

CHALLENGE	SOLUTION	BENEFIT
Provide a diverse research community with high-performance computing (HPC) systems while lowering the costs and efforts of maintaining those systems	Several multipurpose, large-scale Dell™ clusters using Intel® processors and running Microsoft® software	A robust, easy-to-manage, state-of-the-art HPC environment that provides a range of computing resources, scalability, cost-efficiency, flexibility, and outstanding reliability

A new theory in high-performance computing

Dell provides the Cornell Theory Center with the high performance of HPC without the high costs of proprietary systems

Frequently at the forefront of technology, universities have the opportunity to explore, test, and develop new technology. However, academic research can require significant computing resources, often far beyond those of the typical enterprise. At Cornell University, the Cornell Theory Center (CTC) is a center of excellence in high-performance computing (HPC) and interdisciplinary research. CTC supports faculty and staff from more than 100 different research areas as well as corporate clients that require leading-edge computational resources. Since its founding in 1985, CTC has been a leader in national and international high-performance computing.

As one of the four original National Science Foundation super-computing centers, CTC ran proprietary UNIX®-based systems from IBM, SGI, and others for more than 10 years. But high costs and the extra effort for maintaining these proprietary systems prompted CTC to look for another solution that would provide its users with state-of-the-art computing resources.

In 1999, CTC migrated to Intel® processor-based Dell™ clusters running Microsoft® Windows® Server software. Since then, these clusters have met the center's specific research requirements while providing a standardized management model. CTC achievements have demonstrated that a high-performance cluster system can be built with commercially available tools and that administrators can manage such a system for a fraction of the effort committed to proprietary UNIX-based systems.

CTC brings Dell into the equation

CTC worked with software vendors to pull together the essentials for a high-performance Dell cluster that would provide Cornell users with the right tools, such as message-passing libraries, performance tools, math libraries, and compilers.

The first large-scale Dell cluster system that CTC deployed was Velocity1, which comprises 64 Dell PowerEdge™ 6350¹ servers with quad Intel Pentium® III Xeon™ processors at 500 MHz, 2 MB cache per processor, and 4 GB RAM per node. Velocity was ranked among the TOP500 Supercomputer Sites within a few months of installation and was saturated with users soon after. Velocity achieved 99.9986 percent availability (as independently verified by the Massachusetts Institute of Technology) during the first three months of running Microsoft Windows 2000; by late 2001, CTC had reached 99.99999 percent across the entire machine room.

CTC provided a variety of support systems for users during the migration, including training, consulting, and tool development. CTC staff helped many researchers who had custom applications port several million lines of code from UNIX to Windows. The center also provided scientists with a dedicated collaborative environment that included all the tools they needed at their fingertips.

With success comes challenge

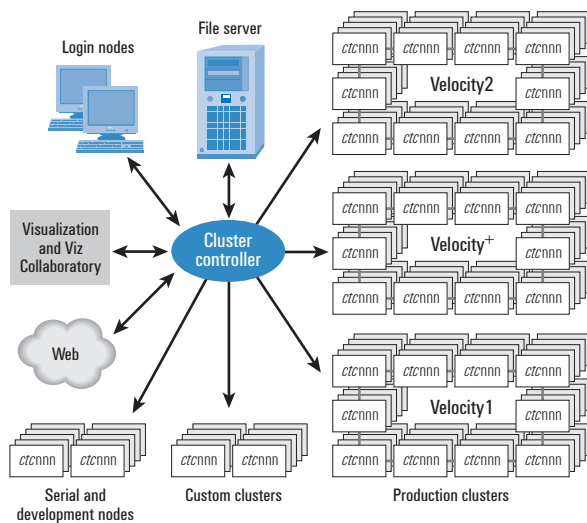
Although the initial challenge was straightforward—provide a diverse community with excellent computing systems while



maintaining a lean and efficient staff—CTC's success led to expanded opportunities and new challenges.

Once the word was out across campus, several research groups came to CTC with an interest in a large-scale cluster designed to meet their specific computing needs. These needs ranged from statistical analysis of large demographic databases to materials science simulations of fracturing helicopter gears. The CTC staff identified the appropriate configurations to meet the needs of these strategic applications and then built another cluster: Velocity⁺.

As a center of expertise, CTC encourages scientists and engineers to advance their research through the use of parallel computing. The CTC staff consults on parallel code porting and offers training in designing parallel applications. Several Cornell research groups quickly moved to parallel Windows computing, and Velocity⁺ was built to meet their needs.



The Cornell Theory Center Velocity cluster complex

Velocity⁺ consists of 64 Dell PowerEdge 2450¹ servers with dual Intel Pentium III processors at 733 MHz and 2 GB RAM per node. Each processor has 256 KB cache, each node has 27 GB disk capacity (RAID-0), and the system uses Emulex[®] cLAN[™] interconnect technology. Velocity⁺, with half the processors of the original Velocity cluster and one-third of the cost, achieved 50 GFLOPS on the MP-Linpack benchmark.

As the queue of parallel users grew for this system, CTC implemented a special-purpose cluster of 36 serial nodes for code development and serial users, so that the parallel users received priority access to the Velocity systems. CTC also installed a small development cluster of eight Dell PowerEdge 1550¹ servers with dual

Pentium III processors. The short time limit set on this system provides researchers with quick turnaround on debugging and tuning runs.

Everyone wants a cluster

Realizing they could afford to purchase their own custom systems for applications, research groups soon came to CTC for guidance and support. For example, the Cornell Institute for Social and Economic Research (CISER) deployed a 32-processor cluster dedicated to running SAS—both parallel and serial—against secure census data. One of the groups that drove the purchase and installation of Velocity⁺, the Computational Materials Institute (CMI) at CTC, later purchased a dedicated 64-processor parallel system designed for compute-intensive simulations in computational materials science.

When integrated into the overall computing environment at CTC with secure access for specific user groups, these custom clusters allow for the best use of shared resources and demonstrate the flexibility of Dell server systems to meet a variety of needs. CTC has also made cluster management more cost-efficient. For example, fewer administrative resources are now required for upgrading the standard operating system and software tools used among the clusters.

Velocity2 enhances state-of-the-art systems

Computational researchers are never satisfied for long. Not only do they constantly push the limits of their high-performance systems, they also must demonstrate to their funding sources that they have access to the best available systems, and that funds are well invested in state-of-the-art systems.

Velocity2, CTC's most recent Dell cluster system, became available November 4, 2002. At the same time, Velocity1 and Velocity⁺ were repurposed. The Velocity2 cluster now handles large-scale parallel applications while Velocity⁺ is the general-purpose system and Velocity1 is a Microsoft .NET cluster.

Velocity2 comprises 128 Dell PowerEdge 2650 servers with dual Intel Xeon processors at 2.4 GHz, 2 GB RAM per node, 50 GB disk capacity (RAID-0) per node, and Gigabit Ethernet² and InfiniBand[™] interconnects.

Dell and CTC provide tools for the research trade

The Dell-based infrastructures also have allowed researchers to take advantage of new tools developed by CTC's CMI and running within the Velocity complex—directly from their desktops. Not only can researchers perform production simulation runs from a

¹ Newer models are available at <http://www.dell.com>.

² Gigabit Ethernet indicates compliance with IEEE[®] 802.3ab and does not connote speeds of 1 Gbps.

Web-based interface that leverages the Microsoft .NET framework, but they can also visually explore the results of those simulations, which are stored in SQL databases.

In its first implementation, this system has integrated a Dell server with 2 GB RAM for visualization and control and four-way Intel Xeon processors (IA-32) with 2 GB RAM each. This server runs the fracture mechanics simulation. A server with four-way Intel Itanium® 2 (IA-64) processors and 4 GB RAM runs Microsoft SQL Server 2000 as the back-end database.

CTC uses a variety of Dell systems to provide innovative scientific visualization tools for fields ranging from bioinformatics to computational finance. For example, visualization of the fracture simulation and genomics data mining applications are run on a workstation. However, any HPC environment can benefit from a scientific visualization system such as the Cave Automatic Virtual Environment (CAVE) virtual reality system. These environments allow researchers to explore raw data and simulation results.

CTC has recently made a new cluster system available for research that takes advantage of the 64-bit Intel Itanium architecture. This next-generation system comprises 32 Dell PowerEdge 7150 servers with quad Intel Itanium processors at 1 GHz, 8 GB RAM per node, 50 GB disk capacity (RAID-0) per node, and Gigabit Ethernet for the interconnect.

CTC launches new HPC initiative

When the migration to Dell clusters began, the CTC team believed that if cluster computing were truly an industry-standard system, institutions and companies should be able to purchase the software they need, install a cluster quickly, and support the system with a lean and efficient staff. On October 1, 2002, CTC launched CTC High-Performance Solutions, a collaboration with Dell, Intel, and Microsoft to reach out to business, industry, academia, and government institutions and to help bring high-performance computing to Main Street.

According to Thomas F. Coleman, CTC director and Cornell computer scientist, this collaboration has the potential to benefit many organizations. “With our expanded relationships and combined strengths, we can show companies, government agencies, and academic institutions how to expand their technical computing environments while reducing their overall IT budget. They can take their existing expensive, proprietary systems—often islands of performance requiring extra staff—and replace them with a more flexible, scale-out clustered environment that can be expanded and one that fits into the overall Windows-based office environment.”

The CTC High-Performance Solutions initiative drove the delivery of Velocity2 and nearly doubled the resources of the Velocity complex on the Cornell campus. CTC has installed two showcase


“With our expanded relationships and combined strengths, we can show companies, government agencies, and academic institutions how to expand their technical computing environments while reducing their overall IT budget.”

— Thomas F. Coleman
CTC Director and Cornell Computer Scientist

systems at CTC-Manhattan in New York City’s financial district so that clients can explore and test HPC using special clusters for benchmarking, test runs, or application development. The first of these systems comprises 16 Dell PowerEdge 2650 servers with dual Intel Xeon processors at 2.4 GHz and 5 GB RAM per node. Each node has a 50 GB disk capacity (RAID-0) and the cluster uses Gigabit Ethernet and InfiniBand interconnects. The second system has four Dell PowerEdge 7150 servers with quad Intel Itanium processors at 1 GHz, 8 GB RAM per node, 50 GB disk capacity (RAID-0) per node, and Gigabit Ethernet as the interconnect.

Dell redefines high-performance computing

CTC has significantly benefited from its move to Windows-based cluster computing on Intel processor-based Dell systems. “With our Dell high-performance computing clusters we have led breakthroughs in the social sciences, introduced biologists to an entirely new class of computational tools, as well as developed new analytic methods for financial markets, driving the transition from art to science,” says Hunter L. Rawlings III, president of Cornell University.

The performance of these Dell clusters meets the requirements of its demanding users, the systems are reliable and flexible, the latest technologies are available, and the total cost of operation is a fraction of that for large proprietary systems. In addition, CTC has provided its users and staff with a fully integrated environment—from desktops to high-end server-based clusters and visualization systems. Dell has helped CTC breathe new life into high-performance computing. 

FOR MORE INFORMATION

<http://www.dell.com/servers>

<http://www.dell.com/hpcc>

<http://www.intel.com>