Revitalization of High-End Computing

David B. Nelson, Ph.D.

Director

National Coordination Office for

Information Technology Research and Development

George Mason University

April 30, 2004

Topics

- **High-End Computing Revitalization Task Force**
  - Application Needs
  - Task Force Response
- **Architectural Problems in High-End Computing**
- **President's Information Technology Advisory Committee**
  - Study on Computational Science

High-End Computing Revitalization Task Force (HECRTF)

- **Inter-agency planning group**
  - **Develop 5-year plan/roadmap to improve how the Federal government develops, purchases, and provisions HEC**
  - **Participants include DoD (DARPA, ODUSD (S&T), HPC Modernization Program, NSA), DOE (NNSA and Science), EPA, NASA, NIH, NIST, NOAA, NSF, OMB, OSTP, NCO (approx. 60 people)**
  - **Focus on advancing agency/end-user needs in HEC**
- **Established by OSTP, under the auspices of the National Science and Technology Council, in March 2003**
  - **Charge to Task Force available at www.nitrd.gov**
- **Commissioned CRA Workshop held June, 2003**
  - http://www.cra.org/Activities/workshops/nitrd/
- **Plan now in final concurrence before public release**

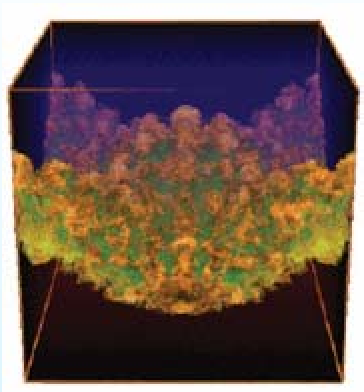
HECRTF Goals

- **Overarching: Revitalize U.S. leadership in high-end computing as a key tool for science and technology.**
  - **Make high-end computing easier and more productive to use.**
  - **Make high-end computing readily available to Federally funded missions that need it.**
  - **Sustain the development of new generations of high-end computing systems.**
  - **Effectively manage and coordinate Federal high-end computing.**

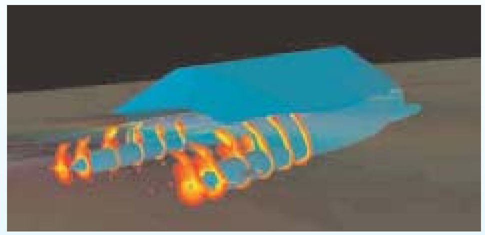
High-End Computing Revitalization Task Force – Guidance in FY04 Budget

<u>OMB Analytical Perspectives, Budget of the United States Government, FY2004, Page 177:</u> "… Due to its impact on a wide range of federal agency missions ranging from national security and defense to basic science, high end computing—or **supercomputing** —capability is becoming **increasingly critical**. Through the course of 2003, agencies involved in developing or using high end computing will be engaged in planning activities to guide future investments in this area, coordinated through the NSTC. The activities will include the development of an **<u>interagency R&D roadmap for high-end computing core technologies</u>**, a **<u>federal high-end computing capacity and accessibility improvement plan</u>**, and a discussion of issues (along with recommendations where applicable) relating to federal **<u>procurement</u>** of high-end computing systems. The knowledge gained from this process will be used to guide future investments in this area. Research and software to support high end computing will provide a foundation for future federal R&D by improving the effectiveness of core technologies on which next-generation high-end computing systems will rely."

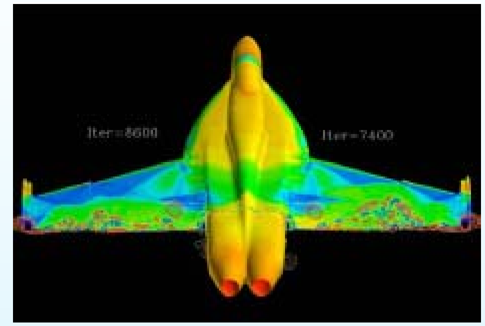
Applications of High-End Computing:
Big Problems with Big Impacts



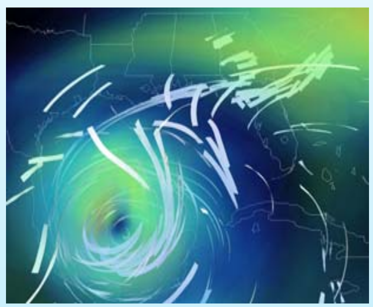
Nuclear Stockpile Stewardship



Ship Design

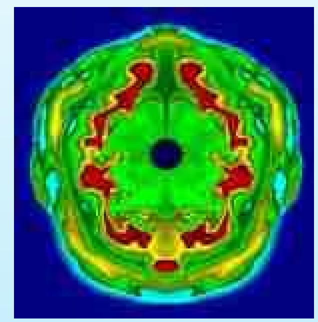


Aeronautics

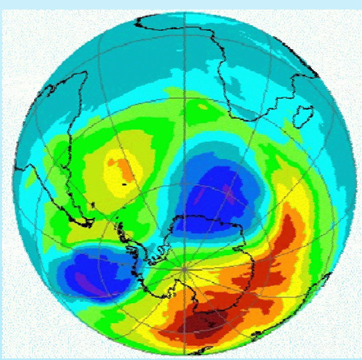


Weather Prediction

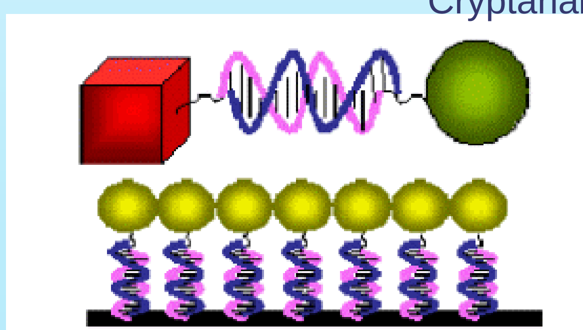
Cryptanalysis



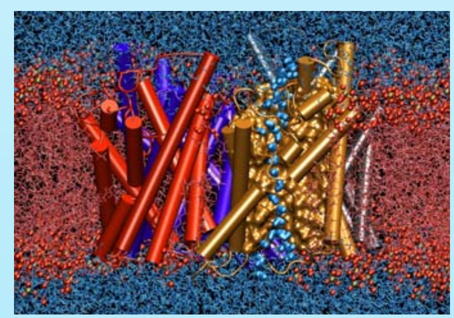
Astrophysical Simulation



Climate Modeling



Nano-Science



Biology

Opportunities for Advancing Science

Application	Science Accomplishment	Required Capability Multiple	Benefit to Nation
Magnetic Fusion Energy	Optimize balance between self-heating of plasma and heat leakage caused by electromagnetic turbulence.	**100**	Underpins U.S. decisions about future international fusion collaborations. Integrated simulations of burning plasma crucial for quantifying prospects for commercial fusion.
Combustion Science	Understand interactions between combustion and turbulent fluctuations in burning fluid.	**100**	Understand detonation dynamics (for example, engine knock) in combustion systems. Solve the "soot" problem in diesel engines.
Astrophysics	Realistically simulate the explosion of a supernova for the first time.	**1000**	Measure size and age of Universe and rate of expansion of Universe. Gain insight into inertial fusion processes.
Structural and Systems Biology	Simulations of enzyme catalysis, protein folding, and transport of ions through cell membranes.	**1000**	Ability to discover, design, and test pharmaceuticals for specific targets and to design and produce hydrogen and other energy feedstocks more efficiently.
Catalyst Science/ Nanoscale Science and Technology	Calculations of homogeneous and heterogeneous catalyst models in solution.	**1000**	Substantial reductions in energy costs and emissions associated with chemicals manufacturing and processing. Meeting federally mandated NOx levels in automotive emissions.
Nanoscale Science and Technology	Simulate the operation of nanoscale electronic devices of modest complexity.	**1000**	Takes miniaturization of electronic devices to a qualitatively new level enabling faster computers, drug delivery systems, and consumer and military electronics.
Nanoscale Science and Technology	Simulate and predict mechanical and magnetic properties of simple nanostructured materials.	**1000**	Enables the discovery and design of new advanced materials for a wide variety of applications potentially impacting a wide range of industries, including the high-tech industry that generated more than $900 billion in sales and accounted for 4 million jobs in 1999 and the $34 billion disk drive industry.

Energy and Physics

Biology and Nanoscience

Opportunities for Advancing Science

Application	Science Accomplishment	Required Capability Multiple	Benefit to Nation
Signals Intelligence	Model, simulate, and exploit foreign codes, ciphers and complex communications systems.	**1000**	Supports U.S. policy makers, military commands and combat forces with information critical to national security, force protection and combat operations.
Directed Energy	To advance the directed energy systems design process out of the scientific research realm into the engineering design realm	**1000**	Ability to efficiently design next generation directed energy offensive and defensive weapon systems. Change the design process from years to days.
Signals Image Processing & Automatic Target Recognition	To replace electromagnetic scattering field tests of actual targets with numerical simulations of virtual targets	**1000**	Creates the ability to design more stealthy aircraft, ships, and ground systems and creates the ability to rapidly model new targets enabling more rapid adaptation of fielded weapon systems' ability to target new enemy weapon systems.
Integrate Modeling and Test of Weapon Systems	To model complex system interaction in real time with precision	**1000**	Creates the ability to replace many expensive, dangerous and time consuming ground tests with virtual tests resulting in lower test costs and more rapid development of weapon systems.
Climate Science	Resolve additional physical processes such as ocean eddies, land use patterns, and clouds in climate and weather prediction models.	**1000**	Provide U.S. policymakers with leading-edge scientific data to support policy decisions. Improve climate and weather prediction skill at timescales from minutes to decades.
Weather and Short-term Climate Prediction	Enable dynamical prediction of frequency and intensity of occurrence of hurricanes/typhoons and severe winter storms 90 days in advance.	**1000**	Provides critical support to deployed naval, air and land forces in local, regional and global combat environments. Lives saved and economic losses avoided due to better severe weather prediction.
Solid Earth Science	Dynamic earthquake forecasting with 5 year lead time.	**100**	Provide prioritized retrofit strategies. Reduced loss of life and property. Damage mitigation.
Space Science	Realistically simulate explosive events on the sun, the propagation of the energy and particles released in the event through the interplanetary medium, and their coupling to Earth's magnetosphere, ionosphere, and thermosphere.	**1000**	Provide decision makers (both civilian and military) with status and accurate predictions of space weather events on time scales of hours to days.
Subsurface Contamination Science	Simulate the fate and transport of radionuclides and organic contaminants in the subsurface.	**1000**	Predict contaminant movement in soils and ground water and provide a basis for developing innovative technologies to remediate contaminated soils and ground water.

National Security

Environment

User and Agency Views on High-End Computing

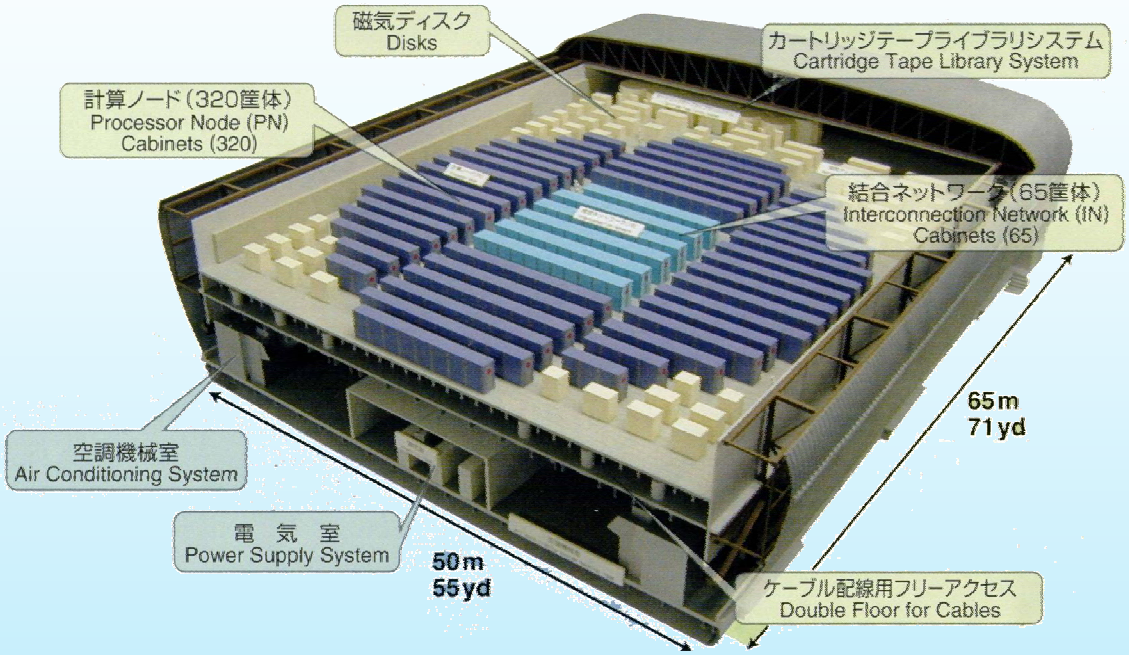
- **Research pipeline dry**
  - Funded research ideas
  - Expertise
- **Industrial base issues**
  - Market size and forces
  - Company interests and attention
  - Technology development and use
- **Technology improvement**
  - Radical improvements in time-to-solution;
  - Significant improvements to system bandwidth, reliability, ease of programming
  - Diversity of architectures
- **User demands exceeds available resources for both capacity and capability**

Opportunities

- Move from event-driven investments to one based upon a strategic planning process
- Leverage various Agency strengths, capabilities, and resources via improved cooperation
- Rebuild and sustain critical-mass funding levels to attack national challenges
- Rebuild teaming within government, academia, and industry to address pervasive issues

The Japanese Earth Simulator

One Scientist's View



- **" … for the first time in my scientific career I find myself having to go outside the U.S. in order to do cutting-edge science."**
  - **Jeroen Tromp, Director & McMillan Prof. of Geophysics, Caltech, May 2003**

The Japanese Earth Simulator
A Center Director's View (Dr. Dan Reed, UNC)

- **Long term planning matters**
  - Strategic, rather than tactical objectives
  - Multiple generations of implementation
    - Lessons learned influence next generations
- **Goals tied to policy and citizenry needs**
  - Societal benefits and rewards
    - Climate, health, safety/security, jobs, …
- **Resources commensurate with goals**
  - Strategic resources needed for strategic goals

The Japanese Earth Simulator
Task Force View

- **Japanese Earth Simulator entered service March 2002.**
  - **Took the #1 spot on the Top 500 Supercomputer list.**
- **What it wasn't:**
  - **A revelation.**
  - **A demonstration of US loss of scientific leadership.**
- **What it was:**
  - **An attention-getting event that elevated the prominence of what was already considered to be an important issue.**
  - **A demonstration of the value of sustained planning and support**
  - **A challenge that provides added incentive for action.**
- **Federal high-end computing planning was historically not as well coordinated as it should have been.**

RD&E Strategy

Activity	Purpose	Performers
Basic and Applied Research	*Refill the academic pipeline with new ideas and people*	*Academia and government labs*
Advanced Development	*Develop component and subsystem technologies*	*Mostly industry led, partnering with academia and government labs*
Engineering and Prototype Development	*Integration at system level and development of Serial No. 1*	*Industry*
Test and Evaluation	*Reduce risk for development, engineering, and government procurement*	*Government labs and HEC centers*

RD&E - Key Technologies

Hardware

- ⇨ Microarchitecture
- ⇨ Memory
- ⇨ Interconnect
- ⇨ Power, cooling, and packaging
- ⇨ I/O and storage

Software

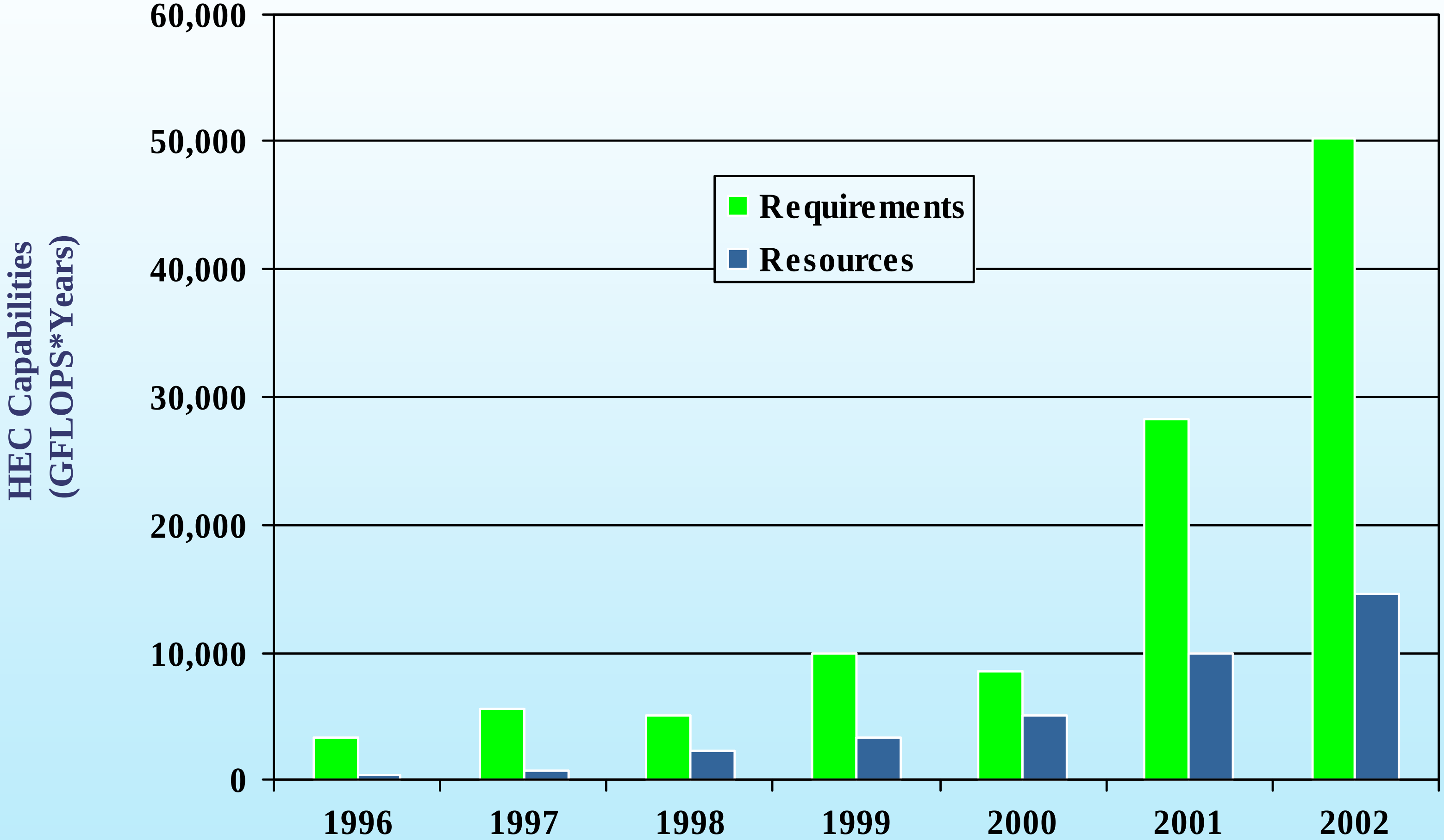
- ⇨ Operating systems
- ⇨ Languages, compilers, and libraries
- ⇨ Software tools and development environments
- ⇨ Algorithms

Systems

- ⇨ System architecture
- ⇨ Reliability, availability, and serviceability (RAS)
- ⇨ System modeling and performance analysis
- ⇨ Programming models
- ⇨ System modeling and performance analysis

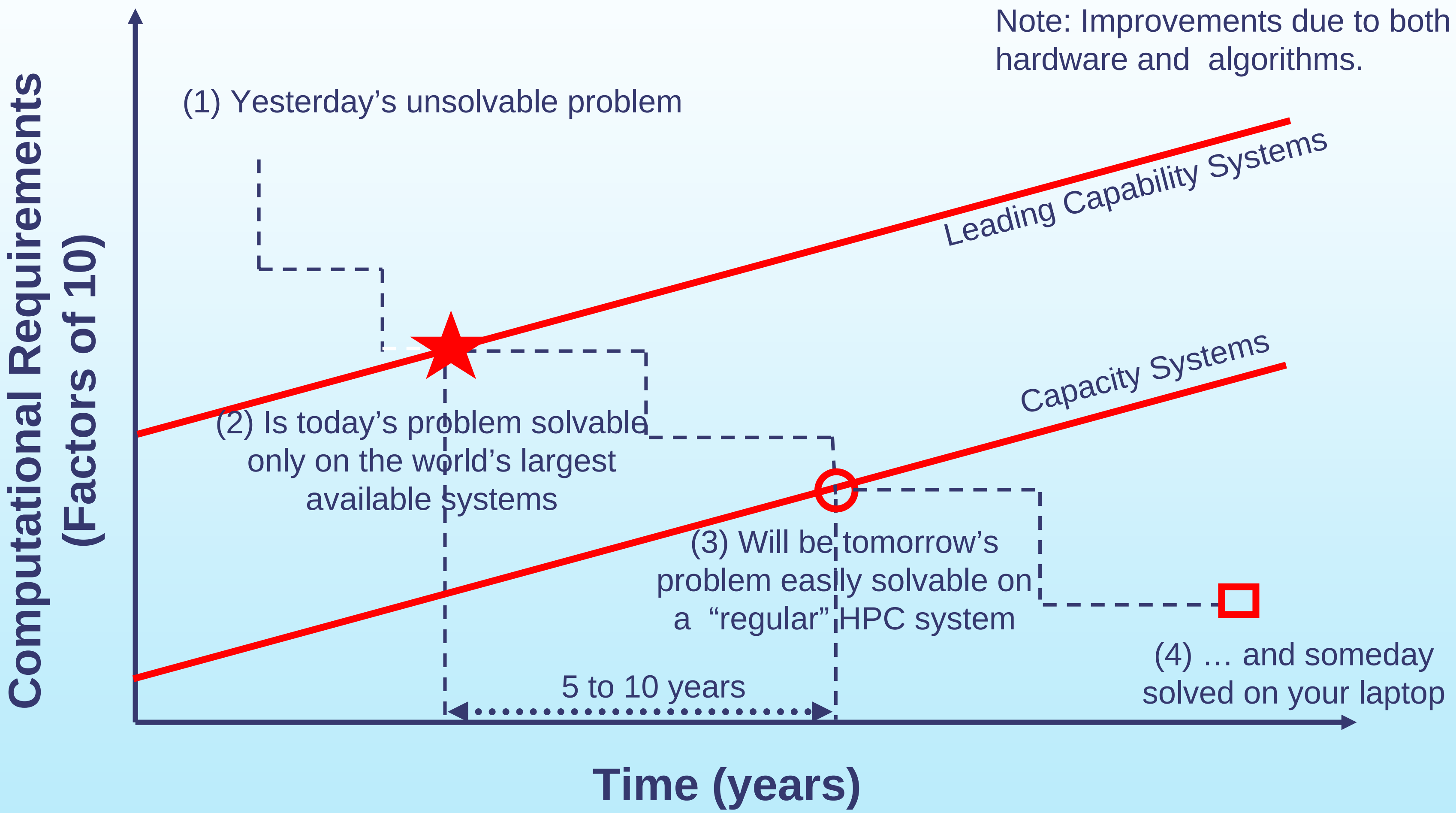
Growing Gap Between Resources and Requirements

Example: DoD HPC Modernization Program



Trend is Consistent with Other Agencies and Studies

Balancing Capability and Capacity



High-end computing enables U.S. to solve "unsolvable" problems first.

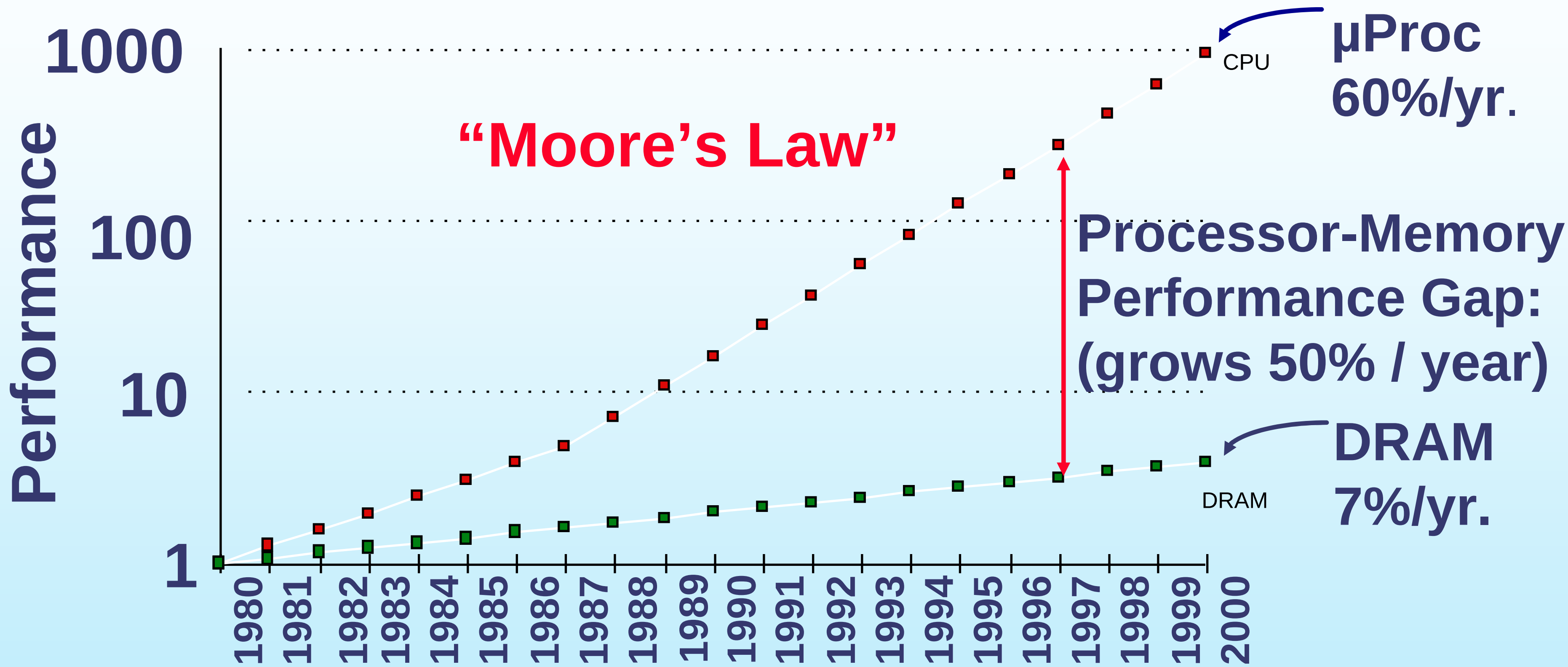
Acquisition Strategies

- **Need for better mapping of acquisition metrics to user needs with consistent, multi-year, indication to venders**
  - Architectural characteristics
  - Ease of use
  - Role of benchmarks
  - Multi-agency approach
- **Sharing, and perhaps cooperation, among agencies on acquisition strategies**

Architectural Problems in High-End Computing

- Processor-Memory Performance Gap
- Cache and Data Layout
- Inteconnect Fabric
- Role of Benchmarks

Processor-Memory Performance Gap



- Alpha 21264 full cache miss / instructions executed: 180 ns/1.7 ns =108 clks x 4 or 432 instructions
- Caches in Pentium Pro: 64% area, 88% transistors

*Taken from Patterson-Keeton Talk to SigMod

Processing vs. Memory Access

- **Doesn't cache solve this problem?**

  It depends. With small amounts of contiguous data, usually. With large amounts of non-contiguous data, usually not.

  In most computers the programmer has no control over cache.

  Often "a few" Bytes/FLOP is considered OK.

- **However, consider operations on the transpose of a matrix (e.g., for adjunct problems)**

  $\boldsymbol{Xa = b}$ $\qquad\qquad$ $\boldsymbol{X^T a = b}$

  If ***X*** is big enough, 100% cache misses are guaranteed, and we need at least 8 Bytes/FLOP (assuming a and b can be held in cache).

- **Latency and limited bandwidth of processor-memory and node-node communications are major limiters of performance for scientific computation**

Testing Processing vs. Memory Access with Benchmarks

- **Simple benchmark: Stream Triad**

$$a_i + s \times b_i = c_i$$

a_i, b_i, and c_i are vectors; s is a scalar. Vector length is chosen to be much longer than cache size.

Each execution includes 2 memory loads + 1 memory store and 2 FLOPs, or 12 Bytes/FLOP (assuming 8 Byte precision)

No computer has enough memory bandwidth to reference 12 Bytes for each FLOP!

Testing Processing vs. Memory Access with Benchmarks

- **Another Benchmark: Linpack**

$$A_{ij}\, x_j = b_i$$

Solve this linear equation for the vector *x*, where *A* is a known matrix, and *b* is a known vector. Linpack uses the BLAS routines, which divide A into blocks.

On the average Linpack requires 1 memory reference for every 2 FLOPs, or 4Bytes/Flop.

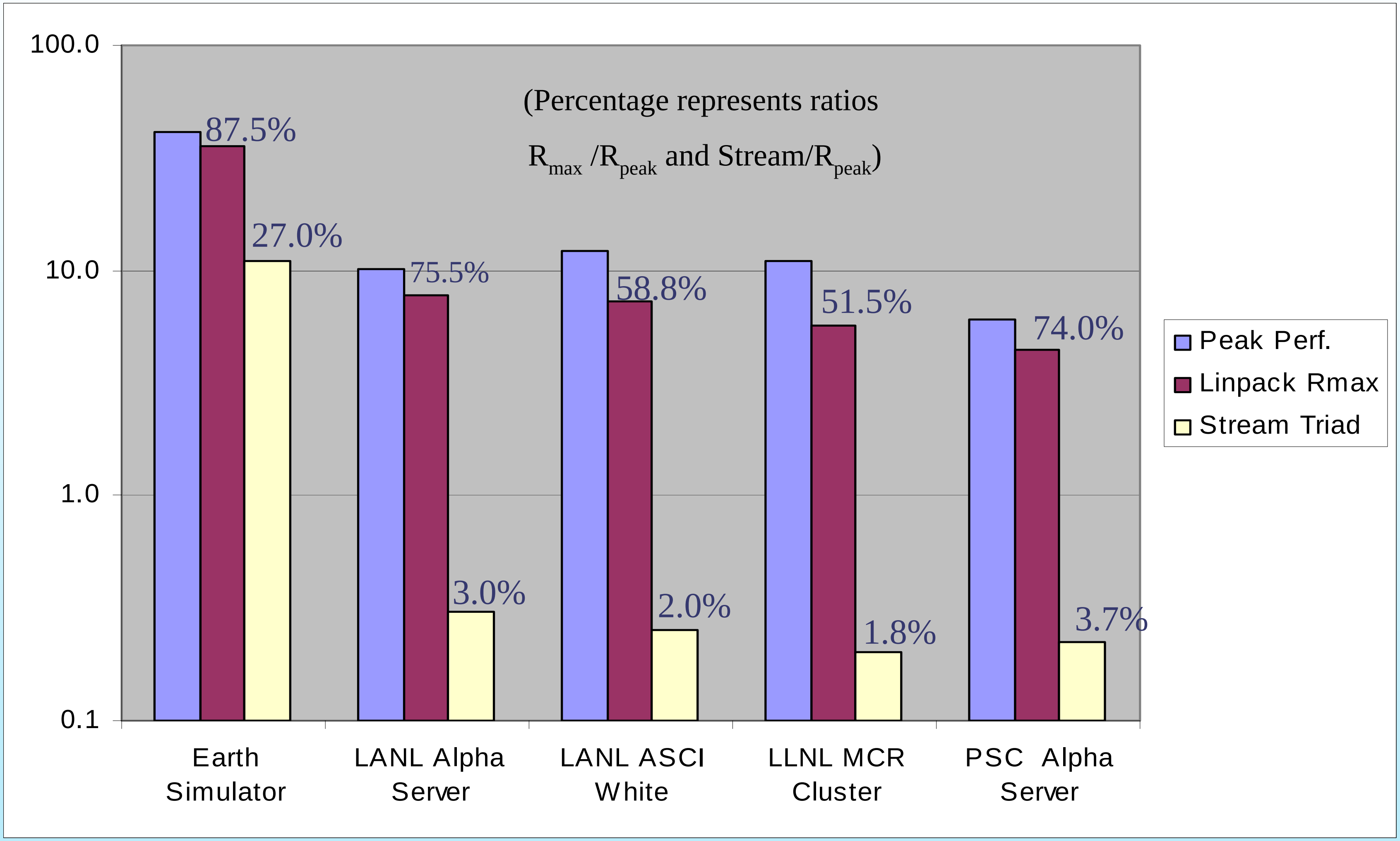
Many of these can be cache references.

Selected System Characteristics

		Earth Simulator (NEC)	ASCI Q (HP ES45)	ASCI White (IBM SP3)	MCR (Dual Xeon)	32 Node Cray X1 (Cray)
Year of Introduction		2002	2003	2000	2002	2003
Node Architecture		Vector SMP	Alpha micro SMP	Power 3 micro SMP	Xeon micro SMP	Vector SMP
System Topology		NEC single-stage Crossbar	Quadrics QsNet Fat-tree	IBM Omega network	Quadrics QsNet Fat-tree	2D Torus Interconnect
Number of Nodes		640	2048	512	1152	32
Processors	- per node	8	4	16	2	4
	- system total	5120	8192	8192	2304	128
Processor Speed		500 MHz	1.25 GHz	375 MHz	2.4 GHz	800 MHz
Peak Speed	- per processor	8 Gflops	2.5 Gflops	1.5 Gflops	4.8 Gflops	12.8 Gflops
	- per node	64 Gflops	10 Gflops	24 Gflops	9.6 Gflops	51.2 Gflops
	- system total	40 Tflops	30 Tflops	12 Tflops	10.8 Tflops	1.6 Tflops
Memory	- per node	16 GB	16 GB	16 GB	16 GB	8-64 GB
	- per processor	2 GB	4 GB	1 GB	2 GB	2-16 GB
	- system total	10.24 TB	48 TB	8 TB	4.6 TB	
Memory Bandwidth (peak)						
	- L1 Cache	N/A	20 GB/s	5 GB/s	20 GB/s	76.8 GB/s
	- L2 Cache	N/A	13 GB/s	2 GB/s	1.5 GB/s	
	Main (per processor)	32 GB/s	2 GB/s	1 GB/s	2 GB/s	34.1 GB/s
Inter-node MPI						
	- Latency	8.6 µsec	5 µsec	18 µsec	4.75 µsec	8.6 µsec
	- Bandwidth	11.8 GB/s	300 MB/s	500 MB/s	315 MB/s	11.9 GB/s
Bytes/flop to main memory		4	0.8	0.67	0.4	3
Bytes/flop interconnect		1.5	0.12	0.33	0.07	1

Most of this data is from Kerbyson, Hoisie, Wasserman; LANL; unpublished. Additional data from Jack Dongarra.

Performance Measures of Selected Top Computers



Note Logarithmic Y axis

What About Synthetic Benchmarks?

- Peak performance – nuf said
- Linpack –only measures performance of cache-friendly code
- Stream – only measures contiguous communications with memory, but good measure of bandwidth
- GUPS – really tough benchmark because it makes random memory access; may exceed requirements of most codes
- IDC balanced benchmarks – good compilation, but somewhat artificially combined
- Effective System Performance Benchmark – promising, but not widely used
- NAS Parallel Benchmarks – disused, but may be coming back
- Livermore Loops – obsolete
- Your own workload - ??

Resurgence of Performance Analysis Is Promising

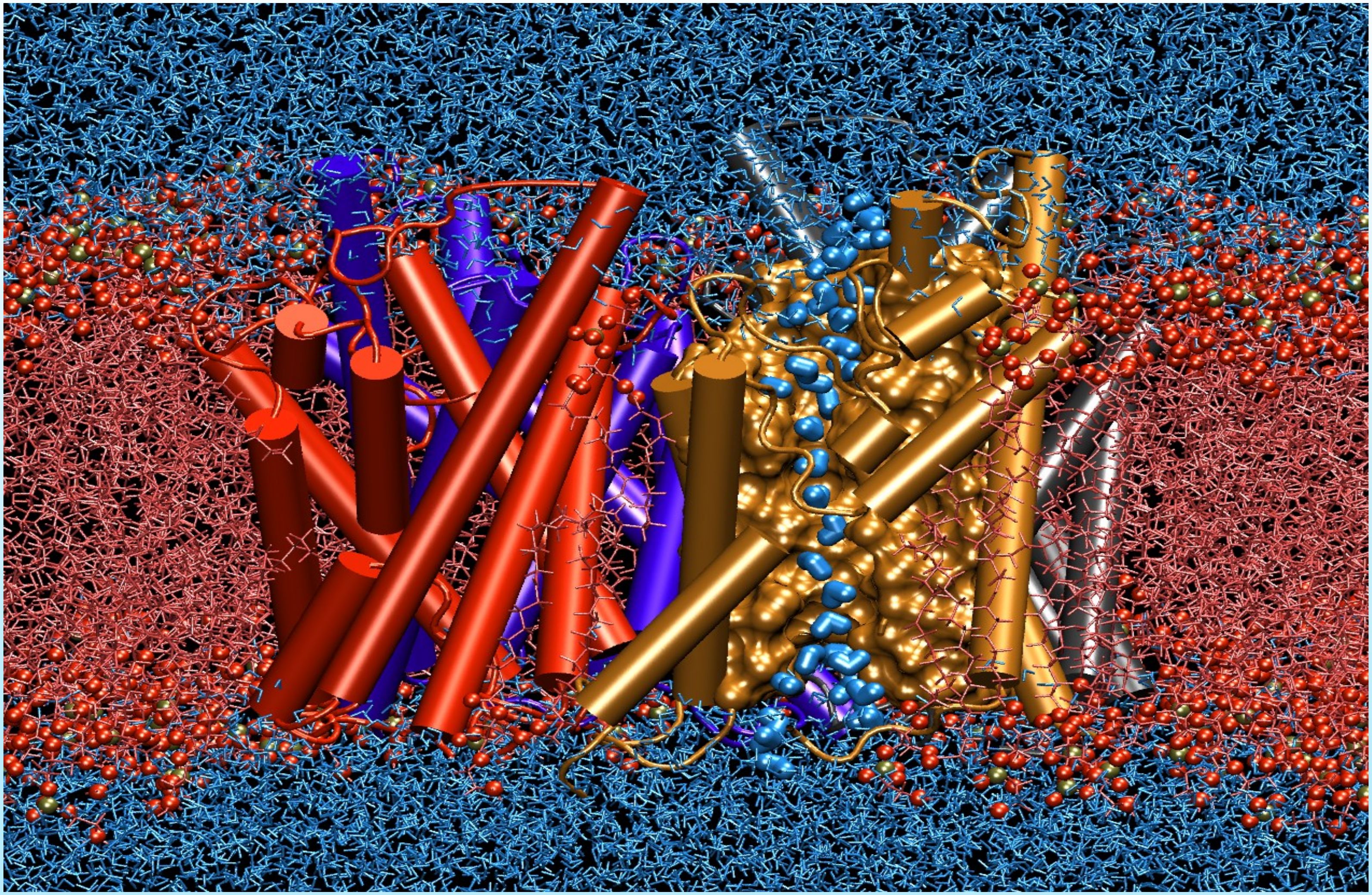
- LANL Performance and Architecture Lab: http://www.c3.lanl.gov/par_arch/
- Performance Evaluation Research Center: http://perc.nersc.gov/
- IDC User Forum: http://64.122.81.35/benchmark/
- Performance Modeling and Characterization: http://www.sdsc.edu/PMaC/Benchmark/
- NAS Parallel Benchmarks: http://www.nas.nasa.gov/Software/NPB/
- Recent High End Computing Workshop offered recommendations for performance evaluation: http://www.cra.orgActivities/workshops/nitrd/
- Great opportunity for agencies to cooperate on performance evaluation.

President's Advisory Committee on Information technology (PITAC)

- **PITAC authorized by law**
- **Reports to President through Office of Science and Technology Policy**
- **Co-chaired by Prof. Ed Lazowska, Univ. of Washington and Marc Benioff, CEO, Salesforce.com**
- **About to launch study of needs of computational science, chaired by Prof. Dan Reed, UNC**
  - Startup in June
  - Community input sought
  - Forum at SC 2004 in November
  - More info at www.itrd.gov/pitac

Simulation of Aquaporin Protein Inside a Cell (PSC Alpha Cluster)

Visualization shows transport of water molecules into cell.



For Further Information

Please contact us at:

nco@itrd.gov

Or visit us on the Web:

www.itrd.gov