Background

Contact: Kevin Lane at (412) 848-8345 or at KLane85579@aol.com

Supercomputing in the Pittsburgh Region

Subject matter experts available for comment on this topic include:

David Moses
Executive Director
Pittsburgh Supercomputing Center

Supercomputing can be described as a search for ever faster, more powerful computing capabilities via groundbreaking advances in hardware, software, memory, data storage and networking equipment. But while the technical achievements are often unimaginable and awe-inspiring, it is the important applications that both demonstrate the value and dictate the growth of global advances in supercomputing.

Applications often made possible solely through supercomputing can include complex scientific calculations, visualization, simulation and modeling, data collection, processing and storage and biomedical discovery tools of an unprecedented range. Researchers nationwide use supercomputers for a range of projects that include AIDS research, astrophysics, fluid dynamics, weather prediction, and materials science.

Known biological projects include the folding of a relatively small protein molecule, three-dimensional explorations of the visible human nervous system and DNA research. Simulation capabilities address everything from earthquakes and hurricanes to colliding galaxies, atoms, icecaps or cars. Nearly every aspect of research and industry is affected by supercomputing capacity; and with it comes an always-increasing need for more.

(more)

Competing Globally

In late 2004, the U.S. reclaimed its position as the global leader of supercomputing resources, following a two-and-half-year period of Japanese dominance.

According to the Top 500, a biannual survey of global supercomputing rankings based on the Linpack Benchmark, the fastest computer in the world is currently the Roadrunner, a computer built by IBM for the Los Alamos National Laboratory. The Roadrunner features a capability of performing more than 1 quadrillion calculations per second, or one petaflop. A petaflop is one thousand times faster than a teraflop, and IBM's BlueGene/L at 360 teraflops previously was the fastest; it now ranks fourth.

The National Academies' National Research Council reported that a 1,000-fold increase in computing power is needed almost immediately and a one-million-fold increase ultimately will be required for applications, such as drug discovery, climate prediction and automobile collision simulations. To that end, the next milestone being pursued is the exaflop, which is 1,000 times faster than the petaflop.

High Performance Computing

The Council on Competitiveness is the nation's leading organization of CEOs, university presidents and labor leaders committed to promoting U.S. economic growth, success in global markets and raising the standard of living for all Americans, and it has made High Performance Computing (HPC) one of its top priorities. The Council on Competitiveness fully recognizes the importance of HPC and the need to make it accessible to private businesses.

The Council has an HPC systems initiative intended to stimulate and facilitate wider usage of HPC across the private sector to propel productivity, innovation and competitiveness. To do this, the Council has brought together a national brain trust of industrial HPC users to gain insights into how the private sector currently uses advanced computing capabilities.

Conventional wisdom is that the U.S. is a now service economy and no longer can be a leader in manufacturing. However, HPC is enabling a renaissance in advanced manufacturing where technology can be used for rapid prototyping to negate the labor cost advantages of other countries.

High-performance computing can help companies reduce costs by minimizing the need to build physical models, by allowing more thorough testing of designs before building and by creating the ability to develop more robust processes and higher quality products.

(more)

By performing more calculations per time unit and moving from design to production in a shorter timeframe, HPC can help speed time to market, which is a critical factor to compete successfully in the global market.

Pittsburgh is playing a dramatic role in HPC with the Pittsburgh Supercomputing Center (PSC).

PSC

In 1985, two Pittsburgh physics professors, Ralph Z. Roskies of the University of Pittsburgh and Michael J. Levine of Carnegie Mellon University, collaborated with Jim Kasdorf, then vice president of supercomputing at Westinghouse Electric Corporation, to develop the proposal that led to the creation of the Pittsburgh Supercomputing Center (PSC). Established in 1986, the PSC is a joint venture of Carnegie Mellon and the University of Pittsburgh, together with Westinghouse Electric Company.

Laboratories such as Oak Ridge, Lawrence Berkeley and Los Alamos housed early supercomputer facilities and were reserved almost exclusively for classified research. As one of the first public research supercomputing facilities, the PSC has become a leading edge site in the National Science Foundation's (NSF) TeraGrid programs, which provide U.S. academic researchers with support for and access to high-end computing infrastructure and research.

The PSC mission is to:

- enable solutions to important problems in science and engineering by providing leading-edge computational resources to the national community
- advance computational science, computational techniques and the national information infrastructure
- educate researchers in high performance techniques and their utility
- assist the private sector in exploiting high performance computing for their competitive advantage

Jim Kasdorf, a high-performance computing operations and hardware expert at Westinghouse, helped secure the original National Science Foundation funding which brought the PSC into existence. Today, the organization stands as one of the region's most successful experiments in collaboration.

With computer room facilities housed at Westinghouse, the PSC is administered from a building on South Craig Street in Oakland owned by Carnegie Mellon. Approximately 125 staff members serve the organization, and Roskies continues to serve as scientific director.

(more)

Funding

Currently, there are four supercomputing centers funded by the NSF, which include the University of Illinois at Champaign-Urbana, the University of California at San Diego, Cornell University and the PSC.

In addition to NSF funding, the PSC also receives funding from the U.S. Department of Energy, the National Institutes of Health, the Commonwealth of Pennsylvania and private industry. Over its nearly 25-year history, the PSC has received nearly $30 million from the state. Using the state's funding as leverage, the Center has received more than $378 million from the federal government and industry grants.

In 2004, the PSC and the University of Pittsburgh received a research grant of $900,000 from IBM for a three-year regional research project to develop a software tool, called the Standardized User Monitoring Suite, or SUMS. The software will quantify and analyze the programming time required for next-generation supercomputing. IBM received $53 million from DARPA as one of three contractors pursuing the research. Rami Melhem, chairman of Pitt's computer science department, directed the local effort.

The NSF awarded a five-year grant ending in 2010 totaling $52 million to support the PSC as a leading partner in the TeraGrid, NSF's program to provide national cyberinfrastructure for education and research. Built over the last four years, the TeraGrid is the world's largest, most comprehensive distributed cyberinfrastructure for open scientific research. The PSC also had received about $5 million to continue its role in user support and security.

Much as physical infrastructure, such as power grids, telephone lines and water systems enables modern life, cyberinfrastructure makes possible much of modern scientific research. Through high-performance network connections, the TeraGrid integrates high-performance computers, data resources and tools and high-end experimental facilities at eight partner sites around the country.

State-of-the-Art in Pittsburgh

Among the three remaining NSF-funded supercomputing centers, Pittsburgh maintains a reputation for providing the "big iron," the largest and most powerful systems, along with particular expertise in maximizing the productivity of these systems. Pennsylvania researchers routinely use upwards of seven million hours of processor usage, approximately 30 percent of the time on PSC's four major computing platforms.

The $52 million to support the PSC operations is in addition to $9.7 million NSF awarded in 2004 to help PSC obtain its newest, most powerful system, a 10-teraflop Cray Inc. XT3 nicknamed Big Ben. Based on the Cray Research "Red Storm" architecture, Pittsburgh has reasserted its reputation

(more)

as a premier resource in high-end supercomputing. The system was awarded the 2004 Reader's Choice Award for Most Innovative Implementation by *HPCWire*.

Late in 2006, PSC more than doubled the capability of Big Ben, to 21.5 teraflops. At 21.5 trillion calculations per second, this increase improves the ability of U.S. scientists and engineers to address the most large-scale, demanding computational science projects.

The PSC replaced the AMD Opteron, 2.4 GHz processors of the 2,090 processor Big Ben system with Opteron's top-end dual-core (2.6 Ghz) chip, doubling the processor count to 4,180, with a corresponding boost in peak performance, while also doubling memory (from two to four terabytes).

Big Ben became operational in July of 2005. More than sheer processor speed, Big Ben's primary technological advance has been its superior inter-processor bandwidth, the speed at which processors share information. This is a large advantage for projects that demand hundreds or thousands of processors working together. Over the past year, because of this capability, Big Ben has demonstrated performance as much as 10 times better than prior tightly-coupled systems on a number of applications. Because of this capability also, Big Ben has proven itself to be a champion at "scaling," which is the ability to use a large quantity of processors without seriously reducing the per-processor performance.

To grasp the scale of performance of Big Ben, if every one of the 6.5 billion people on earth held a calculator and did one calculation per second, they would all together still be 3,000 times slower than the upgraded Cray XT3 system. One example of the type of problem that can now be examined with the expanded computing power is global warming, which defies experimentation in a lab. Now people can describe these processes in mathematical language and then put all those equations in a computer program.

Because of its exceptional inter-processor bandwidth, Big Ben already has demonstrated nearly 13 times better performance than LeMieux, PSC's six teraflop HP-Alpha processor terascale system, on key applications when 1,000 or more processors are used.

The system is based on a similar machine built for the Sandia National Laboratory. The Sandia system will be capable of up to 41.5 teraflops. The overall Red Storm architecture delivers superior scalable application performance and value across a range of configurations from 200 to 30,000 processors, with peak performance of up to 144 teraflops.

LeMieux has been operational at the PSC since October 2001. Upon introduction, it was ranked as the second-fastest supercomputer in the world, at six teraflops.

(more)

Many PSC systems benchmarked important firsts in the field of supercomputing. Jaromir was a 512-processor SGI Cray T3E 900, featuring peak-performance rating of 460 billion floating-point operations per second (flops).

The PSC pursued MPP in 1993 with the first Cray T3D installed anywhere in the world. The $15 million machine featured 540 processors. It is now decommissioned along with Mario, the first nongovernmental Cray C90 installed in the United States. Purchased for $35 million in 1992, the vector-processing machine used 16 high-speed processors, each arrayed on 70-pound circuit boards.

Other notable hardware at the PSC includes Rachel and Jonas. Named after Rachel Carson and Jonas Salk, these computers are two twin Hewlett-Packard GS 1280 Alphaserver shared memory systems, each with 128 processors.

At the time of its founding, The PSC acquired a Cray X-MP with NSF support, which was set up at Westinghouse under Kasdorf's supervision.

NIH Grant for Biomedical Supercomputing

During 2006, the PSC received $8.5 million from the National Institutes of Health to renew its program in biomedical supercomputing. Through this program, the National Resource for Biomedical Supercomputing (NRBSC), PSC scientists pursue research in the life sciences and foster exchange nationwide among experts in computational science and biomedicine. The renewal award supports NRBSC's research in three core areas: spatially realistic cellular modeling, large-scale volumetric visualization and analysis and computational structural biology.

Established in 1987, the PSC's biomedical supercomputing program, renamed NRBSC, was the first such program in the country external to NIH. Along with core research, NRBSC develops collaborations with biomedical researchers at many centers around the country and provides computational resources, outreach and training. The current award from NIH's National Center for Research Resources renews NRBSC through 2010.

The TeraGrid

High-performance computers often work with large data sets, and often the data and the processing power are not in the same location. To bring together supercomputing resources with users and data sets across the country, the NSF has created the TeraGrid, a trans-continental high-performance network.

TeraGrid's unified user support infrastructure and software environment allow users to access storage and information resources, as well as more than a dozen major computing systems via a single allocation, either as stand-alone resources or as components of a distributed application using

(more)

Grid software capabilities. The multi-year effort builds and deploys the world's largest, most comprehensive distributed infrastructure for open scientific research.

The PSC involvement with the TeraGrid began in October 2002, when it announced that the TeraGrid has entered full production mode, providing a coordinated set of services for the nation's science and engineering community.

Through TeraGrid, more than 100 teraflops of computing power are available to scientists across the country. The TeraGrid also offers storage, visualization, database and data collection capabilities. Hardware at multiple sites across the country is networked through a 30-gigabit per second (gbps) backplane, the fastest research network on the planet.

Other sites connected to the TeraGrid include:

- Indiana University
- the National Center for Supercomputing Applications at the University of Illinois Urbana-Champaign
- the National Center for Atmospheric Research, Boulder, Colorado
- Oak Ridge National Laboratory
- Purdue University
- the San Diego Supercomputer Center at the University of California, San Diego
- the Texas Advanced Computing Center at The University of Texas at Austin.
- the University of Chicago/Argonne National Laboratory

Advanced Networking

In 2004, the PSC staff members demonstrated that real-world data transmission at 40 gbps is now attainable over a single light wave, or lambda. The link was established with two next-generation Cisco CRS-1 routing systems, fitted with OC-768 interfaces. One OC-768 will support the same bandwidth as four OC-192s, the current standard.

The PSC and the University of Pittsburgh share membership and a seat on the board of the National LambdaRail (NLR), a national network infrastructure supporting experimental and production networks for the U.S. research community. The consortium joins leading U.S. universities and companies in deploying an advanced, nationwide fiber-optic infrastructure to encourage next-generation applications in science, engineering and medicine. Through NLR, many different networks will exist side-by-side in the same fiber-optic cable, but will be independent of each other, each supported by its own lightwave or lambda.

(more)

The PSC's Advanced Networking Group conducts research on network performance and analysis in support of high performance computing applications. The group also develops software to support distributed supercomputing applications and to implement high-speed interfaces to archival and mass storage systems. Research is focused on such areas as TCP implementations, tools to tune TCP for better performance and software to monitor and improve network performance.

Work addresses the maximization of usable network bandwidth for any office or research computing system. National projects such as the Net100 and Web100 seek to improve "real" network performance for each network host and to provide tools to diagnose problems between the host and the network that might limit the host's available bandwidth.

Partnerships

From its inception, PSC has fostered a spirit of collaboration throughout the Pittsburgh Region.

In August 1999, the Pittsburgh Supercomputing Center joined with the Department of Energy's National Energy Technology Laboratory, Carnegie Mellon University, West Virginia University and the West Virginia Governor's Office of Technology, in creating the Supercomputing Science Consortium. Known simply as $(SC)^2$, the regional partnership acts to advance energy and environment technologies through the application of high performance computing and communications. Since its establishment, the University of Pittsburgh, Duquesne University, Waynesburg College, the Institute for Scientific Research and the NASA Independent Verification and Validation facility also have joined the partnership.

Research by the life sciences community is supported by the PSC's high-performance computing resource as part of the NRBSC. The internal users group has operated at the PSC offices for nearly 20 years, and it is funded primarily through the National Institutes of Health, instead of by the PSC's core NSF grant.

Other users in the region include PPG and the Bettis Atomic Power Laboratory.

The reach of the PSC now extends throughout the Pittsburgh region in the form of a network called the Three River's Optical Exchange (3ROX). Administered by PSC's Advanced Networking Group, 3ROX provides high-bandwidth Internet access to area educational institutions, including the University of Pittsburgh, The Pennsylvania State University, Carnegie Mellon University, West Virginia University and several Pittsburgh area public schools.

(more)

Supercomputing Conferences

The PSC won the award for "Best Demonstration at TG08" during the annual conference of the TeraGrid, the National Science Foundation's program of cyberinfrastructure for U.S. science and education. A PSC team of two scientists and a University of Pittsburgh student received the award for "WiiMD," an innovative project that merges the video-game technology of the Nintendo Wii with interactive supercomputing.

Pittsburgh hosted a previous supercomputing conference, the SC2004, which drew more than 7,000 attendees interested in high-performance computing.

An annual feature of the conference includes SC Global, an Access Grid-enabled component that provides remote participation in conference events. This year marked the first ever demonstration of a simultaneous connection of AG nodes on all six inhabited continents.

Organizers of the event also assembled the largest data storage capacity on the planet. StorCloud provided an unprecedented one petabyte of memory, storage for more than 1,000 trillion bytes of computer information or the equivalent of 100 times the contents of the Library of Congress, the largest library in the world. Thirty-two tons of equipment worth about $80 million and donated by 22 vendors were required to create StorCloud on the exhibition hall floor. Some 300 kilowatts of power were consumed by StorCloud, making it the warmest place in the building.

Accessing Supercomputing

Users may apply to use the PSC's other supercomputing resources through the National Science Foundation's TeraGrid program, through the Corporate Affiliates program and through biomedical or starter grants. Any academic researcher is eligible to use the PSC facility under the NSF funding, and non-classified corporate research also is supported for a fee.

The PSC Corporate Affiliates Program is designed to bring resources to bear on helping businesses solve their most challenging information-processing and research problems. The program combines training, consulting and access to high-performance systems to meet the needs of each participant. Each affiliate relationship is uniquely designed to meet the needs of the corporate partner.

The technology is useful to anybody who needs to visualize large volumes of data in a three-dimensional space — perhaps a jet engine, skyscraper or human heart — with potential customers including engineering firms and medical laboratories. The PSC staff actively participates in research, including co-authoring papers, and mobilization of the staff's expertise is an added benefit that Pittsburgh region researchers enjoy.

(more)

Research suitable for supercomputing resources often shares a common chicken-or-egg complexity: a researcher needs to know enough to create a model, but the problem is complicated enough that he or she cannot see their way through it, short of creating the model. In general, however, Roskies notes that proposals that justify that "the science is worthy" are sufficient for PSC work.

Visit: www.psc.edu

www.nlr.net

www.3rox.net

###

Backgrounders in this series featuring technology centers of excellence in the Pittsburgh region include:

Cybersecurity
Data Storage
Electro-Optics
Energy Technology
Entertainment Technology
Micro-electromechanical Systems
Nanotechnology
Robotics
Specialty Metals
Supercomputing
System-on-a-Chip
Tissue Engineering