Optimizing High-Resolution Climate Variability Experiments on the Cray XT4 and XT5 Systems at NICS and NERSC

John Dennis and *Richard Loft*, NCAR

ABSTRACT: *Simulations supporting the scientific consensus that human activity is changing the Earth's climate have been derived from models run at O(100 km) resolutions. The impact of unresolved scales on these predictions is not precisely known: indeed it has been hypothesized that noise in the climate system (fluctuations on short spatial and temporal scales) could be "reddened," thereby influencing the low-frequency components of the climate signal. To test this hypothesis, we need to run high-resolution, century-scale simulations of the Earth System: this is the primary goal of our use of the Cray XT4 and XT5 systems at the National Institute of Computational Science (NICS) and the National Energy Research Scientific Computing Center (NERSC). These large-scale, resource-intensive climate experiments require careful tuning at scale to achieve a reasonable compromise between integration rate and efficiency. This paper presents preliminary performance and scaling data from a variety of Cray XT systems for a high-resolution (0.5° atmosphere and land surface coupled to 0.1° ocean/sea ice) development version of the Community Climate System Model (CCSM) in configurations capable of running efficiently on up to 5,844 processors. We have achieved integration rates of 2.3 simulated years/day for CCSM4_alpha on the Franklin XT4 and approximately 2.0 years/day on the Kraken XT5 in benchmarks with I/O turned off. An 18-fold variability in output times when CCSM4_alpha writes monthly history and restart files to Kraken's multi-petabyte Lustre file system during the first seven simulated years of production is also presented and discussed.*

KEYWORDS: Cray XT4, XT5, climate variability, CCSM, scaling

1. Science Background

The Earth's climate is characterized by a number of complex, dynamic, and highly interactive components – including the atmosphere, oceans, land surface, and cryosphere – that evolve through a rich collection of interlinked phenomena occurring on a wide range of time and length scales. Because of the complexity of this system, predictions of global and regional climate changes and understanding of the limit of climate predictability rely entirely on numerical experiments performed on supercomputers, typically using a coupled set of climate system model components.

To date, our understanding of Earth's climate has been derived from models run at relatively coarse, O(100 km) resolutions. Now, petascale "Track-2" computers on the National Science Foundation's TeraGrid and the Department of Energy's "leadership class" systems offer an important opportunity to further advance Earth System science. Of particular interest is the notion that these systems might allow climate modelers to run and validate dramatically higher resolutions that resolve mesoscale features of ocean and ice dynamics. By directly forcing coupled climate models with noise statistics from high-resolution models, scientists could test the long-standing hypothesis (Hasselmann, 1976) that low-frequency fluctuations of unpredictable climate noise in one component (e.g., the atmosphere) can be spectrally "reddened" by the longer-time-scale fluctuations in another component (e.g., oceans with higher heat capacity). Similarly, there may be unpredictable fluctuations in the oceans (small-scale eddies) that force low-frequency variations in the atmosphere. If true, the impact on future climate research could be enormous: it would mean that modeling improvements, such as better physical parameterization of unresolved scales, perhaps combined with higher resolution, would be necessary to model climate variability correctly. That conclusion could increase the computational cost of future climate studies by many orders of magnitude. If the hypothesis is proven false, i.e., if increased resolution does not change climate variability significantly, then climate scientists can proceed with much of the current low-resolution research program intact. As is typical in exploration, someone will have to go there to find out: experience with high-resolution, century-scale simulations of the Earth System is needed to test the importance of noise at unresolved scales. This is the initial goal of our research team's effort, and it is the point of our work with the Cray XT systems described in this paper.

Our team's first priority has been to enhance the resolution of the baseline coupled climate system to adequately resolve the noise statistics of the atmosphere and the ocean and complete a high-resolution control run. For the atmosphere, our experience shows that the weather statistics are reasonably well resolved at approximately 0.5° (50 km); for the ocean, 0.1° (10 km) resolution is sufficient to capture the bulk of the eddy statistics. It is the nature of climate system modeling that very long and resource-intensive runs are required to both validate models and to achieve equilibrium states that can establish baselines for future work. For example, our team's first baseline experiment will require a total of 200 simulated years. For such a simulation to occur within the timeframe of a single supercomputer resource allocation (typically one year), an integration rate of multiple simulated years per day is required. A high-resolution, century-long climate experiment of the type contemplated here requires enormous amounts of computer time – on the order of 8 million CPU-hours per simulated century. Thus, from the perspective of resource stewardship, it is important to achieve reasonably efficient model execution.

2. Climate Application Description

The modeling platform used is a development version of the Community Climate System Model Version 4 (denoted here as CCSM4_alpha). CCSM is an important and widely used community application for climate research. After decades of development by staff at the National Center for Atmospheric Research (NCAR), the Department of Energy (DoE), and the university community, CCSM has matured into a true, multi-component Earth System model with dynamic and fully interacting atmosphere, land, ocean, and sea-ice components (Collins et al., 2006). CCSM has an impressive resume of scientific achievement, including contributing to the Nobel Prize-winning Fourth Assessment report by the Intergovernmental Panel on Climate Change (IPCC AR-4) (Parry et al., 2007).

CCSM is a multi-component Earth System model. Each component uses its own distinctive numerical technique, grid, and resolution. Climate system components interact via fluxes exchanged at the interfaces, e.g., at the surface. Because components operate on different grids, fluxes passed between them must, in general, be interpolated. These and related coordination and control operations are performed by a mediating component called the flux coupler.

All CCSM4_alpha components are parallelized using MPI. The atmosphere, land surface, and sea ice components also support the use of hybrid MPI/OpenMP. The grids of each component are decomposed in ways designed to optimize performance on the assigned number of processes. The salient algorithms and computational characteristics of CCSM4_alpha's component models include:

The ocean component, POP, is an ocean general circulation model developed at Los Alamos National Laboratory that uses a finite-difference discretization of primitive equations on a curvilinear orthogonal grid that allows the pole singularities to be remapped out of the ocean computational domain and into continental landmasses. (Smith and Gent, 2004). The high-resolution configuration uses the tripole grid (Murray, 1996) is used with a nominal resolution of 0.1°. The ocean grid dimensions are 3,600 "longitude" points, 2,400 "latitude" points, and 42 vertical levels.POP solves primitive equations for the ocean under hydrostatic and Boussinesq approximations, and uses depth as the vertical coordinate. POP's time integration scheme separates the slow (baroclinic) and fast (barotropic) modes of the model. The baroclinic includes a computational piece plus a 3D halo update to compute finite difference spatial derivatives. In contrast, the barotropic invokes a 2D solver update of the surface pressure. The solver used in the barotropic component of POP is based on the single inner product preconditioned conjugate gradient method (D'Azevedo et al., 1993). Nevertheless, the scalability of global reductions (i.e., MPI_ALL_REDUCE) and the 2D halo update within the conjugate gradient solver is typically very latency sensitive at high processor counts.

The atmosphere component, CAM3.5, is based on a variant of the Lin (2004) finite-volume dynamical core developed by Mirin and Sawyer (2005). The CAM grid is roughly 0.5° × 2/3°, with 576 longitude points, 384 latitude points, and 26 vertical levels. The physical processes (e.g. cloud physics) represented in CAM are quite complex, and can produce both static and dynamic load imbalances. See Collins et al., (2006) for a complete description of the numerical methods applied in CAM.

The sea ice component, CICE4, uses the elastic-viscous-plastic ice dynamics and treats thermodynamics and deformation with a subgrid-scale ice-thickness distribution (Hunke and Lipscomb, 2008). CICE is run on the same horizontal grid as the ocean component with four ice layers, a single snow layer, and five ice-thickness categories plus one open-water category.

The land component, CLM3.5, is closely related to CLM3.0; see Oleson et al., (2004) and Hoffman et al., (2005). CLM represents spatial land surface heterogeneity via a nested subgrid hierarchy in which grid cells are composed of multiple land units, snow/soil columns, and plant functional types. The land model is embarrassingly parallel but not perfectly load balanced. It is run on the same horizontal resolution as the atmosphere component, with a vertical grid containing 10 vertical soil layers, five snow layers, and 17 plant functional types.

The flux coupling component, CPL7, is a single executable design that provides flexibility in running the CCSM model components sequentially, concurrently, or in a mixed sequential/concurrent mode. It achieves this flexibility through a driver that runs on every CCSM process and controls the time sequencing, processor concurrency, and exchange of fluxes between components. In CPL7, all model components and the coupler methods can run on potentially overlapping processor subsets. This design permits the model system to have greatly increased flexibility in achieving a climate component layout that optimizes the overall performance of the model. However, overlapping components running MPI and hybrid MPI/OpenMP on the same hardware is generally not allowed, as the operating system cannot turn processes into threads or threads into processes.

Output fields for analysis or restart are currently gathered to a single processor within each component and written to disk, typically once per simulated month.

3. Description of Cray XT Systems Used

The specific supercomputing architectures and systems referenced in this paper include:

3.1 The Cray XT4

The basic architectural plan of the Cray XT4 compute node is shown in Figure 3.1. Each compute node consists of a single AMD socket connected via the high bandwidth/low latency AMD HyperTransport (HT) layer to a dedicated SeaStar2 router, and these routers are in turn interconnected to one another by six, 7.6-GB/sec-peak bidirectional network links, forming a 3D torus topology.

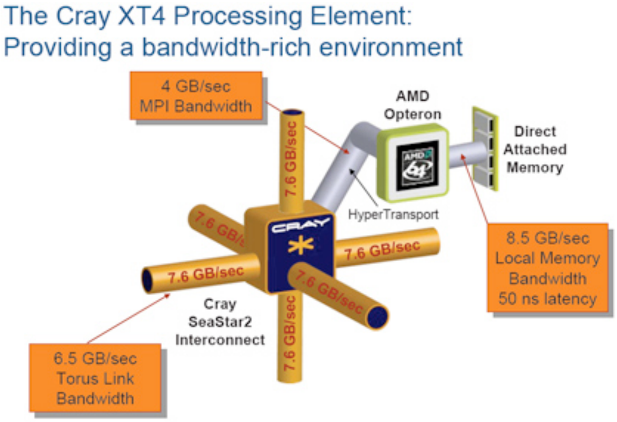

Figure 3.1. Cray XT4 compute node architecture.

3.2 The Cray XT5

The basic architectural plan of the Cray XT5 compute node is shown in Figure 3.2. Each compute node consists of dual AMD sockets sharing a connection via the AMD HT layer to a SeaStar2+ router, and these routers are in turn interconnected to one another by six, 9.6-GB/sec-peak bidirectional network links in a 3D torus topology.

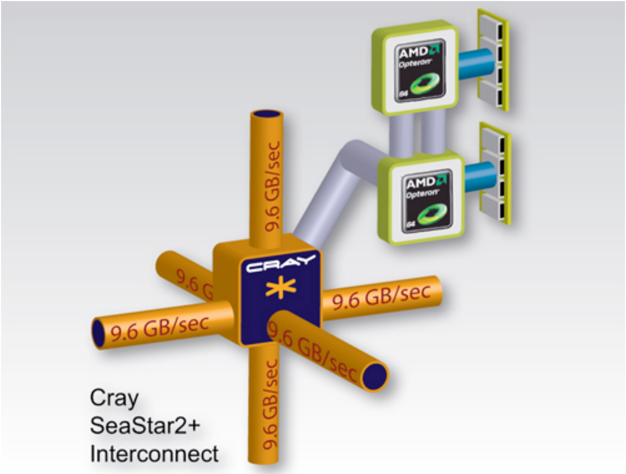

Figure 3.2. Cray XT5 compute node architecture.

3.3 "Franklin" XT4 at NERSC

Frankin, the Cray XT4 system at the National Energy Research Supercomputing Center (NERSC), is the seventh fastest supercomputer in the world according to the November 2008 Top 500 List, having achieved an RMax (HPL) rating of 266.3 TFLOPS. Franklin is composed of a total of 9,660 XT4 compute nodes connected in a (17x24x24) 3D torus. Each compute node consists of a single 2.3 GHz AMD Opteron quad core socket, each core of which is capable of executing 4 floating point operations per clock cycle. Each compute node has 8 GB of memory (2 GB/core). In aggregate, Franklin has 38,640 cores, a total of over 77 TB of memory, and a theoretical peak of 355 TFLOPS.

A computer allocation of 2.3M CPU-hours was obtained on Franklin to port, debug, and performance-tune the high-resolution CCSM4_alpha to Cray XT systems. An initial 10-year simulation of the high-resolution CCSM4_alpha identified several problems in the high-latitude atmospheric circulation patterns as well as problems in the sea ice coverage. This simulation ended with an unrealistic polar climate, but it did improve understanding of high-resolution climate and allowed for significant improvements in simulation performance.

3.4 “Kraken” XT4 at NICS

The XT4 version of Kraken, recently retired from service at the National Institute for Computational Science (NICS), was composed of 4,512 compute nodes arranged in a 3D torus with dimensions of (12 x 16 x 24). Each compute node had a single 2.3-GHz quad core AMD socket and 4 GB of memory (1 GB/core). In aggregate, this system had 18,048 cores, 18 TB of system memory, and a theoretical peak speed of 166 TFLOPS. The principal difference between the Kraken and Franklin XT4 systems that we tested was the version of Compute Node Linux Environment (CNL) deployed: Kraken used CLE 2.0.62, and Franklin used version 2.1.56HD.

The Kraken Cray XT4 system at NICS provided our team an opportunity to perform initial porting, testing, and optimization of the CCSM4_alpha configuration via a small 200,000 CPU-hour TeraGrid Start-up Account.

3.5 “Kraken” XT5 at NICS

The XT5 version of Kraken currently at NICS has a total of 8,256 XT5 compute nodes connected in a (22 x 16 x 24) 3D torus. Each compute node consists of dual, 2.3-GHz AMD Opteron quad core sockets. Kraken is a heterogeneous system from the point of view of memory: 3,840 nodes have 8 GB (1 GB/core), and 4,416 nodes have 16 GB (2 GB/core). In aggregate, Kraken has 66,048 cores, over 100 TB of system memory, and a theoretical peak speed of 608 TFLOPS.

4. Climate Application Performance Measurements

Our team recently received a 35M CPU-hour allocation from the TeraGrid on the Cray XT5 version of Kraken at NICS. This enormous allocation of computer time comes with the responsibility to maximize the efficiency of integration while satisfying memory constraints and achieving the integration rates necessary to complete the planned experiments within the one-year resource allocation period. This section describes efficient component task counts for CCSM that allow us to meet these goals, and we report the performance achieved.

4.1 Overall CCSM Model Scalability on Cray XT4 and Cray XT5

As part of the NSF PetaApps project, our team is particularly focused on achieving scientifically useful climate results at unprecedented levels of resolution and parallelism. Over a number of years, a development branch of the CCSM code base has migrated from the O(100)-processor, O(100-km) model operating space of today to an O(10-km) version of the system capable of running on many thousands of processors. Recent runs under a DOE Grand Challenge computing grant, which produced 12 years of simulated climate on 5,400 CPUs, have demonstrated the feasibility and advantages (in terms of scalability and integration rate) of using very large Cray XT systems such as the NERSC Franklin system at Lawrence Berkeley National Laboratory to conduct ultra-high-resolution runs. Further, we have run performance studies on the existing Track-2 Cray XT systems at NICS, Kraken XT4 and Kraken XT5.

We next describe the performance of high-resolution CCSM4_alpha used for this project: it is composed of five different component models, each with its own scalability characteristics. The scalability of POP has been studied intensively, and it has been successfully scaled to 30,000 cores (Dennis, 2007). Versions of CICE have also been successfully scaled beyond 10,000 cores on Cray XT equipment (Dennis and Tufo, 2008). The scalability and performance of CAM has also been studied extensively (Mirin and Sawyer, 2005) and (Mirin and Worley, 2007). The performance of the entire CCSM coupled system has not been formally studied, but it has many of the same characteristics of its component models.

As noted in Section 2, CCSM4_alpha has the flexibility to assign cores to components at runtime. We key our decomposition configurations from the 0.5° CAM component, which has the least flexibility in terms of processor counts. We configure the remaining component models to minimize idle time. An example of our load balancing strategy is provided in the next paragraph. For our high-resolution Kraken simulations, we have identified five common processor configurations: extra small (XS), small (S), medium (M), large (L), and extra large (XL). Table 4.1 contains the current best processor counts for each of these configurations. Figure 4.1 illustrates the most efficient XL configuration; it uses 5,844 cores.

Table 4.1. CCSM component processor configurations for the high-resolution baseline configuration.

Config size	**Type**	**Processor counts**					
		POP	**CAM**	**CICE**	**CLM**	**CPL7**	**Total**
XS	Sequential	480	480	480	480	416	480
S	Sequential	1024	1024	1024	1024	416	1024
M	Hybrid	1232-1376	480	480	480	416	1712-1865
L	Hybrid	2448-2618	1024	1024	16	416	3488-3658
XL	Hybrid	3476-4028	1664	1800	16	416	5292-6380

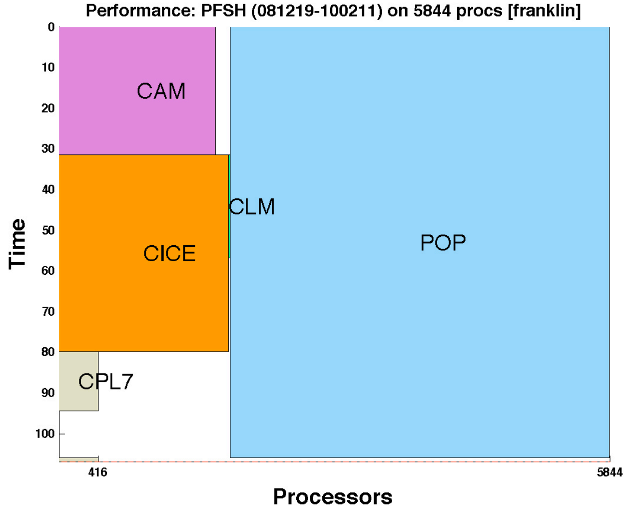

Figure 4.1. Performance of high-resolution CCSM on 5,844 processors of Franklin, a Cray XT4.

Figure 4.1 represents the computational cost to simulate a single model day, excluding disk I/O costs. The x-axis shows the number of processors, and the y-axis shows the time each component takes to execute. In this configuration, CAM is placed on the first 1,664 processors and takes a total of 31 seconds to execute a wall-clock day. The CICE model uses the first 1,800 processors and takes 49 seconds to execute a wall-clock day. The CAM and CICE models execute sequentially on the same group of cores. The CLM model executes concurrently with CICE on 16 processors. CPL7 currently takes 25 seconds to complete a model day on 416 processors. We typically load-balance CCSM by adjusting the processor count for POP such that it takes approximately the same time as CAM, CICE, and CPL7 combined. In this case, POP is assigned the remaining 4,028 cores to simulate a model day in 107 seconds.

This particular XL configuration, based on the scaling characteristics and limits of the decomposition of each component model, appears to be optimal on Franklin, the Cray XT4 system at NERSC. Figure 4.1 offers an easy way to calculate the computational cost of each component, as it is the area of each component's rectangle. Due to limitations in the flexibility of each components decomposition approach, the total cost of idle cores in this configuration is 8%.

Figure 4.2 is a plot of the scalability of CCSM4_alpha for all five processor configurations on the NERSC Franklin system, the NICS Kraken XT4 system, and the NICS Kraken XT5 system. Note that we do not include the time to write history and restart files. We describe disk I/O performance later in this section. The top panel of Figure 4.2 shows the simulated years per wall-clock day on Franklin, Kraken XT4, and Kraken XT5. The bottom panel of Figure 4.2 is the computational cost to simulate a model year for each configuration. With perfect scaling, the computational cost would be independent of processor count. However, load imbalance in the coupled system and scalability issues with the component models increase the cost of the system as processor count is increased.

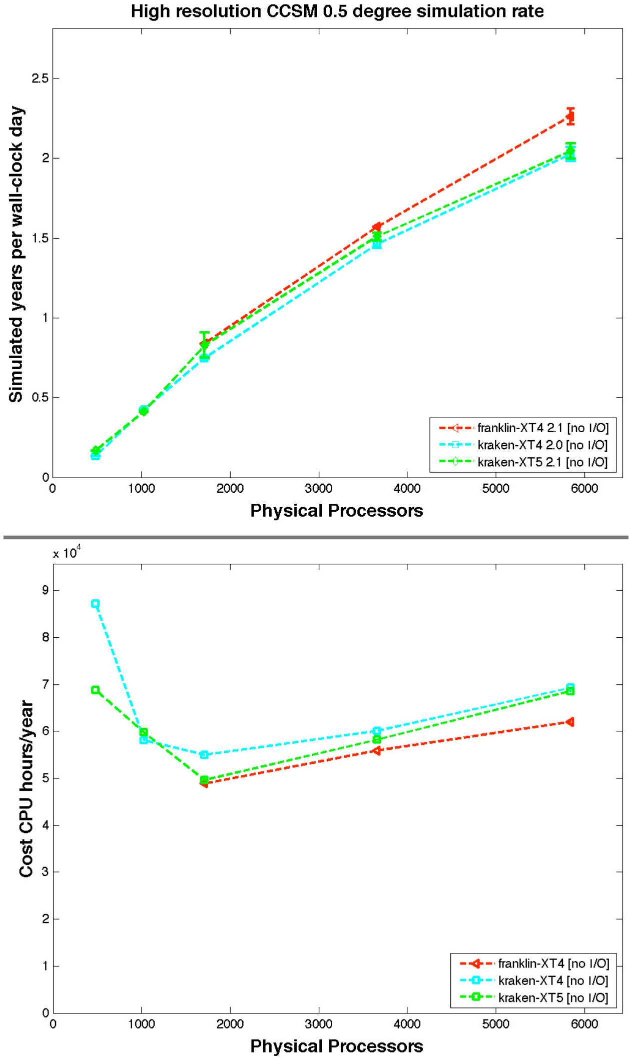

Figure 4.2. Scalability of high-resolution CCSM4_alpha on Franklin, Kraken XT4, and Kraken XT5. Top panel is the simulation rate. Bottom panel is the computational cost.

We have completed 7.25 years of high-resolution CCSM_alpha on Kraken XT5 using the XL configuration. The execution time for each day of the simulation is measured. As noted previously, history and restart files are written at the end of each month of the simulation, approximately every 70 minutes of wall clock time. Ten files that range in size from 95 MB to 24 GB are written, resulting in a total of 57.9 GB of output per simulated month. While the execution time for simulated days in which no disk I/O is performed is relatively consistent, the execution time for model days in which files are output varies significantly, from a low of 759 seconds to a high of 11,889 seconds. It should be noted that these times are an aggregate that include the cost of gathering data to the master processor, as well as the time required to perform write operations. These write sizes in CCSM vary from 864 KB in the land model (CLM) to 1.4 GB in size in the ocean model (POP). Figure 4.3 shows the measured execution time for output days between 4/21/2009 and 4/28/2009. Note that the variability in output times is particularly severe on 4/23/2009.

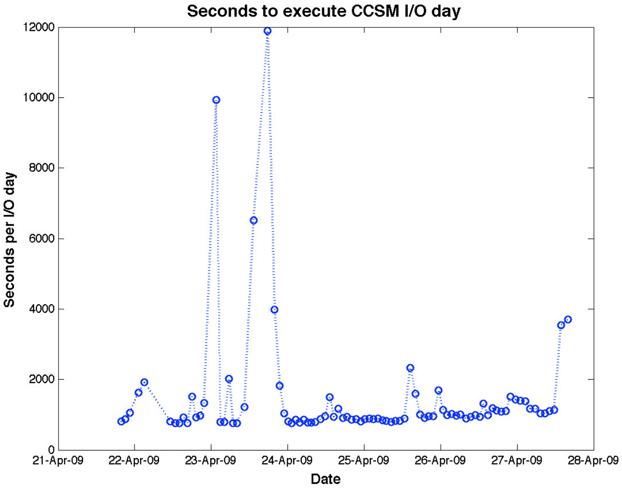

Figure 4.3: The time for CCSM to perform I/O at the end of each simulated month on Kraken XT5. Note the large variability in output time on April 23rd.

Figure 4.4 shows sustained bandwidth for each model output day, which varies from a low of 5 MB/sec to a high of 92 MB/sec. As noted in Section 2, the current production version of CCSM supports I/O for each component through a single master processor. This feature accentuates the importance of CCSM I/O operations relative to computation. For that reason, a version of CCSM that supports parallel I/O through MPI-IO is now being developed. Until that scheme is available, however, we must optimize the performance of the model during simulated days with file output. In that regard, a sustained rate of 92 MB/sec is an acceptable for our current simulation program. However, a large variation in output bandwidth is problematic, due to the idle processor time incurred while serial monthly file output is performed.

The impact of current file output time variability on the cost of our simulation is provided in Table 4.2. Of the total 790,000 CPU hours used during April 21 to April 28, 605,000 CPU hours were consumed by the computational aspects of CCSM, while 185,000 CPU hours were due to performing monthly output. If the system provided a consistent bandwidth of 92 MB/sec, then the overhead of serialized file output in CCSM4_alpha would have only consumed 89,000 CPU hours. The observed overhead

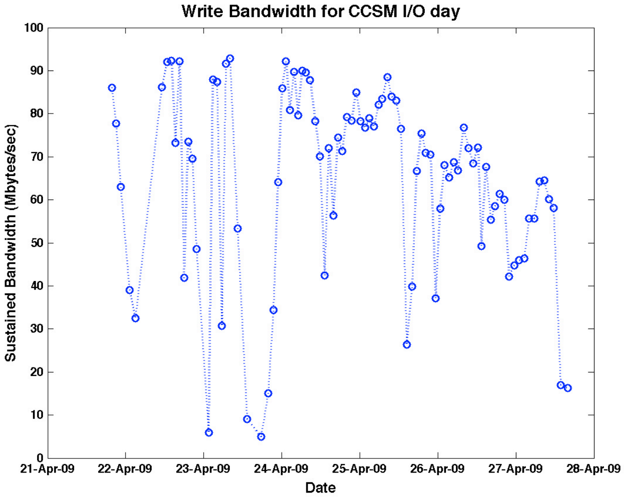

Figure 4.4: The sustained output bandwidth from CCSM varies from a low of 4.9 to 92.8 MB/sec over a six-day period.

impact of system variability is 96,000 CPU hours, or 12% of the total cost of the simulation.

While Lustre performance can vary by a factor of five (Crosby, 2009), we are observing up to 18-fold variations. The reason for this abnormality is unclear. Possible sources include pathologies in the MPI-based data gathering schemes of one or more of the component models, latency sensitivity of some of the smaller write operations, e.g. in CLM, or issues within the Lustre file system itself. Regardless, this excessive variability and cost is problematic, and it may impact our ability to successfully complete the science objectives of our PetaApps project.

Table 4.2: Various I/O and computational costs for existing 7.25-year simulation.

Cost to simulate 7.25 years	**CPU hours**	**% of cost**
Computational Cost	605K	76.6%
Output Overhead [@92 MB/sec]	89K	11.2%
Output Variability Overhead	96K	12.2%
Total Output Overhead	185K	23.4%
Actual Total Cost	790K	100%

5. Conclusions and Future Work

This paper presents preliminary performance and scaling data from various Cray XT systems for a high-resolution (0.5° atmosphere and land surface coupled to 0.1° ocean/sea ice) development version of the Community Climate System Model (CCSM) in configurations capable of running efficiently on up to 5,844 processors. We have identified efficient multi-component layouts in a hybrid-coupling framework that leave processors idle only 8% of the time. The largest of these configurations has achieved integration rates as high as 2.3 simulated years/day for CCSM4_alpha on the Franklin XT4 running CLE 2.1, and approximately 2.0 years/day on the Kraken XT-4 running CNL 2.0.62 and the Kraken XT5 running CLE 2.1, in benchmarks with I/O turned off. We cannot account for the 15% better scaling and performance observed on Franklin for these benchmarks, but kernel level noise is an obvious suspect. We are also working with Cray engineers to better characterize the problems with Kraken XT5 scaling.

The sustained output performance of coupled model while writing monthly history files to Kraken's multi-petabyte Lustre file system was found to range between 5 and 95 MB/sec. In the future, we intend to investigate the I/O bandwidth variability problems we have encountered in four ways. First, we intend to investigate the question of potential MPI gather pathologies, for example due to buffer overflows, in the component gathers. Buffer overflow issues could be addressed by introducing handshaking code designed to prevent such conditions from occurring. Second, we intend to profile the writes in CCSM to uncover and address any potential latency overhead issues. Third, we plan to work with NICS and Cray experts to determine if there are problems in the Lustre file system itself contributing to this problem. Finally, we intend to replace the current serial I/O scheme in CCSM with MPI-IO-based parallel I/O capabilities currently under development. In these ways we hope to characterize and ultimately reduce the cost, memory footprint, and variability of CCSM4_alpha history and restart file I/O on Kraken.

Acknowledgements

Earth System modeling is the ultimate interdisciplinary problem, requiring the close collaboration of highly qualified climate scientists, computer scientists, and software engineers to harness massive amounts of computing power. The authors would like to acknowledge the contributions of their NSF PetaApps collaborators, most notably Mariana Vertenstein of the Climate and Global Dynamics Division and Nathan Hearn of CISL at NCAR, Kathy Yelick at NERSC, and the climate science team for these experiments including: Ben Kirtman, a climate researcher at the University of Miami, Jim Kinter, a senior research scientist and director of the Center for Ocean-Land-Atmosphere Studies (COLA, the lead institution for the PetaApps proposal), Cristiana Stan at COLA, Cecelia Bitz, an assistant professor of Atmospheric Sciences at the University of Washington who specializes in sea ice and climate interactions, and Bill Collins, a climate scientist at UC Berkeley.

This effort has also leveraged substantial interagency support for related research in scalability of the Community Climate System Model. In particular, computer time for this research was provided by the National Science Foundation through startup and research allocations of TeraGrid resources on the Kraken system at the National Institute for Computational Science. Access to Department of Energy resources at the National Energy Research Supercomputing Center was provided through Warren Washington's INCITE proposal "Climate-Science Computation End Station Development and Grand Challenge Team," and Frank Bryan's Director proposal "Ocean-Atmosphere Interaction in the Eddy-Resolving Regime." Computational resources at Lawrence Livermore National Laboratory were provided by David Bader's LLNL Third Institutional Grand Challenge proposal "Ultra-High-Resolution Coupled Climate Simulations." NSF Award OCI-0749206 and OCE-0825754 supported high-resolution model development and performance enhancement work. Development of the POP optimizations utilized in this paper would not have been possible without access to the IBM Blue Gene/L system at NCAR, which is funded through NSF MRI Grants CNS-0421498, CNS-0420873, and CNS-0420985, and through the IBM Shared University Research (SUR) Program with the University of Colorado. The authors would like to thank Cray Research for providing the images for Figures 3.1 and 3.2.

About the Authors

John Dennis is a computer scientist in the Computer Science Section of the Computational and Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR), 1850 Table Mesa Drive, Boulder, CO, 80305, 303-497-1809, dennis@ucar.edu. Richard Loft is the Director of CISL Technology Development Division at NCAR, 1850 Table Mesa Drive, Boulder, CO, 80305, 303-497-1262, loft@ucar.edu.

References

Collins, W.D., P. Rasch, B.A. Boville, J. Hack, J.R. McCaa, D.L. Williamson, B.P. Briegleb, C.M. Bitz, S-J. Lin, and M. Zhang, 2006: The formulation and atmospheric simulation of the Community Atmosphere Model Version 3 (CAM3). *J. Climate,* 19(11), 2144-2161.

Crosby, L., 2009: Personal conversation with J. Dennis about Lustre performance variability.

D'Azevedo, E.F., Eijkhout, V.L., and Romine, C.H., 1993. Conjugate gradient algorithms with reduced synchronization overhead on distributed memory multiprocessors. Technical Report 56, LAPACK Working Note, August 1993.

Dennis, J.M., 2007: Inverse Space-Filling Curve Partitioning of a Global Ocean Model. International

Parallel and Distributed Processing Symposium 2007, Long Beach, Calif. doi: 10.1109/IPDPS.2007.370215

Dennis, J.M. and H.M. Tufo 2008: Scaling climate simulation applications on the IBM Blue Gene/L system. *IBM Journal of Research and Development:* Applications for Massively Parallel Systems, 52(1/2), 117-126.

Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. *Tellus*, **28**, 473-485.

Hoffman, F.M., M. Vertenstein, H. Kitabata, and J.B. White III, August 2005: Vectorizing the Community Land Model (CLM). *Inter. J. of High Perf. Comp. Appl.,* 19(3), 247-260.

Hunke, E.C. and W.H. Lipscomb, 2008: CICE: the Los Alamos Sea Ice Model, Documentation and Software User's Manual, version 4.0. Tech Report LA-CC-06-012, Los Alamos National Laboratory.

Lin, S-J. 2004: A vertically Lagrangian finite-volume dynamical core for global models. *Mon. Wea. Rev.,* 124(10), 2293-2307.

Mirin, A.A. and P. Worley, 2007: Extending Scalability of the Community Atmosphere Model. SCiDAC Conference, Westin Copley Place, Boston, Mass, (http://www.csm.ornl.gov/~worley/talks/index.html).

Mirin, A. and W. Sawyer, 2005: A scalable implementation of a finite-volume dynamical core in the community atmosphere model. *Inter. J. of High Perf. Comp. Appl.,* 19(3), 203-212.

Murray, R.J., 1996: Explicit generation of orthogonal grids for ocean models. *J. Comp. Phys.,* 126, 251-273.

Oleson, K.W., Y. Dai, G. Bonan, M. Bosilovich, R. Dickinson, P. Dirmeyer, F. Hoffman, P. Houser, S. Levis, G.-Y. Niu, P. Thornton, M. Vertenstein, Z.-L. Yang, X. Zeng, May 2004:Technical Description of the Community Land Model (CLM). Technical Note TN-461+STR, NTIS no. PB2004-105836, 183 pp.

Parry, M., O. Canziani, J. Palutikof, P. van der Linden, C. Hanson, et al., 2007: Climate Change 2007: Impacts, Adaptation and Vulnerability. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* 976 pp., (http://www.ipcc.ch/ipccreports/ar4-wg2.htm).

Smith, R.D. and P.R. Gent, 2002, revised 2004: Reference manual for the Parallel Ocean Program (POP), ocean component of the Community Climate System Model (CCSM 2.0 and 3.0). Technical Report LA-UR-02-2484, Los Alamos National Laboratory, (http://www.ccsm.ucar.edu/models/ccsm3.0/pop).