Petascale Computing: Its Impact on Geophysical Modeling and Simulation

NCAR Theme of the Year Workshop #2, Boulder, CO, 5-7 May 2008

Joseph Werne, Keith Julien, Bjorn Stevens, co-organizers

The goal of the workshop was to establish a roadmap for most-productive use of petascale computing systems for improving our knowledge of important geophysical dynamical processes. The intent of the meeting was to help the geophysical dynamics community position itself to make the best use of these resources when they are available. Given the large range of length and time scales accessible with such systems, before the meeting we anticipated significant changes would be required to models, algorithms, and analysis methods so that new discoveries would be possible. These pre-meeting expectations were confirmed and amplified by the presentations made at the workshop.

The workshop had five themes that explored different aspects of petascale computing and its potential impacts on geophysical modeling, simulations and scientific discovery. These included: 1) global-scale, 2) regional-scale, and 3) small-scale phenomena, 4) coupled dynamics, spanning more than one of these spatial regimes, and 5) logistical issues impacting scientific discovery with systems this large.

The workshop agenda, list of participants, their abstracts and presentations are available online at http://www.image.ucar.edu/Workshops/TOY2008/focus2/agenda.shtml. Important points made by participants are touched on below. For greater detail, their abstracts and presentations can be consulted online.

The major findings of the workshop, as expressed by participants through their presentations, comments, and ensuing discussions, were as follows:

1. Moore's Law is not sufficiently fast to keep pace with the near-future demands of Earth Systems Modeling (ESM). ESM is inherently an exascale problem, which remains several decades away. Rich Loft emphasized this in the workshop's opening presentation, in which he provided a detailed overview and `some' answers to the petascale computing challenge. Loft addressed many important issues, including computational challenges (e.g., stalled thread speeds and algorithmic scalability and parallelism), software complexity (e.g., the inherent interdisciplinarity of the earth system and model complexity), data challenges (e.g., complex workflows and data volumes that break tools), and human-resource problems (e.g., talent entrainment and training).

2. Supporting technologies (e.g., hard drive data transfer rates, analysis and visualization tools, etc.) mature at a slower pace than compute-engine capacity. The disparity increasingly taxes researchers who use high-performance systems, and by the time petascale systems are available, updated tools and different modes of analysis and discovery will be required. For example, with disk access speeds

and data storage capacity lagging raw processor performance, we are currently at a crossroads where it may make more sense to save only coarsely sampled data or subsets of data, and plan to recompute previously computed solutions to recover data that was not archived initially. The talks by John Clyne and Mark Rast presented data and trends that convincingly advocate shifts like these in numerical-simulation-based scientific research. Also, John Clyne's presentation illustrated a new wavelet-based compression, analysis, and visualization tool called VAPOR (www.vapor.ucar.edu) that employs progressive data-access strategies that have the potential to decrease the analysis and visualization demands on storage, communication, processing, and graphical analysis by orders of magnitude. In a complementary presentation highlighting the use of java-based software and an Ajax interface, David Yuen demonstrated how to perform real-time volume-rendering visualizations that could be captured remotely on platforms ranging from desktops to hand-held devices. Finally, in addition to shifting how we access and analyze simulation results, significantly increased system performance also offers an opportunity to reduce the end-to-end problem-completion cycle by claiming gains in simplified and automated code development. Damian Rouson addressed techniques to accomplish significant reductions in code-development time, defining several metrics for evaluating progress, and he posed the following questions: a) Are performance reductions expected or necessary in this new coding paradigm? and, if so, b) What performance reductions, if any, are acceptable given potentially significant savings in the code-development cycle.

3. The higher resolutions possible with these systems will be incompatible with existing subgrid-scale (SGS) parameterizations and some numerical techniques used in all of the community global-scale and regional-scale atmospheric models employed today.

    a. Regarding numerical methods, old trade-offs chosen decades ago when model resolutions were necessarily coarse are inappropriate with the fine spatial scales that will be possible with petascale systems. Scott McRae presented analyses of regional scale models, and he identified currently used numerical techniques and algorithms that lead to numerical damping of the smallest scales of motion when mesh spacings below 1km are used to simulate atmospheric motions. If the same methods are applied using petascale systems (capable of meter-scale resolution for regional models), little additional spectral content would be realized unless and until the numerical damping is significantly reduced. Both Scott McRae and Tom Lund discussed adaptive-mess refinement methods for mesoscale models. McRae emphasized the need to overhaul the dynamics cores of existing community-based models, and Lund advocated the use of energy conserving low-order unstructured grids, as has been a recent trend in turbulence and aerospace code development. Piotr Smolarkiewicz discussed various anelastic and pseudo-incompressible (and pseudo-compressible) approaches that eliminate damping associated with acoustic-mode splitting from current mesoscale and global-scale codes, but which add challenges for computation in complex geometries.

b. With regard to the need to update SGS techniques, John Wyngaard described the modeling *terra incognita* to which petascale systems will provide us access. In particular, Wyngaard's central message was that resolved scales between the traditional boundaries of meso-scale models and large-eddy simulation (LES) models will require SGS techniques which differ from those used today, and we, as a community, absolutely must address this modeling challenge because existing algorithms cannot simply be applied to the finer resolutions petascale systems will make possible. Wyngaard then went on to present very promising *a priori* analyses using atmospheric observational data of an improved SGS model that adds important tilting terms to traditional scalar gradient-diffusion approaches. Several other speakers addressed Wyngaard's modeling *terra incognita* and offered contributions which help prepare us for this new length- and time-scale regime. For example, Hassan Hassan presented multi-scale SGS methods that smoothly transition between fine-scale (LES) approaches and coarser-scale (RANS or Reynolds Averaged Navier-Stokes) approaches by using a novel blending-function technique. Also, Joe Werne described how the probabilistic methods presented by Mark Berliner could be used to develop new SGS methods which forecast both solution and model uncertainty, and which permit integration with observational data, laboratory data, and high-resolution numerical simulation results of isolated process studies. Peter Sullivan demonstrated exciting results from a probabilistic SGS approach to multi-scale simulations of an ocean boundary layer. And Steve Krueger presented a multi-scale modeling approach addressing turbulence and cloud microphysics. The exciting aspect of all of these presentations is that they represent a paradigm shift in traditional LES and direct-numerical simulation (DNS) work in which modeling improvements predominantly target the LES regime, i.e., within the inertial range of turbulence. Instead, in the work presented on this topic at the workshop, significant thought was devoted to using LES and DNS techniques and results to directly improve the SGS methods for regional and global scale models, i.e., in the energy containing range or even at larger scales of motion.

4. Predictability limitations of resolvable small scales will make computed individual realizations difficult to interpret and use. This problem is amplified by the finer resolutions possible with petascale systems. Therefore probabilistic methods should be developed to provide forecast predictions of large-scale quantities as well as their uncertainty and likely variation. Existing ensemble methods may be augmented by other techniques for deducing model and physical uncertainties (e.g., Bayesian Hierarchical Modeling or BHM). Mark Berliner gave a review of BHM methods and numerical techniques for implementing them.

5. Provided the necessary improvements to SGS schemes and numerical methods are made, emerging petascale systems will bring immediate access to millennial time-scale simulations and a greater capacity to resolve eddies in ocean-simulation models. Phil Jones and Eric Chassignet reviewed the state of the art for ocean modeling and forecasting, and they stressed that enhanced resolution, like that possible with petascale systems, will lead to more accurate simulations of current, sea-edge ice interactions, and deep-water formation. Both indicated a `fierce urgency' for improved SGS modeling to satisfy requests from policy makers. They specifically mentioned the decadal prediction of ice sheet and sea-level rise, dynamic vegetation, the hydrological cycle, and biochemical interactions. And as ocean modeling is truly an exascale-class problem, several participants (Jones, Chassignet, Rich Loft) looked beyond the petascale era to the increased capacity of exascale machines, and they anticipated future paradigm shifts. For example, non-hydrostatic models in climate simulations will be possible in exascale environments. Finally, on the practical side of global-scale computing, reaping the benefits of improved SGS methods or the power of peta- and exascale systems requires the careful marriage of code structure with large-scale system architecture. Amik St. Cyr addressed some of these issues by discussing efficient multi-scale time-stepping schemes, unstructured adaptive mesh methods, and optimal solvers for global codes.

6. To address specific improvements to SGS models, six talks presented results from concentrated studies of isolated processes that may need to be included when parameterizations are developed. Yukio Kaneda presented the most fundamental of these studies, concentrating on enormous simulations of isotropic turbulence with 4096^3 spectral modes using Japan's Earth Simulator. Kaneda also presented some preliminary results for anisotropic effects, such as stratification, shear, and magnetic fields. Ed Lee also described results from simulations involving magnetic fields, concentrating on reconnection physics for very thin current sheets and stressing the need to obtain still higher resolution, as these events occur on very fine spatial scales. Yoshifumi Kimura discussed the transition in the energy spectrum for forced stratified turbulence, and emphasized that petascale systems will offer significant advantages for resolving the Ozmidov scale separating anisotropic motions from the (hoped for) equilibrium range of fine-scale turbulence. Joe Werne presented results demonstrating that stratified wind shear exhibits significant variations in behavior as the ratio of stratification and shear time scales are varied, and he sketched a probabilistic approach to SGS development for larger scale models based on compiled catalogs of his high-resolution simulations (on domains as large as 4000×2000^2 spectral modes)[1] combined with similarly cataloged observational results. Peter Sullivan presented results from a multi-scale simulation of the ocean boundary layer with stochastic wave breaking that employed a probabilistic component for its SGS prescription, similar in ways to what Werne advocated, demonstrating the feasibility of such

[1] Werne demonstrated perfect asymptotic scaling on a 4000 processor Cray XT4.

approaches.[2] Finally, Steve Krueger presented a different approach to multi-scale modeling appropriate for addressing cloud microphysics and atmospheric turbulence. In addition to these talks on isolated processes that can be used to develop improved SGS descriptions and techniques, Chenning Tong presented work that formalizes procedures for testing and evaluating new SGS models once they are developed.

7. Emerging petascale systems and the higher-wavenumber-mode inclusion possible with them can amplify known difficulties with existing methods on global domains (e.g., pole singularities). New numerical methods are being developed which have the potential to alleviate these and other difficulties. To introduce the geophysics high-performance computing community to one very promising approach, a series of presentations was made at the workshop on Radial Basis Functions (RBFs), which can be used to develop a class of numerical implementations that embody both the accuracy characteristics of spectral methods and simultaneously the unstructured meshes of existing flexible low-order techniques (like those described by Lund). The following participants presented work on RBF methods: Bengt Fornberg, Natasha Flyer, Edward Kansa and Grady Wright.

In addition to these issues that grow directly from opportunities and challenges associated with petascale computing systems, indirect ramifications were also discussed during Tuesday's round-table discussion, led by Dr. Junping Wang. In particular, because petascale computations are necessarily collaborative in nature, several participants commented on the challenges and frustrations associated with the review process (citing examples from their CMG and PetaApps experiences) of proposed multi-disciplinary projects. Specific comments made included:

1. Multi-disciplinary projects perforce require panels that represent multi-disciplinary skill sets. Panels should take full advantage of the expertise of individual reviewers, but at the same time should avoid the risk of weighing unduly reviewers' comments on topics for which they are not expert, i.e., a mathematician commenting on the geophysics content of a proposal or a geophysicist harshly evaluating a proposed new mathematical avenue. This leads to inconsistent reviews that could be avoided by requiring greater discipline by the panel during the review process of multi-disciplinary efforts.

2. Often multi-disciplinary solicitations advertise a desire for innovative science projects and the potential for breakthroughs in a particular field, but the process inherently reduces the likelihood that these types of projects will be selected. If the NSF truly desires large inter-disciplinary projects to provide the potential for breakthrough science, the review process needs to consciously and seriously consider the qualities in a project that will more likely lead to breakthroughs, and

---

[2] Sullivan demonstrated how to get excellent parallel scaling up to 16,000 processors using a `long-brick' domain decomposition on DOE's Franklin (Cray XT4) at NERSC.

procedures to emphasize this need to be developed and followed, otherwise the size of panels and the number of reviewers will naturally tend to the opposite outcome. Dr. Rich Loft mentioned an example during discussion pertaining to the PetaApps program in which very few award recipients explored communication paradigms other than MPI, when other cutting-edge avenues clearly exist.