THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Division on Engineering and Physical Sciences
Computer Science and Telecommunications Board

500 Fifth Street, NW
Washington, DC 20001
Phone: 202 334 2605
Fax: 202 334 2318
www.cstb.org

The Future of Computing Performance: Game Over or Next Level? A Symposium

AGENDA

March 22, 2011

Venable LLP Conference Center
575 7th Street NW
Washington, DC 20004

Overview: This symposium will begin with a briefing and discussion of the recently-released National Research Council report, *The Future of Computing Performance: Game Over or Next Level?* Subsequent panel sessions will explore issues raised by the report and consider opportunities and challenges for sustaining growth in computing performance. Panels will focus on software (parallelism, programming models, and so on), hardware and power (chip multiprocessors, computer architectures, energy constraints), computer science and engineering education, and federal and industrial research efforts.

9:00 AM	-	9:05 AM	**Welcome** *Jon Eisenberg, The National Academies* *Director, Computer Science and Telecommunications Board*
9:05 AM	-	10:15 AM	**The Future of Computing Performance: Game Over or Next Level?** **Report Briefing and Discussion** *Samuel H. Fuller, Analog Devices, Inc.* *Chair, Committee on Sustaining Growth in Computing Performance*
10:15 AM	-	10:25 AM	**Break**
10:25 AM	-	11:25 AM	**Session 1 – Parallelism and Innovative Programming Models, Algorithms, and Languages** Panel moderator: *Kathryn McKinley, University of Texas at Austin*

For many of today's applications, the underlying algorithms in use do not assume or exploit parallel processing explicitly. Instead, software creators typically depend, implicitly or explicitly, on compilers and other layers of the programming environment to parallelize where possible, leaving the software developer free to think sequentially and focus on higher-level issues. The

report makes clear that that method now leads to a dead end. The intellectual keystone of the new way forward is rethinking programming models. A fundamental focus on parallelism will be needed when designing future solutions.

Questions to consider:

- What is the current state of the art for sequential programming and parallel programming?
- Why is parallel programming fundamentally different? It is really intrinsically harder? Why?
- What are promising program models, algorithms, and languages that could effectively enable parallelism?
- To what extent will typical software developers need to grapple with parallelism directly, and to what extent could it be managed through automated tools and techniques lower in the stack?
- What impacts will an increased focus on parallelism have on software assurance, security, reliability, and so on?
- Where are the key opportunities for programming model experts and computer architecture design experts to collaborate to achieve breakthroughs that will push computing performance forward?
- What will the computing "stack" of the future look like? Is stack even the right metaphor? ?

Panelists:

David Grove, IBM Corporation
Keshav Pingali, University of Texas at Austin
Guy Steele, Oracle Labs
Katherine Yelick, University of California, Berkeley

11:25 AM	-	11:40 AM	**Break and Gather Lunch**
11:40 AM	-	12:10 PM	**Lunchtime Talk** **Perspective on Investment and Resources to Support Continuing Innovation in Computing Performance** *David Liddle, U.S. Venture Partners*
12:10 PM	-	1:10 PM	**Session 2 – Computing in a Power Constrained World** Panel Moderator: *Mark A. Horowitz, Stanford University*

While incremental advances in computing performance will continue, overcoming power constraints is difficult, potentially impossible, and likely requires a radical rethinking of computation and the logic gates used to build computing systems. Limitations imposed by power consumption are responsible for the present shift toward parallel computing; nevertheless, such limitations will constrain the computation performance of even parallel computing systems unless computer designers take fundamentally different approaches. If we cannot scale down the energy per function as fast as we scale up the performance (functions per second), the power (energy per second) consumed by the system will rise as will the cost of the overall system. Moreover, form-factors such as smart phones are in increasingly widespread use and have strict power limits. Designers today must find the best performance that can be achieved within a specified power envelope.

Questions to consider:

- Will we be able to effectively harness chip multiprocessors on desktops and even in mobile phones?
- Under what circumstances should some cores be more capable than others or even use different instruction set architectures?
- How will challenges associated with power and energy be addressed and at what level?
- What will the computing "stack" of the future look like? Is stack even the right metaphor?
- Where are the most promising avenues for research: computing vs. communication, software vs. hardware, breakthroughs in fundamental electronics, better cooling, better mobile power options, or something else?
- What is the low hanging fruit? What short-term power efficiency improvements can be made?
- Even past the simpler power inefficiency problems, how can we make the fundamental computations more energy efficient through new technologies, customization of hardware, different instruction set architectures, and so on?

Panelists:
Robert Dennard, IBM Corporation
Daniel Dobberpuhl, Consultant
Kevin Nowka, IBM Corporation
Partha Ranganathan, Hewlett-Packard Company

1:10 PM	-	1:20 PM	**Break**
1:20 PM	-	2:20 PM	**Session 3 – Reaching the Next Level in Computer Science and Engineering Education** Panel Moderator: *Mark D. Hill, University of Wisconsin–Madison*

Computing performance represents a serious and significant set of challenges, not only for the computing industry, but for the many sectors of society dependent on advances in IT and computation. Similarly, a slowdown in the growth of computing performance will have global economic repercussions and potentially impact U.S. competitiveness. These pressing issues call for: sustained research investment—the best science and engineering minds must be brought to bear on these challenges; changes in practice—how we go about developing computer hardware and software *today* will form a foundation for future performance gains; and changes in education—the emerging generation of technical experts will need to understand quite different (and in some cases not-yet-developed) models of thinking about IT, computation, and software.

Questions to consider:

- How are the new skills and tools needed for the next generation of computer scientists and engineers different from the skills and tools of today?
- What are the implications and trends for programming languages and methods?
- Can current knowledge in education and cognitive learning theory contribute to improving computer science education to address parallelism and related emerging needs?
- How can the portability/legacy software challenge best be managed?

- What are the implications for the core curriculum? How can parallelism and parallel thinking best be incorporated?
- What can the software field learn from the natively-parallel world of hardware designers?
- Can parallel programming skills really be taught separately from power-constrained system design?

Panelists:
Guy Blelloch, Carnegie Mellon University
Daniel Ernst, University of Wisconsin–Eau Claire
David Kirk, NVIDIA
Marcia Linn, University of California, Berkeley

2:20 PM - 3:20 PM **Session 4 – Exploring the Terrain: Research Directions, Priorities, and Strategies**
Panel Moderator: *Samuel H. Fuller, Analog Devices, Inc.*
Chair, Committee on Sustaining Growth in Computing Performance

This is a golden time for innovation in computing architectures, software, and hardware. We have already begun to see diversity in computer designs to optimize for such metrics as power and throughput (e.g. green-computing) and the substantial centralization of computational capability in the cloud-computing paradigm. Our computing models are likely to continue to evolve quickly in the foreseeable future (e.g. internet-enabled embedded devices that run complex, sophisticated software). A prudent research portfolio must include concomitant efforts to advance all systems aspects, lest they become tomorrow's bottlenecks or crises. In addition to performance, the next generation of discoveries will require advances in many other aspects of computer design to keep systems balanced. Memories and networks should be larger, faster, and less expensive. All components must be designed for energy-efficient operation. To salvage value from the nation's current, substantial IT investment, we should seek ways to bring sequential programs into the parallel world as well as developing tools and strategies to enhance code creation, maintenance, verification, and adaptation. Implementing a research agenda targeting the issues discussed here, although crucial for progress, will take time and dedicated effort in the part of industry, academia, and government.

Question to consider:
- What research and development efforts are underway and what new efforts and initiatives are needed—and from whom—to meet these challenges?

Panelists:
Susanne Hambrusch, National Science Foundation
Norman Jouppi, Hewlett-Packard Company
Keith Marzullo, National Science Foundation

3:20 PM - 3:30 PM **Wrap-up and Adjourn**
Samuel H. Fuller, Committee Chair