CAPRICORN TECHNOLOGIES

Vanderbilt University's ACCRE Selects Capricorn Technologies and AMD64 for High Performance Storage Solution

"HyperTransport has the high memory bandwidth that we need to handle extremely difficult computing problems."
-Alan Tackett, Technical Director, ACCRE

Vanderbilt University's Advanced Computing Center for Research & Education (ACCRE) is busy around the clock crunching data on some of the world's most difficult computing problems: from reconstructing the Big Bang to single-cell neurophysiology to improving the reliability of integrated circuits in space. This diversity of usage differentiates ACCRE from many of its peers and at the same time, complicates its computing requirements.

ACCRE's mission is to enhance the research efforts of all the colleges and schools at Vanderbilt, and ensure that all those who need access to the Center's computational power to advance their particular academic discipline can take advantage of the resource. Alan Tackett, ACCRE's technical director, puts it this way, "We have an open community and we let everyone on [the cluster.]"

ACCRE's HPC cluster includes a mix of servers based on the PowerPC platform, legacy x86 processors and the high bandwidth, low-latency AMD64 platform. The integration and performance of the computing cluster have allowed the Center to grow and support both on-going research projects and advance reservation times for classes or labs. For Tackett and

his team, questions arose when it came time to build out the Center's storage capabilities.

He describes the ACCRE environment in this way, "Our users are really varied. We may have folks running Matlab, Perl scripts doing data mining, and traditional HPC apps all at once. With our cluster file system storage, we face the challenge of providing both throughput and simultaneous sustained connections. Another issue in our environment is that we have proteomics researchers generating tens of millions of tiny files each less than 2000 bytes, while others may need to store files on the order of a few gigabytes. We evaluated different types of file system solutions—both parallel and traditional. We realized that a file system did not exist that could meet the varied needs of our community and that a combination of a file system and another storage system would be needed. The additional storage systems needed to be flexible enough to handle the gamut of small files, large files, high bandwidth, and high fault tolerance – much greater than traditional RAID6. We needed a tunable approach to storage and set about solving these issues on the software side."

L-Store is designed to provide virtually unlimited scalability in both raw storage and associated file system metadata; a decentralized management system; security; fault tolerant metadata support; user controlled replication and striping of data on a file and directory level; scalable performance in both raw data movement and metadata queries; a virtual file system interface in both a web and command line form; and support for the concept of geographical locations for data migration to facilitate quicker access. These features are accomplished using two fundamental technologies designed to handle such wide-ranging data storage and access needs: Distributed Metadata and Logistical Networking.

Both resourceful and determined to get the storage piece of the Center's computing puzzle just right, ACCRE, supported by both internal Vanderbilt funding and a grant from the NSF, developed their own Logistical Storage Software, or "L-Store," which was first demonstrated at the Supercomputing 2005 conference.

L-Store's strength is that it provides a clear, functional delineation between data layers and each data management function is designed to be independently scalable. Built on the varied requirements of ACCRE's users, the software supports distributed and secure access via client initiations on IP-based networking protocols and employs distributed metadata, which uses distributed hash tables originally developed for peer-to-peer systems.

Tackett explains, "When you see a news website, for example, you may see content on a single displayed page that is stored and then pulled from maybe 10 separate servers. The url is just a placeholder. This is how our storage works."

Beyond architecturally addressing ACCRE's fundamental storage issue, L-Store is high-performance software, helping deliver sustained data transfers of 3 Gigabytes per second at the Supercomputing 2006 conference on specially designed Capricorn Technology Power Store™ hardware.

"Once we had the software situation under control, taking care of the metadata synchronization and requests for space and fault tolerance, we needed the right hardware to run it. When I went searching for the right platform, I knew that I didn't want to pay for frills."

ACCRE's storage hardware requirements for L-Store can be met by a number of hardware solutions on the market. But the investment of time and resources needed to get the right balance of performance, scalability, accessibility, low power consumption, and price-performance mandated that ACCRE give the hardware side a very careful look.

As Tackett mentioned, "frills", or unnecessary options, certainly weren't going to improve their storage scenario and, for a University-run center, acquisition and implementation costs, as well as on-going administration costs factor heavily into decision-making. The team decided they wanted a hardware solution that was as custom-tailored to ACCRE's unique demands as the software.

"In our research, we came across Capricorn Technologies and realized their primary market direction was very much in line with what we needed. Capricorn is focused on delivering storage hardware that is affordable and energy efficient. They were able to meet our need for high capacity, lower cost storage systems that give on-going cost savings in power usage and management but at the same time are highly performant."

With its unique storage architecture, Capricorn Technologies' PetaBox delivers high density, highly reliable, highly scalable, and highly available data storage in a flexible environment with very low total cost of ownership. A single 19-inch PetaBox rack can support up to 120TB of raw disk space. This density is achieved through an economy of design that consumes as little as 27 watts per terabyte.

In fact, Capricorn's concentration is on delivering low cost per terabyte while maintaining a high degree of performance and scalability – exactly what research and high performance computing institutions like ACCRE need. The architecture of Capricorn's PetaBox™ product family is designed for use in a "clustered" manner, similar to many of the highly-scalable application server environments available today. Their newest product in the PetaBox line, the PowerStore™, was specifically developed in response to customer requests for faster throughput, and achieves that, in part, through the unique building block architecture of the AMD Athlon™ 64 X2 processor and its HyperTransport™ technology interconnect. Additionally, AMD64 processor technology has been at the forefront in helping vendors and institutions decrease their systems' power consumption and thus, costs.

"When we met with Capricorn, we let them know our concern with achieving a balance of all the hardware components – disk, processor, memory, and network - and realizing a low power draw. Capricorn had hardware that matched that criteria."

"On the processor side, we saw performance differences between single and dual core. We witnessed a scenario where one core covered data transfer and operating system overhead while the second handled data encryption, fault tolerance encoding, or other data filtering functions. There was a reason why we wanted AMD64. A 386 could probably do the raw data movement, but it is the balanced architecture overall, the I/O throughput, memory capabilities, dual core, and a low power processor that gives us options. For example, we can compute on the box or use the client for erasure encoding

and fault tolerance. We can use the processor for modifying or filtering data before sending it down to the client. Then our system is doing more than just storage – it's dynamic and we can determine how best to use it."

"Capricorn helped us design a rack with 24 PowerStore devices as disk depots which we then connected to 30 clients. When we used the rack to test L-Store, the results were impressive to say the least. We achieved sustained transfer write speeds of 3GB per second for over four hours. To put that in perspective, we were storing data at a rate equivalent to 4.6 CD-ROMs per second."

"While we continue development on the L-Store system to meet the needs of our diverse user base, we are very happy with the low power options that the AMD64 processor-based solution can provide. We now have the ability to meet a wide range of user storage requirements while, at the management level, achieving a credible option for maximizing our storage resources with fewer dollars. The AMD64 processor-based Capricorn solution delivered exactly what we wanted."

To learn more about Capricorn Technologies' PowerStore™ products, please visit www.capricorn-tech.com.

Information on AMD64 solutions for the storage market can be found at www.amd.com/amd64embedded.

The online home for Vanderbilt University's Advanced Computing Center for Research and Education is at www.accre.vanderbilt.edu.

More information about L-Store is available at www.lstore.org.

AMD64 technology with Direct Connect Architecture addresses the top performance issue in the networked storage environment: throughput. Direct Connect Architecture can improve overall system performance and efficiency by eliminating traditional bottlenecks inherent in legacy architectures, which restrict and interrupt the flow of data. Slower data flow means slower system performance. Interrupted data flow means reduced system scalability. With AMD64 technology, there are no front-side buses. Instead, the processors, memory controller, and I/O are directly connected to the CPU and communicate at CPU speed. AMD64 technology features an integrated memory controller, dramatically reducing memory latency and with HyperTransport™ technology delivering the industry's highest possible I/O bandwidth, data can speed through the system without encountering the front side bus bottleneck that plagues competing x86 platforms.