

Comprehensive Scientific Support Of Large Scale Parallel Computation



David Skinner, NERSC

Overview



- **What is NERSC? An insider's view.**
- **What is Comprehensive Scientific Support?**
- **Summary of INCITE 2004**

Focus on service to science projects

Facility provides:

Identification of research compute needs

Experience with real scientific applications



New tools and solutions

Research provides:

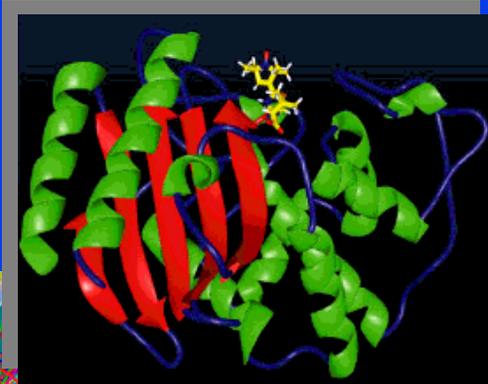
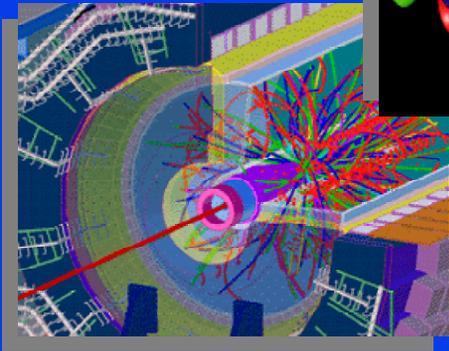
Rapid introduction of new technology

Perspectives on emerging architectures

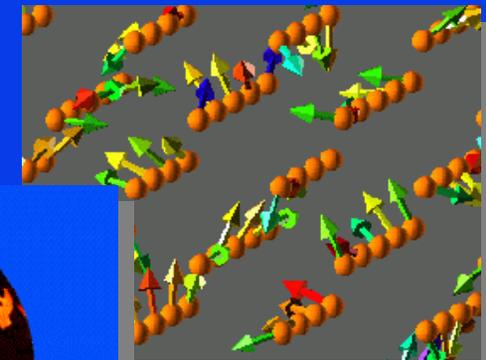
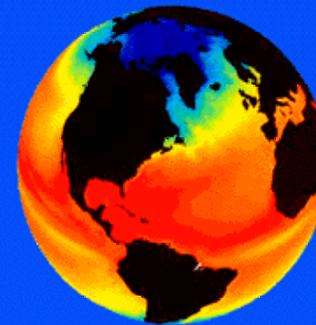
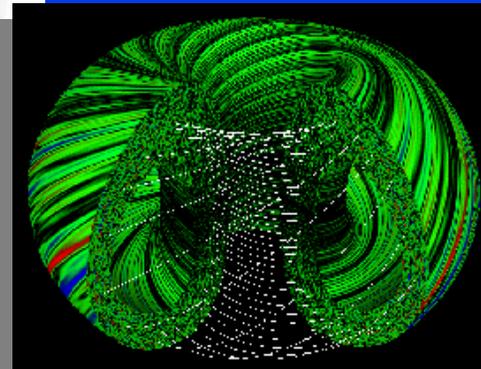


National Energy Research Scientific Computing Center

- Serves all disciplines of the DOE Office of Science
- ~2000 Users in ~400 projects



- Focus on large-scale computing

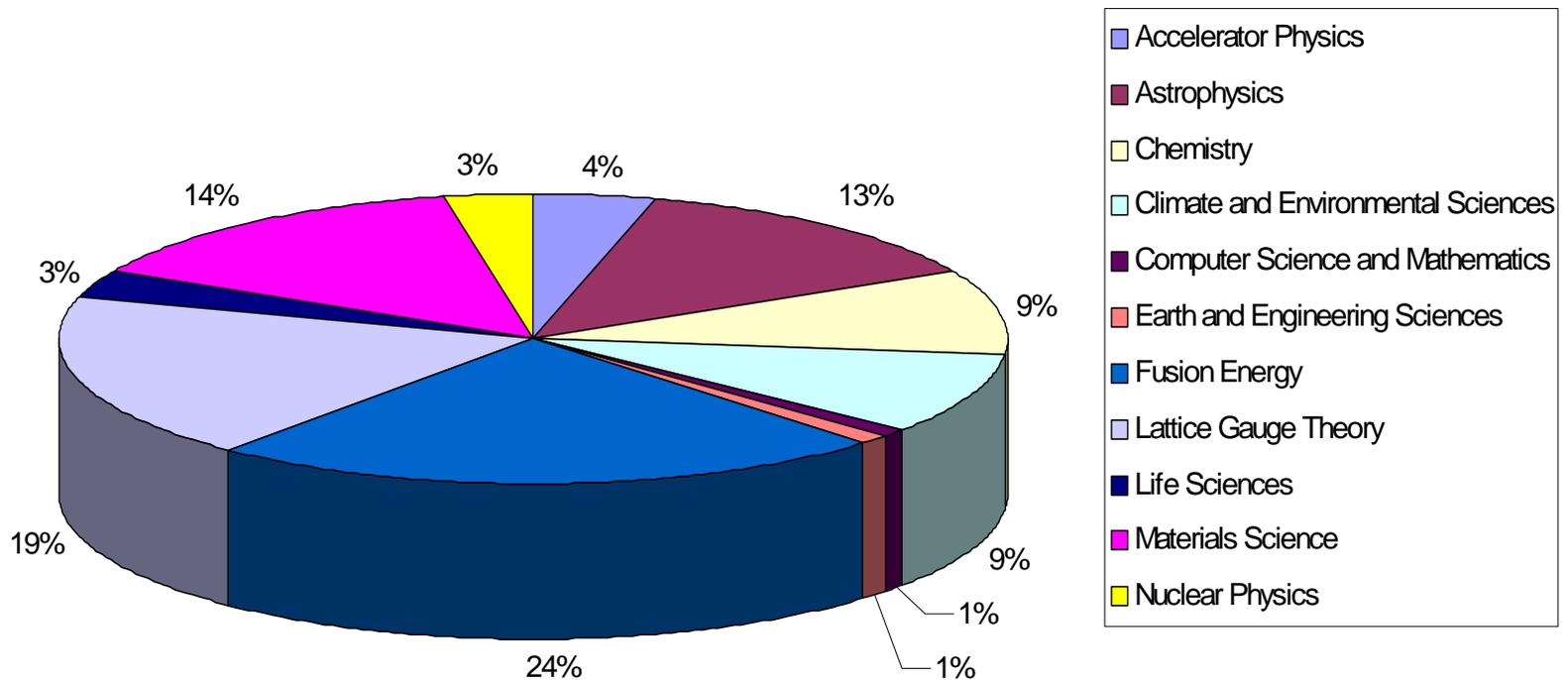


NERSC: Mission and Customers



NERSC Usage by Scientific Discipline, FY02

NERSC provides reliable computing infrastructure, HPC consultancy, and accurate resource accounting to a wide spectrum of science areas.



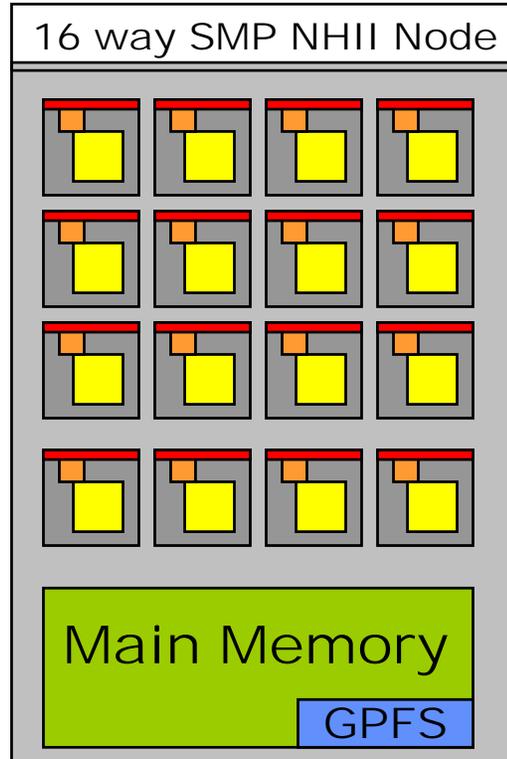
Comprehensive Scientific Support



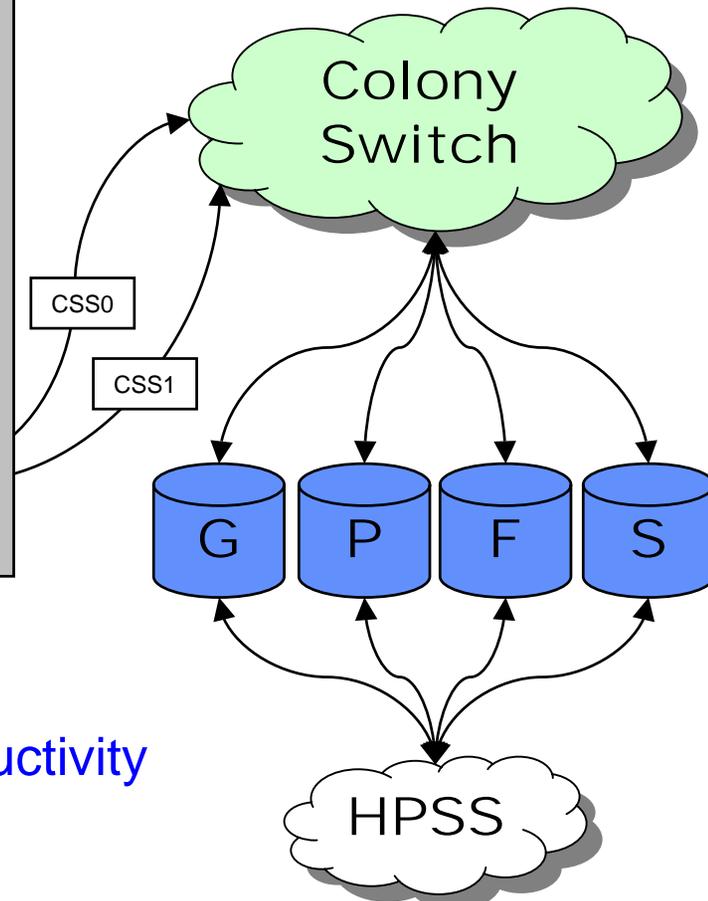
Here's the machine. What more do you need?

IBM SP

380 x



Resource	Speed	Bytes
Registers	3 ns	2560 B
L1 Cache	5 ns	32 KB
L2 Cache	45 ns	8 MB
Main Memory	300 ns	16 GB
Remote Memory	19 us	7 TB
GPFS	10 ms	50 TB
HPSS	5 s	9 PB



- Many Time Scales, Many Choices
- NERSC Center is Focused on Scientific Productivity
- Removing Bottlenecks Speeds Research

Focus on Achieving Scientific Goals



- **Hardware and Software are only the start**
 - Queue Policies, Allocation, Environment
 - User Training, Code analysis and Tuning
 - HW and SW testing, Reliability
- **An integrated approach enables rapid scientific impact.**
- **Feedback from Researchers is valued:**
 - NUG (NERSC Users' Group)
 - Collaborative contact with NERSC

Not just optimizing codes, NERSC optimizes the HPC process

Comprehensive Support



Why

- **So researchers can spend more time researching**
- **For many projects NERSC is a hub for HPC services : data storage , visualization, CVS, web**
- **Quick Answers = Highly Productive Computing**

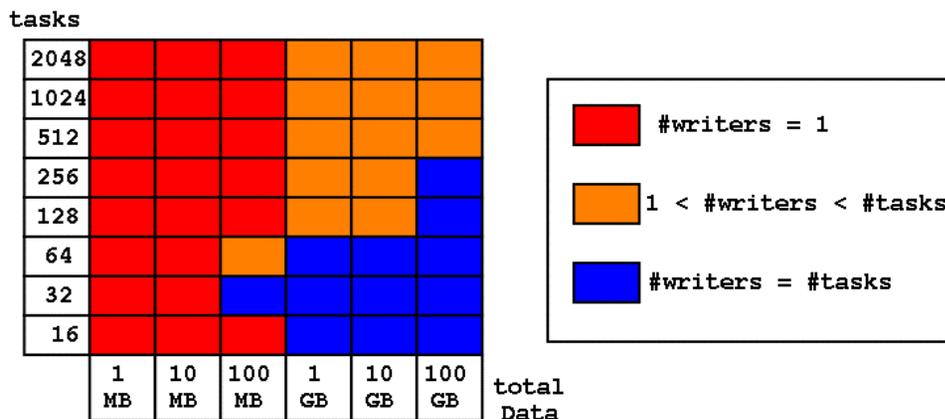
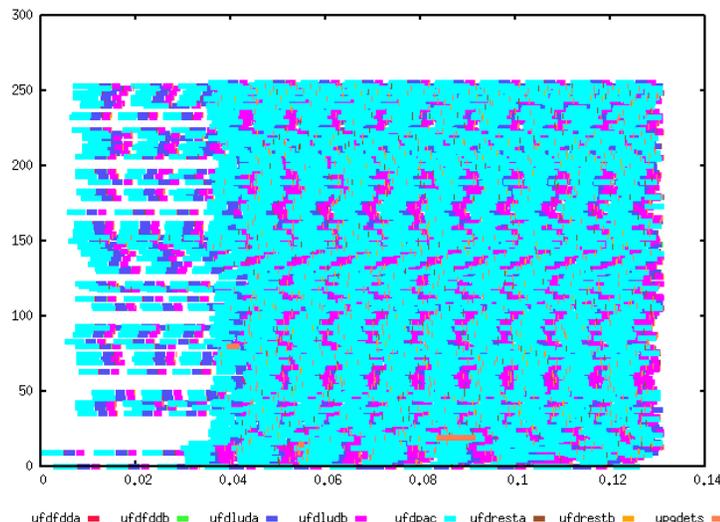
How

- **HPC consultancy and visualization resources**
 - **A Breadth of Scientific and HPC Expertise**
- **Systems Administration and Operation**
 - **High availability, RAS Expertise**
- **Networking and Security**
 - **Secure access without jumping through hoops**

Examples of Comprehensive Support: Performance Tuning



- Analysis of Parallel Performance
 - Detailed view of how the code performs
 - Little effort by researcher
- Code Performance Tuning
 - Optimally tuned math libs
 - MPI Tuning
 - Scaling and Job structuring
- Parallel I/O strategies
 - Tuning for seaborg
 - Using optimal I/O libs



Examples of Comprehensive Support: Information Services



- An increasing number of information services are being offered through www.nerisc.gov
 - Online documentation and tutorials
 - Machine Status / MOTD
 - Queue look, history, job performance records
 - Project level summaries via NIM

Send us your ideas for improvement.
Tell us what works for you.

Examples of Comprehensive Support: Information Services

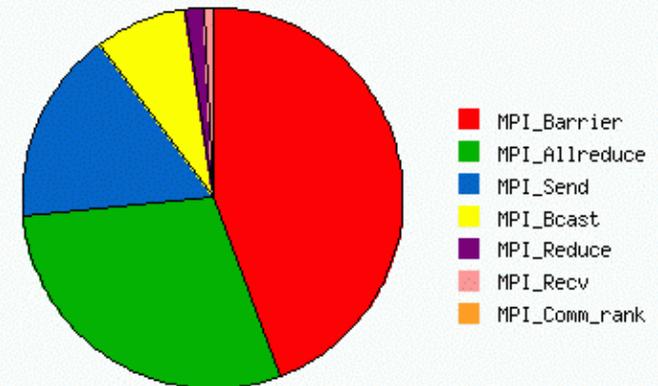


Login to www.nersc.gov

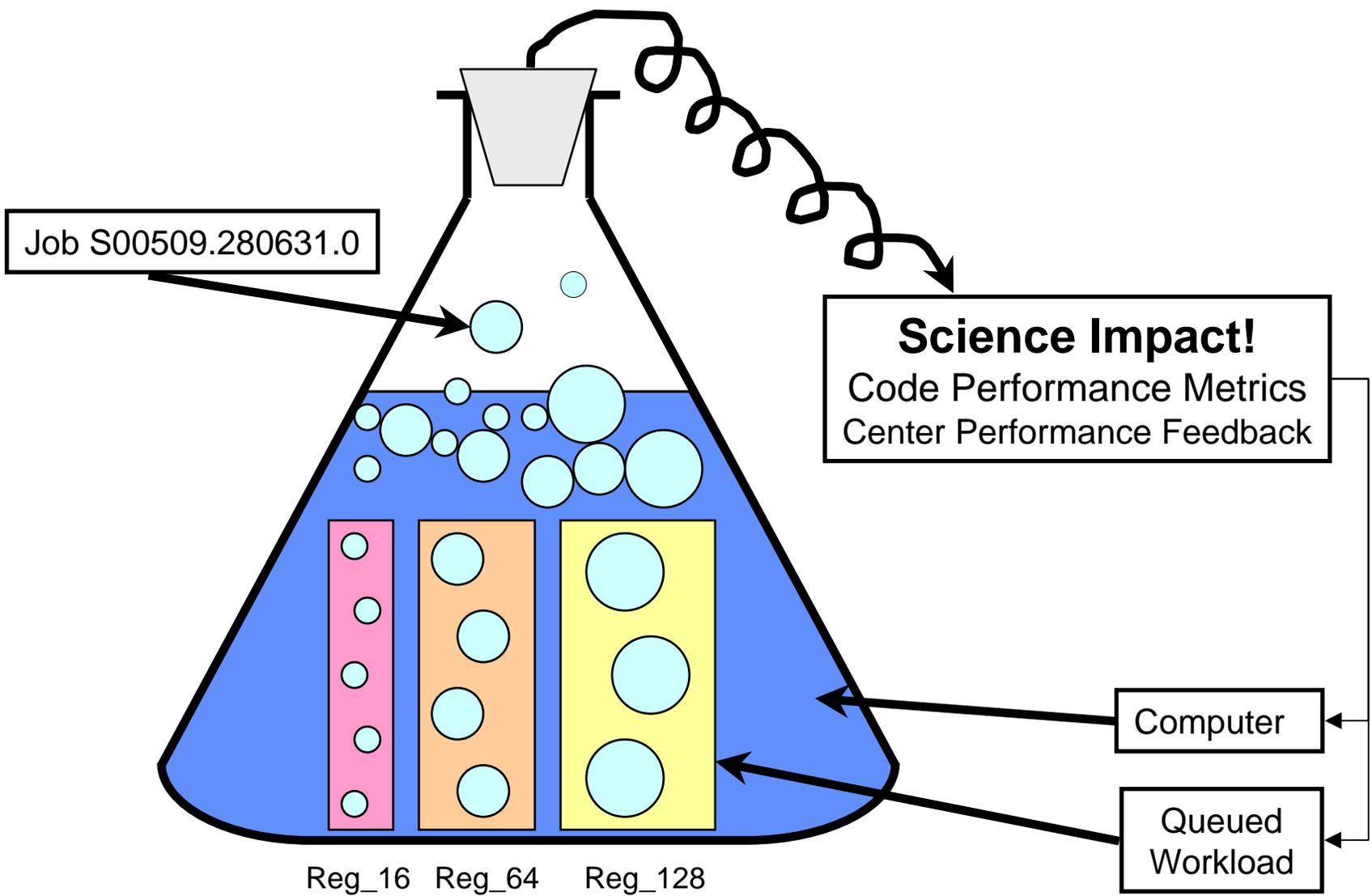
Batch job history and performance data are web accessible

Function	Total calls	Total time (sec)		Total buffer size (MB)	Avg. Buffer Size/call (Bytes)
MPI_Barrier	6.02e+05	3.48e+05	44.23%	0	0
MPI_Allreduce	3.18e+07	2.31e+05	29.33%	3.61e+05	11,936
MPI_Send	1.29e+08	1.29e+05	16.36%	5.24e+04	426
MPI_Bcast	5.73e+07	6.08e+04	7.73%	5.39e+04	987
MPI_Reduce	1.08e+08	1.24e+04	1.58%	1.66e+05	1,620
MPI_Recv	1.29e+08	6.11e+03	0.78%	5.24e+04	426
MPI_Comm_rank	1.14e+03	5.92e-01	7.52e-05%	0	0
MPI_Comm_size	6.66e+02	0	0%	0	0

Percent of MPI Time



The HPC Center: A Chemist's View



Expanding Scientific Understanding

A Quantum Mechanical understanding of how cells protect themselves from the photosynthetic engine.

Cosmological understanding of the distribution of atomic species generated by Supernovae.

Fundamental understanding of Turbulence and Mixing processes from astrophysical to microscopic scales.

INCITE 2004: How we got there



- **MPI tuning (15-40% overall improvements)**
 - Scalability and Load Balance
 - Advanced Math and MPI libraries
- **Parallel I/O**
 - Optimizing concurrent writes
 - Implementing Parallel HDF
- **Misc. Programming**
 - Batch “time remaining” routines
 - Threading and math library tests
- **Networked Data Transfer**
 - from scp to hsi (0.5 → 70 MB/s)

HPC Support
Yields Science
Productivity

NERSC: Ready for New Challenges



- **Emerging Software and Algorithms**
 - Early testing of new HPC software
 - Performance analysis of new HPC applications
- **Emerging HPC Systems Technology**
 - Parallel File Systems, Parallel I/O profiling
- **Emerging Architectures**
 - “Performance Characteristics of a Cosmology Package on Leading HPC Architectures”, HiPC 2004
 - “Scientific Computations on Modern Parallel Vector Systems”, Supercomputing 2004
 - “A Performance Evaluation of the Cray X1 for Scientific Applications”, VECPAR 2004
- **Partnerships to Improve HPC**
 - BluePlanet / ViVA / Workload Analysis / PERC

INCITE 2005!