The Challenge and Promise of Scientific Computing

Jeff Nichols, Director
Computer Science and Mathematics
Oak Ridge National Laboratory

March 4, 2003
Outline

- Oak Ridge National Laboratory’s Center for Computational Science (CCS)
- Promise of the Computational Scientist; Their Software Development Activities and Application
- Challenge of Delivering Leadership Class Petascale Computers
- Promise of Massively Parallel Vector Architectures (Cray X1, X2, …)
Center for Computational Sciences
DOE’s Advanced Computing Research Testbed
Center for Computational Sciences
Established 1992 as part of DOE HPCRC program

- Partnership in Computational Science
  - A proposal to meet challenges of DOE HPCC program
- One of two HPCRC centers (along with Los Alamos)
- User Facility designation (1994)
- Principal resource for SciDAC

Dr. Marburger, Dr. Decker, and Dr. Nelson at the dedication of CCS
Goals of the Center for Computational Sciences

- Evaluate new computer hardware for science
- Procure the largest scale systems (beyond vendors design point) and develop software to manage and make them useful
- Deliver leadership-class computing for DOE science
  - By 2005: 50x performance on major scientific simulations
  - By 2008: 1000x performance
- Educate and train next generation computational scientist
Focused on grand challenge scientific applications

- SciDAC Astrophysics
- Genomes to Life
- Nanophase Materials
- SciDAC Climate
- SciDAC Fusion
- SciDAC Chemistry
ORNL’s Computational Materials Research Program

- Move from modeling bulk materials and film heterostructures modeling to modeling of 0D and 1D nanostructures
  - Predictive modeling of new data storage technology (magnetic / molecular)
  - Exploring and applying quantum effects
- Solve “big” materials science problems
  - Mechanism of high temperature superconductivity
  - Fundamental question on transport in 1D systems
- Move towards open source community codes / reorganize and integrate major codes and methods (multiscale modeling, physics at all length scales, integrate manybody and DFT methods)
- Expected growth in nanosciences within DOE
  - 2003 $133M
  - 2004 $197M
  - 2005 $217M
  - 2006 $239M
- CNMS-NTI and Computational Nanoscience call are part of this

Carbon nanotubes are quasi 1D quantum wires
Spin structure of FeMn in proximity of Co 1D quantum wires
Nanoscience projections - unique opportunities for modeling

Spallation Neutron Source (SNS)
Center for Nanophase Materials Science (CNMS)

Synthesis & Characterization
~ 12,000 atoms
~ 4,000 surface + sub

Surfaces and interfaces are important; new quantum confinement phenomena require quantum description

First principles simulation size
Current largest 2176-atoms (3 TF)

Real device size
Nano dot: 5x5x5 (6TF)
Nano-wire: 10x10x60 (250TF)
New algorithms

OAK RIDGE NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY
Climate (CCSM) simulation resource projections

At current scientific complexity, a century simulation requires 12.5 days. Single researcher transfers 80Gb/day and generates 30TB storage each year.

Science drivers: regional detail / comprehensive model

Machine and Data Requirements

- Blue line represents total national resource dedicated to CCSM simulations and expected future growth to meet demands of increased model complexity.
- Red line shows data volume generated for each century simulated.

CCSM Coupled Model Resolution

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmosphere</td>
<td>230kmL26</td>
<td>30kmL96</td>
</tr>
<tr>
<td>Land</td>
<td>50km</td>
<td>5km</td>
</tr>
<tr>
<td>Ocean</td>
<td>100kmL40</td>
<td>10kmL80</td>
</tr>
<tr>
<td>Sea Ice</td>
<td>100km</td>
<td>10km</td>
</tr>
<tr>
<td>Model years/day</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>National Resource</td>
<td>3</td>
<td>750</td>
</tr>
<tr>
<td>(dedicated TF)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage (TB/century)</td>
<td>1</td>
<td>250</td>
</tr>
</tbody>
</table>

**Science drivers:** regional detail / comprehensive model

**Machine and Data Requirements:**

- Blue line represents total national resource dedicated to CCSM simulations and expected future growth to meet demands of increased model complexity.
- Red line shows data volume generated for each century simulated.

**CCSM Coupled Model Resolution:**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmosphere</td>
<td>230kmL26</td>
<td>30kmL96</td>
</tr>
<tr>
<td>Land</td>
<td>50km</td>
<td>5km</td>
</tr>
<tr>
<td>Ocean</td>
<td>100kmL40</td>
<td>10kmL80</td>
</tr>
<tr>
<td>Sea Ice</td>
<td>100km</td>
<td>10km</td>
</tr>
<tr>
<td>Model years/day</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>National Resource</td>
<td>3</td>
<td>750</td>
</tr>
<tr>
<td>(dedicated TF)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage (TB/century)</td>
<td>1</td>
<td>250</td>
</tr>
</tbody>
</table>

GTL resource projections

- Comparative Genomics
- Constraint-Based Flexible Docking
- Genome-scale protein threading
- Community metabolic regulatory, signaling simulations
- Molecular machine classical simulation
- Constrained rigid docking
- Cell, pathway, and network simulation
- Protein machine interactions
- Molecule-based cell simulation

Current U.S. computing: 1 TF*

* Teraflops
**Proposed GTL Facilities: Computational Needs and Tools**

**Facility 1: Production and Characterization of Proteins ($200M)**

- genome annotation
- regulatory element and operon identification
- metabolic pathway analysis

**Facility 2: Whole Proteome Analysis ($175M)**

- mass spec data analysis
- expression analysis and clustering
- metabolic and regulatory network modeling

**Facility 3: Characterization and Imaging of Molecular Machines ($225M)**

- image analysis
- mass spec analysis
- protein / machine modeling
- docking and molecular dynamics

**Facility 4: Analysis and Modeling of Cellular Systems ($285M)**

- new tools & infrastructure for analyses of cell state coupling experiment, analysis, and theory in a recursive fashion to reveal networks, function
- metabolic simulation
- regulatory simulation
- cell modeling and simulations

**Center for Computational Sciences**

**OAK RIDGE NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY**
Compact stellarator science discoveries enabled by CCS computers

- Advanced computation and modeling plays a critical role in fusion research
  - Historical usage of super-computers, solving “Grand Challenge” problems. NERSC began existence as MFECC (Magnetic Fusion Energy Computer Center)
  - Fusion funded projects in SciDAC
- CCS is already providing an increasing fraction of fusion computing – plasma turbulence problems (General Atomics, PPPL), MHD stability (PPPL, U. Wisc.), stellarator analysis (ORNL, PPPL), plasma waves (ORNL, MIT, PPPL)
- We expect the fusion program to grow (~FY05) and fusion computing to grow
  - Presidential announcement of commitment to fusion energy
  - US rejoining ITER negotiations
  - OFES planning a new round of SciDAC projects ($3M in FY04)
  - Planning a major initiative – Fusion Simulation Project (FSP), ~$20M/yr for 5 yr, joint between OFES and OASCR
  - Ann Davies has announced that OFES SciDAC projects will be oriented toward laying basis for FSP
- ORNL has strengths and can play major roles (fusion physics, computer science, computer hardware) in ongoing fusion computing, SciDAC, and FSP

Massively parallel particle simulations for physics analysis

Physics/engineering based parallel optimizations used to synthesize innovative compact QPS stellarators

OAK RIDGE NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY
TeraScale Supernova Initiative

- Find the supernova explosion mechanism(s).
- Predict supernova element synthesis.
- Significantly advance the nuclear theory/particle theory frontier.
- Make it possible to use supernovae as nuclear and particle physics labs.
  - Supernova models are bridge between observations and such discovery.
- Predict supernova gravitational wave emission.
  - The first LIGO detection may be of waves from a supernova!
- Motivation and guidance for existing and proposed experimental facilities.
  - RIA; to better understand the nuclear physics of the r-process
  - National Underground Science Laboratory; site of a next generation supernova neutrino detection capability (1/2 Mton, extra-Galactic)
  - SNO; pivotal role in future supernova neutrino detection and analysis
  - RHIC; to characterize the properties of high density nuclear matter critical to our understanding of stellar core bounce dynamics
Chemistry projections: advanced methods for electronic structure

• Objectives:
  – O(N) methods free of basis set error
  – Functioning DFT, HF, gradients, TDDFT
  – Investigating correlated models
• Distinguishing features
  – Multiresolution in multiwavelet bases
  – Near-spectral precision
  – Efficient computation in high dimensions
• Impact
  – Fast methods with guaranteed precision
  – Reliable description of excited states, response properties, …
  – Possible path to O(N) electron correlation
  – New framework for chemical computation
• Requirements
  – Fully quantitative coupled cluster for catalysis or combustion
  – O(10^2) TB memory, O(10^3) TB disk
  – Decades of TFLOP/s of computation

1992
First Paragon XP/35
KSR1-64
CCS formed

1993
PVM used to create first International Grid

1994
PVM wins R&D 100

1995
Install Paragon XP/150
Worlds fastest computer
Connected by fastest network OC-12 to Sandia

1996
ORNL-SNL create first high-performance computational Grid

1997
R&D 100 Award for successful development and deployment of HPSS

1998
Developed first application to sustain 1 TF

1999
NetSolve wins R&D 100
ATLAS wins R&D100

2000
Longstanding climate simulation milestone first met on CCS Compaq

2001
First IBM Power 4
SciDAC leadership Human Genome

2002
IBM Blue Gene CRADA to develop super scalar algorithms begins

2003
Design changes for X2 based on ORNL-Cray partnership

Construction starts on new CCS building World class DOE facility

Partnership with Cray on X1 begins
U.S. HPC community: Where are we? CoGS

- Clusters of General-purpose SMPs
- Clusters
  - Multiple independent systems
  - Interconnected
- General purpose
  - Not designed specifically for HPC
- SMPs
  - Multiple processors per system
  - Shared memory
CoGS dominate U.S. HPC

- Largest DOE systems
  - NNSA: LANL (HP), LLNL (IBM, Intel)
  - SC: LBL (IBM), ORNL (IBM), ANL (Intel), PNL (Intel)
- Largest NSF systems
  - PSC (HP), NCAR (IBM), SDSC (IBM), NCSA (Intel)
- Largest (known) DOD systems
  - NAVO (IBM), ARL (IBM)
- Largest of other US agencies
  - NOAA (Intel), NASA (HP)
- Largest state systems
  - LSU (Intel), SUNY (Intel), FSU (IBM), NCSC (IBM)
CoGS example

- ORNL CCS Cheetah
- 27 IBM p690 systems
  - 32 processors each (1.3 GHz Power4)
  - 32-128 GB of shared memory
- IBM Switch2 interconnect
- 4.5 TF of peak performance
U.S. scientific computing capability is 20x behind the Japanese - impacting our ability to do science

Ouch!
What is the Earth Simulator?

- A building
- An SMP cluster!
  - 640 NEC SX-6 systems
  - 8 processors each
  - 16 GB of shared memory each
- Not general purpose
  - Vector processors
  - Ultra-high memory and interconnect bandwidths

http://www.es.jamstec.go.jp/esc/eng/
Where have we been?

- HPC crisis in late 1980's
- HPCRC program advanced massively parallel processing (MPP)
- CoGS took over in 1990's
  - Adaptation of general-purpose systems
  - Attrition and consolidation of vendors
- Behind!
Where has HPC been for U.S. open science?

- Leadership scale computer
- Leading DOE-SC computer

- Earth Simulator
- Top DOE-SC machine
- Extreme symptom
- Chronic problem
- Japanese or ASCI
Even though clusters of general purpose SMPs dominate U.S. HPC....

- Largest DOE systems
  - NNSA: LANL (HP), LLNL (IBM, Intel)
  - SC: LBL (IBM), ORNL (IBM), ANL (Intel), PNL (Intel)

- Largest NSF systems
  - PSC (HP), NCAR (IBM), SDSC (IBM), NCSA (Intel)

- Largest (known) DOD systems
  - NAVO (IBM), ARL (IBM)

- Largest of other US agencies
  - NOAA (Intel), NASA (HP)

- Largest state systems
  - LSU (Intel), SUNY (Intel), FSU (IBM), NCSC (IBM)
the research community we serve and our users found:

- Increasing
  - Processor speed
  - Parallelism
  - Algorithm efficiency
  - Computational requirements for scientific simulation
  - Relative memory and interconnect latencies
  - Power consumption
  - Heat generation
  - System complexity
  - Software complexity

- Decreasing
  - Relative memory bandwidth
  - Relative interconnect bandwidth
  - Relative I/O speed
  - % of peak performance

Our users requested a balanced, leadership-class system based on science needs
How do we get there? DOE is the agency that has the experience to provide leadership in scientific computing for the nation.

- 30 years of experience operating the world’s largest computing facilities
- Mission goal to solve computational grand challenges in nanotechnology, biology, and climate among others
- Office of Science Labs foster multidisciplinary teams required for large science
- SC culture provides the organizational framework needed for development of community codes
- SciDAC initiative supplies the needed software infrastructure for large scale scientific computing
Charge from Dr. Orbach

- U.S. cannot afford to be behind in an enabling tool with broad impact in U.S. Scientific Leadership in areas such as nanotechnology, biotechnology, fusion and climate by a factor of 10 to 50 or more

- Review “. . . current state of the national computer vendor community relative to high performance computing”
- “. . . Vision for what realistically should be accomplished in the next five years within the Office of Science in high performance computing”
### Office of Science Vision: FY04 - FY08

#### Scientific Discovery Through 21st Century Computation
- Develop computer architectures that will dramatically improve hardware performance on DOE scientific problems.
- Develop scientific simulation codes to fully exploit the capabilities of terascale computers for DOE problems.
- For Simulation of:
  - Climate
  - Nano-Materials
  - Protein Folding
  - Cell Functions via Genomes to Life
  - Origins of Mass (QCD)
  - Quark-Gluon Plasma
  - Fusion Confinement
  - Combustion

#### Revolutionary New Materials Through Nanoscience
- Five Nanoscale Research Centers linked to large scientific research instruments at the DOE National Labs to enable:
  - High Efficiency energy storage & conversion.
  - Miniature sensors.
  - Nanocatalysts with enhanced specificity and reactivity.
  - Novel materials that are light weight, strong and conductive.
  - Low cost, high-efficiency photovoltaic cells.
- Harnessing microbes, microbial communities and other organisms to produce energy, sequester carbon, and remediate hazardous waste sites.

#### National Security, a Clean Environment & Energy Security Through Basic Research
- Demonstrate the scientific and technical feasibility of fusion energy on ITER by 2020.
- New materials for lighter weight vehicles, more efficient engines, more efficient photovoltaic cells.
- Understand the origins of the Universe:
  - Mass
  - Accelerating Universe
  - Beginning of Time
  - Dominance of Matter over Anti-matter
- Create the quark-gluon plasma that existed immediately after the “Big Bang”, providing fundamental insights in the evolution of the early universe.
- Nature of Quarks and Gluons: internal structure of protons and neutrons.

#### Uncovering the Origins of Time and Matter
- By 2004: Initiate pilot “Laboratory Science Teacher Professional Development Program”.
- By 2006: Provide hands-on experience in science and math research to 2,500 K-14 teachers each year.
- By 2006: Complete Spallation Neutron Source for improved drugs and materials.
- By 2008: Complete 5 unique Nanoscience Research Centers, providing the tools for nanoscale machines, designer materials & medical advances.
- By 2009: Construct a Linear Coherent Light Source, providing the ability to image atoms.
SciDAC provides the foundation
Community Models on SciDAC Computing Resources

Hardware Infrastructure

Software Infrastructure

Operating Systems

Collaboratories

Data Grids

Computing Systems Software

Data Analysis & Visualization

Programming Environments

Scientific Data Management

Problem-solving Environments

Mathematics

Scientific Simulation Codes

BES, BER

FES, HENP

OAK RIDGE NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY
Office of Science proposed a plan to provide leadership in scientific computing for the Nation

**GOALS:** Deliver UltraScale Scientific Computing Capability (USSCC):

- By 2005: Deliver computational performance 50 times greater than now achieved for selected scientific modeling problems
- By 2008: Deliver computational performance a 1,000 times greater than is now achieved for selected scientific modeling problems

Graphical depiction of increased spatial resolution in climate models - a map of DC area showing today’s grid in black and the USSCC grid in red.

USSCC Funding

Deliver factor 50 Improvement in delivered Performance over FY 2002

Deliver factor 1000 Improvement in delivered Performance over FY 2002

Begin Cray X-1 Evaluation

Begin Installation First USSCC System

Begin Installation Second USSCC System

Begin Installation 3rd USSCC System - 1petaflop

Retire First USSCC System
DOE-SC has a unique opportunity and responsibility to provide leadership in computational science

Goals
• Deliver Leadership Class Scientific Computing Capability to:
  – By 2005: Deliver computational performance 50 times greater than now achieved for selected scientific modeling problems
  – By 2008: Deliver computational performance a 1,000 times greater than is now achieved for selected scientific modeling problems (Petascale)

Implementing Actions
• Begin Cray X-1 Evaluation at ORNL
• Begin Installation First Leadership Class System
  – Deliver factor 50 Improvement in delivered Performance over FY 2002
  – Cray X-1 is the best U.S. option for such a system
Our users requested a balanced system based on science needs

- 12.8 gigaflops processor
- 64 processors per cabinet
- Cray X1 offers best opportunity for delivered performance in scientific computation
- Limitations: Unproven system
  - Scaling problems to overcome
  - Stability of system software
## Preliminary X1 evaluation promising

<table>
<thead>
<tr>
<th></th>
<th>SX6 (MFlops)</th>
<th>X1 (MFlops)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Climate code</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ocean Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kernel A</td>
<td>3,900</td>
<td>4,400</td>
</tr>
<tr>
<td>Kernel B</td>
<td>2,500</td>
<td>2,700</td>
</tr>
<tr>
<td>Kernel C</td>
<td>1,000</td>
<td>700</td>
</tr>
<tr>
<td>Kernel D</td>
<td>1,800</td>
<td>1,400</td>
</tr>
<tr>
<td><strong>Matrix Multiply</strong></td>
<td>4,000</td>
<td>10,000</td>
</tr>
<tr>
<td><strong>Ocean Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FE Weather Code</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kernel A</td>
<td>84</td>
<td>540</td>
</tr>
<tr>
<td>Kernel B</td>
<td>2,000</td>
<td>5,900</td>
</tr>
<tr>
<td>Kernel C</td>
<td>150</td>
<td>5,300</td>
</tr>
<tr>
<td><strong>FE Weather Code</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kernel A</td>
<td>2,700</td>
<td>4,900</td>
</tr>
</tbody>
</table>
Preliminary climate and materials performance on Cray X1 promising

**Climate**

Co-Array Fortran

SHMEM

32 processors, one p690

**POP Performance (higher is better)**

**Materials**

QMC Performance (smaller is better)

55% of time spent generating formatted output!
Cray X1/Black Widow
4 phase evaluation and deployment

BW architecture per DOE apps

Phase 1
3.2TF, 1TB, 20TB

Phase 2
8.192TF, 2.621TB, 24TB

Phase 3
40.96TF, 13.107TB, 102TB

Phase 4
120TF, 40TB, 400TB

Simulator

Phase 1
3TF (256 CPU)

Phase 2
8TF (640 CPU)

Phase 3
40TF (3200 CPU)

Phase 4
120TF

ORNL/Cray proposed plans and path forward for a petascale computing capability for science

- Cray X1/X1e
- Cray Red Storm
- Cray X2/X2+
What does a petascale system require in 2008?

• Based on estimates of IBM, Cray, HP
• 20,000 to 33,000 processors (ES has 5,120!)
• 200 to 300 cabinets
• 30,000 ft\(^2\) to 35,000 ft\(^2\) including CPUs, disks, cooling room, archival storage
• 12 to 18 MW of power
Requirements drove construction of a new world class facility capable of housing petascale computers

- Space and power for world class facilities
  - 40,000 ft² computer center
  - 36” raised floor; 18 ft. deck-to-deck
  - 8 megawatts of power (expandable)
- Office space for 400 staff members
- Classroom and training areas for users
- High ceiling area for visualization lab (Cave, Power Wall, Access Grid, etc.)
- Separate lab areas for computer science and network research
- Strong university partnerships
University partnerships help us achieve our goals research, education and training

- Establish a strong student and postdoctoral program; graduate program in computational science
- Develop new initiatives/capabilities
- Hire scientists and engineers jointly
- Expand Research Alliance for Minorities
- Build the Joint Institute for Computational Sciences for expanded collaborations

- Computational Sciences Initiative with UT (~$1M)
  - Faculty release time
  - Graduate students and post-docs
- 12 Joint faculty appointments with UT
- 12 Joint faculty appointments with Ga. Tech
- 12 Joint faculty appointments with Va. Tech
State-of-the-Art network connectivity to the Nation’s three principal scientific networks (ESnet, Internet 2, NSF Backbone)
In Summary:

- DOE Office of Science has a unique opportunity with the people, programs, facilities and partners to deploy a leadership class petascale computer.
- Computational scientists have been developing software tools and are uniquely poised to apply them against DOE’s grand challenge problems.
- Leadership class petascale computers are on the horizon.
- Cray X1/X2 massively parallel vector architectures are arriving soon.