UT-BATTELLE

AMES LABORATORY

ARGONNE NATIONAL LABORATORY

Los Alamos

Lawrence Livermore National Laboratory

NASA

NCAR

Pacific Northwest National Laboratory

PPPL

Sandia National Laboratories

Leadership Computing Facility at Oak Ridge National Laboratory

Studham RS, Kuehn J, White JB, Fahey M, Carter S, Bland A, Nichols JA, Zacharia T

CUG
May 17th 2005

Outline

- What is "Leadership" computing
- What is the "Leadership Computing Facility at ORNL"
- Overview of the LCF computing architectures and long term plans
- Benchmarks of LCF computing systems
- Example science results from LCF systems.
- LCF proposal process and challenges.

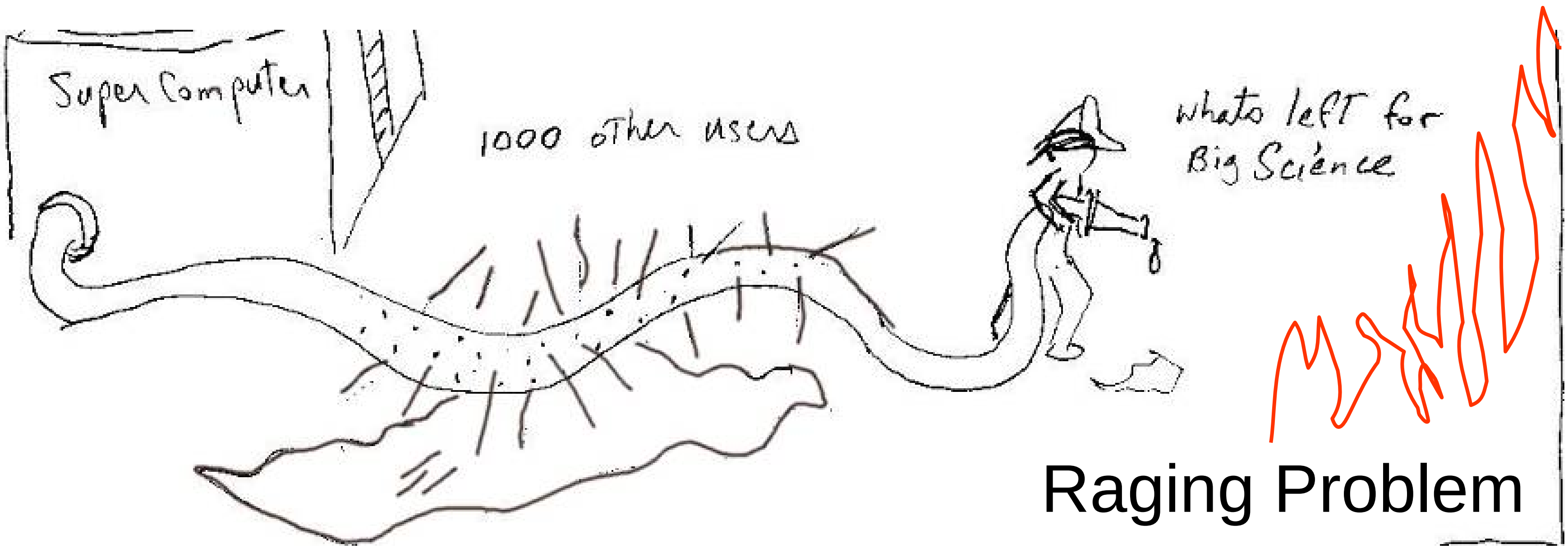
CCS first Cray
1985 -1992
1 Processor
8 MB memory
180 MFLOPS

The user paradigm must change to do Big Science (AKA "Supercomputing")

DOE-SC's primary computing resource is a 10TF (peak) system that supports ~363 projects

The average a project gets:

$$\frac{\text{10TFyears}}{\text{363 Projects}} = \text{27 GFyears/Project}$$



Artwork courtesy of Al Geist

What is "Leadership Computing"?

Leadership computing systems will provide computational capability that is at least **100 times greater** than what is generally available for advanced scientific and engineering problems.

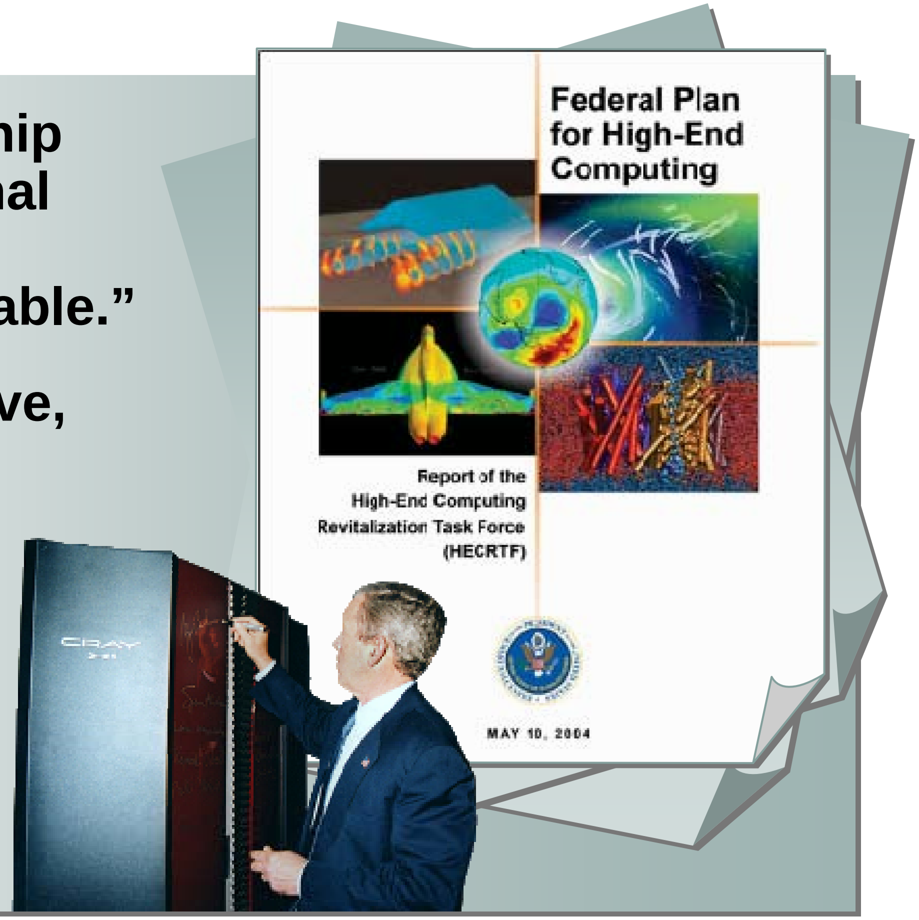
	Generally Available TFy/Project	Leadership Computing TFy/Project	TF at NLCF	Number of Leadership Computing Projects
2005	0.027	2.7	6.8	2
2006	0.055	5.5	44	8
2007	0.110	11.0	118	11
2008	0.220	22.0	368	17
2009	0.441	44.1	1000	23

Leadership computing is a White House priority

"The goal of such systems [leadership systems] is to provide computational capability that is at least 100 times greater than what is currently available."

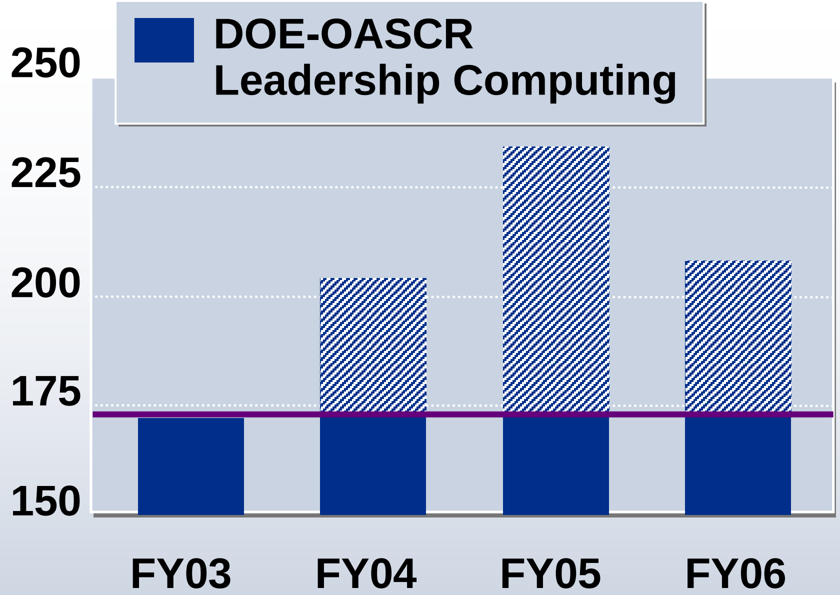
"…Leadership Systems are expensive, typically costing in excess of $100 million per year...."

– Page 29, Federal Plan for High-end Computing



Leadership computing is a congressional priority

- **Appropriated $30M in FY04 for leadership computing**
- **Additional $30M appropriated in FY05**
- **Public Law 108-423 Department of Energy High-End Computing Revitalization Act of 2004**



Leadership computing is a State of Tennessee priority

- **$9M State of Tennessee Investment in the Joint Institute for Computational Sciences**
- **$10M for National Academy Level Joint Faculty**
- **$12M for high speed networks for research and education**
- **$1M/year for Computational Science Initiative for graduate student training and outreach**

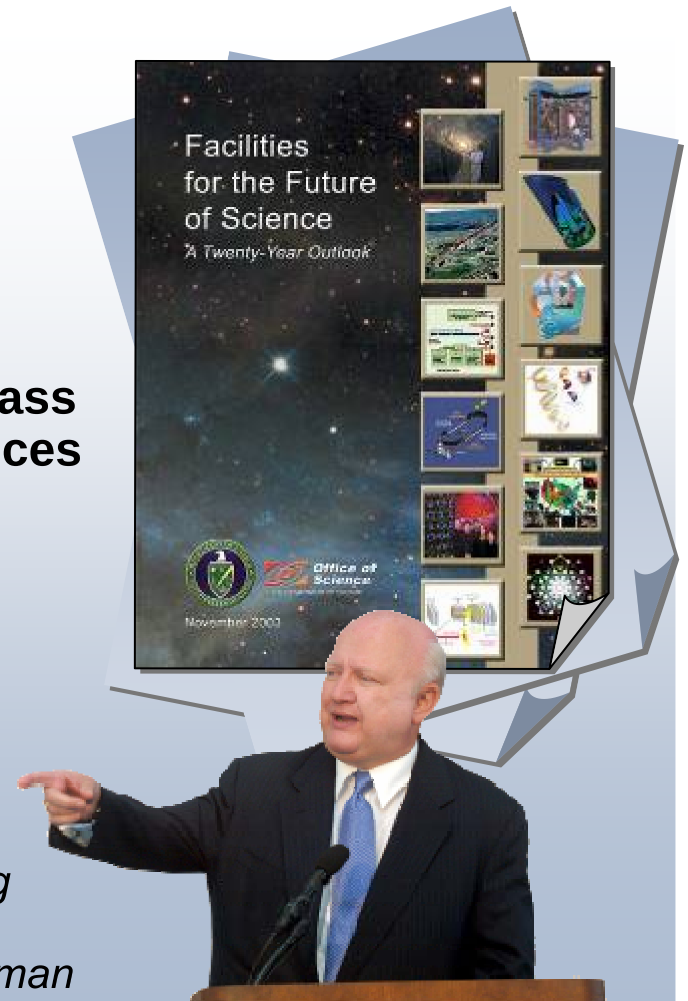
"I have recommended funds …to attract more nationally-recognized faculty members (jointly with ORNL)…. There is an opportunity today…. to rapidly become world class in some areas like supercomputers, materials science, and nanotechnology.

….Our pioneer ancestors wouldn't have known what supercomputers were, but I believe they would have understood our aspirations perfectly."

- Gov. Bredesen, State of the State Speech, January 31, 2005

Leadership computing is the highest domestic priority for Office of Science

- **Ray Orbach has articulated his philosophy for the SC laboratories**
  - **Each lab will have world-class capabilities in one or more areas of importance to Office of Science**
  - **ORNL: SNS and NLCF will underpin world-class programs in materials, energy, and life sciences**
- **20-year facilities plan being used to set priorities among projects**



"I am committed to the concept of a Leadership Class Computing facility at Oak Ridge National Laboratory. The facility will be used to meet the missions of the Department and those of other agencies. I can assure you that I understand the important role supercomputing plays in scientific discovery."

Secretary Bodman

Office of Science notice to SC Laboratories

Leadership-Class Computing Capability for Science

"The focus of the proposed effort should be on capability computing in support of high-end science – rather than on enhanced computing capacity for general science users . . . "

"The proposed effort must be a user facility providing leadership class computing capability to scientists and engineers nationwide independent of their institutional affiliation or source of funding."

"Proposals must include specific information on architecture or architectures to be provided over the life of this project as well as a list of targeted applications."

Our plan of action to deliver leadership computing for DOE

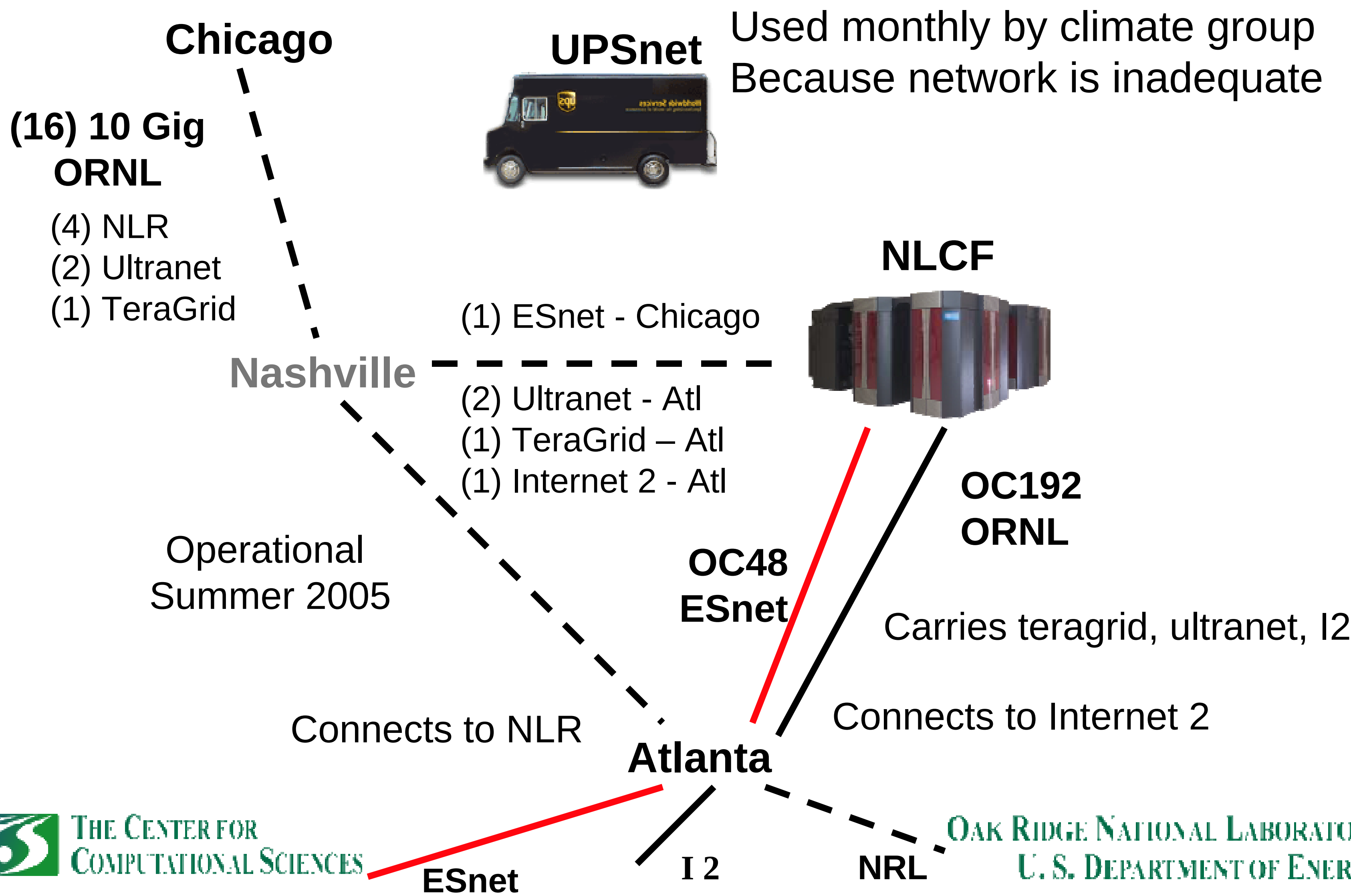
- **Rapidly field most powerful open capability computing resource for scientific community**
  - **Providing clear upgrade path to at least 100 teraflop/s (TF) by 2006 and 250 TF by 2007/2008**
- **Deliver outstanding access and service to research community**
  - **Utilizing most powerful networking capability extant coupled with secure and highly cost-effective operation by proven team**
- **Deliver much higher sustained performance for major scientific applications than currently achievable**
  - **Developing next generation models and tools**
  - **Engaging computer vendors on hardware needs for scientific applications**
- **Engage research communities in climate, fusion, biology, materials, chemistry, and other areas critical to DOE-SC and other federal agencies**
  - **Enabling high likelihood of breakthroughs on key problems**
- **Conduct in-depth exploration of most promising technologies for next-generation leadership-class computers**
  - **Providing pathways to petaflop/s (PF) computing within decade**


National Leadership Computing Facility
May 12, 2004
CRAY X1
Office of Science
FCW.COM
Oak Ridge gets super win
"Feds spark high-end computing resurgence"
CNN
U.S. Moves to Build Top Supercomputer
ITmedia News
ComputerWeekly
U.S. Energy Dept. Awards $25M Grant for Supercomputer Build
COMPUTERWORLD
Energy Department Grant Spurs Work on Supercompute
BusinessWeek
Where No Computer Has Gone B
U.S. to Build Civilian Supercomputer
The Atlanta Journal-Constitution
Computer to be more than a bit faster
Supercomputer-generated "power wall" at Oak Ridge, Tenn., National Laboratory illuminates climate change
CBS 2
cbsnewyork.com
Uncle Sam Wants World's Fastest Computer
Want To Surpass Japanese Supercomputer
TechWeb
Tennessee is a Rising Star in Supercomputing
The Oak Ridger
ORNL Nabs Supercomputer Deal
LAB CHIEF: 'Oak Ridge will help lay the foundation for our country's next generation of scientific discovery.'
WASHINGTON TECHNOLOGY
DOE Orders World's Fastest Supercomputer
San Francisco Chronicle
SFGate
U.S. Government Moves to Build World's Fastest Civilian Computer
矽谷科技
全球最快速民用科研電腦
GLOBEANDMAIL.COM
U.S. Aims to Regain Computer Speed Lead
DOE Leadership-Class Computing Capability for Science will be Developed at Oak Ridge National Laboratory
CENTER FOR COMPUTATIONAL SCIENCES
OAK RIDGE NATIONAL LABORATORY
Office of Science
U.S. DEPARTMENT OF ENERGY

Leadership-class Computing Requires Leadership-class Facilities

- $72M private sector investment in support of leadership computing
- Space and power:
  - **40,000 ft^2 computer center with 36-in. raised floor, 18 ft. deck-to-deck**
  - **8 MW of power (expandable) @ 5c/kWhr**
- High-ceiling area for visualization lab (Cave, Powerwall, Access Grid, etc.)
- Separate lab areas for computer science and network research

Leadership-class Computing Requires Leadership-class Connectivity



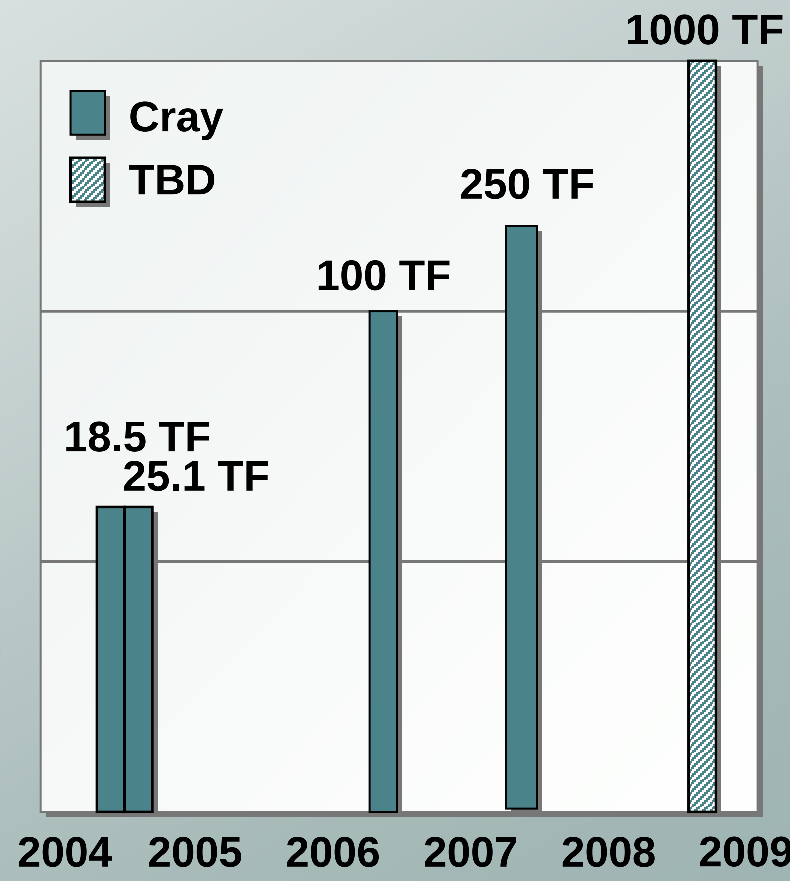
NLCF hardware roadmap

In 2005,

- **Deploy 18.5TF Cray X1E and 25.1TF Cray XT3 systems**
- **Cray forms and supports "Supercomputing *Center of Excellence*"**
- **Develop/deploy complementary software environment**
- **Full operational support of NLCF as a capability computing center**
- **Deliver computationally intensive projects of large scale and high scientific impact through competitive peer review process**

In 2006, deploy 100TF Cray XT3

In 2007-8, deploy 250+TF system



Phoenix – The CCS Cray X1/X1e

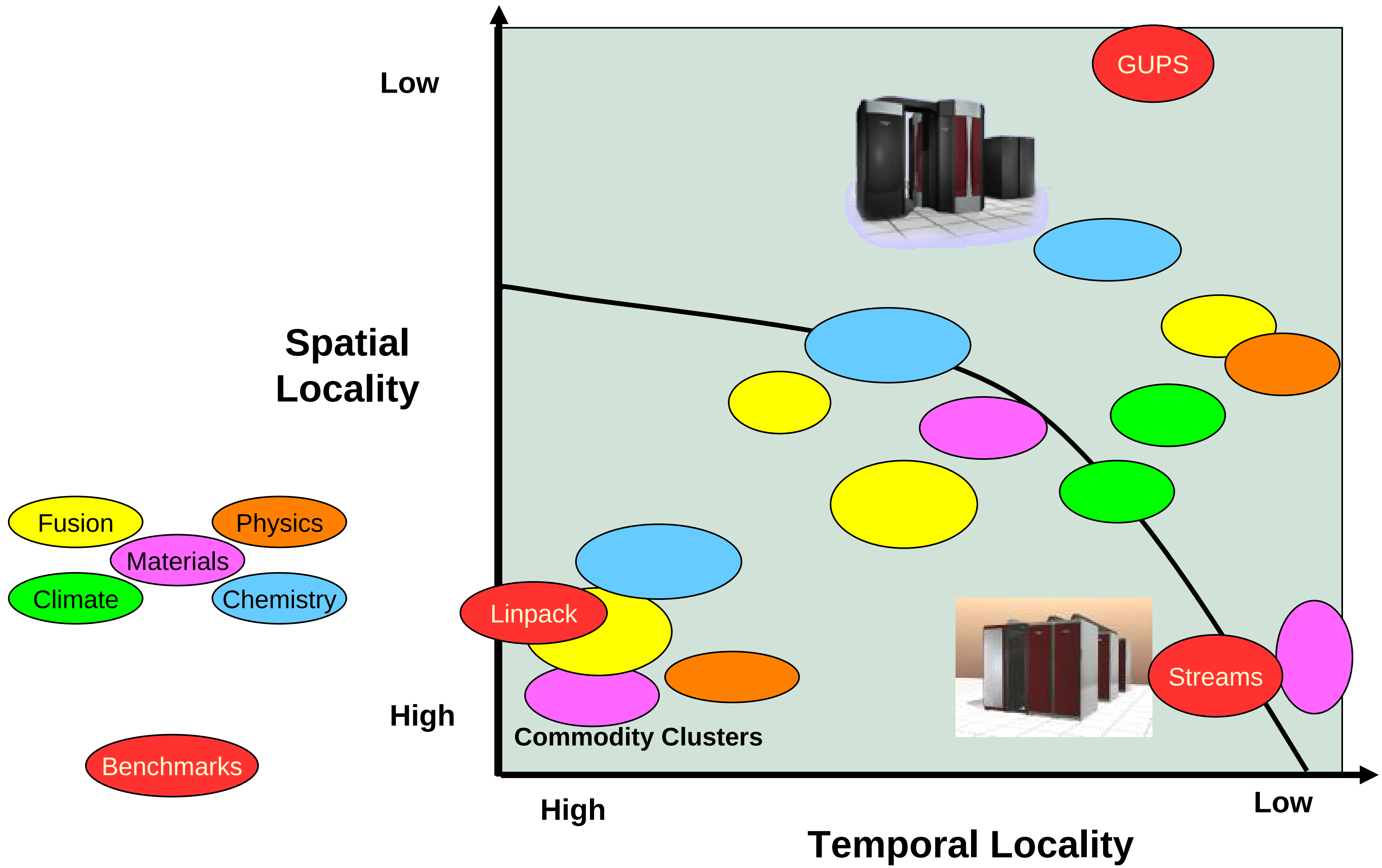
- Largest Cray X1 in the world. Upgrading to 18TF X1E this summer.
- Highest bandwidth communication with main memory

- Highly scalable hardware and software
- High sustained performance on real applications

Jaguar – The CCS XT3 system

Status	Performance	Processors	Memory	I/O Bandwidth
Summer 05	25 TF	5,304	10.5 TB	15 GB/s
Fall 05*	50 TF	11,374	23 TB	30 GB/s
Spring 06*	100 TF	22,748	46 TB	60 GB/s

Different applications have different memory access patterns – “Once size does not fit all”

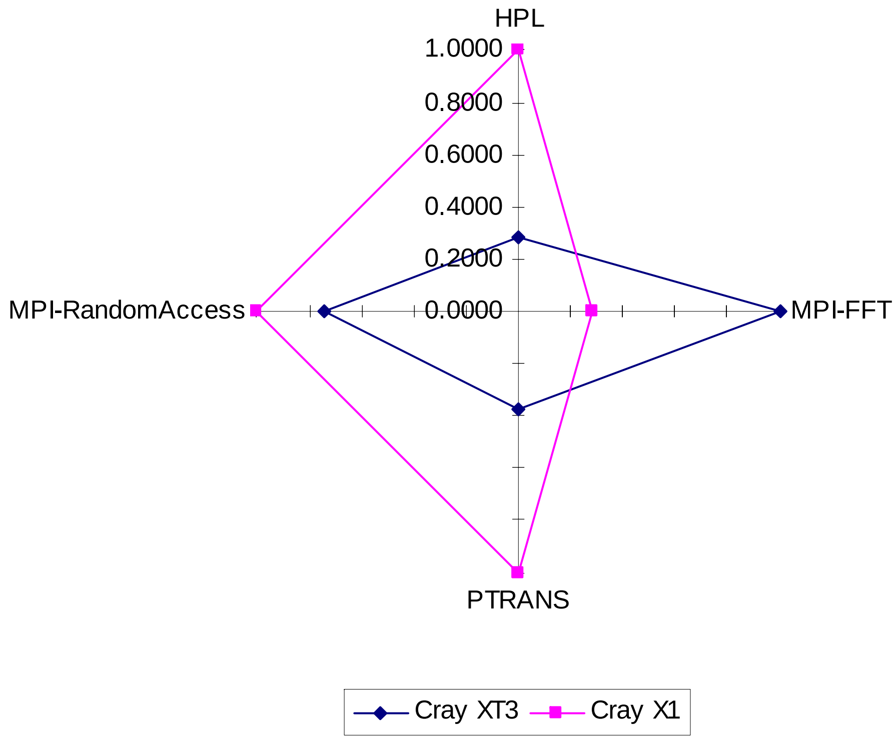


This is a qualitative example. Studies are being done on applications to make this more quantitative.

X1/XT3 Relative Global Benchmark Comparison

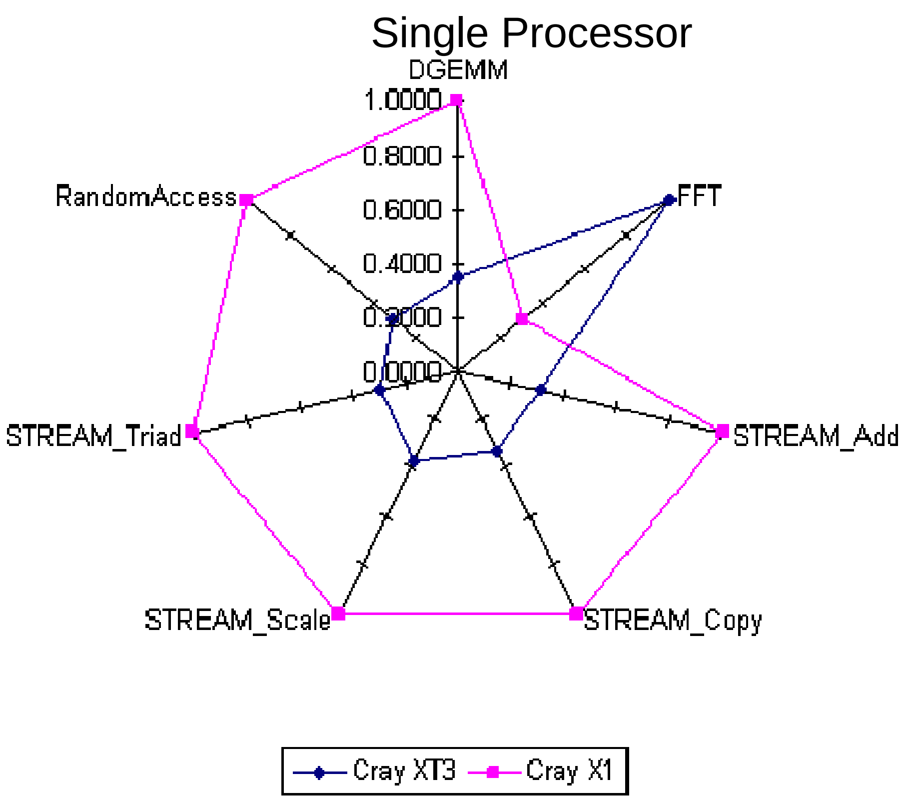
Relative performance on 64 CPUs

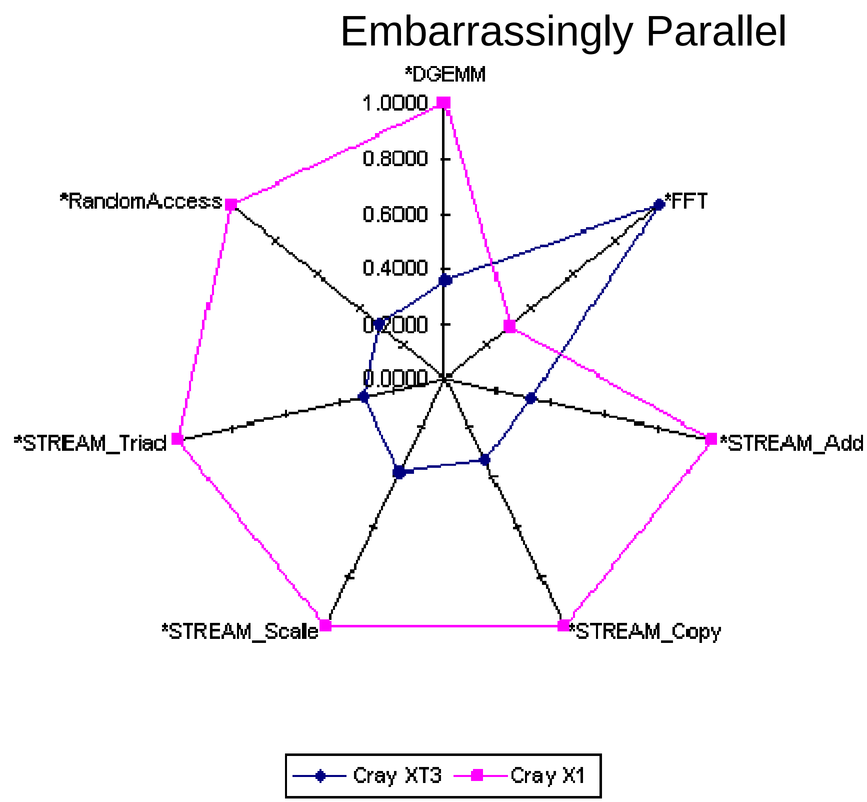
- Opteron scalar performance dominates MPI-FFT for XT3
- X1 bandwidth dominates MPI-RandomAccess and PTRANS
- X1 processor performance dominates HPL



XT3 and X1 Relative Processor Performance

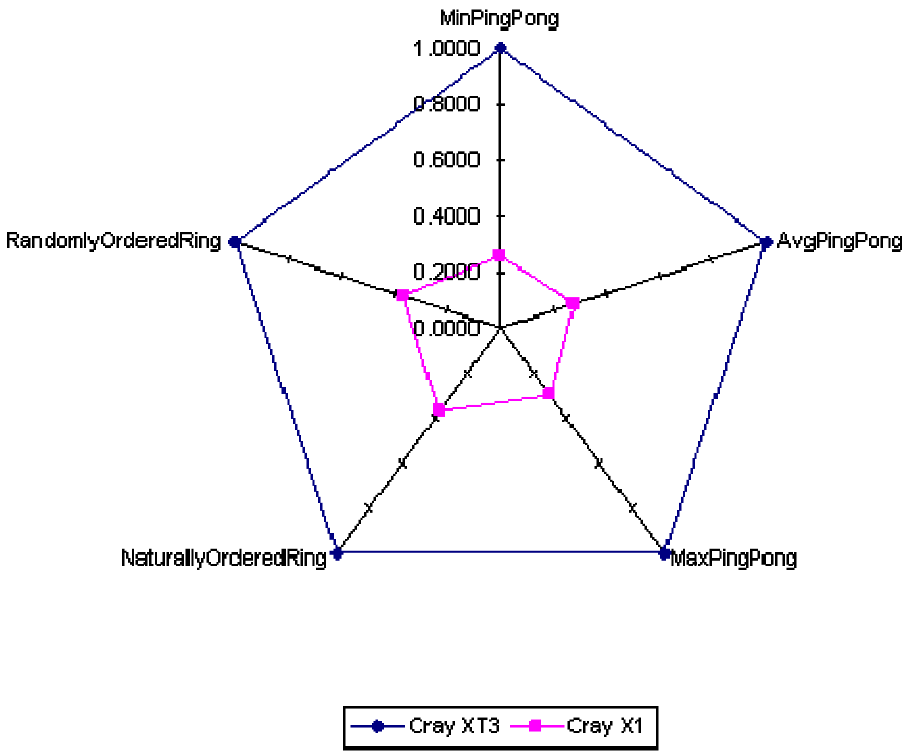
1D FFT can't be vectorized so the Opteron's 2.4GHz scalar performance is highlighted over the 400MHz X1 scalar performance

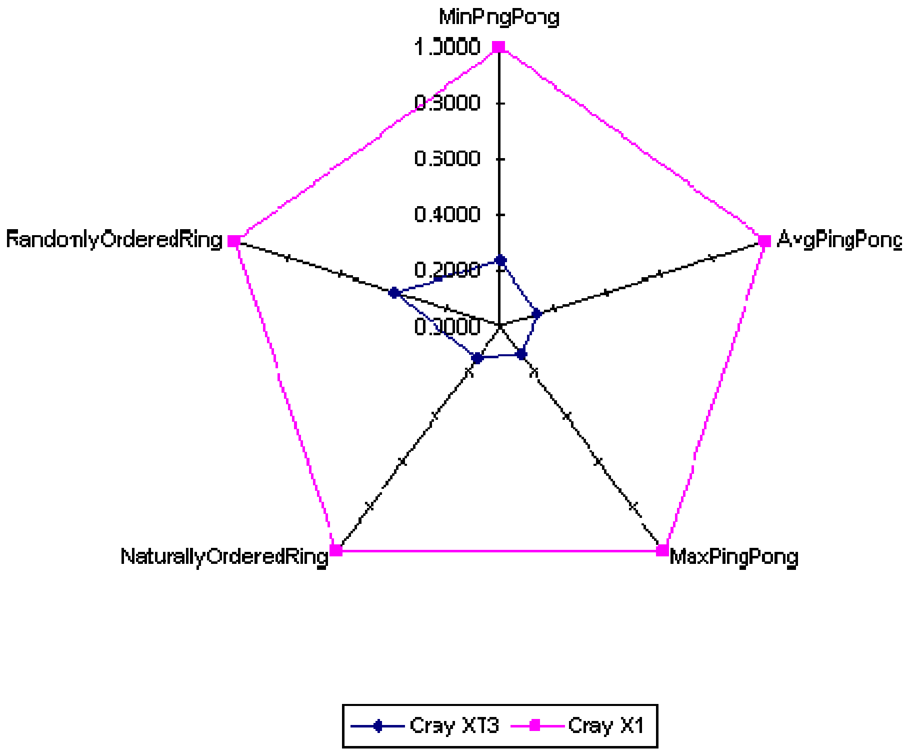




Results by Jeff Kuehn

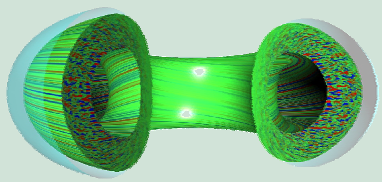
XT3/X1 Relative interconnect performance





Results by Jeff Kuehn

GYRO: A 5-D gyrokinetic-Maxwell solver



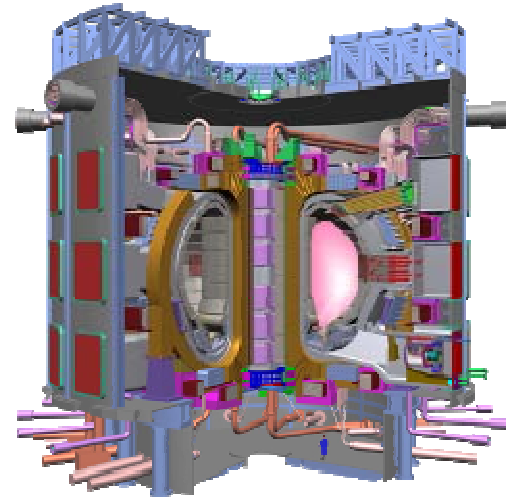
$$\frac{\partial f}{\partial t} = L_a f + L_b \Phi + \{f, \Phi\} \text{ where } F\Phi = \iint dv_1 dv_2 f$$

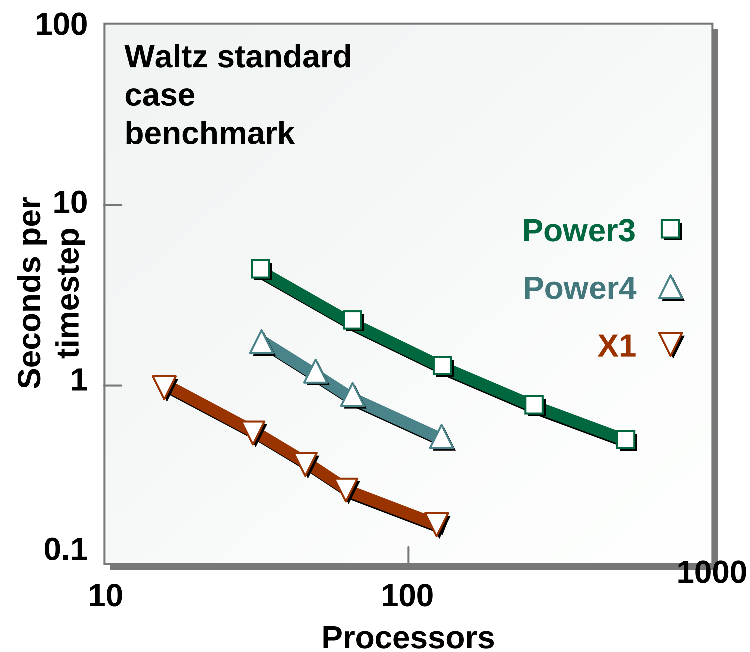
GYRO employs Implicit-Explicit Runge-Kutta scheme discretized on Eulerian grid

For typical simulations, GYRO performance is at least 4x faster on Cray X1 than SGI Altix or IBM Power4;

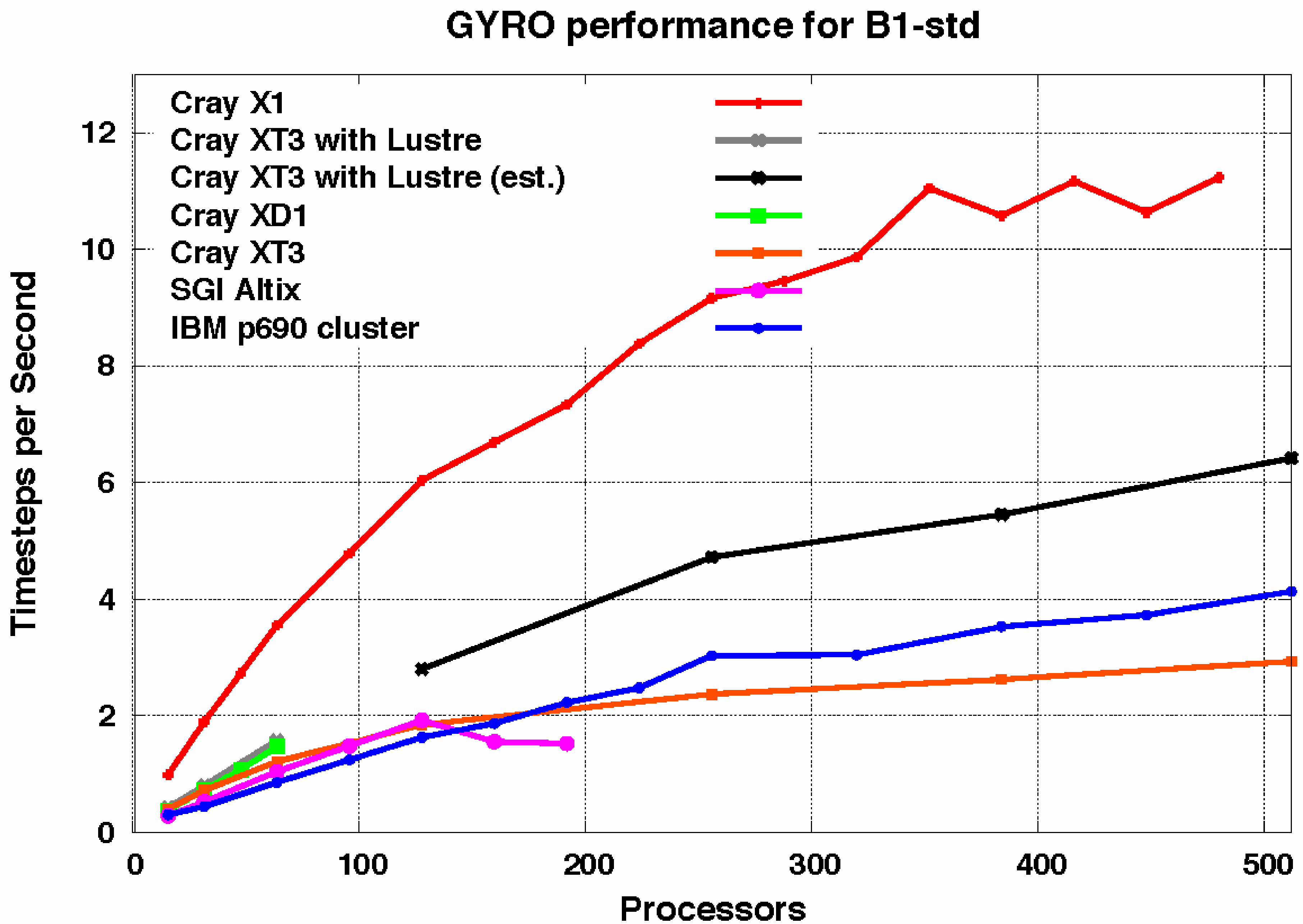
X1 is significantly faster for larger problems.

Recent accomplishments include: completed the most comprehensive tokamak turbulence simulation, showed transport is smooth across an s=0 (minimum-q surface) due to the appearance of gap modes, found that χ_i matched the “Cyclone value” at small ρ_* resolving an earlier conflicting report while simultaneously finding there is a long transient period for which χ_i exceeds the statistical average in GK simulations.





Impact of using NFS for IO on XT3

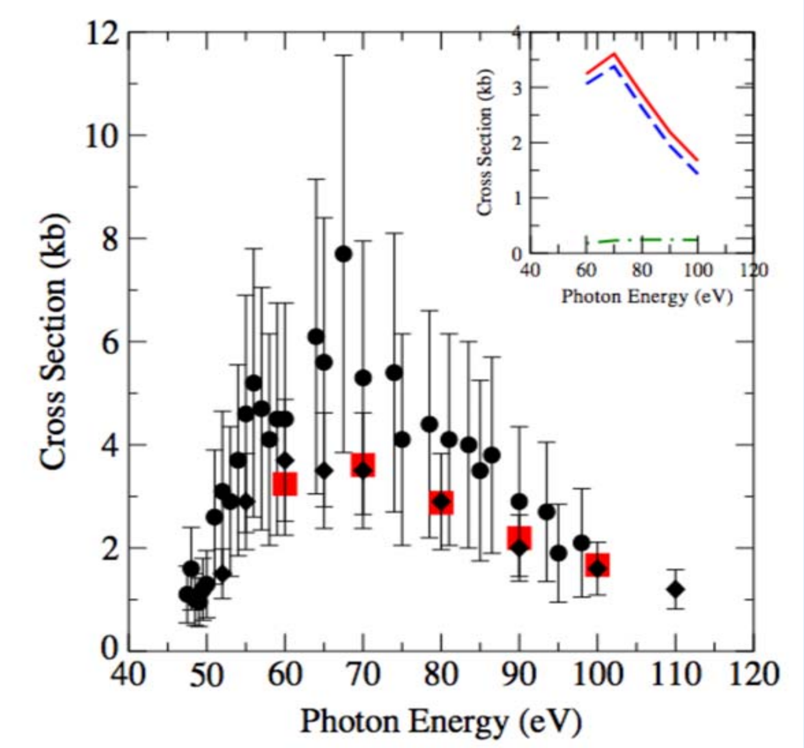


IS
r

Double photoionization of H_2

James Colgan (LANL) and Mitch Pindzola (Auburn University)

- First ever fully *ab initio* calculation
- Important step in exploring the physics inherent in molecular break-up by a single photon
- 4D wave function propagated in time using Schroedinger equation
- Required hundreds of Cray X1 MSPs

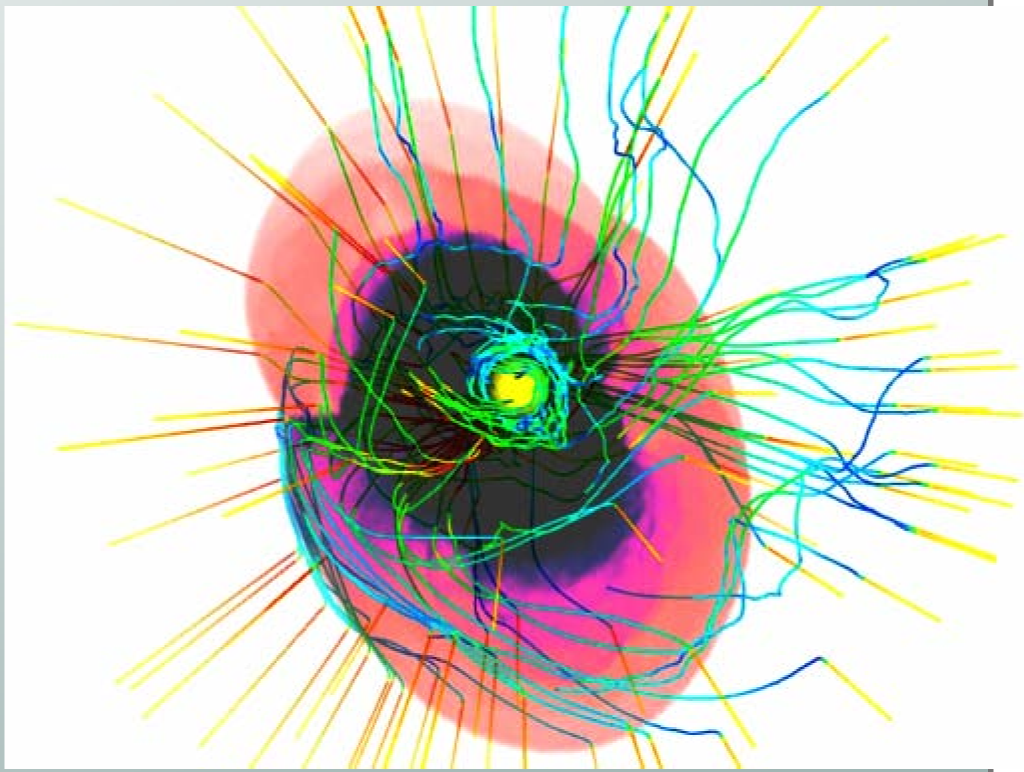


"Time-dependent close-coupling calculations for the double photoionization of He and H2." J Colgan, M S Pindzola, and F Robicheaux. *J. Phys*. B: At. Mol. Opt. Phys. 37 (2004) L377–L384.

"Neutron star spin-up discovered with 3D simulations on Cray X1"

http://astro.physics.ncsu.edu/TSI/

- Terascale Supernova Initiative
- 3D hydrodynamics using 600 million zones
- Shading shows extent of supernova shock
- Streamlines show collapsing stellar material
  - Flow decelerates at shock
- Angular momentum of flow creates spinning neutron star



Some NLCF Science Highlights

- Fusion: Gyrokinetic simulation
  - **“GYRO Performance on a Variety of MPP Systems” (yesterday)**
- Climate
  - **“Porting and Performance of the CCSM3 on the Cray X1” (8:30 AM Thursday)**
- High-Temperature Superconductors
  - **Significant results submitted for publication**
- Chemistry
  - **Full-Configuration Interaction at 5.5 TF on 432 MSPs (60% of peak)**
  - **The new code already enables computations 40+x larger and 100+x faster than previous work**
- Combustion
  - **38% performance improvement using targeted Co-Array Fortran**

NLCF Applications at CUG 2005

- Early Evaluation of the Cray XD1 (2PM)
- Early Evaluation of the Cray XT3 at ORNL (3PM)
- Comparative Analysis of Interprocess Communication on the X1, XD1, and XT3 (2PM Wednesday)
- High-Speed Networking with Cray Supercomputers at ORNL (2:30PM Wednesday)
- Comparisons of the X1 to X1E on Major Scientific Applications (3PM Wednesday)
- Optimization of the PETSc Toolkit and Application Codes on the Cray X1 (11:30AM Thursday)

National Leadership Computing Facility
2006 Call for Proposals

- **Largest, most computationally intensive scientific endeavors to be conducted on high capability Leadership systems of DOE's National Leadership Computing Facility (NLCF)**
- **Seeks high priority, challenging, high-payoff, and heretofore intractable, computationally-intensive experiments, where capability of NLCF systems can enable new breakthroughs in science**
- **Expectation that Leadership systems will enable U.S. to be "first to market" with important scientific and technological capabilities, ideas, and software**
- **Selection of limited set of scientific applications (perhaps 10 per year) and given substantial access to leadership system**

NCCS Proposal Review Process

DOE SC issues a call for proposals

↓

Users submit proposals to NCCS

↓

Onsite Technical Review of Proposals

Panel members selected by by NCCS Advisory Committee and appointed by NCCS program manager

- O(3) reviewers from each of SC domains
- Domain experts review and present to full panel (NSF ITR style)
- NCCS scientific consulting team will be available

↓

Panel ranks proposals on

- Scientific Merit
- Technical Merit
- H/W Readiness

Panel recommends allocations

↓

Panel sends rankings and recommendations to HQ for further review

- Ranked by Program Offices for programmatic priority

↓

Final rankings and recommendations collected by HQ

↓

DOE issues the awards

Summary

- The “Leadership” computing facility at ORNL is deploying systems to provide 100X more computational resources to critical problems.
- The Cray X1 and Cray XT3 complement each other by servicing different types of calculations.
- The NLCF at ORNL has the necessary support and is ready to deliver leadership computing to the nation.
- The largest challenge is proving to be getting scientists to think about problems that are 100x larger than what they are currently solving.

Thank You

To join our team please go to:
http://computing.ornl.gov/Employment/postings.html