Supercomputing's Exaflop Target

The twin challenges of parallelism and energy consumption are enlivening supercomputers' progress.

ASIA HAS COME out swinging at the top of June 2011's Top500 list, which rates the world's fastest computers based on the LIN-PACK benchmark. Leading the list is the K Computer, which achieved 8.2 quadrillion floating-point operations per second (petaflops) to give Japan its first appearance in the much-coveted number-one position since November 2004. It knocks China's Tianhe-1A, at 2.6 petaflops, to second place. The U.S.'s Jaguar (1.75 petaflops) was pushed from second to third place. China's Nebulae (1.27 petaflops) dropped from third to fourth, and Japan's Tsubame 2.0 (1.19 petaflops) slipped from fourth to the fifth position.

Asia's success comes at a time when national governments are reconsidering the value of these ratings—and what defines supercomputing success. The U.S. President's Council of Advisors on Science and Technology (PCAST) recently warned that a focus on such rankings could "crowd out … fundamental research in computer science and engineering." Even Jack Dongarra, the list's founder and a computer science professor at the University of Tennessee, believes the rankings need to be seen in a larger context. "You can't judge a car solely by how many RPMs its engine can do," he says, pointing to the Graph500 list and HPC Challenge Benchmark as other sources of comparative supercomputing data.

Computer scientists are also skeptical about the value of such rankings. Petaflops are not the same as useful work; practical applications need to both preserve and take advantage of their power. Xuebin Chi, director of the Supercomputing Center at the Chinese Academy of Sciences, points out that "programming [for the CPU/GPU

The Top500's leading supercomputer, the K Computer, is more powerful than its five closest competitors combined.

combinations now popular in super-computing] is more difficult than for commodity CPUs," and that "changing from sequential to parallel code is not easy." He predicts that such a transition would take three to five years. But even with superb programming, real-world aspects of data delivery and error correction could significantly reduce application speeds from those reported by the Top500.

Regardless, the November 2010 Top500 list's release, with China's Tianhe-1A in first place, spurred political discussions about national commitment to high performance computing (HPC) throughout the world. In the U.S., 12 senators cited Tianhe-1A in a letter to President Obama warning that "the race is on" to develop super-computers capable of 1,000 petaflops

(1 exaflop). In asking for funding, they wrote that "Our global competitors in Asia and Europe are already at work on exascale computing technology … we cannot afford to risk our leadership position in computational sciences."

David Kahaner, founding director of the Asian Technology Information Program, believes that Tianhe-1A is the leading edge of a Chinese push to not only increase supercomputing speeds, but also domesticate production. "It represents a real commitment from the Chinese government to develop supercomputing and the infrastructure to support it," he says. "A Chinese domestic HPC ecosystem is evolving. Domestic components are being developed and incorporated, and their use is likely to increase." Dongarra agrees, noting, "The rate at which they're do-

"If you peek a little bit further into graphics problems, they look a lot like supercomputing problems," says Sumit Gupta. "Modeling graphics is the same as modeling molecule movement in a chemical process."

ing that is something we've not seen before with other countries."

CPUs vs. CPU/GPU Hybrids

If the June Top500 list had been a foot race, the K Computer would have lapped the competition. At 8.2 petaflops, it wields more power than the next five supercomputers combined. The K Computer's name alludes to the Japanese word "Kei" for 10 quadrillions, and represents the researchers' desired performance goal of 10 petaflops.

Aside from national aspirations, the Top500 list reveals technical trends in HPC research worldwide. Most notable is an increased use of general-purpose graphics processing units (GPUs) in a hybrid configuration with CPUs. The June 2011 list includes 19 supercomputers that use GPU technology; the June 2010 list contained just 10.

GPUs contain many more cores than CPUs, allowing them to perform a larger number of calculations in parallel. While originally used for graphics tasks, such as rendering every pixel in an image, GPUs are increasingly applied to a wide variety of data-intensive calculations. "If you peek a little bit further into graphics problems, they look a lot like supercomputing problems," says Sumit Gupta, manager of Tesla Products at NVIDIA. "Modeling graphics is the same as modeling molecule movement in a chemical process."

In today's supercomputers, GPUs provide the brute calculation power, but rely heavily on CPUs for other tasks. For example, the number-two Tianhe-1A contains two six-core Intel Xeon X5670 CPUs for each 448-core Tesla M2050 GPU (14,336 to 7,168); it also contains a much smaller number of eight-core Chinese-built Feiteng CPUs (2,048). Altogether, GPUs in Tianhe-1A contribute approximately three million cores—30 times as many as are in its CPUs.

But speed is not simply a matter of throwing more cores into the mix, as it is not easy to extract all of their processing power. First, data must be queued and managed to feed them—and to put the results together when they come out. "You need the CPU to drive the GPU," Chi explains. "If your problem can't be fit into the GPU itself, data will need to move frequently between the two, hurting performance." Dongarra correlates this, saying, "The speed of moving data to the GPU and the speed of computing it once it's there are so mismatched that the GPU must do many computations with it before you see benefits."

The K Computer, though, bucks this trend. Unlike Tianhe-1A and other recent large supercomputers, it does not utilize GPUs or accelerators. The K Computer uses 68,554 SPARC64 VIIIfx CPUs, each with eight cores, for a total of 548,352 cores. And the Japanese engineers plan to boost the K Computer's power by increasing the number of its circuit board-filled cabinets from 672 to 800 in the near future.

Asia's Ascension

Asian researchers appear to be well positioned to exploit GPUs for massively

Top500 List, June 2011.

Rank	Computer	Nation
1	K Computer	Japan
2	Tianhe-1A	China
3	Jaguar	U.S.
4	Nebulae	China
5	TSUBAME 2.0	Japan
6	Cielo	U.S.
7	Pleiades	U.S.
8	Hopper	U.S.
9	Tera-100	France
10	Roadrunner	U.S.

Source: http://www.top500.org

ACM-ICPC World Finals

Students from Zhejiang University won the 2011 ACM International Collegiate Programming Contest World Finals in Orlando, FL, in May, beating student teams from 104 universities in the IBM-sponsored competition.

En route to its championship, Zhejiang University students used IBM's large-scale analytics and cloud computing technologies to solve eight out of 11 programming problems in five hours.

The only other team to solve eight problems, the University of Michigan, won a gold medal for finishing second.

The University of Michigan team's success was due to several factors, says coach Kevin Compton, associate professor in the school's electrical engineering and computer science department. "First, they are very talented programmers and problem solvers," he says. "A couple of them began entering programming competitions in high school." In addition, from the time the regional contests were held in October 2010 and the World Finals, the Michigan team practiced intensely.

The Michigan team also created an efficient strategy for solving problems, with team members having clearly defined roles, Compton says. "Each member had a fairly specialized role, either as a coder or debugger and tester," he says. "By following this strategy, they were able to avoid getting bogged down on a particular problem."

Tsinghua University and St. Petersburg University finished in third and fourth place, respectively, winning gold medals. The other top 12 finalists, each of which received a silver medal, were Nizhny Novgorod State University (5th), Saratov State University (6th), Friedrich-Alexander-University Erlangen-Nuremberg (7th), Donetsk National University (8th), Jagiellonian University in Krakow (9th), Moscow State University (10th), Ural State University (11th), and University of Waterloo (12th).
—*Bob Violino*

parallel supercomputing. Kahaner believes China's relative isolation from Western influences may have led to economics that favor such innovations. "They're not so tightly connected with U.S. vendors who have their own perception of things," he says. "Potential bang for the buck is very strong in Asia, especially in places like China or India, which are very price-sensitive markets. If your applications work effectively on those kinds of accelerator technologies, they can be very cost effective."

According to Satoshi Matsuoka, director of the Computing Infrastructure Division at the Global Scientific Information and Computing Center of the Tokyo Institute of Technology, China's comparatively recent entry into HPC may help them in this regard. "Six years ago, they were nowhere, almost at zero," he says. "They've had less legacy to deal with." By contrast, Gupta says, programmers in more experienced countries have to undergo re-education. "Young programmers have been tainted into thinking sequentially," he notes. "Now that parallel programming is becoming popular, everybody is having to retrain themselves."

These issues will only get more complicated as time progresses. Horst Simon, deputy laboratory director of Lawrence Berkeley National Laboratory, says a high level of parallelism is necessary to progress past the 3GHz–4GHz physical limit on individual processors. "The typical one-petaflop system of today has maybe 100,000 to 200,000 cores," says Simon. "We can't get those cores to go faster, so we'd have to get a thousand times as many cores to get to an exaflop system. We're talking about 100 million to a billion cores. That will require some very significant conceptual changes in how we think about applications and programming."

Matters of Energy
Hybrid architectures have historically had another advantage besides their parallelism. They have also usually used less energy than comparable CPU-only systems. In the November 2010 list, hybrid systems generally delivered flops more efficiently than the CPU-only systems.

But the new Top500 list shows that the architectural battle over energy efficiency is still raging. The CPU-based

---

In today's hybrid CPU/GPU supercomputers, GPUs provide the brute calculation power, but rely heavily on CPUs for other tasks.

K Computer attains an impressive 825 megaflops (Mflops) per watt even as the third-place, CPU-based Jaguar ekes out a so-so 250 Mflops/watt. By comparison, the hybrid Tianhe-1A achieves 640 Mflops/watt, Nebulae gets about 490 Mflops/watt, and Tsubame 2.0 gets 850 Mflops/watt. (The list's average is 248 Mflops/watt.)

The most energy-efficient system is the U.S.'s CPU-based IBM BlueGene/Q Prototype supercomputer, which entered the Top500 in 109th place, with an efficiency of 1,680 Mflops/watt. The IBM BlueGene/Q tops the Green500, a list derived from the Top500 that ranks supercomputers based on energy efficiency. But despite BlueGene/Q's supremacy, eight of the Green500's top 10 are GPU-accelerated machines.

Energy is no small matter. The K Computer consumes enough energy to power nearly 10,000 homes, and costs $10 million a year to operate. These costs would significantly increase in an exaflop world, notes Simon.

Looking Ahead
Despite the headlines and U.S. senators' statements, Dongarra and colleagues are quick to dismiss the supercomputing competition as a "race." At the same time, he expects to see an increase in Top500 scores, and notes that several projects are aiming for the 10-petaflop target, which could be realized by the end of 2012. But the real prize is the exaflop, which the U.S. government, among others, hopes to achieve by 2020.

Matsuoka believes this goal is possible, but it will be "a very difficult target,"

---

especially when compared with traditional expectations. "Look at Moore's law," he says. "Computers will get about 100 times faster in 10 years. But going from petascale to exascale in 10 years is a multiple of a thousand." Having said that, he notes that it is been done before—twice. "We went from gigaflops in 1990 to teraflops in about 10 years, and then to petaflops in another 10 years. Extrapolating from this, we could go to exascale in the next 10 years."

But Dongarra warns that we won't reach that stage solely by focusing on hardware. "We need to ensure that the ecosystem has some balance in it. Major changes in the hardware will require major changes in the algorithms and software," he says. "We're looking at machines in the next few years that could potentially have billions of operations at once. How do we exploit billion-way parallelism?"

The payoffs could be enormous. Supercomputing is already widely used in fields as diverse as weather modeling, financial predictions, animation, fluid dynamics, and data searches. Each of these fields embodies several applications. By way of example, Matsuoka says, "You can't do genomics without very large supercomputers. Because of genomics, we have new drugs, ways of diagnosing disease, and crime investigation techniques." While exaflop computers will spawn now-unimagined uses, any current increases in speed as we race toward that goal will greatly benefit many existing applications. ⓒ

Further Reading

Anderson, M.
Better benchmarking for supercomputers,
IEEE Spectrum 48, 1, Jan. 2011.

Endo, T., Nukada, A., Matsuoka, S., and Maruyama, N.
Linpack evaluation on a supercomputer with heterogeneous accelerators, 2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS), Atlanta, GA, April 19–23, 2010.

Miller, C.
Most popular supercomputing videos, July 13, 2010. http://www.datacenterknowledge. com/most-popular-supercomputing-videos/

Top500 list
http://www.top500.org

Tom Geller is an Oberlin, OH-based science, technology, and business writer.