NERSC NEWS

News from the National Energy Research Scientific Computing Center August 2004

LLNL Scientists Use NERSC to Advance Global Aerosol Simulations

While "greenhouse gases" have been the focus of climate change research for a number of years, DOE's "Aerosol Initiative" is now examining how aerosols (small particles of approximately micron size) affect the climate on both a global and regional scale.

Scientists in the Atmospheric Science Division at Lawrence Livermore National Laboratory (LLNL) are using NERSC's IBM supercomputer and LLNL's IMPACT (atmospheric chemistry) model to perform simulations showing the historic effects of sulfur aerosols at a finer spatial resolution than ever done before. Simulations were carried out for five different decades, from the 1950s through the 1990s.

The results clearly show the effects of the changing global pattern of sulfur emissions. Whereas in 1950 the United States emitted 41 percent of the world's sulfur aerosols, this figure had dropped to 15 percent by 1990, due to conservation and anti-pollution policies. By contrast, Asian emissions jumped six fold during the same time, from 7 percent in 1950 to 44 percent in 1990.

Under a special allocation of computing time provided by the Office of Science INCITE (Innovative and Novel Computational Impact on Theory and Experiment) program, Dan Bergmann, working with a team of LLNL scientists including Cathy Chuang, Philip Cameron-Smith, and Bala Govindasamy, was able to carry out a large number of calculations during the past month, making the

(continued on page 3)

NERSC News

Welcome to the August 2004 issue of NERSC News, highlighting achievements by staff and users of DOE's National Energy Research Scientific Computing Center Division. NERSC News is distributed every other month via email and may be freely distributed. NERSC News is edited by Jon Bashor, JBashor@lbl.gov or 510-486-5849.

NERSC Helps Climate Researchers Get Results Faster to Meet Deadline

When experts on the Earth's environment join forces with experts on the high performance computing environment, the future of our global climate comes into focus faster — at least in the results of model simulations.

Jonthan Carter

That's what happened this summer when researchers from the National Center for Atmospheric Research (NCAR) asked NERSC consultants to help them improve the throughput of their simulations so that they could present the results at an upcoming meeting of the Intergovernmental Panel on Climate Change (IPCC).

The research team, led by Warren Washington, is carrying out a substantial portion of the climate change scenarios for the U.S. DOE/NSF contribution to the IPCC Fourth Assessment Report. This multi-institutional effort involves scientists and software engineers at NCAR, Los Alamos National Laboratory, Oak Ridge National Laboratory, and Lawrence Berkeley National Laboratory.

The NERSC portion of this project is a foundational segment of a coordinated international effort aimed at understanding the global and regional impacts of human-induced climate change. The climate change scenarios are being simulated with the Community Climate System Model (CCSM), which consists of four models — atmosphere, ocean, sea ice, and land — coupled through the "flux coupler." Climate change experiments are performed by running five CCSM jobs as an ensemble, in which each job is run using different initial conditions.

The NCAR team had been running the CCSM ensemble on NERSC's Seaborg system as five separate jobs; but with the IPCC meeting deadline looming, they asked NERSC's consultants if there was an efficient way to combine them into a single job. A single job running on 512 or more processors could take advantage of

(continued on page 2)

PDSF to Boost Processors, Performance

Researchers using the PDSF cluster managed by NERSC will soon have access to more processing power and benefit from a higher speed network connection for accessing archived data.

New hardware is being shipped and is expected to be installed by the end of September. The additions include 48 nodes of dual

(continued on page 2)

Climate Researchers (continued from p.1)

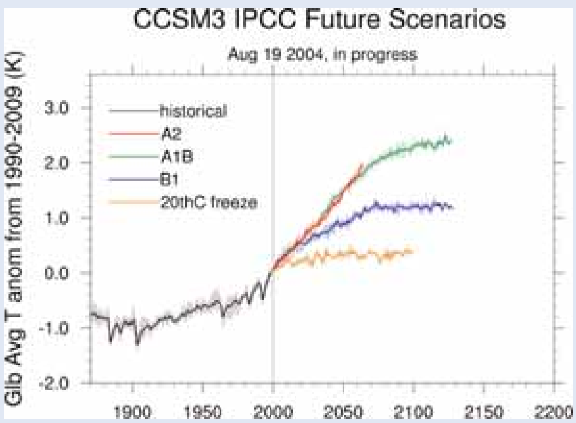
NERSC's incentives for highly parallel jobs — preferential scheduling and a 50% discount on allocation charges.

But merging five separate jobs into one is easier said than done. The first obstacle involved IBM's LoadLeveler job scheduling software, which is used on Seaborg. LoadLeveler allows only one parallel job in a batch job and can have only one set of environment variables — data associated with a running program that tells the computer about the user's system configuration and preferences, and how to execute the parallel program. But each job in the CCSM ensemble has its own large set of environment variables. Merging the five jobs into one job would involve redefining all of those environment variables in various script files as well as in the CCSM source code.

Another obstacle was combining the instructions for task geometries and communication from all five jobs into one big set of instructions. Task geometries determine how tasks are assigned or distributed among the nodes and processors, and communication instructions set the rules for how the computational results are communicated between nodes. Combining these instructions would be a major effort involving modification of the source code. And any change in the job run configuration meant more modifications to the code.

Harsh Anand

Fortunately, NERSC consultant Jonathan Carter had successfully tackled similar challenges of running multiple parallel jobs as a single job. He solved the problem by using Internet Protocol (IP) as the combined job's communication protocol instead of the default User Space (US) communication subsystem, which is designed to take advantage of the IBM SP's high performance switch. Although using IP instead of US results in a slight performance degradation, IP does offer several advantages:



CCSM3 DOE/NSF IPCC scenario runs as of August 19 2004. Observed atmospheric/solar forcings are used during the historical period, from years 1870–2000. A spectrum of future forcing scenarios (20th century freeze, B1, A1B, and A2) are used from years 2000–2200 to simulate the most likely range of future climates.

IP allows the user to run multiple parallel jobs.

IP does not require renaming of environment variables, since each component job can be executed by a subshell.

The task geometries for all five jobs can be handled by building hostlist files, which list the processors on which the tasks should run.

No source code modification is required. Any change in job run configuration would require only modification of the scripts to build new hostlist files corresponding to the job task geometries.

NESRC consultant Harsh Anand, who specializes in providing support for climate applications, demonstrated the feasibility of this approach for CCSM by setting up a job script that combined five CCSM jobs into a single job. She then worked with the NCAR researchers on a longer, tougher test: they ran a combined job for a month, then compared the results with the output obtained from individual job runs. The two results were identical.

Finally, Harsh Anand helped the NCAR team implement the new and modified scripts in their production runs, and CCSM is now running on 720 processors. Although there is a slight loss of efficiency (~10–15%) for the IP implementation of CCSM, the NCAR team thinks that trade-off is made worthwhile by the faster overall turnaround time, the availability of more processor hours, and the reduced human workload associated with fewer job runs — not to mention the fact that they did not have to rewrite their code while trying to meet a deadline.

PDSF (continued from p.1)

Xeon processors and 10 nodes of dual Opteron processors. These will be added to the existing 550 processor cluster.

The system's connection to NERSC's internal network will be upgraded from a 2 gigabit switch to a 10 gigabit switch. This will improve the transfer rate of data between PDSF and NERSC's HPSS archival storage.

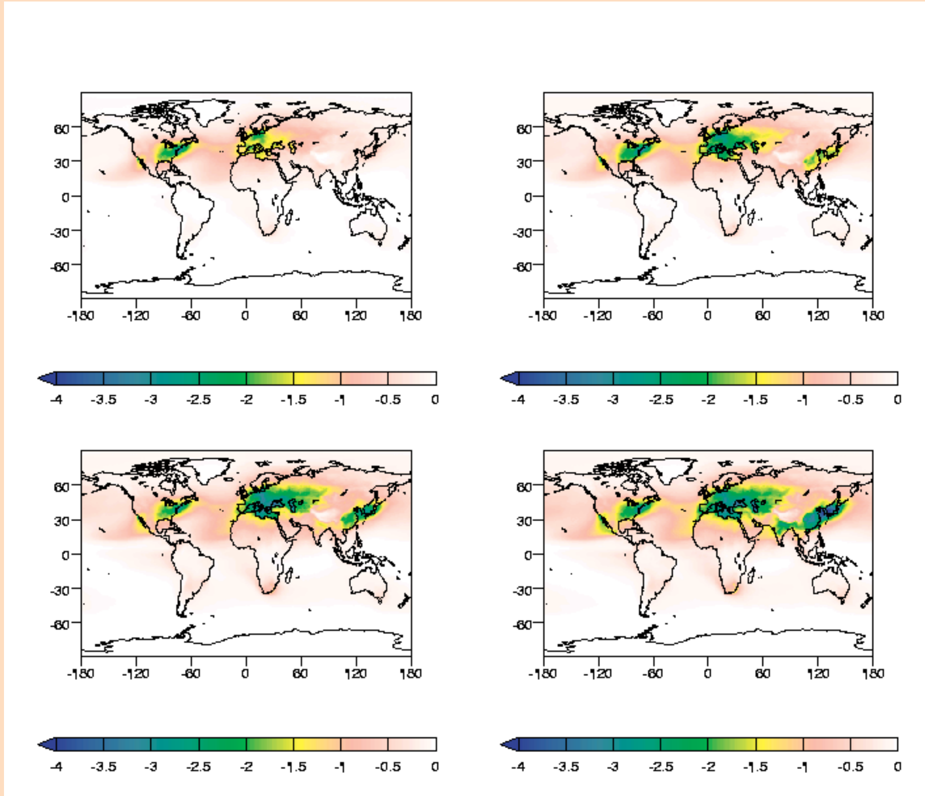
Finally, some of the PDSF disk drives are also being upgraded. This is expected to add 12 terabytes of disk capacity to the cluster.

Global Aerosol Simulations (continued from p.1)

aerosol project one of the largest users of NERSC resources. The applications ran on 128 and 256 processors.

The objective was to assess the effects of anthropogenic (man-made) sulfate aerosols. The IMPACT model calculates the rate at which SO_2 (a gas emitted by industrial activity) is oxidized and forms particles known as sulfate aerosols. These particles have a short lifespan in the atmosphere, often washing out in about a week. This means that their effects on climate tend to be more regional, occurring near the area where the SO_2 is emitted.

To accurately study these regional effects, Bergmann needed to run the simulations at a finer horizontal resolution, as the coarser resolution (typically 300km by 300km) of other climate models is insufficient for studying changes on a regional scale. Livermore's use of the Community Atmosphere Model (CAM3), a high-resolution climate model developed at NCAR (with collaboration from DOE), allows a 100km by 100km grid to be applied. NERSC's terascale computing capability provided the needed computational horsepower to run the application at the finer level.



Depending on their composition, aerosols can either scatter or absorb sunlight, thereby cooling the Earth and acting as a counter to the warming effects of greenhouse gases. Greenhouse gases such as carbon dioxide are much more longer-lived, so they stay in the atmosphere and have more uniform distribution. But since global warming has not increased as much as some computer models predict based on greenhouse gases alone, scientists have become more interested in aerosols and their possible role in countering some effects of greenhouse gases.

The atmospheric concentrations of both aerosols and greenhouse gases have increased over the past century and, because of their effects on shortwave and longwave radiation (which are partially offsetting), have presumably upset to some degree the thermal equilibrium of the climate system. While satellite measurements and field studies provide vital information to document the global and regional climate impacts by aerosols and greenhouse gases, accurate climate model simulations are an equally important tool in quantifying the radiative forcing as well as identifying and analyzing the climate response.

By running numerous calculations at finer resolution, the team was able to see the cooling effects by region, and to see more differences in the effects of varying concentrations.

The simulations also gave the team an opportunity to test new meteorological data generated using NCAR's CAM3. This model calculates the various factors influencing climate, such as winds, temperatures and precipitation. This data, which bordered on the terabyte scale, was also generated at NERSC.

The LLNL researchers then used the climate data to drive their chemistry model known as IMPACT. By varying the SO_2 emissions levels while running IMPACT, the team was able to simulate different scenarios for each decade, analyze aerosol distributions from each simulation and then examine their effects on radiative forcing at regional scales.

The figure above shows the predicted effects on shortwave radiation (watts/m2) due to anthropogenic sulfate aerosols in the 1950's (top left), 1970's (top right), 1980's (bottom left), and 1990's (bottom right). Note the increased regional cooling effects in central Asia, India, and particularly eastern Asia.

For more information, contact Dan Bergmann at dbergmann@llnl.gov.