Brown University studies the world's scientific problems

With an IBM System x iDataPlex and ScaleMP vSMP Foundation for Cloud solution

Overview

The need

Brown University needed a flexible, cost-effective high-performance computing solution that could provide a very large memory footprint needed to run complex scientific applications.

The solution

Brown University deployed 16 IBM® System x® iDataPlex® servers featuring intelligent Intel® Xeon® processors (192 cores) and 1.5 terabytes of addressable memory, along with ScaleMP vSMP Foundation for Cloud.

The benefit

It's a flexible, virtualized environment with an optimal balance of high-performance computing and bandwidth in a large memory footprint.

What happens when one of the nation's most prestigious universities needs a versatile new high-performance computing solution to help answer some of the world's most complex scientific problems? They turn to IBM for solutions.

Founded in 1764, Brown University in Providence, Rhode Island, is the seventh-oldest college in the United States. Brown's vibrant, diverse community includes 6,000 undergraduates, 2,000 graduate students, 400 medical school students, and nearly 700 faculty members. Brown students come from all 50 states and more than 100 countries.

Brown offers degrees in more than 70 concentrations, ranging from Egyptology to cognitive neuroscience. In recent decades, the Ivy League school has moved to bolster its supercomputing capabilities to support advanced research requiring complex numerical simulation, modeling, and data analysis in fields such as genomics.

Extending research capabilities

These high-performance computing capabilities became formally known as the Brown University Center for Computation and Visualization in 2006. The center's stated goal? Provide the resources to help researchers at Brown and across Rhode Island pursue cutting-edge computational science, visualization and virtual reality research.

Overtime, the need to augment Brown's existing high-performance computing capabilities with a solution capable of making very large amounts of memory available for researchers involved in genomics and other scientific studies became evident, says Sam Fulcomer, associate director for the Center for Computation and Visualization.

"We have a very wide, diverse community of computational scientists and their applications have very different needs," says Fulcomer. "What we began to realize is that we needed a larger memory system for our genomics researchers, for example, who are looking at sequence data from a couple of high throughput sequencers to understand the DNA of entire organisms."

In more exacting terms, explains Fulcomer, Brown "needed a large memory system that had more than a terabyte of shared memory. That was really the biggest factor."

Many needs, one solution

Beyond a large memory footprint, the solution needed to have enough flexibility to handle highly distributed applications requiring raw computational processing. Typically, supercomputers are tailor-made to handle either large memory or distributed applications, not both—leading to a more complex infrastructure, supporting multiple machine types. However, the diversity of research at Brown necessitated a versatile, less complex and cost-effective solution, explains Fulcomer.

"We have many different types of applications, some of which are very I/O intensive and some of which are very compute intensive," says Fulcomer. "One option would have been to build a physical machine with shared memory with sufficient bandwidth to keep a large number of processors running, but that's very expensive. And plus, shared memory systems are complex to build."

The task for Fulcomer was to find a cost effective solution system that could effectively address researchers' broad computational requirements. It turned out that part of the answer already existed at Brown's high-performance facility in the form of IBM System x iDataPlex servers. Fulcomer says Brown already had used iDataPlex before, but not on the scale they were about to implement in a new solution.

The power trio: IBM, Intel and ScaleMP

The new IBM solution was installed in just five days. It includes 16 iDataPlex servers featuring intelligent Intel Xeon processors and 96 gigabytes of memory each, for a total of 128 processing cores and 1.5 terabytes of addressable memory. The systems are inter-connected with QDR InfiniBand, and connected to storage over multiple gigabit Ethernet interfaces. There are plans to bolster this connectivity by adding 10GigE interfaces soon.

Solution components

Hardware

- IBM® System x® iDataPlex® servers
- Intel® Xeon® processors
- IBM System Storage®
- InfiniBand

Software

- IBM General Parrell File System (GPFS™)
- IBM Tivoli® Storage Manager
- IBM xCAT
- ScaleMP vSMP Foundation Advanced Platform
- Linux

IBM General Parrell File System (GPFS™) provides the high I/O for diskless nodes, allowing for more versatile aggregate performance. For large-memory and shared-memory applications the solution runs ScaleMP vSMP Foundation for Cloud virtualization solution, provisioned by IBM xCAT open source system, giving Fulcomer the virtualization capabilities needed to rapidly provision and manage research projects on the cluster. Linux OS is running within the virtual machine.

This high-performance computing solution is connected via GPFS to an IBM System x server and 12 IBM System Storage® disk systems, as well as a IBM Tivoli® Storage Manager system with co-located LTO tape libraries to facilitate physical backup of data and disaster recovery.

Says Brown, "It is really the first well-balanced high-performance computing platform that Brown University has had, both in terms of having a large number of floating-point operations per second and the communications interconnect bandwidth and I/O subsystem performance to support a wide variety of applications."

Powerful performance, commercially available

The overall IBM/ScaleMP combined solution now functions as the heart of Brown University's flagship computer cluster. Dubbed Oscar, the cluster also includes GPU-based systems and older iDataPlex offerings with several hundred multi-core nodes sharing a high-performance interconnect and file system.

The new solution has already delivered solid performance for a variety of workloads. In one example, researchers are using a set of algorythimgs called Velvet which is helping in the study of genes as well as MATLAB. Fulcomer says what's great about the new solution is that it allows scientists to run Velvet code that is nearly a terabyte in size.

"This IBM solution, paired with ScaleMP, gives us the ability to handle applications that we just couldn't run before because of the memory constraints," says Fulcomer.

The facility is making an impact well beyond the Brown campus, says Fulcomer.

"It's the largest system around and it is being used by researchers from all over, including the state of Rhode Island, which is conducting genomics research on marine life," explains Fulcomer. "It has also been instrumental in helping us to attract new faculty hires and graduate students who know that Brown has the high-performance computing solution they will require for their research."

For more information

Contact your IBM sales representative or IBM Business Partner. Visit us at: **ibm.com**/systems/x

For more information about Brown University, visit: www.brown.edu

For more information about ScaleMP, visit: www.scalemp.com

XSC03105-USEN-00