Strategic Plan for CHAMMP-ANL/ORNL Model Development Effort

J.B. Drake and I. Foster

May 31, 1995

Abstract

In support of the Model Development Effort of the CHAMMP program, two projects are being pursued. First is the continued support of the Parallel Community Climate Model, PCCM2, and follow on atmospheric climate models. Previous funding has supported the development of this climate model for scalable, distributed memory computers. Maintenance and incremental modification of this code to the changing computing environments will be necessary so that an optimal implementation is available for climate researchers. Major modifications to the PCCM2 will result in a new atmospheric code that will be required for the second project. The second project proposed involves the coupling of the PCCM2, or follow on model, with the Parallel Ocean Program (POP) and other climate component models. The development of a coupled model running on massively parallel computers is a major goal of the CHAMMP program in the 1994-97 time frame. The development of the coupled model will be in collaboration with the climate scientists at the National Center for Atmospheric Research who developed the CCM2 and our colleagues at Los Alamos who have worked with the Princeton Geophysical Fluid Dynamics Laboratory on the development of POP. ANL and ORNL's role will be to implement these models on the Intel Paragon XPS150 and other massively parallel computers and to provide support for simulations required in the development of a coupled model.

Use of the models that are developed for climate research will be supported on the Intel Paragon at Oak Ridge National Laboratory. This development work is closely related to research proposals to improve existing climate models. The proposal includes the following tasks:

- support for PCCM2 on the Intel Paragon;
- enhancements to the portability of PCCM2;
- optimization of computation and I/O for POP on the Intel Paragon;
- development of software for coupling CCM2 and oceanic general circulation models in a message passing environment; and
- development of a coupled version of POP and a modified PCCM2.

These tasks will be addressed in collaboration with researchers from the National Center for Atmospheric Research and from Los Alamos National Laboratory.

1 Purpose

These projects are intended to support the development of a coupled ocean-atmosphere-terrestrial climate simulation system on parallel supercomputers. The project is one component of an effort to develop an advanced climate model with throughput capability significantly greater than that of current models and with more comprehensive representation of key climate processes. The CHAMMP program is described in the document

Building an Advanced Climate Model: Program Plan for the CHAMMP Climate Modeling Program, DOE/ER-0479T, (1990) This plan supports the CHAMMP program and describes the research approach and interdisciplinary, interagency team for studying modeling issues of a global climate system. The CHAMMP program is designed to accelerate the identification, evaluation, and implementation of long-range energy policies that may be needed to improve global standard-of-living while assuring environmental quality. An interdisciplinary team at each of the three national research institutions, Oak Ridge National Laboratory, Argonne National Laboratory and the National Center for Atmospheric Research, has been established and will jointly perform research in parallel algorithms for atmospheric models, advanced numerical methods, and the evaluation of numerical methods. The projects proposed are focused around the improvement and use of a parallel version of the Community Climate Model (PCCM2) which is available on ORNL's Intel Paragon and other message passing distributed memory computers. A project to couple the PCCM2 with the Parallel Ocean Program, developed by Los Alamos and the Princeton Geophysical Fluid Dynamics Lab has also been undertaken.

The long range goal of these development projects is to provide the advances in numerical methods and parallel processing techniques required to support the modeling of the atmospheric component of the climate system on advanced computing architectures, i.e., massively parallel computers. The ultimate worth of the climate models will depend on the inclusion of key climate processes and the ability to interpret model results and compare with earth climate data. The project, when combined with other DOE research efforts in the CHAMMP program, has the ultimate goal of developing the technology required to produce an advanced climate model with improved resolution and representation of physics that will be appropriate for use in addressing national energy issues on a firm scientific basis.

2 Background

The latest version of the NCAR Community Climate Model, CCM2, is in wide use by researchers not only for climate and weather studies but also for a wide class of studies including pollutant transport, acid rain, and ozone distribution. Every other year NCAR convenes a two-week workshop in which graduate students and junior faculty receive training on the use and application of the CCM [1]. In the global modeling community it serves the function of a model that is easily coupled with other modules yet computationally efficient to allow parameter studies and sensitivity analysis [6].

The parallel versions of the model have been developed by the CHAMMP program at ANL and ORNL in collaboration with NCAR. Two major efforts at parallel implementations of the CCM are now complete. An effort at NCAR has constructed an implementation in a data parallel style appropriate for machines such as the Thinking Machines CM-5. Another effort by researchers at Argonne and Oak Ridge national labs has constructed a distributed memory, message passing version of the model. The target machines of this effort are the Intel Paragon, IBM SP2, and Cray T3D. A public release of this code along with documentation is now available.

The coupling project follows and complements the NCAR Climate System Model (CSM) project which aims to develop a complete coupling framework and to provide low resolution models of both atmosphere and ocean. Significant modifications to the CCM2 are being made at NCAR as a result of the CSM, with major code restructuring to accommodate a general "flux box" style coupling, new land surface models and enhanced physics to improve the surface values of model fields. In addition, a major algorithmic change is being considered for inclusion in the atmospheric model. The new atmospheric

model, while a direct descendant of the CCM2, is no longer the same model either in algorithmic detail or in simulated climate. For the coupling studies it is understood that a new atmospheric model will take the place of CCM2 and in our development of the MPP version, that this new model will replace PCCM2.

3 Current Projects

3.1 Modifications to PCCM2

The PCCM2 will be upgraded and maintained as necessary to support climate research on current parallel systems. Our goal is the production use of the Community Climate Model, CCM2, on both the Intel Paragon and on the Thinking Machines CM-5. New computing systems, such as the IBM SP2 and CRAY T3D, will be included as appropriate to support climate modeling on massively parallel computers. This will provide important input to the manufacturers of massively parallel computers as well as the HPCC computing centers. In addition, the availability of the model on a number of platforms will encourage climate research to begin using massively parallel computers.

The ability to port the PCCM2 to new machines is important to this development project. Ideally, we would like a single version of the code able to run on multiple platforms of different types. As a first step in this direction, we have isolated I/O and message passing structures in low level compatibility libraries. An implementation using the standard Message Passing Interface (MPI) [4] provides portability across virtually all MPPs and workstation systems. The Parallel Virtual Machine (PVM) system [2] is also being used, offering similar benefits for debugging and code development and testing in a workstation environment as well as execution on a wide class of massively parallel machines.

A significant remaining obstacle to portability is that PCCM2 uses different data and code structures to the CCM2 code from which it was developed. These differences have hindered the timely deployment of a production quality PCCM2, as each new version of CCM2 requires a significant recoding, testing, and validation process. In order to improve the efficiency with which CHAMMP human resources are applied to parallel CCM2 development, we will investigate the techniques required to develop a single CCM2 code that is able to achieve good performance on both vector and MPP computers. While this approach is known in principle to be feasible (it has been adopted by ECMWF, for example), we need to determine how to achieve it in the context of CCM2. One proposed activity in this area is the holding of a workshop on model portability; others include investigations of code structuring techniques, the development of a prototype common code, and (assuming good results are obtained) the eventual transition of this code to the production version of the follow-on OMEGA model.

The PCCM2 is being used for a high resolution atmospheric climate study. We are attempting to answer the question, "At what resolution does the simulated atmospheric climate converge." A T170 study is underway. On 1024 processors of the XPS150 the PCCM2 can simulate 1 day in about 15 minutes. Thus a 1 month simulation requires 8 hours of computing time. The standard tuning run is from September to the end of January in order to get a reasonable January average. This 153 day simulation can be done in 40 hours on the XPS150. We forsee several tuning runs and then a longer 2-3 year simulation. A two year simulation will require over two weeks of dedicated machine time.

3.2 Development of Coupled Models

Previous to this year, we had assumed that the PCCM2 itself would be used for the coupling work. But in an important interagency collaborative agreement that emerged

this FY, CHAMMP and NCAR have agreed to work very closely on the study of the coupled climate system. This means that the new developments of the NCAR CSM program will be available to CHAMMP and that the atmospheric model of choice will be the model which follows CCM2 rather than CCM2 itself.

At ORNL, the CHAMMP MDE coupling project provided an initial implementation of POP on the Intel Paragon and investigated the logistics of running a coupled model on the Intel Paragon. It was found that a message passing version of POP can be run effectively on the Paragon. Work with POP has now moved into another phase where we are updating the code to be compatible with the current LANL version and developing the input routines to use real ocean topography on the pole-shifted orthogonal grid. We are also investigating the use of new Paragon High Performance Fortran (HPF) compilers that will allow for a single version of the code to exist between LANL and ORNL. The immediate goal is to provide the option to the CHAMMP program of using POP on the Paragon for simulations in support of the coupling efforts.

ORNL will work closely with both LANL and NCAR on the development of a coupled model. First, the ocean code, POP, was ported to the Paragon. In conjunction with Climate System Modeling project (CSM) at NCAR a general parallel flux coupling scheme is being developed and we are implementing this on the Paragon. The POP code and the PCCM2 will be modified to accept input and to channel output through this flux coupler.

Following the NCAR design, the flux box is based on a message passing paradigm. On the Paragon, this requires communication between two compute partitions, if the ocean code is running in one and the atmosphere in another. This simple configuration has proved difficult to support. Our first exercise has been to run the PCCM2 with the flux box providing sea surface temperatures which the box reads from a file. With the box running as a master PVM process, this is possible. But due to a limitation of the Paragon OS, PVM has difficulty spawning two partitions. We are working with the PVM developers to provide this capability. An alternative solution using the MPI communicator facility is also being investigated.

Features of the coupling software will include a conservative flux formulation, accurate, advective interpolation with daily time averaging of fluxes input to the ocean code. The flexibility to run on multiple machines or on multiple partitions of the same parallel computers is also being built into this software. Initially, the PCCM2 and POP will be run in a semi-coupled mode where the POP Sea Surface Temperatures are used to drive the PCCM2. Then with atmospheric climatology established, the ocean will be driven using the PCCM2 wind stress. For the fully coupled runs we anticipate that the PCCM2 will be changed into the fully semi-Lagrangian OMEGA model that NCAR is developing. This will allow the CHAMMP program to contribute more significantly to the science agenda elaborated by the CSM. The Oak Ridge contribution will focus on the use of message passing parallel computers to provide cycles for the numerical experiments. The high resolution ocean simulations will compliment the low resolution studies undertaken by the CSM project at NCAR.

Computational experiments will be run to establish the climatology of the new models and to investigate any problems or inconsistencies that might arise.

4 Technical Progress

4.1 Staff and Lead Responsibilities

ORNL:

J. B. Drake (0.2 FTE) Atmospheric model adaptation and support;

B. D. Semeraro (0.5 FTE) POP adaptation and support;
J. A. Holmes (0.5 FTE) Coupling runs and flux box development.

ANL:

J. Michalakes (0.3 FTE) Atmospheric model structuring;
J. Mogill (0.5 FTE) Atmospheric model adaptation and support.

4.2 Model Development Schedule for FY 1995

First a qualifier. Since this is a collaborative project and depends strongly on favorable outcome of research projects at NCAR in the CSM and on development projects at multiple sites, we can only estimate the completion date of each project. In the following, an "A" or "O" in the first column denotes tasks for which ANL or ORNL, respectively, are primarily responsible. Currently, ORNL is assigned primary responsibility for ocean model and flux box tasks, although based on previous experience it seems likely that there will be sharing of effort in these areas also.

The message passing version of the Community Climate Model, PCCM2, was completed, validated and released this FY. Emphasis for version 1.0 was on correctness and functionality. The released code was updated with the final "commented and cleaned up" CCM2 physics routines and monthly averaging history tape ability. In addition, the load balancing schemes were incorporated and message passing consolidated to facilitate the portability of the code between the Intel Paragon, IBM SP2, Cray T3D and networks of workstations using MPI or PVM. We do not plan to release any further optimization work or future versions of the PCCM2. Some optimization work will be necessary to support production runs and to adjust to system changes at various installations but this can be done in the course of these runs. Documentation in report form and HTML WWW format are available in the June timeframe.

The production use of the code for a high resolution (T170) climate simulation study was begun in FY95. Some adjustments were required in response to user input. In particular, some restructuring of the I/O was required as the Intel PFS system developed. Intel added another I/O mode to their parallel file system at our request. Mode M_ASYNC has now been incorporated in the history file output. The code was also multi-threaded for use on the Intel Paragon XPS/150 system that arrived at ORNL in January 1995. Tuning was performed on both the Intel Paragon and IBM SP2, and extensive benchmarking was performed to characterize performance on these systems [3].

The POP code was run on the Intel Paragon XPS/150 based on Fortran 77 with message passing. Scaling experiments were done and a performance study to determine the effectiveness of the "stencil" message passing. For high resolution cases, the message passing cost is small compared with the computational costs. Further blocking of the message passing was not seen as important at this time and work on the I/O routines has begun. (Dave Semeraro is performing this work and is coordinating with Matt Multrand at LANL.)

4.3 Model Development Schedule for FY 1996

With the shift in emphasis at NCAR away from the CCM2 model and toward the fully semi-Lagrangian OMEGA model, we will be spending much time updating the atmospheric code. This is the code that will eventually be used for coupled climate studies. A related

[1] Support from NCAR Core Group required
[2] Coordinated with LANL MDE

Expected	Atmospheric Model	Ocean Model	Flux Box
Jan 1	PLX22 physics integrated (O)		
Feb 1	Double precision default	Message passing POP run on XPS150 (O)	Simple PVM flux box run with PCCM2 (O)
Mar 1	Load balance short wave calculations (A)		
May 1	Portable, tunable message passing routines installed in PCCM2.1 (O)		
June 1	New I/O structures with monthly averaging (A)		
July 1	Validation of PCCM2.1 [1] (O)	HPF POP evaluation [2] (O)	PVM demonstration of flux box between two Paragons (O)
July 15	Release of PCCM2.1 and completion of UserGuide (A&O)	POP I/O structures on Paragon tested (O)	Multiple partitions (O)
Aug. 1	Publication of PCCM2.1 UserGuide (A&O)	POP validation on Paragon (O)	Interpolation routines POP $\leftarrow$ PCCM2 (O)
Sept 1	Begin incorporation of CSM model changes in PCCM2 [1] (A&O)	Begin POP mods for coupled runs (O)	Integration of full flux box with atmospheric code ,2[1] (O)
Sept 1	Portable code structure studies (A)		
Oct 1	Change land surface model [1]	Begin POP mods for coupled runs (O)	Testing of minimal coupling with full models (O)

Figure 1: Time Table for FY 95

task will address the question of how to structure this code so as to permit the use of a single portable model on both vector and MPP systems.

Development and optimization of the coupled model will continue with emphasis on establishing an efficient model with reasonable climatology. For this many simulations will be necessary of varying duration. Examination of the interacting processes in the coupled model will be necessary in short runs as well as detailed examination of simulations on the 10 - 1000 year time frame.

We also expect to begin interactions with CHAMMP science team members that require the resources of a parallel computer for atmospheric modeling.

Expected	Atmospheric Model	Ocean Model	Flux Box
Jan 1	Atmospheric model fully implemented (A&O)	Spin up runs performed in collaboration with LANL (O)	Integration of other component models, e.g., ICE [2]
Feb 1	Workshop on code portability (A)		
Mar 1	Begin fully coupled runs	Begin fully coupled runs	Begin fully coupled runs
May 1	Refinement, optimization, etc..		
Oct 1	Prototype portable vector/MPP code (A)		

Figure 2: Time Table for FY 96

4.4 Model Development Schedule for FY 1997

The coupled model should have been completed and either development of an improved model underway and/or a series of compute intensive climate runs underway. We will make every effort to work in concert with the CHAMMP Model Development Teams at LANL and NCAR and to be responsive to changes in the situation with respect to computer resources and scientific progress achieved by CHAMMP Science Team members.

5 Planning Assumptions

The two major planning assumptions made in this proposal are that

- computer resourses on MPP machines are available to both ORNL and ANL for development purposes with production level support available on the ORNL Paragon systems, and
- the NCAR CSM project will provide the necessary scientific expertise to establish climatology of a coupled model.

If current levels of support for computing systems are maintained the first assumption will hold.

6 Relationship to Other Projects

Links with other CHAMMP projects. This research is complementary to the CHAMMP research efforts to develop advanced atmospheric models on parallel computing platforms.

Ties with non-CHAMMP community. The close tie between our research and other efforts at NCAR provides the potential for a much improved PCCM2. Although our efforts in this proposal are concentrating on the underlying algorithms associated with the dynamical and numerical components of the model, we will strive to keep the PCCM compatible with the standardized physical parameterization interfaces adopted by the CCM and by most of the international modeling community [5]. As new parameterizations are proven in CCM2 under CHAMMP or under other programs such as NCAR's CSM they can be immediately implemented in PCCM2 with minimal technical effort. As PCCM2 is developed through the various phases of our proposal it will always be state-of-the-art rather that lagging a year or so behind relevant concurrent developments on conventional supercomputers.

Ties with ORNL Laboratory Director Research Projects. In order to support the type of computing environment conducive to research using large scale computational experiments the ORNL Laboratory Director is funding a project dealing with the analysis of large data sets. The CHAMMP project and global climate modeling in general is one focus and is providing example datasets for the project. Current work is focused on building analysis modules (filters) that can be interactively combined to perform a given analysis. These analysis and the data management tools are using the CCS equipment as target platforms. The staff of the ARM project are also providing consultation on this project.

References

[1] R. Anthes. 1986 Summary of Workshop on the NCAR Community Climate/Forecast Models. *Bull. Amer. Meteor. Soc.*, 67:194–198, 1986.

[2] A. Beguelin, J. Dongarra, A. Geist, R. Manchek, and V. Sunderam. A Users' Guide to PVM (Parallel Virtual Machine). ORNL Tech. Report ORNL/TM-11826, Oak Ridge National Laboratory, Oak Ridge, TN, 1991.

[3] J. Drake, I. Foster, J. Michalakes, B. Toonen, and P. Worley, Design and performance of a parallel Community Climate Model, *Parallel Computing*, 1995 (to appear).

[4] W. Gropp, E. Lusk, and A. Skjellum, *Using MPI: Portable Parallel Programming with the Message Passing Interface*, MIT Press, 1995.

[5] E. Kalnay, M. Kanamitsu, J. Pfaendtner, J. Sela, M. Suarez, J. Stackpole, J. Tuccillo, L. Umscheid, and D. Williamson. Rules for Interchange of Physical Parameterizations. *Bull. Amer. Meteor. Soc.*, 70:620–622, 1989.

[6] M. C. MacCracken and F. M. Luther. Projecting the Climatic Effects of Increasing Carbon Dioxide. DOE Tech. Report DOE/ER-0237, US Department of Energy, Washington, DC, 1985.