

Establishing a High-Performance Cloud With Mellanox 56Gb Ethernet Solutions

Background

When it comes to advanced scientific and computational research in Australia, the leading organization is the National Computational Infrastructure (NCI). NCI was tasked to form a national research cloud, as part of a government effort to connect eight geographically distinct Australian universities and research institutions into a single national cloud system.

Funding for this project comes from the NeCTAR initiative, the National eResearch Collaboration Tools and Resources Project, which aims to support researchers by providing access to a full suite of digitally-enabled data, analytic and modeling resources, as well as the performance and interconnectivity to use those resources to their fullest. NeCTAR designed a cohesive cloud environment, providing a quick provisioning of a Virtual Machine (VM) and any required support to researchers affiliated with Australian universities.

NCI decided to establish a high-performance cloud, based on Mellanox 56Gb/s Ethernet solutions. NCI, home to the Southern Hemisphere's most powerful supercomputer, is hosted by the Australian National University and supported by three government agencies: Geoscience Australia, the Bureau of Meteorology, and the Commonwealth Scientific and Industrial Research Organisation (CSIRO).

Whether executing data analytics or modeling scientific phenomena like climate change, marine biodiversity, and human genomics, researchers are empowered to share knowledge and data and to rapidly, securely, and cost-effectively deploy and access software applications without the burden of operating their own computer servers, instead drawing

from the shared computational capabilities of a cloud environment.

NCI splits its cloud into two components: the NeCTAR public cloud, open to any Australian researcher; and the OpenStack on Red Hat Enterprise Linux (RDO) based private cloud, available only to designated users, such as NCI's research partners or specific NeCTAR virtual laboratories.

High-Performance Cloud Requirements

Cloud computing relies on converged infrastructure and shared services to create economies of scale. While today's data centers are growing to handle the world's exponential data growth, budgets cannot grow at the same exponential pace. Therefore, the scalability of a cloud is paramount.

For a cloud to efficiently scale across a high number of nodes and provide high performance without bottlenecks, the ultimate in today's interconnect technology and network hardware is required. A cloud such as NCI's must offer:

- Exceptionally high bandwidth to enable maximum throughput of data
- The lowest possible latency to ensure lightning-fast data transfers
- Solid-state disk (SSD) storage for high IOPS
- Open, transparent infrastructure that can orchestrate and manage both the hardware and software for the benefit of its end users
- A high level of manageability
- Full integration with OpenStack



OVERVIEW

By implementing Mellanox's end-to-end Ethernet solution, NCI has shown significant gains in performance, scalability, and manageability.

Solution

When NCI searched for an interconnect provider for its high-performance cloud, it was important that their choice be a cloud vendor that has significant experience with scalable environments. Mellanox quickly emerged as the ideal partner.

Mellanox offered its new CloudX platform, and it demonstrated the performance value, scalability, and savings of using 56Gb Ethernet in a cloud environment. Mellanox switches provide higher bandwidth per unit, saving on both physical footprint and IT costs; plus, the higher bandwidth future proofs the infrastructure for even greater demand in data transfer rates.

With 224 nodes running 56GbE across 22 36-port switches (8 aggregation spines and 14 Top-of-Rack leafs), Mellanox connects 3200 cores in a non-blocking architecture. NCI then splits those nodes into its public and private cloud environments.

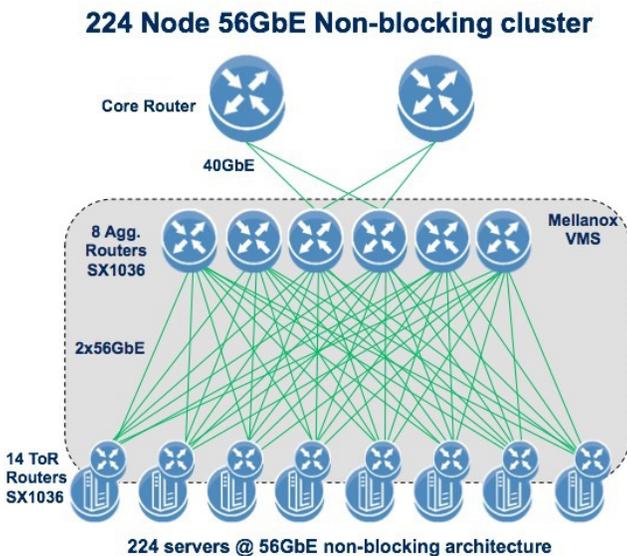


Figure 1. NCI's high performance cloud architecture

The Virtual Modular Switch® (VMS) configuration and bring-up tool can be fully automated, from the early planning stages until it is operational. VMS via puppet provides automated configuration for two Mellanox Ethernet switches connected with an MLAG. These switches can be installed either at the edges of the VMS or in other network topologies.

In the Mellanox solution, VMS provided a full Fat Tree with RDMA over Converged Ethernet (RoCE) and with link aggregation and redundancy, handled via ECMP over OSPF Layer 3 protocols. RoCE allows remote direct memory access over an Ethernet network, and is used for storage acceleration.

The RoCE protocol is used by Mellanox adapters to decrease latency, increase bandwidth, and reduce the CPU consumption required to process traffic, thereby making more CPU available for VM workloads.

NCI recognized that as part of the OpenStack community, Mellanox's RDMA solution is integrated into both the Neutron (communication) and Cinder (storage) aspects of OpenStack, and is at the development forefront of plugins and support for related software and middleware development packages.

Results

By implementing Mellanox's end-to-end Ethernet solution, NCI has shown significant gains in performance, scalability, and manageability.

Benchmarking tests of the cloud have shown remarkable results:

- 47.6 Gb/s bandwidth when writing from VM to VM using VMs with 1 core each
- 49.5 Gb/s bandwidth when writing from VM to VM using VMs with 14 cores each
- Latency under 45 microseconds for up to 128-bit message size
- Deployment of the entire cloud in only a couple of weeks
- CloudX demonstrated the best IOPS and bandwidth results of all cloud settings
- Superior IOPS results

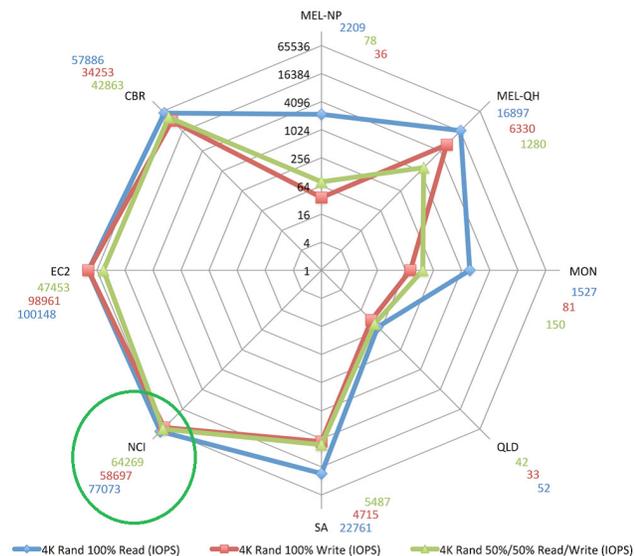


Figure 2. IOPS results using flexible I/O testing tool with local disk node

Conclusion

Mellanox's CloudX offered the ideal components to install a high-performance and scalable cloud environment at NCI. Mellanox provided NCI with convergence of its primary requirements: higher bandwidth, lower latency, lower costs, smaller physical footprint, fewer cables, easier management and scalability, support for more users, and future proofing.

Mellanox offered the open, transparent infrastructure that NCI sought by providing flexible infrastructure

coupled with open source software and ongoing participation in the OpenStack community.

Most of all, CloudX clearly placed above the competition with its superior performance as a scalable high-performance cloud, providing high bandwidth, exceptionally low latency, and an increased number of VMs for the same compute power.

CloudX stood out as the right choice for NCI to bring the most efficient cloud to a large volume of users for their research applications.



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com