resolution is less favorable than the scaling of the remaining parts of the calculation (with respect to mesh resolution), this improvement is critical, particularly for higher resolution simulations.¹

2.2 Code Performance

We have carried out scalability studies in which both the problem resolution and the number of processing elements have been varied [6]. Figure 1 shows the performance of the atmospheric code on both the LLNL Cray-T3D and the NASA Ames IBM-SP2, with the NERSC Cray-C90 as a reference point. On a per-processor basis, the SP2 is faster than the T3D, as expected, but the T3D demonstrates superior parallel scaling as seen by comparing the concavities of the two curves. Using 256 processors on the Cray-T3D, the atmospheric code operates 12 times faster than on a single Cray-C90 processor for a 2-deg \times 2.5-deg \times 9-level resolution. The oceanic code (performance not shown) exhibits perfect parallel performance for up to 256 T3D processing elements when run with a 1/2-deg \times 1/2-deg \times 15-level mesh [7].

2.3 Scientific Studies

Scientific studies have been carried out using both the atmospheric and oceanic codes.

2.3.1 Simulations of the Atmosphere A standard means for validating models of atmospheric general circulation is provided by the Atmospheric Model Intercomparison Project (AMIP). The test problem is defined by simulation of the 10-year period from

¹Since the conclusion of this project, we have obtained even better filtering performance through use of Fast Fourier Transforms instead of convolutions.

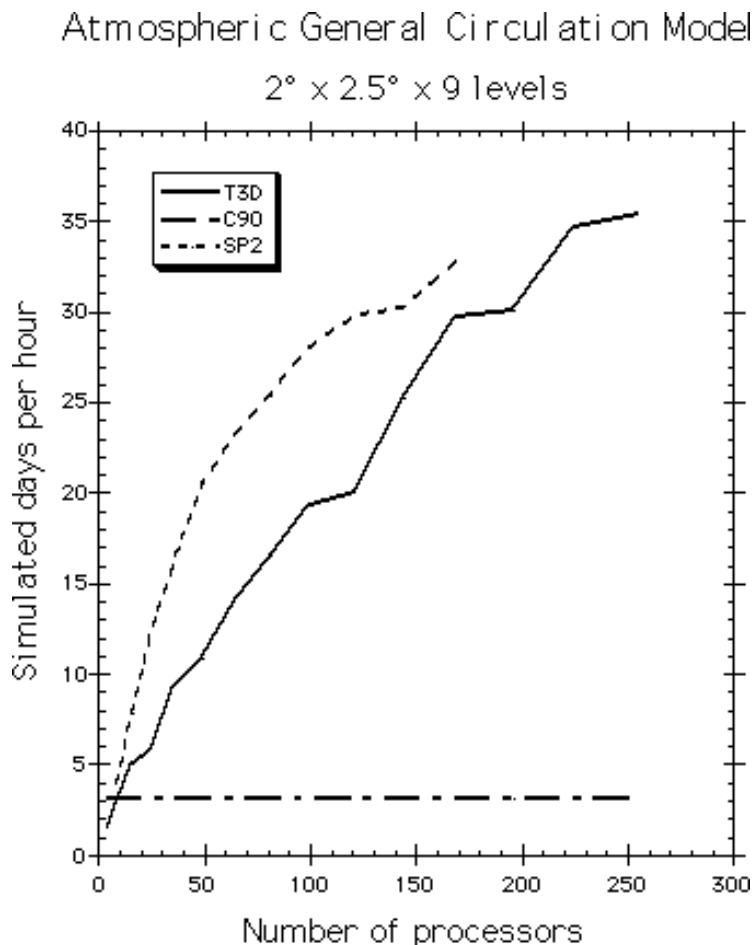


FIG. 1. Performance of the Atmospheric General Circulation Model on the Cray-T3D, IBM-SP2 and Cray-C90, in units of simulated days per machine hour, on a 2-deg (latitude) × 2.5-deg (longitude) × 9-level mesh. The performance on the massively parallel processors is roughly comparable to that of an efficiently multitasked code on the Cray-C90.

1979 to 1989, with prescribed sea surface temperatures. Over thirty atmospheric models worldwide have been compared using this test problem.

Our acquired parallel processing technology has enabled us to execute an ensemble of twenty such realizations and to calculate the model variability of the atmosphere [8]. This provides an estimate as to the natural variability of the atmosphere, without which we could not assess human-induced climate change. We believe this ensemble to be the largest yet performed by any one group. Figure 2 shows the mean surface temperature during the northern hemisphere winter months. A calculation of the average standard deviation of the ensemble mean surface temperature (not shown) indicates that the variation is largest over regions of sea ice and snow-covered land, suggesting a radiative feedback mechanism.

2.3.2 Simulations of the Ocean We have been studying the carbon cycle using the oceanic code coupled with a parallel oceanic biogeochemistry model. This combination provides the most sophisticated carbon isotope treatment in an ocean code to date [9]. One use of this coupled model is to estimate the degree of uptake of anthropogenic (human-produced) carbon dioxide of the ocean. Figure 3 shows a column inventory (i.e., the amount contained in a vertical column) of anthropogenic carbon dioxide as a function of latitude and longitude. This calculation takes into account estimates of atmospheric carbon dioxide emanating from human activities. This computation led to our discovering a heretofore unobserved flux of carbon-14 from the ocean to the atmosphere induced by fossil fuel burning.

3 Conclusions

Our plan was to undertake an ambitious program in optimization, proof-of-principle and scientific study. Those goals have been met. We are now regularly using massively parallel processors such as the Cray-T3D for scientific study of the ocean and atmosphere, and preliminary parallel coupled ocean/atmosphere calculations are being carried out as well. Furthermore, our work suggests that it should be possible to develop an advanced comprehensive climate system model with performance scalable to the teraflops range.

References

- [1] A. Arakawa, *Design of the UCLA General Circulation Model*, Technical Report No. 7, Dept. of Meteorology, UCLA (1972).
- [2] R. Pacanowski, K. Dixon and A. Rosati, *The GFDL Modular Ocean Model Users Guide, Version 1.0*, GFDL Ocean Group Tech. Report No. 2, GFDL/NOAA, Princeton (1991).
- [3] A. A. Mirin, J. J. Ambrosiano, J. H. Bolstad, A. J. Bourgeois, J. C. Brown, B. Chan, W. P. Dannevik, P. B. Duffy, P. G. Eltgroth, C. Matarazzo and M. F. Wehner, *Climate System Modeling using a Domain and Task Decomposition Message-Passing Approach*, Comput. Phys. Commun., 84 (1994), pp. 278–296, LLNL Report UCRL-JC-116232.
- [4] J. C. Brown and A. A. Mirin, *MICA, a Facility to Achieve Portability for Message-Passing and Dynamic Memory Management in Fortran*, NERSC Buffer, 18 (1994), pp. 12-18.
- [5] A. A. Mirin, D. E. Shumaker and M. F. Wehner, *Efficient Filtering Techniques Applied to a Finite-Difference Atmospheric General Circulation Model for Massively Parallel Computers*, to be submitted to Parallel Computing.
- [6] M. F. Wehner, A. A. Mirin, P. G. Eltgroth, W. P. Dannevik, C. R. Mechoso, J. D. Farrara and J. A. Spahr, *Performance of a Distributed Memory Finite Difference Atmospheric General Circulation Model*, Parallel Computing, 21 (1995), pp. 1655–1675, LLNL Report UCRL-JC-120678.
- [7] M. F. Wehner, A. A. Mirin, J. H. Bolstad, U. E. Creach, P. B. Duffy, W. P. Dannevik, P. G. Eltgroth, C. M. Matarazzo and B. Chan, *Climate Systems Modeling on Massively Parallel Computers at Lawrence Livermore National Laboratory*, Next Generation Environmental Models Computational Methods Workshop, Environmental Protection Agency, Bay City, MI (1995), LLNL Report UCRL-JC-121018.
- [8] M. F. Wehner and C. C. Covey, *Description and Validation of the LLNL/UCLA Parallel Atmospheric GCM*, LLNL Report UCRL-ID-123223 (1995).
- [9] P. B. Duffy, J. Amthor, K. Caldeira, P. S. Connell, D. E. Kinnison, J. Southon and D. J. Wuebbles, *The Global Budget of Bomb Radiocarbon*, submitted to Climatic Change (1996), LLNL Report UCRL-JC-120675.

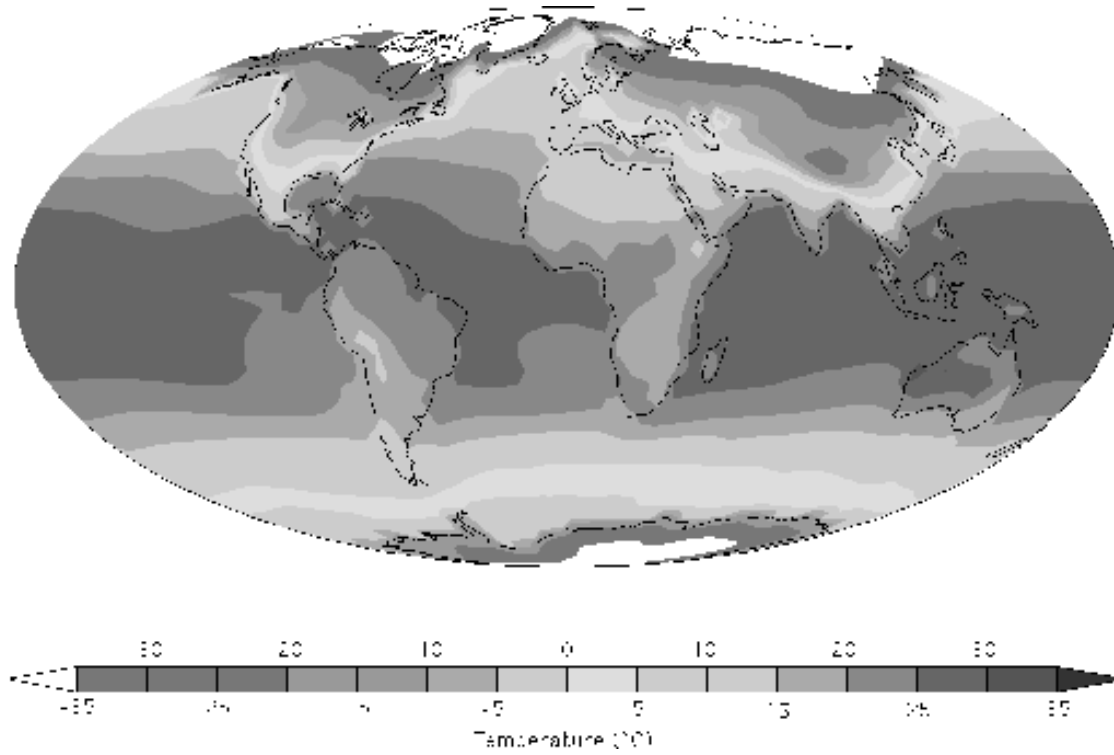


FIG. 2. *Mean surface temperature during the northern hemisphere winter months, as computed by our atmospheric general circulation model.*

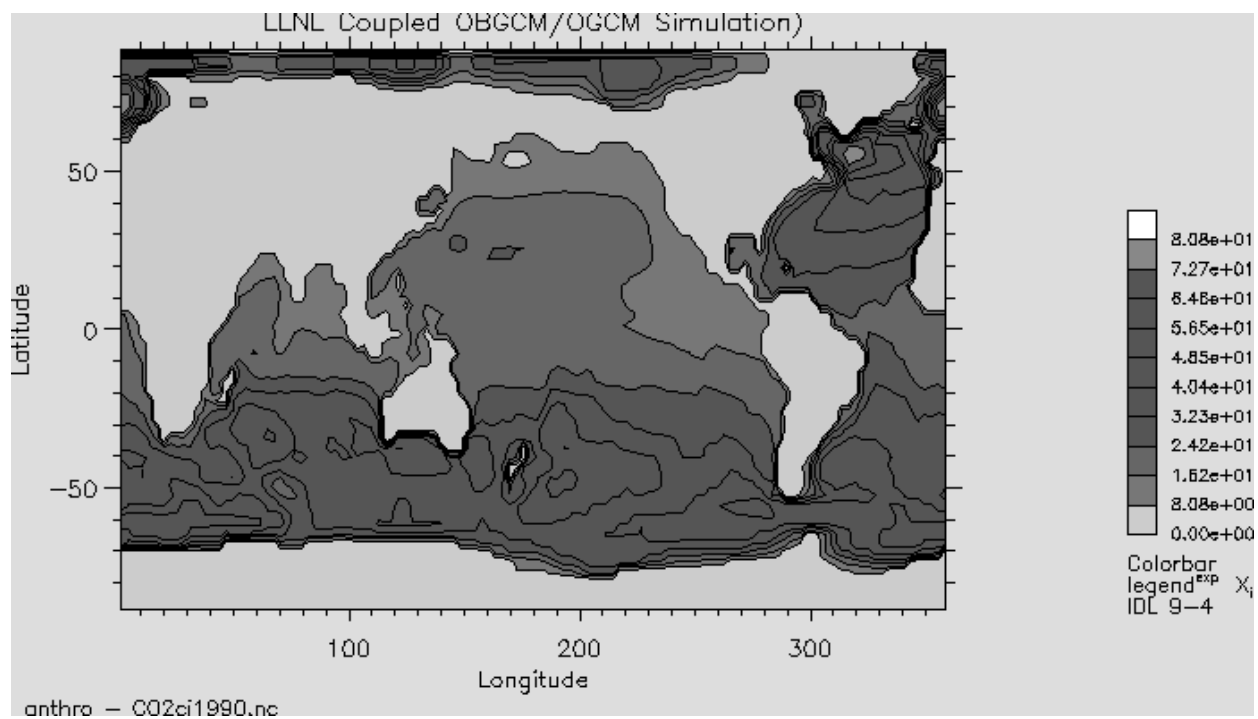


FIG. 3. Column inventory of anthropogenic carbon dioxide (mol/m^2), as computed by our coupled oceanic circulation/biogeochemistry model.

Technical Information Department • Lawrence Livermore National Laboratory
University of California • Livermore, California 94551

