PROJECT DESCRIPTION

Introduction

We propose to assemble under the aegis of the School of Computational Science and Information Technology a team of computer scientists, mathematicians, and geophysical modelers to conduct an interdisciplinary program of research in: i) effecting modeling and algorithmic improvements to the state of the art in simulating global circulation with passive and active transport, and ii) adapting full application codes (from input to output, not just kernels) to global distributed memory, local hierarchical memory environments; and iii) developing and demonstrating global circulation software and broadly applicable partitioning and load balancing tools, and a problem solving environment hospitable to geophysics research in general and weather and climate prediction in particular.

This newly established interdisciplinary School spans three colleges – the College of Arts and Sciences, the College of Social Sciences, and the College of Engineering, which is common to both Florida A & M University and Florida State University. Its mission objectives are i) to educate scientists for the twenty-first century multidisciplinary workplace, ii) to integrate the latest research in areas of initial focus into a graduate curriculum, and iii) to develop a multidisciplinary research atmosphere which transcends department/discipline barriers and is conducive to the collaboration of scientists from national laboratories, industries and academia.

The principal investigator has a track record in conducting and managing interdisciplinary research in numerical mathematics, fluid mechanics and computer science (Hussaini 1995). He now proposes to assemble a team of geophysicists expert in atmospheric and oceanic sciences (Krishnamurti, O'Brien, Dewar, Navon, Chassignet, Busalacchi), computer scientists expert in parallel systems software, parallel algorithms, problem solving environments (Fox, Gallivan, Keyes, Mehrotra, van Engelen) and applied mathematicians expert in numerical methods and fluid physics (Erlebacher, Kopriva, Woodruff, Sussman) to attempt influential strides in the simulation of coupled models. This research program proposes to draw upon the expertise at the Laboratory for Hydrospheric Processes at the Goddard Space Flight Center, the atmospheric and climate sciences research at Los Alamos National Laboratory, and the computer science research at the Institute for Computer Applications in Science and Engineering at NASA Langley Research Center.

The objectives of our investigation reflect the diversity of the research team and the complexity of the scientific task, and include: i) modeling, particularly with regards to turbulent mixing; ii) discretization in space and time, with both p-type and h-type adaptivity, in a locally structured framework; iii) implicit coupling mechanisms, recognizing that there is little value in improving the accuracy of oceanic and atmospheric models within their individual thin domains if the results are dominated by either modeling or numerical errors committed along their vast common interface; iv) efficient parallel implicit solution algorithms, motivated by the multi-scale nature of the global circulation and the frequent desire to focus selectively on coarser scale features, and rendered challenging by the load imbalance caused by adaptivity and the all-to-all coupling which is formally present in any mathematical description of the inverse action of the implicit operator; v) cache-conscious and page-conscious programming implementations, given the pronounced sensitivity of the fastest processors to cache-residency and page-residency and the relative lack of programmer control over the details of interlevel memory mappings; vi) portable, scalable, message-passing implementations, given the volatility of individual hardware platforms (within the context of relatively converged general architectural characteristics); vii) parallel data base management, looking towards the use of the simulation capability we plan to develop in production mode, where the sizes of the input and output are comparable to the workingset size of the application itself.

c nic ro c

The ongoing research program at the Florida State University in the area of the coupled ocean-atmosphere-land surface system involves a critical mass of well-known faculty, postdoctoral fellows and students and is preeminent in the world. This program includes major thrusts in physics, dynamics, and data and computational issues. Current work, which concentrates on the prediction of tropical activity, deals with the issues of storm frequencies (tropical storms, hurricanes and typhoons) and periods and locations of their ocean basin preferences, given seasonal forecasts of the atmosphere and ocean. On shorter time-scale forecasting (from one to ten days), FSU has been engaged in major model and data refinement efforts using a high-resolution global model implemented on a parallel computer platform. Basic issues in model development have also been under consideration by FSU faculty, and methods for advancing model physics have already been articulated.

We briefly mention immediately below a few of the major current issues pertaining to seasonal and decadal climate, which will be the focus of the model development part of our research.

Seasonal Climate: Of interest are questions of precipitation rates, average temperatures, etc., which in turn place a premium on detailed simulation (Cocke and LaRow, 1999). Perhaps the state of the art is represented by the so-called AMIP (Atmospheric Model Intercomparison Project) studies (Gates et. al, 1999). This effort consisted of hindcasts (predictions over prior intervals for which observations exist) from identical conditions over the same historic interval by a collection of member atmospheric general circulation models. From this effort, much has been learned about the current skill of the prediction community and the quality of individual member models. An important recent development is the application of the so-called super-ensemble methodology (Krishnamurti et al., 1999) to the statistical analysis of the member models data. This method develops a statistical model based on the dynamical models that filters the output in an optimal fashion, with the result that the statistical forecast is superior both globally and locally to any of the individual dynamical models. One of the important challenges facing seasonal prediction is the migration of the super-ensemble methodology to fully coupled climate models, thereby generated true climate predictions to be used for the practical applications outlined above.

Another area of sensitivity on the seasonal scale is in the prediction of El Nino events. While ocean models have demonstrated increasing skill in recent years in hindcasting El Nino and ENSO (El Nino-Southern Oscillation) events, coupled models are still in their infancy with regards to such prediction (LaRow and Krishnamurti, 199) At the heart of the matter are subtle issues of model initialization and data assimilation during the model spin up stage, prior to the generation of the true forecast. Correctly accounting for ocean SST anomalies in this phase of model preparation is crucial to getting the proper phasing and amplitude of the El Nino event.

One approach to model initialization via data assimilation is ocean spin-up, where daily wind stresses and the observed SST fields for a decade are used to guide the initialization of the ocean model. This decade-long spin-up communicates the surface forcing to the deeper oceans. In physical initialization, a coupled assimilation technique of Krishnamurti et al. (1991, several), years of rainfall observations are incorporated into the atmospheric model. The structure of the moisture field, diabatic heating and divergence is improved, as well as the rotational wind in data-sparse areas and the definition of the trades, monsoon and the ITC .

Soil moisture initialization and prediction are critical issues. Land modules routinely carry soil moisture as a predictive variable, but little is known about the proper way to initialize it, or for that matter to what values it should be initialized (White et al, 199) Assimilation methods are being developed to address this problem.

a e Climate o e astin : There are slow components of the climate system which do suggest the possibility for useful calculations to be carried out for the decadal time scale forecasting. The emphasis here shifts importantly to getting the ocean dynamics correct because this is probably the most significant slowly varying climate component in the decadal climate problem. Within the ocean, the most critical variable is the SST that responds to a host of inputs of solar, atmospheric and oceanic origin. Computing SSTs accurately is thus a very broad climate problem, although two areas emerge as major current errors in climate models.

First, atmospheric models must necessarily parameterize cloud input to radiation, especially for multi-decade long climate studies (Krishnamurti and Bedi, 19 ; Krishnamurti and Bounoua, 199) Current parameterizations typically underestimate cloud albedo (reflectivity). The result is an abnormal solar radiation penetrating the ocean surface, leading to abnormal warming of the ocean in an integral sense. This radiation error appears to be compensated for by the diffusion error of the verdiffusive ocean models (that conduct heat to the ocean at rates far greater than ocean observations would support (Griffies et. al, 1999)) thereby generating smaller SST errors. There are ideas in the ocean modeling community about tackling the ocean diffusion problem (Bleck and Chassignet, 199 ; McDougall and Dewar, 199 ; Dewar and McDougall, 1999), which will be explored and improved on.

Our objective is to concentrate on developing high-accuracy solutions for the principal partial differential equations of the coupled models, on greater computational efficiency on current computer architecures, and on a problem-solving environment that permits easy incorporation of exogenously developed advances in phenomenological modeling as they become available. It is our conviction that as the fundamental conservation equations are enforced on successively finer resolutions, the need for ad hoc phenomenological closures for diffusion phenomena will decrease, and the sensitivity of the overall coupled models to these phenomenological terms will thereby decrease.

o m ements to t al e o :el Current basic geophysical approximations will be improved by the incorporation of modern turbulence models. These models are such that the large-scale computation automatically becomes a direct numerical simulation (DNS) on fine grids, a large-eddy simulation (LES) on intermediate grids, and a Reynold-averaged Navier-Stokes (RANS) computation on coarser grids. In geophysical codes presently lacking any subgrid-scale turbulence models altogether, this will boost resolving power mesh point as much as any other enhancement we will make. For example, fine-scale details crucial to the accuracy of contaminant tracking can thus be freed from overly diffusive, ad hoc mixing models based on the mean flow only.

etis ation :ss We will employ high-order finite difference and spectral discretizations in space, and low-storage Runge-Kutta in time (Canuto et al., 19 hieve) greater resolving power per mesh point. Attention to interface conditions between subdomains with nonconforming grids will permit locally adaptive refinment (van Rosendale and Thomas, 1995) without contaminating reflections. The use of adaptive refinment in atmospheric modeling, as proposed in the present research program, is new and will engender a major advance in this area.

A useful compromise between structured grids, for which efficient solvers based on multidimensional array data structures exist, and unstructured grids, which may be easily adapted to geometrically complex domains with highly spatially variable resolution requirements, are semi-structured grids, which sacrifice global logical translational invariance for local. This methodology has been applied to both steady-state and unsteady problems (uirk and Hanebutte, 199 ; uirk, 199) The state of the art in semi-structured telescoping models is in the nested regional forecast models that have been used by the NMC and ECMWF developed in association with FSU researchers (Krishnamurti, 1995). To attempt high spatial resolution regional forecasts, this model uses a nested grid approach in which each regional model variable is composed of a base and a perturbation (Cocke, 199 Ba)se values are obtained from a coarse grid global model; one may then solve for the perturbations at the next time step, typically using a semi-implicit time integration scheme. A relaxation technique is used to force the perturbations at the lateral boundaries to be zero. Since the time steps of the global and regional models may be different, linear time interpolation of the global variables is used. Currently only one-way nesting, from global to spectral, is used. This can lead to large errors near the lateral boundaries if the nesting period is too frequent. Two-way nesting to be developed under this proposal should help to alleviate this problem. Both global/local and intrinsically local spectral multidomain approaches more amenable to parallelism will be explored.

Sol tibn o: We *it*will em*pl*oy operator splitting to advance the first-order nonlinear terms in an explicit manner in physical coordinates, using a high-order RK scheme. A global system for the second-order terms, which would otherwise impose an unacceptable explicit limitation on the time-step will then be solved. We will be fully point implicit, so that stiff source terms arising from chemical reaction or energy deposition can be adapted to physical modeling needs without destabilizing the solver. We will employ a Krylov iterative scheme with a parallel multilevel preconditioner to solve the implicit system. There is a large variety of such methods, with a choice of many different Krylov accelerators (e.g., conjugate residual, GMRES, Bi-CGSTAB, TF MR) wrapped around many different domain-decomposed preconditioners. For parallelism, we will concentrate on the acobi-(blk additiv) form of Schwarz-style domain decomposition, which are subdomain-blocked acobi methods (Keyes, 199).

In the context of our coupled ocean-atmosphere models, we will choose the components of the global parallel preconditioner in a modular, object-oriented manner best adapted to the local physics. Physical intuition is employed to choose partitions and preconditioners which preserve or approximate well the strongest coupling, while sacrificing or coarse-graining the weakest. Interdisciplinary, problem-specific collaborations between applications experts, computer scientists, and applied mathematicians are required to identify and capture in some low-dimensional, low-complexity way the key deterrents to fast convergence, and to implement the best algorithm for the architecture, in keeping with the ultimate metric of accuracy per wall-clock time unit.

ist i m lementation: In pursuit of high performance on the coupled ocean-atmosphere models, we will exploit three forms of parallelism: (1) an outer task parallelism, permitting the atmospheric and oceanic domains to be processed concurrently, except for synchronization to enforce interfacial compatibility, ( ) parallelism within each domain, based on domain decomposition, and ( ) inner task parallelism de from lightweight threads and as automated by compilers for multifunctional unit processors. The outer task parallelism has a granularity of just two, and appears at the level of the ocean-atmosphere coupling. The inner task parallelism is likewise of small granularity, and does not enter the discussion on distributed implementation except insofar as it will permit the overlap of communication with useful computation whenever the hardware support exists. The most interesting parallelism is the parallelism within the atmospheric and oceanic domains, in which each processor is responsible for applying essentially *all* of the operations to *a t* of the data. This paradigm can be applied at almost any granularity.

o lem Sol oinment o n Climatel: The problem solving environment (PSE) coordinates numerical simulations of the atmospheric, oceanographic, geologic, and other sub-models and the sub-model boundary interface interactions involved. Sub-model interface interaction is an important issue in climate modeling. Atmospheric, oceanographic, and geologic models have different spatial and temporal scales, units of measurement, initial and boundary conditions, etc. In addition, simulation codes are based on a range of different numerical methods that demand specific grids and data structures, thereby severely hampering the development of sub-model data interchange routines. The PSE includes a workbench for interactive experimentation of parameter settings for each sub-model to relieve model designers from the burden of application (re)development. Parameter selections include choices of mathematical models, numerical methods, discretization methods, discretization densities, data structures, boundary interaction methods, geographic features to be modeled, modeling scenarios, and simulation visualization techniques. The workbench provides application scientists the ability to rapidly test new physical models, new numerical algorithms, or any combination of these applied to their particular problem. Model optimization and the generation and optimization of simulation codes is a function of several, often complementary, factors. For example, higher numerical accuracy often requires more processing time by the simulation code. Further, sub-model integration requires compatible data structures to enable effective data interchange. The workbench aims to automate the method selection, the optimization, and the code generation aspects of a climate model to a high degree.

The workbench user interface is based on computer algebra technology enabling scientists to conveniently enter, manipulate, and browse equations. An algorithm template database serves for the automatic selection of solution methods, provided by the properties of the algorithm and model characteristics. The selection and composition of templates is orchestrated by a knowledge-based system exploiting methods for case-based reasoning. For example, optimal discretization methods are determined by the target hardware architecture (e.g. shared or distributed memory system), memory latencies, network latencies, and cache/memory sizes.

An additional consideration is the generation of run-time data visualization routines by the workbench to visualize certain aspects of the evoluation of the climate model over time. The programming of visualization routines by hand would be very time consuming, because the routines have to be redevelopped for different types of grids and data structures.

The use of Ctadel (van Engelen, 199) in this project will permit important advances to be made quickly. Ctadel provides an integrated symbolic processing, compilation, automatic restructuring and performance optimization and text/web-page processing system. From a symbolic mathematical specification of the equations and numerical algorithms, it generates parallel code for the solution of PDEs on regular grids that is competitive with hand-optimized code. Ctadel has been used successfully in weather prediction and hydrology (van Engelen et al., 199) based on second order explicit algorithms and will be a solid foundation on which to build the PSE of the project.

duc tion

The School has the mandate and mechanism to develop a multidisciplinary graduate curriculum which integrates the latest research techniques and results into courses for web-based teaching and distance learning. Publishing agreements with commercial publishers are also being pursued. This mechanism will be exploited to develop specifically multidisciplinary courses in seasonal scale to decadal scale climate dynamics with emphasis on high-performance computing aspects.

I ct

It is clear that successful demonstration of a coupled ocean-atmosphere circulation model in a high performance environment will place a tool in the hands of the geophysical community that will permit revolutionary advances in our ability to understand and to predict our complex, ever-changing environment. In addition, the effort will create many software solutions that are transferable to other partial-differential-equations-based Grand Challenge applications of interest to the nation.