DELL Technologies AMD

A next-generation network

The University of Alabama at Birmingham bridges three data centers with a 100-gigabit network fabric from Dell Technologies.

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Scientific Research | United States

Needs

The University of Alabama at Birmingham needs a reliable, high-speed network to power scientific research workloads across three data centers.

Solutions at a glance

- Servers with 2nd Gen AMD® EPYC™ processors
- Dell EMC PowerSwitch networking
- Dell EMC SmartFabric OS10 network operating system

Results

- Making HPC resources accessible to a wide range of users
- Providing a foundation for an envisioned statewide consortium
- Supporting groundbreaking scientific discoveries
- Accelerating time to results for research workloads

The new UAB network fabric moves data across three data centers at a rate of

100 Gb/s

The UAB private research cloud moves data at an estimated rate of up to

6.4 Tb/s

The rise of a next-generation network

In data-intensive high performance computing, balance is everything. To deliver top performance, an HPC cluster requires equilibrium among compute, storage and network resources. In this spirit, storage and networking systems must be designed to keep pace with the evolving processing speed of the compute components, or the performance of the entire system suffers.

Too often, the network can emerge as a weak spot in an HPC architecture — for example, when newer, faster compute resources are added or when data and application resources are distributed over multiple locations. In such cases, HPC system administrators and their users might find that throughput for workloads is constrained by network limitations.

This was an issue that the research computing leaders at the University of Alabama at Birmingham encountered when they were laying the groundwork for a major expansion of their flagship Cheaha supercomputer, which powers computational research for more than a thousand scientists and researchers. In this process, the team members realized their legacy network architecture wasn't up to the challenges of a system that was growing to encompass new GPU-accelerated components, along with data and applications spread over three data centers.

This realization put the UAB research computing team on the path to a next-generation network fabric based on hardware, software and services from Dell Technologies.

A growing supercomputer

For the UAB research computing team members, the move to a next-gen network architecture began when they were in the process of adding new NVIDIA® DGX™ A100 GPU systems to the Cheaha supercomputer. The DGX A100 system harnesses PCIe 4 support, leveraging the high core count and the performance of 2nd Gen AMD EPYC™ processors to create what NVIDIA describes as "the world's most advanced AI system."

Delivering five petaflops of artificial intelligence performance, the elastic architecture of the NVIDIA DGX A100 enables organizations to accelerate diverse AI workloads, such as data analytics, training and inference. The installation at UAB incorporates four DGX A100s, each with 128 CPU cores and eight NVIDIA A100 Tensor Core GPUs. Each DGX system can allocate up to 56 GPU instances using NVIDIA MIG technology. Each compute node can be allocated to different virtual machines via an OpenStack cloud layer.

For multiple reasons, including power constraints in existing facilities and expected growth over time, the research computing team decided to deploy the DGX A100 system in a Tier III commercial data center in Birmingham, called DC Blox.

"Along the way, as we spec'd the system out, we discovered some limitations in the network architecture on the HPC side," says Ralph Zottola, Ph.D. and Assistant Vice President for Research Computing at UAB. "Earlier on, we had intentionally decided to make a condensed or collapsed network fabric in our HPC space, and that had served us well as we did incremental expansions over the years. But now we realized we had a problem with that component."

The UAB research computing team laid the groundwork to rebuild its entire network infrastructure to connect three data centers via a high-speed fabric.

"We moved from what was a hodgepodge of 40Gb connectivity across two data centers to a 100Gb backplane that unifies our commercial data center and two on-campus data centers," Dr. Zottola explains. "We essentially bridged our compute, which is our HPC, with our storage, which is archives and near-line storage, with our core network and our DGX A100, which is the new piece in the puzzle. This ability to turn each DGXA100 into 56 different MIG devices allows us to serve more users and a plethora of different workloads."

Transforming the network

For this transformation, the UAB research computing team chose Dell EMC PowerSwitch networking with the Dell EMC SmartFabric OS10 network operating system. This solution combines the best of the Linux® operating system, open computing and networking technology to advance open networking disaggregation.

Features of the SmartFabric OS10 platform include:

- Hardware abstraction through common APIs to enable consistency across data center compute and network resources, incorporate compatibility into existing network operations, and enhance programmability across an environment
- Consistency across compute and network resources for system operators groups that require server-like manageability — extending the Linux experience from compute to the network, lowering costs with common operations for compute and network infrastructure, and simplifying network management, orchestration and automation

- Traditional networking integration that allows network operators to take advantage of existing network configurations and designs, continue using command line interface commands and scripts, and assist operations staff in transitioning from multiple groups to just one as the network migrates to full open networking disaggregation, and
- Full programmability to enhance the integration and control allowed to development and operations (DevOps) teams, down to identifying an object as an individual, manageable entity within the platform.

The new spine-leaf network configuration includes redundant 100Gb, long-range 30-kilometer optics that connect UAB's new GPU-accelerated systems in the DC Blox Data Center to SUSE® Ceph storage in the university's Rust Data Center, along with a 100Gb connection to HPC systems in the UAB 936 building. The network spines are all 100Gb across and down to the leaves, and then a variety of connectivity leads to the individual nodes, depending upon what those nodes do. The nodes themselves communicate via an HDR InfiniBand 100Gb interconnect.

The network configuration also includes new 400Gb switches and a 200Gb non-blocking fabric for the private cloud, which provides access to the DGX A100s systems. At full 100-percent capacity, the UAB research computing team estimates performance of 6.4T/bs (terabits per second) for networking traffic that is internal to the private cloud itself.

"We have a lot of reliability across those network switches — not just performance but reliability," Dr. Zottola says. "The new Dell OS10 platform is built on top of a Debian Linux operating system, which opens the door to an enormous amount of very interesting scenarios that we can build into our network. Not just for running the systems that we have, but for gathering telemetry information about the health of our environment on a continuing basis.

"For example, we will be able to do network performance tests, in line with our operations, using the switches themselves as part of the data gathering and telemetry testing environment. We don't have to build a separate harness to do our testing. We can use the network gear itself to leverage its capabilities to participate in the testing — not just as a passthrough wire that packets run across."

The big vision: a statewide consortium

In the course of designing and building their new 100Gb network fabric, the UAB team had a bigger vision in mind: creation of a statewide consortium that would make research computing resources available to universities across Alabama. One of the foundational elements for this consortium is a reliable, high-bandwidth network fabric.

"We have a dream of creating what we call the Alabama Advanced Research Computing Consortium," Dr. Zottola says. "The first part of that vision has been to learn what the priorities are for the users. To that end, I have been meeting with people on the UAB campus and at universities across the state."

This consortium, which is now in the planning stages, would require a reliable, high-speed network fabric — which is exactly what the team put in place when they built the new 100Gb network fabric.

John-Paul Robinson, HPC Architect at UAB, points out that the central resources for the consortium would be based on university-neutral ground — the commercial DC Blox Data Center. That sense of neutrality is important for a resource that will be shared by multiple universities.

Dell EMC PowerSwitch networking in the data center

Dell EMC data center switching solutions are cost-effective and easy to deploy at any scale, from 1GbE to multi-rate 100GbE, for optimum connectivity within the rack or modular compute chassis, between racks, and between data centers themselves. These switching solutions feature a choice of software options, including Dell EMC OS10 Enterprise Edition and Open Edition, as well as options from the Dell Technologies Open Networking software ecosystem and open source communities — to address virtually any enterprise or service provider use case or environment at any scale.

“With everything we have in place now, we can build it in such a way that our relationship with UAB would be the same as our relationship with any other campus in the state,” Robinson says. “We’re really well positioned now to deliver a service. We took away all those little kinks that were left over from the early days of network expansion and got ourselves to a new platform — one that is ready to support version 1.0 of the consortium.”

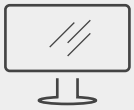
Working with Dell Technologies

In its work to deliver leading-edge HPC resources to the University’s researchers, the UAB research computing team works closely with Dell Technologies.

“We have made giant strides. The collaboration is leading to research that changes the world,” says UAB Vice President and Chief Information Officer Curtis A. Carver Jr., Ph.D.

“It really is a partnership,” he says. “It’s not always even Dell equipment. Dell has the ability to work holistically, to take a big-picture engineering approach. It’s not just about the hardware. They work to identify the right type of resources, connections and services that we will need. But most importantly, they are a partner who helps us think through problems, and find ideal solutions.”

“In today’s world, no one person is an island,” he says. “You need a team to get things done.”



Learn more about Dell EMC advanced computing

Unlock the value of data with artificial intelligence

Share this story