Exponential Improvement in Center's Computational Research with HP Servers

The 464-node "Titan" cluster at Columbia University uses HP BladeSystem to boost performance by 10x and reduce annual power costs by $95,000

"The analysis that consisted of 20,000 single processor jobs—the one that took the previous cluster 50 days to complete—now gets done in just five and a half days with the HP BladeSystem."

John Wofford, director of IT, Center for Computational Biology and Bioinformatics, Columbia University

HP customer case study: High-performance computing

Power and cooling

HP ProLiant BL2x220c server blades

HP ProCurve switches

Industry: education and research

Objective

Deploy a single, large, general-purpose cluster while minimizing space, power and cooling requirements

Approach

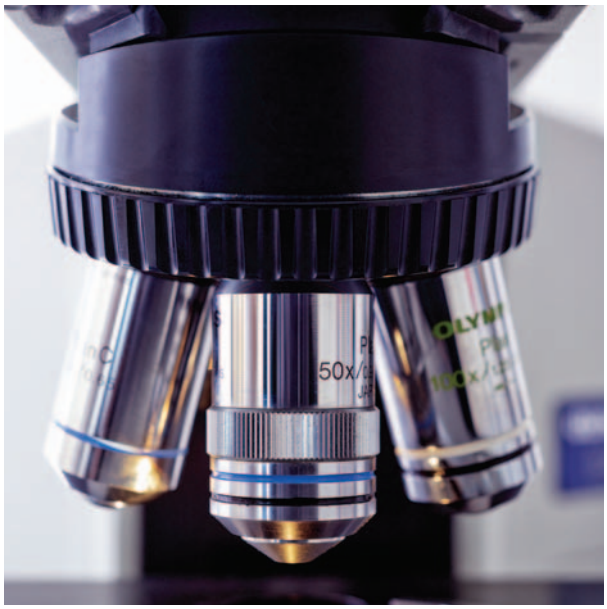
Standardize on HP ProLiant BladeSystem server blades, HP ProCurve switches, and HP Integrated Lights-Out (iLO2) for remote management

IT improvements

- Hundreds of IT administrator hours saved on basic server management tasks
- 10x performance improvement over next-fastest system
- 15 to 20% performance improvement in jobs
- 100x faster server updates
- 50% reduction in administrator visits to physical clusters
- 40% reduction in power consumption

Business outcomes

- $95,000 USD annual savings in power and cooling costs
- $110,000 USD saved in rack and chiller costs compared to other vendor solutions
- Expense of two full-time hires avoided using HP iLO2
- Enhanced ability to attract and retain top researchers and projects



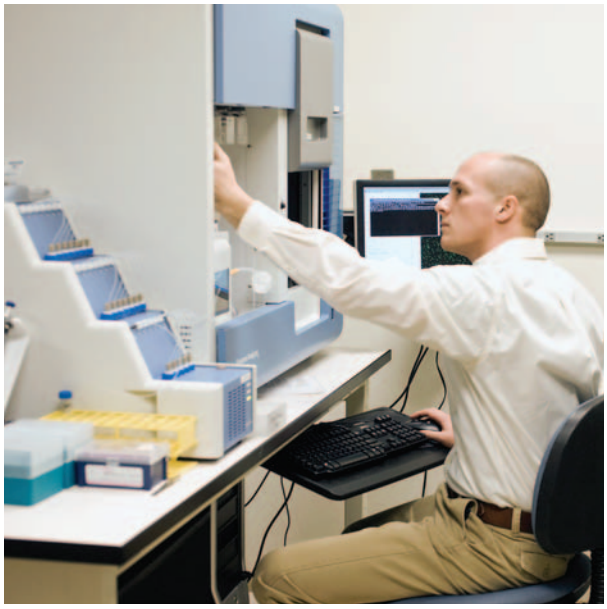
The power to solve medical mysteries

As the H1N1 virus—commonly known as swine flu—spread, researchers at Columbia University wanted to map its precise genetic makeup. How, they wondered, might it have evolved from flu viruses in pandemics past?

To find out, they needed to run extensive computational analyses—and they weren't the only ones seeking high-performance computing time.

At Columbia University's Center for Computational Biology and Bioinformatics (C2B2), large computer clusters are busy unraveling mysteries such as the makeup of the H1N1 virus, how complex traits are inherited, how proteins and DNA interact, and how harmful cell mutations might be predicted.

"The HP BladeSystem is saving us roughly $95,000 in energy costs in a year. The power numbers really sealed the deal."
John Wofford, director of IT, Center for Computational Biology and Bioinformatics, Columbia University



The intent is to gain understanding and uncover insights that might lead toward better treatment and cures.

But until recently, there wasn't enough computer processor capacity to satisfy demand. For example, systems biology is a new field that takes into account complex interactions in whole biological systems. A batch of cellular network calculations in this field can contain upwards of 20,000 single processor jobs. Until recently, that calculation would have taken 50 days to complete using the largest computing cluster that C2B2 had.

To deliver results faster, and continue to be able to attract and keep top researchers, more capacity was needed.

The center already had a dozen specialized compute clusters. But according to John Wofford, director of IT at C2B2, the clusters no longer offered enough scalability and were a challenge to manage. "We hit a point with our old clusters where we had many more jobs to run than we could feasibly fit into the time we had to run them," Wofford explains.

Going big while reducing power consumption 40%

The team evaluated possible next steps. "We had the idea to build a single monster cluster that could accommodate a variety of applications," Wofford explains. "Previous clusters were a bit more specialized for what a particular lab did, so we wanted to create a general purpose cluster."

Because the C2B2 data center is based in Manhattan, power and space efficiency are critical. Therefore, the C2B2's new "monster" cluster, nicknamed Titan, would consist of blade servers to achieve maximum compute density. The C2B2 team evaluated solutions from major vendors and chose the HP BladeSystem because of the power and space efficiency and streamlined management that it offers.

Space is saved because Titan consists of HP ProLiant BL2x220c server blades, a design that packs two servers in each slot. This enables C2B2 to fit 464 nodes, each with a dual-socket Intel® Xeon® processer 5450, into five racks. "That's pretty small for that much performance," Wofford told a representative from Intel. "It saves us about $110,000 in rack and chiller costs up front compared to the 10 racks that other manufacturers required."

The cluster also consists of 16 HP ProLiant BL460c server blades, which each carry more memory than the BL2x220c and have specialized uses. Two rack-mounted HP ProLiant DL360 Servers act as head nodes.

The HP BladeSystem with Intel Xeon processors minimizes power consumption, Wofford reports. "ConEd could give us 800 kW for the entire data center," he points out. "Anything more than that, we would have to do another power feed. Titan has theoretical peak power consumption of 95 kW, which is pretty impressive. It's about 35–40 kW less than the next-most-efficient system we looked at—roughly 40 percent less."

"The HP BladeSystem with Intel Xeon processors minimizes power consumption."
John Wofford, director of IT, Center for Computational Biology and Bioinformatics, Columbia University

Savings are substantial. "The HP BladeSystem is saving us roughly $95,000 in energy costs in a year," says Wofford. "The power numbers really sealed the deal."

Titanic performance improvements
The most important benefit is that Titan's performance is 10 times faster than C2B2's next largest cluster. Work gets done more quickly, says Wofford. "We've seen a 15 percent to 25 percent performance increase per job run per CPU over the previous cluster iteration," he notes.

That can add up to big gains. "The analysis that consisted of 20,000 single processor jobs—the one that took the previous cluster 50 days to complete—now gets done in just five and a half days with the HP BladeSystem," Wofford observes.

Titan's 44.54 teraflops of peak performance ranked it 228th in the world on the June 2009 Top 500 Super Computers list at www.top500.org.

Saving hundreds of hours in system management
Wofford points out that some basic aspects of cluster management would have become unmanageable without the HP SmartStart Scripting Toolkit. "Look, we have roughly 500 computers—464 plus the head nodes," he says. "A basic task like changing a parameter in the BIOS is something I do not have time to perform on all the blades. Doing it manually is not a manageable task—464 blades at 15 minutes per blade—that's 116 hours to change a single parameter."

Using the HP SmartStart Scripting Toolkit, Wofford can create a BIOS template and push out the change to all the blades at once in less than an hour, allowing updates to occur more than 100 times faster.

"The amount of time I save with HP SmartStart scripting—I don't know how I possibly could have put this thing together if I couldn't manage systems en masse like that," shares Wofford. "We chose HP in large part because it had a lot of scripting ability. Compared to some others, for instance, there are a lot of tasks that you can deploy all at once, so in terms of implementation, we saved hundreds of hours."

I love iLO2
C2B2 has also come to rely on HP Integrated Lights-Out 2 (iLO2), a remote management feature of the HP ProLiant BladeSystem that interfaces with the onboard administrator. "We use it all the time," confirms Wofford.

"What's great is that we don't even have to contact every single iLO on the system," Wofford adds. "I can power the entire chassis on and off at once. With the command line interface of the HP Onboard Administrator, which we really appreciate, we write scripts that cycle through all the HP BladeSystem c7000 Enclosures and execute any onboard administrator command. So I can give one command to power off the whole cluster."

Features like this helped enable C2B2 to deploy Titan without adding staff. "I'm not sure how I could live without iLO, unless of course I had a whole team of people to log individual changes," Wofford says. "The HP SmartStart Scripting Toolkit and iLO very likely saved us from having to hire two additional full-time employees."

About the Center for Computational Biology and Bioinformatics

Columbia University's Center for Computational Biology and Bioinformatics (C2B2, **www.c2b2.columbia.edu)** is one the nation's largest academic research centers using advanced computational methods to investigate a wide range of biological phenomena, including adverse drug reactions, the evolution of influenza viruses, the origins of flu pandemics, and next-generation cancer diagnostics and therapeutics.

Customer solution at a glance

Hardware

- HP ProLiant BL2x220c server blades with Intel Xeon e5450 processer
- HP ProLiant BL460c server blades with Intel Xeon e5450 processer
- HP BladeSystem c7000 Enclosures
- HP ProLiant DL360 Servers with Intel Xeon e5450 processor
- HP ProCurve 3500yl Switches
- HP LeftHand P4000 SAN

Software

- HP iLO2 Advanced Pack
- HP SmartStart Scripting Toolkit
- The MathWords MATLAB
- Grid Engine

Operating system

- Fedora Core Linux

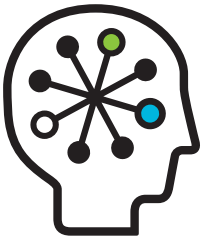
Wofford also mentions that because of HP remote management, his team "has experienced a 50 percent reduction in administrator visits to physical clusters."

Added savings from HP ProCurve

In addition to the ProLiant blades, C2B2 employs 15 ProCurve switches throughout their corporate infrastructure—three are directly related to the large cluster. Wofford explains, "We selected HP ProCurve 3500yl Switches because we have always had solid reliability from them. We like the switch. It works well for us. It's reliable, and there is no annual maintenance contract required–which reduces annual total cost of ownership for switches in the range of 10 percent. We simply have a history of reliability and good support."

C2B2 also uses an HP LeftHand P4000 SAN as storage for several databases, virtualized assets and its mail store. According to Wofford, "The HP LeftHand P4000 SAN is a nice balance: feature-rich and cost-effective. We use its built-in snapshot and replication capabilities. For being a SAN appliance, it is about as plug-and-play as you can get. We had it up and running the day we got it."

Researchers are noticing the improvements as well. According to Wofford, "The consistent performance improvements on the order of 15 percent to 25 percent make everyone happy." And over time, faster research cycles can help in the effort to make people healthier, too.



Technology for better business outcomes

To learn more, visit www.hp.com
www.intel.com

4AA0-4564ENW, April 2010