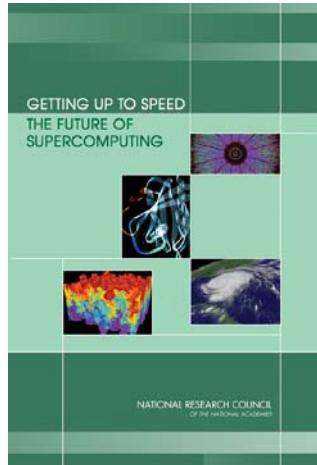


# Getting up to Speed: The Future of Supercomputing

Susan L Graham

Marc Snir

November 12, 2004



1

## Study Process

- Sponsored by DOE Office of Science and DOE Advanced Simulation and Computing
- March 2003 launch meeting
- Data gathering
  - 5 standard committee meetings
  - Applications Workshop (20+ computational scientists)
  - DOE weapons labs site visits (LLNL, SNL, LANL)
  - DOE science labs site visits (NERSC, Argonne/Oak Ridge)
  - NSA supercomputer center site visit
  - Town Hall (SC2003)
  - Japan forum (25+ supercomputing experts)
  - Japan site visits (ES, U. of Tokyo, JAXA, MEXT, auto manufacturer)
- Issuance of Interim report (July 2003)
- Blind peer-review process (17 reviewers); overseen by NRC-selected Monitor and Coordinator
- Dissemination (DOE, congressional staff, OSTP, SC2004)

2

**Study Committee**

- **SUSAN L. GRAHAM**, University of California, Berkeley, *Co-chair*
- **MARC SNIR**, University of Illinois at Urbana-Champaign, *Co-chair*
- WILLIAM J. DALLY, Stanford University
- JAMES DEMMEL, University of California, Berkeley
- **JACK J. DONGARRA**, University of Tennessee, Knoxville
- KENNETH S. FLAMM, University of Texas at Austin
- MARY JANE IRWIN, Pennsylvania State University
- **CHARLES KOELBEL**, Rice University
- BUTLER W. LAMPSON, Microsoft Corporation
- **ROBERT LUCAS**, University of Southern California, ISI
- PAUL C. MESSINA, Argonne National Laboratory
- JEFFREY PERLOFF, Department of Agricultural and Resource Economics, University of California, Berkeley
- WILLIAM H. PRESS, Los Alamos National Laboratory
- ALBERT J. SEMTNER, Oceanography Department, Naval Postgraduate School
- SCOTT STERN, Kellogg School of Management, Northwestern University
- SHANKAR SUBRAMANIAM, Departments of Bioengineering, Chemistry and Biochemistry, University of California, San Diego
- LAWRENCE C. TARBELL, JR., Technology Futures Office, Eagle Alliance
- **STEVEN J. WALLACH**, Chiaro Networks
- CSTB: **CYNTHIA A. PATTERSON** (*Study Director*), Phil Hilliard, Margaret Huynh

## Focus of Study

- Supercomputing – the development and use of the fastest and most powerful computing systems (capability computing).
  - Extends to high-performance computing
  - Does not address grid, networking, storage, special-purpose systems
- U.S. leadership and government policies.
- Market forces.

## Supercomputing Matters

- Essential for scientific discovery
- Essential for national security
- Essential to address broad societal challenges
- Important contributor to economy and competitiveness through use in engineering and manufacturing
- Important source of technological advances in IT
- Challenging research topic per se
- Supercomputing mattered in the past - Supercomputing will matter in the future

## Supercomputing is Government Business

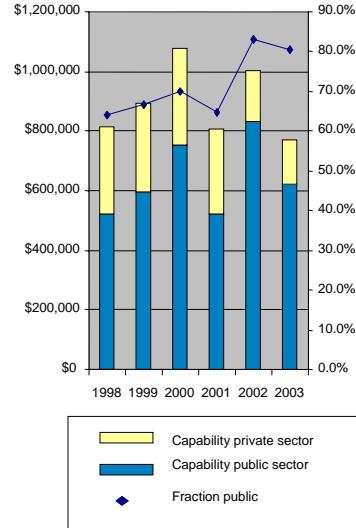
- In 2003 the public sector made > 50% of HPC purchases and > 80% of capability systems purchases (IDC).
- Supercomputing is mostly used to produce “public goods” (science, security...).
- Supercomputing technology has historically been developed with public funding.
  - Spillover to commercial/engineering

## The State of Supercomputing in the U.S. is Good

- As of June 2004 51% of TOP500 systems were installed in the U.S. and 91% of the TOP500 systems were made in the U.S.
- In 2003 U.S. vendors had 98% market share in capability systems and 88% in HPC (IDC).
- Supercomputing is used effectively.
  - Science, ASC, ...
- HPC is broadly available in academia and industry (clusters).

## The State of Supercomputing is Bad

- Companies primarily making custom supercomputers (e.g., Cray, ISVs) have a hard time surviving.
  - Supercomputing is a diminishing fraction of total computer market
  - Supercomputing market is unstable
    - Delayed acquisitions can jeopardize company
    - Private share is decreasing



## Supercomputing is a Fragile Ecosystem

- Small, unstable market, totally dependent on govt. purchases
- Weakened by wavering policies and investments (people leave, companies disappear)
- Recovery is expensive and takes a long time

9

## Current State is Largely Due to Success of Commodity Based Supercomputing

- Supercomputing performance growth in the last decade was almost entirely due to growth in uniprocessor performance (Moore's law). No progress in unique supercomputing technologies was needed and little occurred.
- Increase in parallelism has been modest – top commodity/hybrid system had 3,689 nodes in 6/94 and 5,120 nodes in 6/04.
- As of June 2004, 60% of TOP500 systems are clusters using commodity processors and switches; 95% of the systems use commodity processors.
- **Good:** Commodity clusters have democratized and broadened HPC.
- **Bad:** Commodity clusters have narrowed the market for non commodity systems. Lack of investment has reduced their viability.

10

## Commodity Systems Satisfy Most HPC Needs

- Good parallel performance can be achieved by clusters of commodity processors connected by commodity switches and switch interfaces, e.g., ASC Q.
- For problems with good locality (e.g., bioinformatics) such systems provide better time-to-solution than customized systems at any cost level.

## But Customization Needed to Achieve Certain Critical Goals

- Higher bandwidth and lower overhead for global communication can be achieved by *hybrid systems* (custom switch and custom switch interfaces, e.g., Red Storm).
- For problems with heavy global communication requirements, or when scaling to large node numbers is needed (e.g., climate) such systems provide better time-to-solution at a given cost, or may be the only way to meet deadlines.

## Customization is Becoming Essential

- Higher bandwidth to local memory and better latency hiding can be achieved by *custom systems* (systems with custom processors, e.g., Cray X1).
- For problems with little locality (e.g., GUPS), such systems provide better time-to-solution at given cost or may be the only way to meet deadlines.

## It will be harder in the future to “ride on the coattail” of Moore’s Law.

- Memory latency increases relative to processor speed (the *memory wall*): by 2020 about 800 loads and 90,000 floating-point operations would be executed while waiting for one memory access to complete.
- Global communication latency increases and bandwidth decreases relative to processor speed: by 2020 a global bandwidth of about 0.001 word/flops and global latency equivalent to about 1Mflops.
- Improvement in single processor performance is slowing down; future performance improvement in commodity processors will come from increasing on-chip parallelism.
- Mean Time to Failure is growing shorter as systems grow and devices shrink.

## Software Productivity is Low

- Need high-level notations that capture parallelism and locality.
- Application development environment and execution environment in HPC are less advanced and less robust than for general computing.
- Will need increasing levels of parallelism in future supercomputing.
- Custom/hybrid systems can support a simpler programming model.
  - *But that potential is largely unrealized*

## What Will We Need?

- Fundamentally new architectures before 2010 for supercomputing and before 2020 for general computing.
- New algorithms, new languages, new tools, and new systems for higher degrees of parallelism.
- A stable supply of trained engineers and scientists
- Continuity through institutions and rules that encourage the transfer of knowledge and experience into the future
- Technological diversity in hardware and software to enhance future technological options

## We Start at a Disadvantage

- The research pipeline has emptied.
  - NSF grants decreased 75%, published papers decreased 50%, no funding for significant demonstration systems
- The human pipeline is dry.
  - Averages: 36 PhDs/year in computational sciences (800 in CS); 3 hired by national labs
  - Less focus on supercomputing among other CS/CE disciplines
- Planning and coordination are lacking.

## The Time to Act is Now

- Fundamental changes take decades to mature.
  - Recall vectors, MPPs ...
- Current strengths are being lost.
  - People, companies, corporate memory

## What Lessons Should we Learn from the Japanese Earth Simulator? (1)

- ES demonstrates the advantages of custom supercomputers.
  - First on TOP500 for 2 years
  - Impressive levels of absolute performance on important applications ( $5\times$  to  $25\times$  better than best U.S. systems on climate codes)
- ES shows the importance of perseverance.
  - 5 year development, with government involved from the beginning
  - Successful use of technologies (e.g., High Performance Fortran) that were “out of fashion” in the U.S.
- ES *does not* show that Japan has overtaken the U.S.
  - U.S. had the technology to build a similar system with a similar investment in the same time frame
  - Most of the software technology used on the ES originates from the U.S.

## What Lessons Should we Learn from the Japanese Earth Simulator? (2)

- ES is not a security risk for the U.S.
  - What matters is not who has the largest computer; what matters is who has broad capability to apply supercomputing technology to reduce time-to-solution.
- ES shows how precarious is the worldwide state of custom supercomputing:
  - ES vector architecture is old
  - ES would not have happened without massive government support of industrial R&D
  - ES may not happen again (NEC may not continue with vector processors; government may not repeat such a massive investment)
- U.S. should invest in supercomputing primarily to satisfy its own needs.

## Overall Recommendation

To meet the current and future needs of the United States, the **government agencies that depend on supercomputing**, together with the U.S. Congress, need to take primary responsibility for **accelerating advances in supercomputing** and **ensuring that there are multiple strong domestic suppliers** of both hardware and software.

## Recommendation 1

**To get the maximum leverage from the national effort, the government agencies that are the major users of supercomputing should be **jointly responsible for the strength and continued evolution of the supercomputing infrastructure in the United States, from basic research to suppliers and deployed platforms.** The Congress should provide adequate and sustained funding.**

- Long-term (5-10 years) integrated HEC plan
- Budget requests matched to plan
- Loose coordination of research funding; tight coordination of industrial R&D
- Joint planning and coordination of acquisitions (reduce procurement overheads, reduce variability)

## Recommendation 2

**The government agencies that are the primary users of supercomputing should ensure domestic leadership in those technologies that are essential to meet national needs.**

- Unique technologies are needed (custom processors, interconnects, scalable software); these will not come from broad market
- Need U.S. suppliers because may want to restrict export
- **Need U.S. suppliers because no other country is certain to do it**
- Leadership both helps mainstream computing and draws from it

## Recommendation 3

**To satisfy its need for unique supercomputing technologies such as high-bandwidth systems, the government needs to ensure the viability of multiple domestic suppliers.**

- Viability achieved by stable, long-term government investments at adequate levels
- Either subsidize R&D or support from stable, long-term procurement contracts (UK model)
- Custom processors are a key technology that will not be provided by the broad market
- Other hardware technologies also important (e.g., interconnect)

## Recommendation 4

**The creation and long-term maintenance of the software that is key to supercomputing requires the support of those agencies that are responsible for supercomputing R&D. That software includes operating systems, libraries, compilers, software development and data analysis tools, application codes, and databases.**

- Need larger and more targeted coordinated investments
- Multiple models: vertical vendor, horizontal vendor, not for profit organization, open source model...
- Need stability and continuity (corporate memory)
- Build only what cannot be bought

25

## Recommendation 5

**The government agencies responsible for supercomputing should underwrite a community effort to develop and maintain a roadmap that identifies key obstacles and synergies in all of supercomputing.**

- Roadmap should inform R&D investments
- Wide participation from researchers, developers and users
- Driven top-down (requirements) and bottom-up (technologies)
- Must be quantitative and measurable
- Must reflect interdependence of technologies
- Informs, but does not fully determine research agenda

26

## Recommendation 6

**Government agencies responsible for supercomputing should increase their levels of stable, robust, sustained multiagency investment in basic research.** More research is needed in all the key technologies required for the design and use of supercomputers (architecture, software, algorithms, and applications).

- Mix of small and large projects, including demonstration systems
- Emphasis on university projects - education and free flow of information
- Estimated investment needed for core technologies is \$140M per year (more needed for applications)

## Recommendation 7

**Supercomputing research is an international activity; barriers to international collaboration should be minimized.**

- Barriers reduce broad benefit of supercomputing to science
- Early-stage sharing of ideas compensates for small size of community
- Collaborators should have access to domestic supercomputing systems
- Technology advances flow to and from broader IT industry; fast development cycles and fast technology evolution require close interaction
- No single supercomputing technology presents major risk; US strategic advantage is in its broad capability
- Export restrictions have hurt U.S. manufacturers; some (e.g., on commodity clusters) lack any rationale

## **The U. S. government should ensure that researchers with the most demanding computational requirements have access to the most powerful supercomputing systems**

- Important for advancement of science
- Needed to educate next generation and create the needed software infrastructure
- Sufficient stable funding must be provided
- Infrastructure funding should be separated from funding for IT research
- Capability systems should be used for jobs that need that capability

## **Questions?**

- A prepublication version of the report is available online at:
  - <http://www.sc.doe.gov/ascr/workshop.html>
  - <http://books.nap.edu/catalog/11148.html>
- For more information on receiving a final copy of the report, contact CSTB at 202.334.2605.