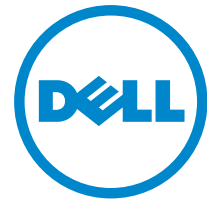# TACC provides researchers with data analysis capabilities that advance discoveries, thanks to partnership with Dell

DELL

• HPC

"Researchers come here to get capabilities that they wouldn't otherwise have access to, and that in turn breeds new science and new discoveries."

Kelly Gaither, Director,
Data and Information Analysis,
Texas Advanced Computing Center

Customer Profile

Company:	Texas Advanced Computing Center
Industry:	Education
Country:	United States
Staff & Students:	90+
Web:	www.tacc.utexas.edu

TACC

Business Need
The Texas Advanced Computing Center (TACC) needed to enable scientists to quickly analyze large datasets to help them solve real-world problems such as forecasting hurricane landfalls and predicting the spread of viruses. It also wanted the ability to view complex data at its native resolution.

Solution
The center partnered with Dell on two projects: Stallion, a high-resolution display wall using Dell™ UltraSharp™ flat-panel monitors; and Longhorn, a GPU-accelerated visualization and data analysis cluster with Dell PowerEdge™ servers and NVIDIA GPU technology sourced through Dell.

Benefits
- Decrease in floor space required for cluster with hybrid CPU/GPU configuration
- Ability to interactively analyze large datasets in real time
- Data can be viewed at native resolution
- Acceleration of new discoveries compared to numerical analysis

As the machines in high-performance computing (HPC) centers increase in processing power, they are able to create more realistic simulations to help study and solve problems that affect the lives of many people. Predicting the spread of viruses and the landfall of hurricanes, for instance, are real-world problems that scientists can analyze using statistical algorithms with computers.

When dealing with large datasets, visualization technology becomes an important factor in enabling scientists to analyze their results. As Kelly Gaither, director, data and information analysis at the Texas Advanced Computing Center (TACC) at The University of Texas at Austin, puts it, "Without visualization, we would need to analyze numerically. We would have to make sense of stacks and stacks of zeroes and ones. We have in our brains the most powerful supercomputer that you can get access to, and visualization technology lets us take advantage of that and leverage our automatic real-time pattern matching, feature recognition and visual acuity to make sense of very large datasets."

Breeding new discoveries

The Texas Advanced Computing Center provides ample opportunities for scientists to crunch numbers and visualize the results in great detail using graphics programs. TACC's visualization laboratory serves researchers at The University of Texas at Austin and across the nation, and it is increasing its impact on science thanks in part to its partnership with Dell. Two of TACC's systems, Stallion and Longhorn, are the largest of their kind in the world.

Stallion, a 307-megapixel display of 75 Dell UltraSharp™ 30-inch monitors, provides users with the ability to perform visualizations on a 15'x5' tiled display.

Powered by Dell XPS 720 tower desktops with Intel Core 2 processors and dual NVIDIA 8800 GTX (G80)

graphics cards, the visualization cluster allows users to access over 36 gigabytes of graphics memory, 108 gigabytes of system memory and 100 processing cores. This configuration enables the processing of datasets of a massive scale, and the interactive visualization of substantial geometries. A large, shared file system is available to enable the storage of terascale size datasets.

"The beauty of Stallion is that we're able to see the imagery at its native resolution," says Gaither. "And Stallion can be used for a variety of things.

Technology at Work

Hardware

Dell™ PowerEdge™ R710 and R610 servers with Intel® Xeon® processors

Dell PowerVault™ MD1000 direct-attached storage arrays

Dell Precision 690 workstations with Intel Xeon processors

Dell UltraSharp™ 30" monitors

Dell XPS 720 desktops with Intel Core™ 2 processors

Mellanox QDR InfiniBand interconnect

NVIDIA® Quadro® Plex S4 with NVIDIA FX 5800 GPUs

Software

Lustre Parallel File System

Red Hat Linux

Certainly we're using it to display very large imagery from simulations when we want to see the resolution of the native data, but we also use it for non-scientific purposes. Stallion is the only place in town where an Austin-based photographer is able to see his digital pictures at native resolution."

For researchers who don't require such a large display, TACC offers a workstation area dedicated to visualization with Dell Precision 690 workstations connected to large LCD displays, and a collaboration room for video-conferencing and small group meetings.

World-class data analysis

TACC's most recent venture with Dell is Longhorn, a cluster designed for remote interactive visualization and data analysis. The system consists of 16 Dell PowerEdge R710 nodes with 144Gb of RAM each, and 240 Dell PowerEdge R610 compute nodes, each with two Intel Xeon 5500 series processors and 48 gigabytes of RAM. Storage nodes are Dell PowerEdge R710 servers with Intel Xeon 5500 series processors. Fourteen Dell PowerVault MD1000 direct-attached storage arrays provide a 210 terabyte global file system, managed by the Lustre Parallel File System. A Mellanox InfiniBand quad-data rate (QDR) fabric provides the interconnect.

To accelerate data analysis and make interactive visualization possible, Longhorn uses 128 NVIDIA Quadro Plex S4 units sourced through Dell, each with four NVIDIA FX 5800 graphics processing units (GPUs), 16 gigabytes of graphics memory and two independent graphics busses (one per GPU pair). Compute nodes are each connected to two dedicated NVIDIA FX 5800 GPUs via the Quadro Plex graphics bus.

"Longhorn is the largest graphics accelerated remote interactive visualization cluster in the world," says Gaither. "We built it with a grant from the National Science Foundation and in cooperation with

Dell. Dell helped us design the cluster architecture using a hybrid approach, combining CPU cycles with NVIDIA graphics processing units."

Traditionally, GPU acceleration in visualization clusters is handled by desktop computers that leverage internal graphics cards. "Dell has a very compact, rack-mountable server footprint in collaboration with NVIDIA that allows us to have rack-mountable nodes while still gaining the benefit of real-time performance for graphics applications," says Gaither. "You can imagine what the footprint would look like for a visualization cluster this powerful using hundreds of desktop machines—the Dell and NVIDIA solution reduced our space requirements for Longhorn."

The Intel Xeon 5500 series processors in the Dell PowerEdge servers are helping TACC solve problems faster. "Although visualization and data analysis are often more memory-bound, we've run a number of CPU-intensive jobs that have been quite successful," says Gaither.

Staying on the bleeding edge

Receiving a grant from the National Science Foundation for a visualization resource of this magnitude was a major step for TACC. It was also critically important for visualization and data analysis as scientific fields— a recognition that visualization and data analysis services are a very important part of the scientific process.

TACC was an early adopter of distributed memory HPC clusters, mostly for reasons of scale. Because analyzing larger and larger problems drives up the cost of processing with shared memory machines prohibitively, distributed memory machines are now the preferred alternative for many HPC centers, using cluster configurations of "commodity" machines.

Dell plays an important part in the visualization facilities at TACC by promptly responding to the center's needs as it provides greater capabilities

> "We communicate with Dell about the kinds of scientific problems that we are interested in, which enables them to respond to our needs."
>
> *Kelly Gaither, Director,*
> *Data and Information Analysis,*
> *Texas Advanced Computing Center*

to scientists at The University of Texas and across the nation. By purchasing servers, storage, visual computing hardware, monitors and networking components through Dell, TACC is able to realize economies of scale, streamline procurement and benefit from Dell's technology expertise.

"We keep an open collaboration with Dell, as we do with all of our technology partners," says Gaither. "We communicate with Dell about the kinds of scientific problems that we are interested in, which enables them to respond to our needs."

Solving real-world problems

Although Longhorn itself does not have a visualization display attached, data can be streamed to Stallion for local viewing or to remote locations to accommodate researchers all over the world. TACC has worked with the National Oceanic and Atmospheric Administration (NOAA) to run simulations to more accurately predict landfall for hurricanes using new techniques with statistical models, and hopes to use Longhorn during hurricane season to perform the visualization in real time.

"Longhorn will allow us to do three-dimensional visualizations as the storm data is being analyzed," says Gaither.

TACC also met with researchers to help predict how certain diseases might spread geographically. "I think we've had an impact on understanding how to prevent the spread of viruses," says Gaither. Researchers are also using TACC systems to produce 2D and 3D simulations of the impact of BP's massive Gulf of Mexico oil spill on coastal areas, and plan for the possibility that a hurricane moving through the Gulf might bring some of the oil inland.

The robust information analysis and visualization capabilities the Dell systems provide are drawing more researchers to TACC. "Stallion and Longhorn are the crown jewels of our visualization resources," Gaither concludes. "Researchers come here to get capabilities that they wouldn't otherwise have access to, and that in turn breeds new science and new discoveries."

TACC

View all Dell case studies at: dell.com/casestudies

DELL