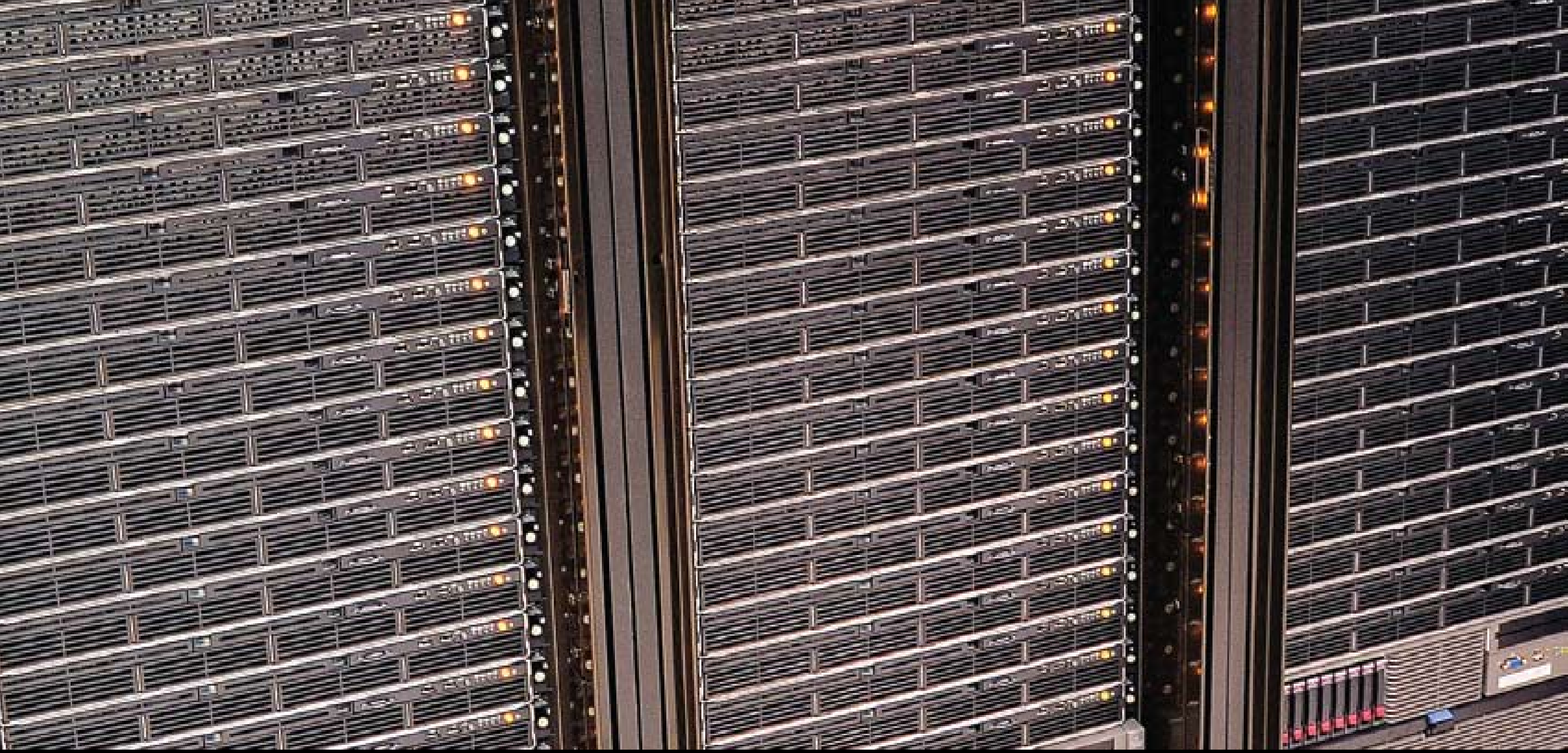


Cluster Computing Helps Researchers Thrive

AMD and HP help Purdue University build a new HPC center

Challenge:

- Support Purdue University researchers with a cost-effective supercomputing solution for demanding simulations and data processing across a range of academic disciplines.
- Provide supplemental High Performance Computing (HPC) support for researchers with the National Science Foundation's TeraGrid.

Solution:

- Purdue's Coates cluster – currently the world's largest entirely academic 10 Gigabit Ethernet cluster with support for a maximum of 1,280 compute nodes, 10,000 processing cores, and an estimated peak performance of 90 teraflops.
- Comprised of HP ProLiant DL165 rack servers featuring AMD Opteron™ processors.

Impact:

- Faculty from a variety of research groups are able to run computations for gene sequencing analysis, nanotechnology circuit design, biological study of virus structures, molecular dynamics in chemistry, and structure simulations for civil engineering.
- Stable and reliable cluster stands up to the heavy demands of faculty, including running jobs that can use dozens of nodes and run for up to thirty days.
- The new cluster is expected to rank in the top 50 on the November 2009 TOP500 Supercomputers List.

AMD
The future is fusion

> "AMD and HP were able to provide Purdue with an extremely competitive solution on this cluster."
>
> —John Campbell, Associate Vice President, Rosen Center for Advanced Computing, Purdue University

Purdue University (West Lafayette, IN) has a long-standing reputation for offering one of the finest engineering educations in the country, with its undergraduate program currently ranked #9 nationally, and graduate program #12. To help drive its success in engineering and other scientific fields, Purdue has relied on large-scale research computing – going back to the monolithic control data systems of the late-1950s.

"In the early days it was with large, centrally located machines that had to be centrally funded because they were beyond the price-range that any individual researcher or even department could afford," says Bill Whitson, Director of Research Support at Purdue's Rosen Center for Advanced Computing.

By the start of the 21st Century, these types of large-scale research systems had evolved into high-performance supercomputers, designed with proprietary architectures and a premium price-tag. "Those machines typically cost 3-, 5-, 10-million dollars. So they were beyond an individual research group's capacity to fund," says Whitson.

As the cost and capabilities of traditional supercomputing systems escalated over the last decade, academic researchers were also finding a rapidly expanding need for time on these types of machines to crunch massive loads of data.

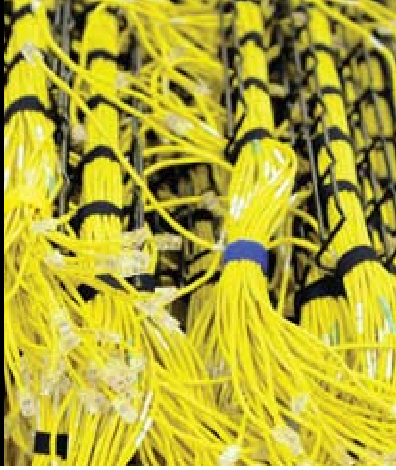
"High-performance computing systems are important because research for discovery at universities has shifted to being more computationally intensive," says John Campbell, Associate Vice President of the Rosen Center. "We're running more simulations; we're running more models than we ever have."

Cooperative Purchasing

Purdue's IT gurus found a solution for their faculty's increasing HPC needs through a combination of best-of-breed technologies and cooperative purchasing across multiple academic disciplines. "Purdue, as well as a number of other institutions, has been working on the concept of a 'community cluster'," says Campbell. "The idea behind it is to pull together a number of faculty to make a larger cluster happen.

"We've been able to come up with a perfect combination of support and services from the central IT organization combined with faculty input. The central IT group provides the racks, provides the facility, and the related power and cooling. It provides the system administration to keep the system going, and a base-level of storage," says Campbell. In turn, participating faculty across multiple disciplines at Purdue share the cost of purchasing the hardware.

"The machines have gotten more and more capable over time, so people are able to increase the size of the problems they're dealing with, and increase the size of the data sets they're dealing with. There has been a significant change in the way people approach computational research just in the past 3 to 5 years," says Whitson.



Bang for the Buck

"The commoditization of fast computers has really helped a smaller institution like us achieve scales that we'd previously only see at a national lab," says Mike Shuey, HPC Technical Architect at Purdue's Rosen Center.

"When we go to procure a new cluster, we basically define a set of requirements based on our faculty needs and then open it up to a number of vendors to see what kind of proposals we can solicit."

In 2009, Purdue looked to add another cluster system. "This time we were very focused on the price for performance. Hewlett-Packard and AMD collaborated together to give us a very attractive solution for this cluster," says Shuey.

The design chosen for Purdue's new "Coates" community cluster is based on HP ProLiant DL165 rack servers and AMD Opteron™ processors. The DL165 is a high performance, low cost, ultra-dense rack server node for HPC environments, web serving and cost-conscious server deployments. The ultra dense 1U rack mount server features advanced processing and memory features for premium performance, and is designed to be cost competitive to enable multi-node configurations.

The AMD Opteron™ processor is also suited for HPC environments with performance features such as Direct Connect Architecture as well as power efficiency features such as AMD PowerNow!™ and AMD CoolCore™ technologies. The high-bandwidth memory controller combined with an optimized cache structure results in superior performance and throughput for memory-intensive applications. At its maximum capacity, the new Coates cluster will include more than 1,200 compute nodes and more than 10,000 processing cores, integrated with 10 Gigabit Ethernet.

"All of those machines and adaptors are tied together with a fairly large 10 Gigabit network. It gives us relatively good bandwidth across the whole cluster and integrates well with the rest of our campus networking," says Shuey. The Coates cluster is currently the world's largest entirely academic 10 Gigabit Ethernet cluster with an estimated peak performance of 90 teraflops. It's expected to rank in the top 50 supercomputers worldwide in November 2009.

"When you talk about science getting done, a lot of that depends on being able to use resources over hundreds of nodes if needed," explains Campbell. "The key behind the community cluster program is to provide a platform which faculty members buy into, that is cost effective and supported by the staff. Faculty members don't have to worry about this hardware."

Stable and Able

Since it was installed July 21, 2009, researchers have placed a near-constant demand on the Coates cluster, with maximal system utilization. "We frequently have people running jobs that are using dozens of nodes and run for thirty days. So for them to be successful, we have to have equipment on the floor that stays up, and stays running for months. Stability and reliability are extremely important for us," says Whitson.

Current projects being run on Coates include gene sequencing analysis, nanotechnology circuit design, biological study of virus structures, molecular dynamics in chemistry, and structure simulations for civil engineering.

Purdue Professor Matthew Huber plans to use Coates to model climate change in ways that could allow him to predict, for example, how global warming will play into the future spread of mosquito-borne diseases like West Nile virus and severe weather events such as hurricanes, among other things.

Huber's modeling frequently spans 1,000 years and can take three years of computer time to run. Coates should allow his models to be run repeatedly, solidifying the predictions, and at higher resolutions, taking the picture from a global scale to the regional and local scales where people actually live with the effects of climate change. Coates also should allow longer models making predictions farther into the future.

"Something that takes me a month today could take me as little as a week," says Huber an associate professor of earth and atmospheric sciences and member of the Purdue Climate Change Research Center. "If you actually want to make climate change predictions that are relevant to human scales, to agriculture, to flooding, to things that people really care about, then you have to get to much finer scales. And if you can run up to four times faster, you can have double or more the resolution and then you can start getting the global-scale models down to scales people actually care about."

A growing trend across Purdue's HPC systems has been taking on projects with ever-increasing quantities of data to process. "Some of our newer researchers are dealing with huge amounts of data, either because they're generating it in simulations or they're collecting it on instruments in their lab," says Whitson. "So we're having to deal with much higher volumes of data than we have in the past. Which means we need clusters and systems that can move data faster, that have better memory bandwidth to the processor."

The unique demands of a community cluster research model requires that hardware vendors provide components that can be counted on for reliability and stability. "Ultimately we have to have a collection of vendors that can provide us with quality equipment. Things that just work, that can maintain quality over the lifespan of the equipment," says Campbell.

"We don't want to be replacing parts all the time. We want it to withstand a regular beating and we keep our machines quite busy. Typically we're at 90-95 percent utilization with these clusters. So they have to stand the test of time with that type of usage."

Partnership Wins the Day

The spirit of partnership helped make the Coates cluster an instant success – from vendors like AMD and HP, to Purdue's IT staff, to the faculty who use the system. Whitson explains, "We don't think of the faculty as our customers. They aren't buying compute time from us. They're buying the equipment and we're working with them to operate it and make it available for them to do their research."

"So we're partnering with the faculty. The faculty are partnering with each other because they're buying in to a shared community resource. And this model has worked out extremely well here over the past several years," says Whitson.

"Besides the cost savings from making a group purchase, the faculty can borrow computing cycles from other faculty when the other clusters are idle," says Campbell. "This gives the researchers more flexibility, and we also have unused computing cycles we can offer to the National Science Foundation's TeraGrid." TeraGrid is an open scientific discovery infrastructure combining high-performance computers, data resources and tools, and high-end experimental facilities at eleven partner sites, Purdue among them, to create an integrated, persistent computational resource.

"In the last 12 months we've had people from more than 100 different research groups across the country using the Purdue systems," says Whitson. "The environment that we provide to TeraGrid makes it very convenient to run hundreds of thousands of jobs, most of which will be relatively short. But it really increases the workflow for many of the researchers who get TeraGrid allocations."

With the success of Coates and other HPC cluster centers, Purdue will continue to expand capabilities with additional systems and explore emerging technologies such as virtualization. "We are anticipating installing a cluster of roughly similar size every year," says Campbell. "So we've been working with our faculty members to understand their needs to develop a program they can count on.

"As the processors evolve and our facilities become a more constrained, energy efficiency is going to become more and more prominent factor in the decision process" says Campbell. "And as machines become available and the price point is attractive, we'll be moving from 8-core to 12-core to 16-core and just following that trend and developing the best solution we can for the faculty."

About AMD

Advanced Micro Devices (NYSE: AMD) is an innovative technology company dedicated to collaborating with customers and technology partners to ignite the next generation of computing and graphics solutions at work, home and play.

For more information, visit http://www.amd.com

AMD
The future is fusion