The Perils of Success: Meeting the Needs for Advanced Computing

William Gropp
www.cs.illinois.edu/~wgropp

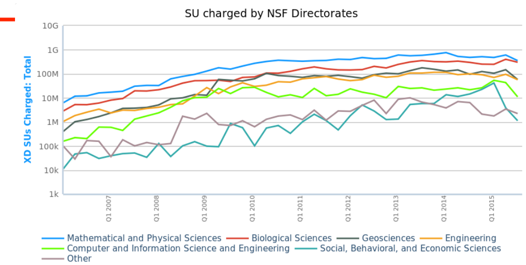
The Take Away

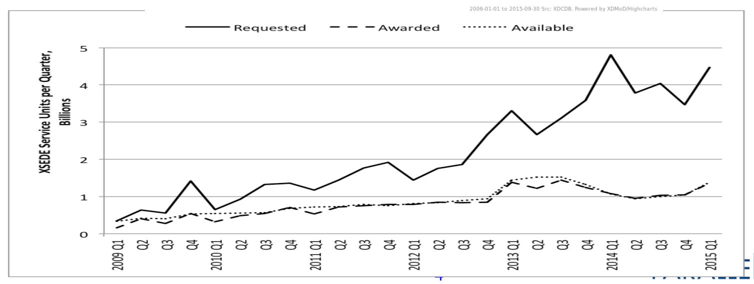
- I'm *not* going to tell you what we recommended NSF do for advanced computing
  - ♦ They can't do it alone
- I *am* asking you to get more involved in quantifying the needs, opportunities, and in making the case for advanced computing *infrastructure*
  - ♦ Not just hardware
- In other words, *you* need to help to realize the recommendations of the report

Computing is a runaway success!

- All fields of science and engineering have come to depend on computing
  - ♦ Many great stories and results this week; more at XSEDE next month
- Demand has been growing exponentially (literally)
- Now limited by supply…

Allocation and request history





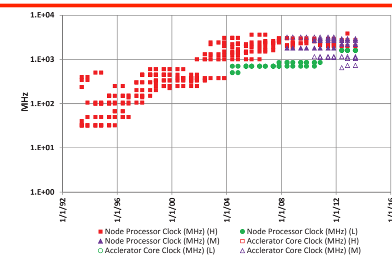
Question

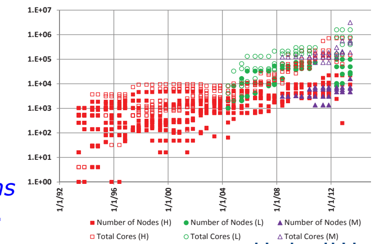
- How do we get more supply and what does it look like?
  - ♦ And what will it cost
    - And how will we be able to use it
      - – And how do we plan for the next generation

- For example…

End of the free ride

- Stagnation in clock rates after 2004 (end of Dennard scaling)
- Rapid rise in concurrency shortly after 2004
- Proliferation of new architectures – light weight/accelerator cores

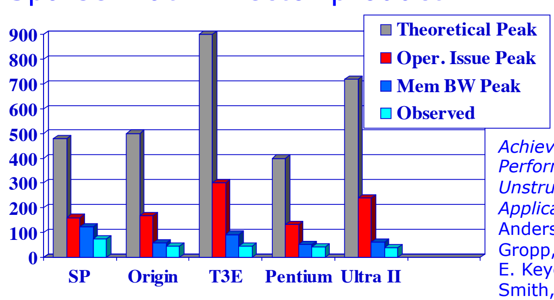




"Yearly Update: Exascale Projections for 2013," Peter M. Kogge, David R. Resnick

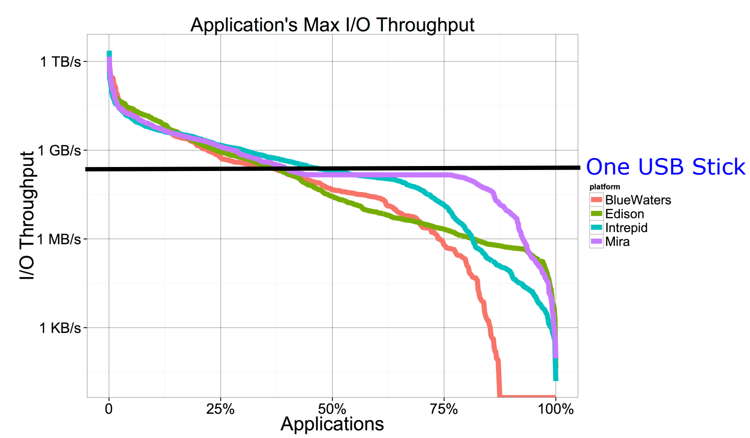
Not just FLOPS

- We've known this for a ***long*** time
- Here's a sample from 1999 showing that sustained (not peak) memory bandwidth is a far better metric for predicting performance for Sparse matrix-vector product



Achieving High Sustained Performance in an Unstructured Mesh CFD Application, W. K. Anderson, William D. Gropp, D. K. Kaushik, D. E. Keyes, and B. F. Smith, SC99

I/O is often forgotten



A Multiplatform Study of I/O Behavior on Petascale Supercomputers, Huong Luu, Marianne Winslett, William Gropp, Robert B. Ross, Philip H. Carns, Kevin Harms, Prabhat, Surendra Byna, and Yushu Yao, HPDC 2015

It will only get harder…

- Proposed Hardware for high-end systems:
  - ♦ Emphasizes FLOPS
  - ♦ Major changes in "execution model" (useful abstraction of machine operation at the HW/SW service level)
    - Does not match current algorithms, programming systems
  - ♦ I/O limited by stagnant Bandwidth to disks (and cost)

DOE's roadmap

System attributes	NERSC Now	OLCF Now	ALCF Now	NERSC Upgrade	OLCF Upgrade	ALCF Upgrades	
Name Planned Installation	**Edison**	**TITAN**	**MIRA**	**Cori 2016**	**Summit 2017-2018**	**Theta 2016**	**Aurora 2018-2019**
System peak (PF)	2.6	27	10	> 30	150	>8.5	180
Peak Power (MW)	2	9	4.8	< 3.7	10	1.7	13
Total system memory	357 TB	710TB	768TB	~1 PB DDR4 + High Bandwidth Memory (HBM)+1.5PB persistent memory	> 1.74 PB DDR4 + HBM + 2.8 PB persistent memory	>480 TB DDR4 + High Bandwidth Memory (HBM)	> 7 PB High Bandwidth On-Package Memory Local Memory and Persistent Memory
Node performance (TF)	0.460	1.452	0.204	> 3	> 40	> 3	> 17 times Mira
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	Multiple IBM Power9 CPUs & multiple Nvidia Voltas GPUS	Intel Knights Landing Xeon Phi many core CPUs	Knights Hill Xeon Phi many core CPUs
System size (nodes)	5,600 nodes	18,688 nodes	49,152	9,300 nodes 1,900 nodes in data partition	~3,500 nodes	>2,500 nodes	>50,000 nodes
System Interconnect	Aries	Gemini	5D Torus	Aries	Dual Rail EDR-IB	Aries	2nd Generation Inte Omni-Path Architecture
File System	7.6 PB 168 GB/s, Lustre®	32 PB 1 TB/s, Lustre®	26 PB 300 GB/s GPFS™	28 PB 744 GB/s Lustre®	120 PB 1 TB/s GPFS™	10PB, 210 GB/s Lustre initial	150 PB 1 TB/s Lustre®

Challenge to the community

- Make sure that there are enough cycles of any kind
  - ♦ Help determine need (not want)
  - ♦ Aside: Cloud cost (currently) ≫ center cost for cycles
- Make sure these are the right kind of cycles/capability/capacity for the science
  - ♦ If other than DGEMM, you maybe unhappy
  - ♦ Help determine details; sensible metrics, not TOP500
- Software/Algorithm support – Who does? Who pays? Who sustains? Where does expertise reside?
  - ♦ Help understand the tradeoffs
    - Software lasts *much* longer than hardware – PETSc is 26, MPICH is 24 years old
  - ♦ Examples: PAID program; XSEDE support options

Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020

Committee on Future Directions for NSF Advanced Computing Infrastructure
to Support U.S. Science in 2017-2020

Computer Science and Telecommunications Board

Division on Engineering and Physical Sciences

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

http://tinyurl.com/advcomp17-20

Committee

- WILLIAM D. GROPP, University of Illinois, Urbana-Champaign, Co-Chair
- ROBERT J. HARRISON, Stony Brook University, Co-Chair
- MARK R. ABBOTT, Woods Hole Oceanographic Institution
- ROBERT L. GROSSMAN, University of Chicago
- PETER M. KOGGE, University of Notre Dame
- PADMA RAGHAVAN, Pennsylvania State University
- DANIEL A. REED, University of Iowa
- VALERIE TAYLOR, Texas A&M University
- KATHERINE A. YELICK, University of California, Berkeley

Reviewers

- Daniel E. Atkins III, University of Michigan
- David A. Bader, Georgia Institute of Technology
- Robert Brammer, Brammer Technology, LLC
- Andrew A. Chien, University of Chicago
- Jeff Dozier, University of California, Santa Barbara
- Dennis Gannon, Microsoft Research (retired)
- Gary S. Grest, Sandia National Laboratories
- Laura M. Haas, IBM
- Tony Hey, University of Washington eScience Institute
- David Keyes, King Abdullah University of Science and Technology
- Michael L. Klein, Temple University
- David A. Lifka, Cornell University
- Jeremiah P. Ostriker, Columbia University
- Terrence J. Sejnowski, Salk Institute for Biological Studies
- Marc Snir, Argonne National Laboratory
- Warren M. Washington, National Center for Atmospheric Research
- John West, Texas Advanced Computer Center

Recommended goals for advanced computing

1. Position the United States for continued leadership in science and engineering
2. Ensure that resources meet community needs (not the same as wants)
3. Aid the scientific community in keeping up with the revolution in computing, and
4. Sustain the infrastructure for advanced computing

Observations about positioning for leadership

- Both large-scale simulation and analysis of massive data revolutionizing many areas of research
  - Rise in volume and diversity of scientific data represents and a significant disruption and opportunity
- Meeting future needs will require systems that support a wide range of advanced computing capabilities, including large-scale high-performance and data-intensive systems
- Increased capability has historically enabled new science. Without continued growth, some research will have difficulty making advances
- Many fields now rely on a greater aggregate amount of computing than a typical university can provide
- “Converged” systems can play a role; more specialized systems may also be needed to meet some requirements

Japan's second tier plans

University	2017	2020	2023
Hokkaido	10+PF/s	10+PF/s	50+PF/s
Tohoku	NEC SX-ACE 800TF/s	30+PF/s	30+PF/s
Tsukuba	PostT2k JHPCA 30PF/s	100+PF/s	100+PF/s
Tokyo	Fujitsu FX10 1PF/s	50+PF/s	50+PF/s
Tokyo Tech.	Tsubame 3 20-25 PF/s	Tsubame 4 (100-200 PF/s)	Tsubame 4
Nagoya	Post FX10 upgrade 3 PF/s	50+ PF/s	50+ PF/s
Kyoto	10+PF/s	50+PF/s	50+PF/s
Osaka	NEC SC-ACE 400TF/s	5+PiB/s	5+PiB/s
Kyushu	10+PF/s	10+ PF/s	50+PF/s

Taken from "Japanese "Leading Machine" Candidates Roadmap of the 9 HPCI University Centers", April 2015. Does not include the "K Computer"

Observations about resources and needs

- Demand for advanced computing growing, changing rapidly
- Gap between supply and demand growing
- Overall planning process for advanced computing
  - Not systematic or uniform
  - Not visibly reflected in NSF's strategic planning
- Ongoing and more regular/structured process would make it possible to:
  - Collect & roll up requirements
  - Prioritize investments based on science and engineering priorities
- To be cost-effective, NSF must secure access to capabilities that represent compromises wrt individual applications but reasonably support the overall portfolio
  - Requirements collection and roadmaps will inform decisions

Observations about understanding costs and benefits

- Better information about the relationship among the cost of roadmap choices, requirements, and science benefits would
  - Help inform program managers about the total costs of proposed research
  - Focus researchers' attention on effective/efficient use of these valuable shared resources

Observations about keeping up with the revolution in computing

- Computer architectures, hardware, program models, are changing rapidly
- Better software tools, technical expertise, and more flexible service models can boost productivity
- Leadership role in defining future advanced capabilities and helping researchers use them effectively will help ensure that:
  - Software and systems remain relevant to science portfolio
  - Researchers are prepared to use current and future capabilities
  - Investments are aligned with future directions

Observations about sustaining infrastructure

- Expertise and physical infrastructure are essential, long-lived assets
- Recent strategy of acquiring facilities and creating centers relies on:
  - Irregularly scheduled competition among institutions roughly every 2 to 5 years
  - Equipment, facility, and operating cost sharing by states, institutions, and vendors
- Challenges with this approach:
  - Relies on cost sharing that may no longer be viable due to mounting costs and budget pressures
  - Repeated competitions can lead to proposals designed to win a competition rather than maximize scientific returns
  - Most importantly, doesn't provide long-term support needed to develop and retain talent needed to manage systems, support users, and evolve software

The future depends on You!

- Read the report. Support it or improve on it!
  - **http://tinyurl.com/advcomp17-20**
- We call for better requirements gathering – help us move past FLOPS so that systems (hardware and software) can be better matched to need
- Roadmaps, both of technology and science, essential to setting directions and priorities

Details

- Recommendations of the Report
- Relationship to NSCI from the Report

Recommendation 1 NSF should sustain and seek to grow its investments in advanced computing—to include hardware and services, software and algorithms, and expertise—to ensure that the nation's researchers can continue to work at frontiers of science and engineering.

Recommendation 1.1 NSF should ensure that adequate advanced computing resources are focused on systems and services that support scientific research. In the future, these requirements will be captured in its roadmaps.

Recommendation 1.2 Within today's limited budget envelope, this will mean, first and foremost, ensuring that a predominant share of advanced computing investments be focused on production capabilities and that this focus not be diluted by undertaking too many experimental or research activities as part of the Foundation's advanced computing program.

Recommendation 1.3 NSF should explore partnerships, both strategic and financial, with federal agencies that also provide advanced computing capabilities as well as federal agencies that rely on NSF facilities to provide computing support for their grantees.

Recommendation 2 As it supports the full range of science requirements for advanced computing in the 2017-2020 timeframe, NSF should pay particular attention to providing support for the revolution in data-driven science along with simulation. It should ensure that it can provide unique capabilities to support large-scale simulations and/or data analytics that would otherwise be unavailable to researchers and continue to monitor the cost-effectiveness of commercial cloud services.

Recommendation 2.1 NSF should integrate support for the revolution in data-driven science into the Foundation's strategy for advanced computing by (a) requiring most future systems and services and all those that are intended to be general purpose to be more data-capable in both hardware and software and (b) expanding the portfolio of facilities and services optimized for data-intensive as well as numerically-intensive computing, and (c) carefully evaluating inclusion of facilities and services optimized for data-intensive computing in its portfolio of advanced computing services.

Recommendation 2.2 NSF should (a) provide one or more systems for applications that require a single large, tightly-coupled parallel computer and (b) broaden the accessibility and utility of these large-scale platforms by allocating high-throughput as well as high-performance workflows to them.

Recommendation 2.3 NSF should (a) eliminate barriers to cost-effective academic use of the commercial cloud and (b) carefully evaluate the full cost and other attributes (e.g., productivity and match to science workflows) of all services and infrastructure models to determine whether such services can supply resources that meet the science needs of segments of the community in the most effective ways.

Recommendation 3 To inform decisions about capabilities planned for 2020 and beyond, NSF should collect community requirements and construct and publish roadmaps to allow the Foundation to set priorities better and make more strategic decisions about advanced computing.

Recommendation 3.1 NSF should inform its strategy and decisions about investment tradeoffs using a requirements analysis that draws on community input, information on requirements contained in research proposals, allocation requests, and Foundation-wide information gathering.

Recommendation 3.2 NSF should construct and periodically update roadmaps for advanced computing that reflect these requirements and anticipated technology trends to help the Foundation set priorities and make more strategic decisions about science and engineering and to enable the researchers that use advanced computing to make plans and set priorities.

Recommendation 3.3 NSF should document and publish on a regular basis the amount and types of advanced computing capabilities that are needed to respond to science and engineering research opportunities.

Recommendation 3.4 NSF should employ this requirements analysis and resulting roadmaps to explore whether there are more opportunities to use shared advanced computing facilities to support individual science programs such as MREFC projects.

Recommendation 4 NSF should adopt approaches that allow investments in advanced computing hardware acquisition, computing services, data services, expertise, algorithms, and software to be considered in an integrated manner.

Recommendation 4.1 NSF should consider requiring that all proposals contain an estimate of the advanced computing resources required to carry out the proposed work and creating a standardized template for collection of the information as one step of potentially many towards more efficient individual and collective use of these finite, expensive, shared resources. (This information would also inform the requirements process.)

Recommendation 4.2 NSF should inform users and program managers of the cost of advanced computing allocation requests in dollars to illuminate the total cost and value of proposed research activities.

Recommendation 5 NSF should support the development and maintenance of expertise, scientific software, and software tools that are needed to make efficient use of its advanced computing resources.

Recommendation 5.1 NSF should continue to develop, sustain and leverage expertise in all programs that supply or use advanced computing to help researchers use today's advanced computing more effectively and prepare for future machine architectures.

Recommendation 5.2 NSF should explore ways to provision expertise in more effective and scalable ways to enable researchers to make their software more efficient, for instance by making more pervasive XSEDE's practice that permits researchers to request an allocation of staff time along with computer time.

Recommendation 5.3 NSF should continue to invest in supporting science codes and in continuing to update them to support new systems and incorporate new algorithms, recognizing that this work is not primarily a research activity but rather is support of software infrastructure.

Recommendation 6 NSF should invest modestly to explore next-generation hardware and software technologies to explore new ideas for delivering capabilities that can be used effectively for scientific research, tested, and transitioned into production where successful. Not all communities will be ready to adopt radically new technologies quickly, and NSF should provision advanced computing resources accordingly.

Recommendation 7 NSF should manage advanced computing investments in a more predictable and sustainable way.

- **Recommendation 7.1** NSF should consider funding models for advanced computing facilities that emphasize continuity of support.
- **Recommendation 7.2** NSF should explore and possibly pilot the use of a special account (such as that used for MREFC) to support large-scale advanced computing facilities.
- **Recommendation 7.3** NSF should consider longer-term commitments to center-like entities that can provide advanced computing resources and the expertise to use them effectively in the scientific community.
- **Recommendation 7.4** NSF should establish regular processes for rigorous review of these center-like entities and not just their individual procurements.

NSCI and NSF

- High-performance computing (HPC) remains critical for science and industry; if anything, the need and value continues to grow. (NSCI Section 1)
- "Increasing coherence between the technology base used for modeling and simulation and that used for data analytic computing."(NSCI Section 2.2)
- Building on its successes in cyberinfrastructure, the National Science Foundation (NSF) has an important role to play in both providing HPC (including data and compute) for basic science and in development of the science needed to advance HPC, including the algorithms, software, and hardware for extreme scale computing. (NSCI Section 3a)
- NSF must also contribute to the development of an HPC workforce. (NSCI Section 3a)
- Public-private partnerships should be explored. (NSCI Section 1.2)
- HPC research must be transitioned into practice. (NSCI Section 1.4) This report's recommendations to NSF echo this need; in particular, NSF needs both to perform research in support of HPC and support bringing that research into practice as needed by the NSF user community.
- Embrace an integrated approach to providing effective HPC, combining hardware, software, and algorithms, as well as addressing the development of an HPC-capable workforce and the whole of HPC, including the mid-range as well as the high-end. (NSCI Section 2.4)