New Computational and Scientific Results at NERSC

Science-of-Scale Applications Achieve Significant Results and up to 68% of Peak Performance on 10 Tflop/s IBM SP

April 7, 2003 (revised May 7, 2003)

Initial results from NERSC's 10 teraflop/s IBM SP supercomputer, which became available for general use in early March 2003, show scientific applications running at up to 68% of the system's theoretical peak speed, compared with the 5–10% of peak performance typical for scientific applications running on massively parallel or cluster architectures. Performance results for four science-of-scale applications are shown in the following table:

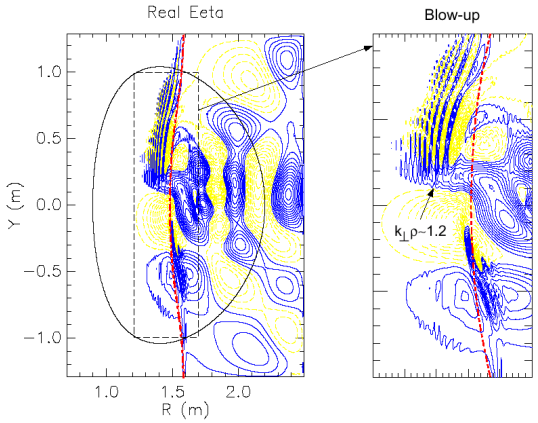
Project	Number of processors	Performance (% of peak)
Electromagnetic Wave-Plasma Interactions	1,936	68%
Cosmic Microwave Background Data Analysis	4,096	50%
Terascale Simulations of Supernovae	2,048	43%
Quantum Chromodynamics at High Temperatures	1,024	13%

NERSC's IBM SP, named "Seaborg," is the most powerful computer in the United States available for unclassified research. Seaborg has 416 16-CPU IBM Power 3+ SMP nodes (a total of 6,656 processors) with a peak performance of 1.5 gigaflop/s per node. The system has the largest amount of aggregate memory available on any unclassified computer in the U.S.—7.8 terabytes—and a Global Parallel File System with 44 terabytes of storage. Calculations for the wave-plasma interaction and cosmic microwave background projects would not have been possible on any unclassified American machine except Seaborg.

The four computational science projects cited above are described in the following pages, along with their latest discoveries. Computational methods and codes are discussed in detail at the end of this paper.

The NERSC (National Energy Research Scientific Computing) Center is the flagship computational facility for the U.S. Department of Energy's Office of Science. For more information, see www.nersc.gov.

Electromagnetic Wave-Plasma Interactions



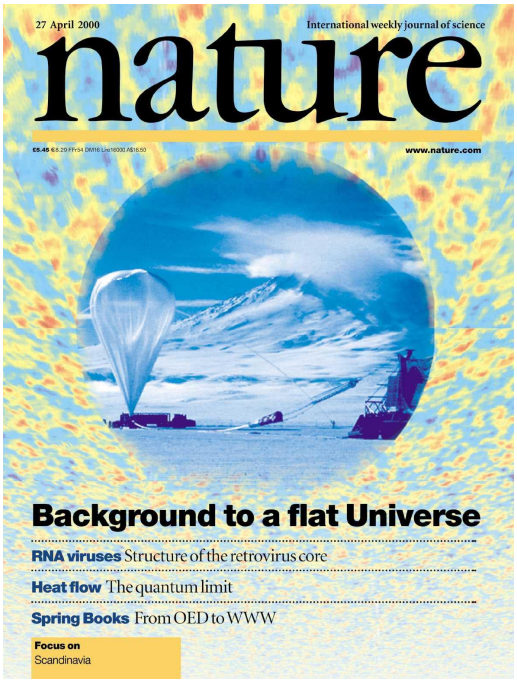
PI: Donald Batchelor, ORNL
Allocation Category: SciDAC
Code: All-orders spectral algorithms (AORSA)
Kernel: ScaLAPACK
Performance: 1.026 Gflop/s per processor (68% of peak)
Scalability: 2 Tflop/s on 1,936 processors
Memory Used: 1.86 TB
Allocation: 1.1 million MPP hours; requested and needs 3 million

Plasma waves are essential processes in systems ranging from the solar corona, to planetary magetospheres, to laboratory experiments, to commercial devices. High-power electromagnetic waves can serve as flexible tools to heat fusion plasmas to temperatures hotter than the sun and to control these plasmas. To understand the role of plasma waves in physics and to realize their potential for practical applications, it is crucial to be able to accurately calculate their behavior.

Solving the equations describing plasma waves in the complex, multi-dimensional geometries of fusion-relevant plasmas is a computational challenge of the highest order. The goal of this research is to use advanced terascale computing to meet this challenge and to obtain quantitatively accurate predictive understanding of electromagnetic wave processes.

The all-orders spectral algorithms (AORSA) code takes advantage of new computational routines in ScaLAPACK to solve the integral form of the wave equation in multiple dimensions without any restriction on wavelength relative to orbit size, and with no limit on the number of cyclotron harmonics retained. These new models give high-resolution, 2D solutions for mode conversion and high harmonic fast wave heating in tokamak geometry. In addition, they have been extended to give fully 3D solutions of the integral wave equation for minority ion cyclotron heating in stellarator geometry. The most recent results include successful simulation of experimentally observed poloidal flows that enhance tokamak confinement regimes (submitted to Physical Review, January 2003).

Cosmic Microwave Background Data Analysis



PI: Julian Borrill, LBNL and UC Berkeley
Allocation Category: Class B
Code: Maximum likelihood angular power spectrum estimation (MADCAP)
Kernel: ScaLAPACK
Performance: 750 Mflop/s per processor (50% of peak)
Scalability: 1.57 Tflop/s on 2,048 processors; 3.02 Tflop/s on 4,096 processors
Memory Used: 3.1 TB
Allocation: 1.1 million MPP hours; requested and needs 2 million

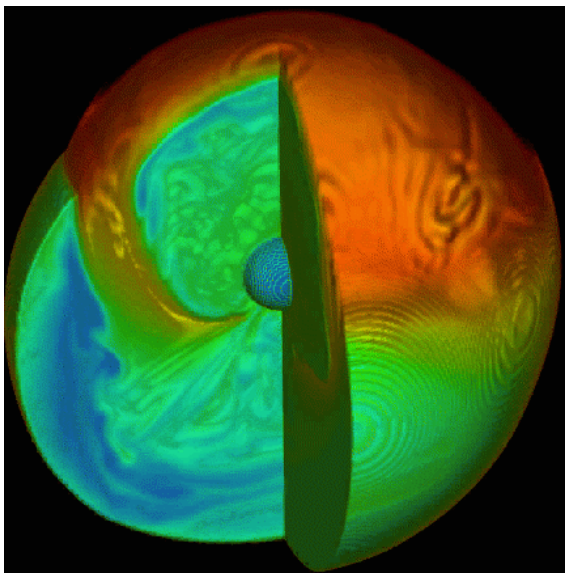
The cosmic microwave background (CMB), a "snapshot" of the Universe as it was only 300,000 years after the Big Bang, provides the most powerful discriminant between different cosmological models. Discoveries based on CMB datasets to date include:

- ruled out an entire class of models (based on topological defects) for the generation of primordial density perturbations in the Universe
- demonstrated the spatial flatness of the Universe as a whole
- coupled with supernova data, provided a first measurement of the overall mass-energy budget for the Universe.

There is much more information still to be mined from CMB data, especially from current and future satellite missions (WMAP and Planck), but the size of the datasets presents new challenges for timely and precise data analysis.

The MADCAP code (Microwave Anisotropy Dataset Computational Analysis Package) was recently extended to enable simultaneous analysis of multiple datasets and CMB polarization—the new frontier. MADCAP was rewritten to exploit extremely large parallel systems, allowing near-perfect scaling from 256 to 4,096 processors. MADCAP++ is being developed using approximate methods to handle extremely large datasets for which matrix multiplications are impractical, such as will be generated by the Planck satellite. Recent results from NASA's WMAP satellite observations of the whole CMB sky confirm the MADCAP analyses of previous partial-sky balloon datasets.

Terascale Simulations of Supernovae



PI: Anthony Mezzacappa, ORNL
Allocation Category: SciDAC
Code: neutrino scattering on lattices (OAK3D)
Kernel: ScaLAPACK (complex double precision)
Performance: 654 Mflop/s per processor (43% of peak)
Scalability: 1.34 Tflop/s on 2,048 processors
Memory Used: 783 GB
Allocation: 565,000 MPP hours; requested and needs 1.52 million

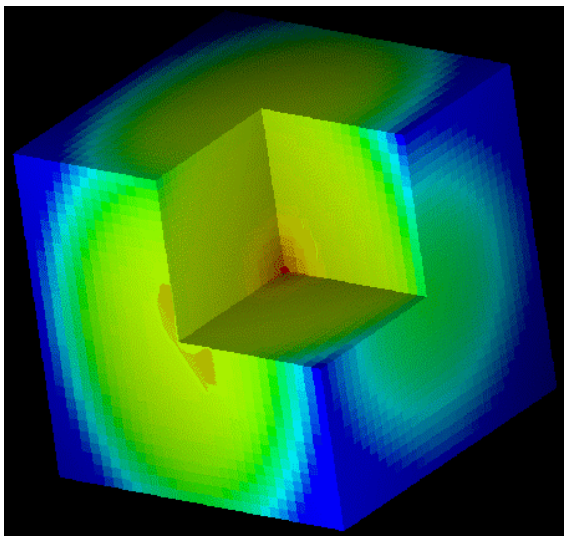
The search for the explosion mechanism of core-collapse supernovae and the computation of the nucleosynthesis in these spectacular stellar explosions are among the most important and challenging problems in computational nuclear astrophysics. Core-collapse supernovae are the most energetic explosions in the Universe, and they produce and disseminate many of the elements in the Periodic Table, without which life would not exist. They are a nexus for nuclear physics, particle physics, fluid dynamics, radiation transport, and general relativity, and serve as cosmic laboratories for matter at extremes of density, temperature, and neutronization that cannot be produced in terrestrial laboratories.

The Terascale Supernova Initiative is a multidisciplinary collaboration to develop models for core-collapse supernovae and enabling technologies in radiation transport, radiation hydrodynamics, nuclear structure, linear systems and eigenvalue solution, and collaborative visualization. This project is the first large-scale systematic effort to remove approximations from the existing supernova models and to produce models that more closely resemble a true supernova.

The project team developed the OAK3D code to study the electron capture and neutrino scattering on lattices of large arrays of nuclei that form during certain phases of star collapse. OAK3D became operational in fall 2002 and has achieved sustained speeds of 1.34 Tflop/s on 2,048 processors. These runs required double-precision complex solutions of linear equations of dimension 524,288. The results provide one more piece in the complex, three-dimensional dynamic puzzle of star collapse.

OAK3D provides the nuclear science input to the radiation transport and radiation hydrodynamics codes that are the compute engines for the supernova simulations themselves. The radiation transport and hydrodynamics codes currently achieve 10–20% of peak performance, and the collaborators are working to boost this performance.

Quantum Chromodynamics at High Temperatures



PI: Douglas Toussaint, Arizona University
Allocation Category: Class A
Code: hybrid Monte Carlo and molecular dynamics (MILC)
Kernel: iterative sparse matrix inversion
Performance: 190 Mflop/s per processor (13% of peak)
Scalability: 200 Gflop/s on 1,024 processors
Memory Used: 23 GB
Allocation: 2.3 million MPP hours; requested and needs 3.4 million

Quantum chromodynamics (QCD) is the component of the Standard Model of High Energy Physics that describes the strong interactions. This project explores the behavior of strongly interacting matter at the high temperatures that are generated in the Relativistic Heavy Ion Collider (RHIC) at Brookhaven and in the planned Large Hadron Collider at CERN.

At very high temperatures one expects to observe a phase transition or crossover from ordinary strongly interacting matter to a plasma of quarks and gluons. The plasma was a dominant state of matter in the early development of the Universe, and may be a central component of neutron stars today. The behavior of strongly interacting matter in the vicinity of the phase transition or crossover is inherently a strong coupling problem, which can only be studied from first principles through lattice gauge theory calculations. Among the issues that can uniquely be addressed by lattice calculations are the nature of the transition, the properties of the plasma, including strange quark content, and the equation of state.

The project team took advantage of NERSC's new 10 Tflop/s supercomputer to start trial runs of a QCD simulation with quark masses that are closer to physical quark masses than could previously be done on a fine grid—specifically, light quark masses at 1/10 the strange quark mass with a lattice spacing of 0.09 fm on a 64,000 × 96 lattice. They were able to run about 17 units of simulation time; 2,000 units will provide more accurate calculations of hadronic properties, topological structures, and the theoretical parameters needed for accelerator experiments.

Computational Methods and Codes

For the first three simulations described in this paper, the heart of the computation involves large, dense matrices, so the application developers chose to use subroutines from the ScaLAPACK library to achieve high efficiencies. In these cases one would expect to achieve high performance, since the arithmetic operations are essentially the same as those used in the Linpack benchmark. As a reference, Seaborg has achieved over 72% efficiency on the Linpack benchmark.

The ScaLAPACK (Scalable LAPACK) library is a follow-on project of the LAPACK Project, both of which were funded in part by the Mathematical, Information and Computational Sciences (MICS) Division in the DOE Office of Science. Both libraries provide application developers with high-quality kernels that they can use in their simulations on high-performance computer architectures. The LAPACK library contains routines that provide robust and efficient implementations of many numerical linear algebra algorithms for dense matrix computation, including solvers for systems of linear equations, least squares problems, and eigenvalue problems. In addition, the ScaLAPACK library provides parallel implementations of many of the LAPACK routines for distributed-memory parallel computers.

Carefully designed algorithms allow the ScaLAPACK library to achieve high performance and efficiency. These algorithms

- attain a favorable computation/communication ratio for large problems
- maintain load balance by intelligently distributing the data to the processors
- minimize network communication
- take advantage of memory hierarchy by choosing optimal block sizes
- use architecture-tuned single-processor Basic Linear Algebra Subroutines (BLAS), such as matrix-matrix multiplication.

Computational methods and codes for each simulation are discussed in detail below.

Electromagnetic Wave-Plasma Interactions

The simulations in the electromagnetic wave-plasma interactions using the AORSA code require the solution of linear systems of equations, for which the coefficient matrices are dense, complex, and nonsymmetric. These dense linear systems arise from the solution of an integral equation relating the plasma current and the wave electric field. The size of the matrices is constrained by the amount of memory that is available. The simulation currently is solving linear systems whose sizes are typically 120,000 for 2D problems, and the execution time is dominated by the solution of these linear systems. For this application, two routines from the ScaLAPACK library are used, PZGETRF and PZGETRS. PZGETRF computes a triangular factorization of the coefficient matrix, while PZGETRS solves a linear system by using the triangular factors.

Cosmic Microwave Background Data Analysis

In the case of cosmic microwave background data analysis, the computational time using the MADCAP package is dominated by dense matrix linear algebra. The core of MADCAP involves two steps. The first step involves the explicit inversion of a large, dense, symmetric and positive definite correlation matrix via a Cholesky decomposition using the ScaLAPACK routines PDPOTRF and PDPOTRI. The dimensions of these correlation matrices range from 10,000 to 100,000. The second step involves multiplication of a set of block-structured matrices by the appropriate blocks of this inverted matrix using the Level 3 PBLAS routine PDGEMM. For the second step, the matrix multiplications are entirely independent of one another. By taking advantage of this, the amount of communication can be reduced substantially by subdividing the total processor set into independent gangs. The matrix multiplications constitute about 90% of the execution time, while the factorization of the correlation matrix and the computation of its inverse only require about 5% of the execution time.

Terascale Simulations of Supernovae

In the terascale simulations of supernovae, OAK3D is a code which tackles the quantum mechanical many-body theory of the structure of nuclei to determine how neutrinos scatter on lattices. It is a parallel, 3D coordinate space, fully time-dependent Hartree-Fock code with realistic (Skyrme) nuclear forces. Mathematically the effort reduces this problem to an iteration over dense complex nonsymmetric linear system solutions coupled with matrix-vector operations to impose time evolution. Several routines from the ScaLAPACK and PBLAS libraries are used to perform this task. Matrix factorizations are performed using PZGETRF, and the triangular solutions are carried out using PZTRSM. The size of the linear systems being solved is determined by the method of discretizing the equations. For a $48 \times 48 \times 48$ space lattice, the size of the matrix is 221,184. Approximately 90% of the computing time on a typical run is spent solving the linear systems. On 2,048 processors of Seaborg, OAK3D took six hours to compute a full iteration at a sustained rate of about 43% of peak.

Quantum Chromodynamics at High Temperatures

At high temperatures, quantum chromodynamics numerical calculations are dominated by the solution of a sparse system of linear equations where the coefficient matrix is Hermitian and positive definite and has highly regular structure. The solution is related to the long-range force in the molecular dynamics evolution of the gluon fields due to the presence of quarks. In the current simulations, the linear system is solved using a custom implementation of the conjugate gradient method with a red-black decomposition. The main reason for this is the special structure of the matrix, where each entry is a 3×3 special unitary matrix. Although this is the same as a complex matrix with three times the dimension, the special properties of unitary submatrices would be lost if it were viewed in this way. The size of the matrices depends on the lattices. For a $28 \times 28 \times 28 \times 96$ lattice, the size of the matrix is 3,200,000. Fortunately, the matrix is quite sparse; each row/column has 96 off-diagonal nonzero entries. The amount of time for the iterative solution of the linear system varies with the quark mass; at the lightest quark mass (which is the most demanding), it is about 70% of the total execution time.