Computing Systems for the 21st Century

Daniel A. Reed
reed@ncsa.uiuc.edu
Department of Computer Science and National Center for Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

Introduction. The history of computing shows that each generation of systems has been partially or totally supplanted by systems that occupied a different point on the price/performance curve, expanding the base of possible owners and users. Mainframes and computer families like the IBM S/360 replaced "one of a kind" research systems and made computing part of the corporate culture. In turn, DEC's introduction of the minicomputer gave laboratory groups direct access to affordable computing. Workstations and PCs, driven by the emergence of powerful microprocessors, made computing broadly available to individual researchers and consumers.

The common theme across these computing generations has been a dramatic decrease in price, an associated increase in performance, and a consequent expansion of the number of units sold. This evolution suggests great opportunities for scaling up, building cost-effective petascale systems scientific and computing research, and scaling down, creating ubiquitous sensor networks for pervasive environments and environmental monitoring.

Petascale Computing. In high-performance computing, these economics have yielded commodity microprocessors, high density memory chips, high capacity disk systems, high performance graphics systems, and high-speed network interfaces, and they now make it economically possible to design and construct multi-teraflop computing systems. Coupled with explosive growth of Linux and open source software, commodity systems have transformed high-performance computing, with deployment of Linux PC clusters for scientific computing now commonplace. Game consoles, with price points below $300, performance rivaling or exceeding that of PCs, and graphics capabilities recently found only on high-end visualization supercomputers, are the vanguard of yet another computing generation – computing on toys.

In March 2001, IBM won the contract to build the processor for the PlayStation3 (PS3), with collaborative support from Toshiba, the designer of the PlayStation2 (PS2) Emotion Engine. Press reports suggest that the performance design point for the PS3 is 1000 times that of the PS2 and will rely on large-scale parallelism to realize its performance goals. The PS3 development, along with the Blue Gene and Blue Light projects at IBM Research, are intended to be the building blocks for petascale systems. Similar opportunities exist using commodity storage components. Terabyte data systems can be assembled for a few thousand dollars and petabyte systems for a few million.

With new price/performance points, designs based on such components could make terascale systems available for a few thousand dollars and petascale systems available for a few million dollars. The critical question is how we leverage such components to create usable and effective computing systems. *Arguably, the United States currently has no active large-scale computer architecture research and prototyping projects. Simply put, we face an architectural crisis, both in software and in hardware. Key challenges include design of hardware and software systems for the torrent of data (i.e., personal petabytes are very near), leveraging scalable systems-on-a-chip designs for realistic applications, and integrating scalable designs with distributed sensors, networks, and user interfaces.*

Ubiquitous Sensor Networks.[1] The confluence of high-speed wired and wireless networks, mobile devices, intelligent environments, high-resolution displays, smart spaces, microelectromechanical (MEMS) devices, and embedded real-time systems promises a new model for distributed interaction and information sharing. Imagine a homeland security environment that integrates these components, enabling researchers to query and control tools – seismographs, environmental monitors, cameras, and robots at remote sites. Combining real-time, remote access to data generated by those tools, along with video and audio feeds, large-scale computing facilities, data archives, high-performance networks, and structural models, researchers and disaster response teams would be able to improve the design of buildings, bridges, utilities, and other infrastructure in the United States.

The computing challenges inherent in developing a distributed sensor infrastructure are both formidable and exciting. As examples, new algorithms are needed for solution of complex problems on distributed systems. Data management and analysis techniques for multidisciplinary, multiple petabyte data collections will be required, along with commensurate computing capabilities and networks. Scalable visualization techniques are needed for feature extraction and display of complex multidisciplinary relationships.

Protocols, standards and interfaces among components will need to be defined and evaluated. New communication protocols, software, and interfaces must integrate high bandwidth wired and wireless communication networks, recognizing device heterogeneity, user needs and constraints, and location. New interaction metaphors and responsive environments are needed that adapt to the needs and preferences of individuals and collaborative groups. Simply put, seminal contributions from virtually all major computer science specializations will be necessary.

[1] Sid Karin, UC-San Diego, and I collaboratively developed many of these ideas and the associated text. I am indebted to him for these insights.

Daniel A. Reed
Edward William and Jane Marr Gutgsell Professor
Director, National Center for Supercomputing Applications (NCSA)
University of Illinois

Dan Reed serves as director of the National Computational Science Alliance (Alliance) and the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign. In this dual directorship role, Reed provides strategic direction and leadership to the Alliance and NCSA and is the principal investigator for the Alliance cooperative agreement with the National Science Foundation. He is one of two principal investigators and the Chief Architect for the recently announced NSF TeraGrid project to create a national infrastructure for Grid computing.

Reed was head of the UI computer science department from 1996 to 2001 and was an Alliance PI and co-leader of the Alliance's Enabling Technologies teams for three years. He is a member of several national collaborations, including the NSF Center for Grid Application Development Software, the Department of Energy (DOE) Scientific Discovery through Advanced Computing program, the DOE Accelerated Strategic Computing Initiative, and the Los Alamos Computer Science Institute. He also serves on the NERSC Policy Board for Lawrence Berkeley National Laboratory, is co-chair of the Grid Physics Network Advisory Committee and is a member of the board of directors of the Computing Research Association. In addition, he is a member of Illinois Governor Ryan's VentureTECH committee, which advises the governor on technology investment in Illinois.

Reed came to the University of Illinois in 1984 from the University of North Carolina at Chapel Hill. Before taking the helm of the UI computer science department, he was an assistant professor of computer science (1984-1988), a senior software engineer (1986-1993), associate professor of computer science (1988- 1991), professor at the UI's Beckman Institute, (1989-1990), and a professor of computer science (1991-present). His work in industry includes time as a visiting scientist at IBM's T.J. Watson Research Center, Yorktown Heights, NY.

Reed's areas of research include parallel computing, system performance and evaluation, and collaborative virtual environments for real-time performance analysis. In these areas, Reed has been involved with major research projects, such as the Pablo Performance Analysis Environment, a project to develop portable performance data capture and presentation tools for scalable parallel systems. Reed has also led jzation tool that dynamically adjusts resources to maximize performance. Other research projects have included Virtue, a real-time toolkit for collaborative, virtual reality presentations, and input/output analysis tools and libraries. Most recently he has led research in Smart Spaces and Pervasive Computing, an effort to explore the possibilities of ubiquitous wireless networking, embedded computing systems, high-resolution displays and virtual reality environments.

During his career, Reed has won numerous awards, including the NSF Presidential Young Investigator Award, the IBM Faculty Development Award, and the Xerox Senior Research Award. He holds one of four University of Illinois Edward William and Jane Marr Gutgsell professorships. Reed received his Ph.D. in computer science in 1983 from Purdue University. He holds an M.S. in computer science from Purdue and a B.S. in computer science from the University of Missouri at Rolla.