CASE STUDY

High-Performance Computing (HPC)
with Intel® Omni-Path Architecture

Architecting a Converged, World-Class, Multi-Mission HPC Cluster

Intel® Omni-Path Architecture gives Bridges supercomputer the performance to support widely ranging workloads

Hewlett Packard Enterprise

At a Glance

PSC built Bridges as a multi-mission supercomputing resource. System includes 908 of advanced compute nodes:

- Two x 12 TB HPE Integrity Superdome* X with 16 Intel® Xeon® processors E7-8880 v3
- Two x 12 TB HPE Integrity Superdome X with 16 Intel Xeon processors E7-8880 v4
- 8 x 3 TB HPE ProLiant* DL580 with 4 Intel Xeon processors E7-8860v3
- 34 x 3 TB HPE ProLiant DL580 with 4 Intel Xeon processors E7-8870v4
- 752 x 128 GB HPE Apollo* 2000 servers with 2 Intel Xeon processors E5-2695v3
- 48 x 128 GB HPE Apollo 2000 servers with 2 Intel Xeon processors E5-2695v3, and two GPUs
- Intel® Omni-Path Architecture (Intel® OPA) 48-port switches create tightly coupled partitions of 1,176 cores at full bisection bandwidth
- Award-wining design

When you look at world-class High Performance Computing (HPC) installations today, more institutions with significant compute demand, such as Pittsburgh Supercomputing Center (PSC), are specifying architectures that serve a wide range of users instead of the traditional scientist running complex simulations. More importantly, these systems are being designed to enable discoveries for users who aren't computer scientists—researchers who write in higher-level languages, such as R* and Python*. This high level of flexibility was one of the design mandates for PSC's Bridges supercomputer, according to Nick Nystrom, Sr. Director of Research at PSC, and Intel OPA played a significant role in achieving the design goals.

Challenge

"The way to reach many people today and provide them the computational facilities they need is to make resources available without requiring them to become HPC programmers," commented Nystrom. Many people would rather work in a familiar desktop environment and with their own tools, rather than having to learn to program in C++ and MPI. And, while some users need these high-performance resources for simulation, others are expressing their problems in different paradigms, leveraging components across HPC, big data, and other domains. For years, traditional HPC architectures have focused on a single domain that could not satisfy all these needs. PSC needed a much more flexible system design, yet one that would not sacrifice performance.

Solution

Hewlett Packard Enterprise (HPE) built and installed Bridges in two Phases – with Phase 1 including 822 servers designed around Intel Omni-Path architecture, while Phase 2 added 86 additional nodes (and 126TB of system memory). Bridges was designed to be partitioned into groups of nodes for different workloads, so the resources could simultaneously service a variety of computing requirements. But, doing so entailed providing very high IO bandwidth. PSC chose Intel OPA fabric to meet these converged computing and IO requirements.

Intel® OPA Connects Bridges Together

Bridges' architecture is formed from 846 advanced compute nodes, which include four large-memory HPE Integrity Superdome X servers with 12 TB of RAM each, 42 HPE ProLiant DL580 servers with 3 TB of RAM each, and 842 HPE Apollo 2000 servers with 128 GB of RAM each. An additional 42 HPE Apollo 2000 servers with 128 GB of RAM each provide persistent databases, web services, and system management functions. The system leverages Intel OPA switches to form an enhanced leaf-spine network layout, interconnecting a heterogeneous architecture that supports Bridges' widely ranging workload mandate.

Results

Intel OPA gives Bridges the performance it needs for the workloads that the system runs. The fabric delivers high bandwidth, extremely low latency, and very high injection rates. According to Nystrom, PSC's first benchmarks revealed an impressive 12.37 GB/s with latency at 930 nanoseconds, which beat expectations.

The Intel OPA Edge switches help create an effective architecture for Bridges. Intel OPA's 48-port Edge switch enables PSC to run tightly coupled applications on approximately 1000 cores, "which is ideal for nontraditional communities and much of traditional HPC," stated Nystrom. The 48-port switches interconnect islands of 42 nodes, each with 28 Intel Xeon processor cores, creating tightly coupled partitions of 1,176 cores, and running at full bisection bandwidth across the 100 Gbps switch. "Had there been only 36-port switches, we would have had to build a much more complicated network to get full bisection bandwidth to 1,000 cores," added Nystrom. Bridges is also able to provide up to 22,400 cores (or around 27,000 cores, including the large-memory nodes) for projects that require higher levels of computational resources, and which are not bisection-bandwidth bound. "Intel OPA allowed us to achieve those complementary goals," he concluded.

Bridges' first mission in February of 2016, with just a few nodes, was running novel and innovative visualization applications in the MIDAS MISSION Public Health Hackathon.

Since then, the system has been used across a wide range of workloads that it was designed for, enabling incredible discoveries across multiple scientific fields.

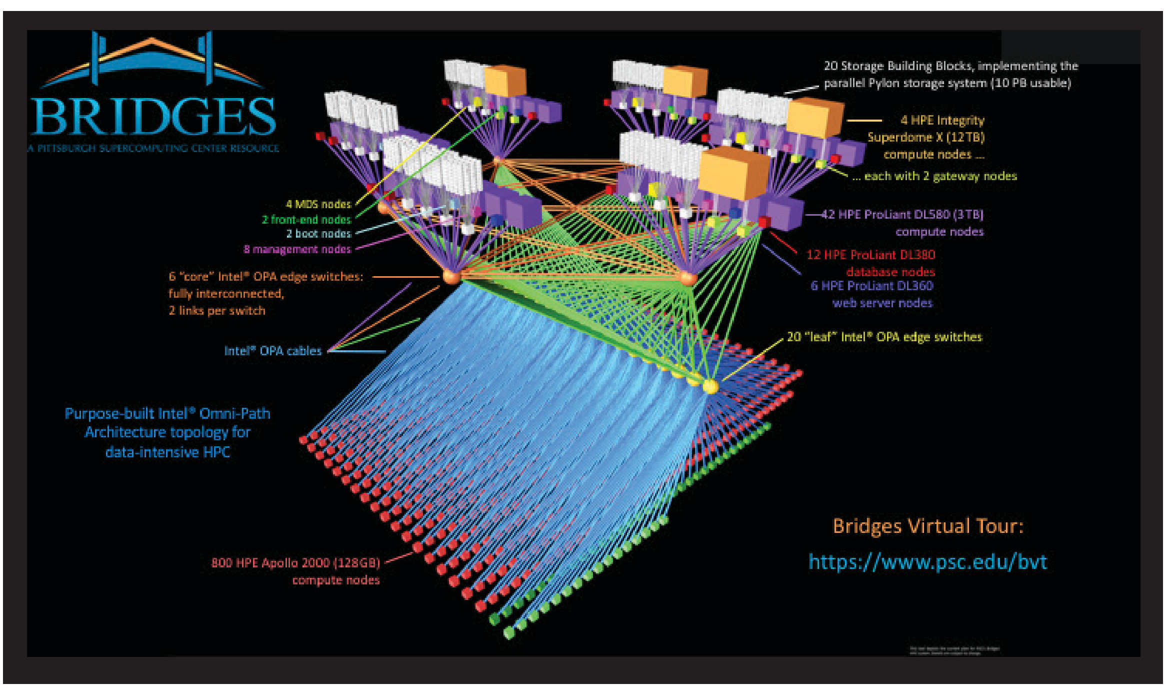
Gambling Against Bridges is a Poor Bet

In January of 2017, Bridges beat four of the world's top poker players at the Rivers Casino in Pittsburgh, PA, U.S.A. in a human against silicon heads-up, no-limit Texas hold-em tournament. Libratus, an artificial intelligence (AI) program developed by the Carnegie Mellon University School of Computer Science, utilized 19 million core hours on Bridges to first calculate its strategy. During the tournament, the program ran on 600 of Bridges' 752 regular memory nodes. Then, after every night of play, Libratus refined its own strategy by adapting to the players' changes in their strategies. For PSC and CMU, Bridges paid off nicely.

The Best of Bridges—Awards from the HPC Community

Bridges garnered two international awards for PSC from HPCwire Readers' and Editors' Choice Awards at SC16.

- **Best Data-Intensive System (End User Focused)**—Bridges' design mandate was around enabling more users on HPC without requiring them to write to familiar HPC programming models. To that end, this award was given for the system's ability to open high-performance com-



puting to more researchers in the physical, social, and computer sciences.

- **Best Use of High Performance Data Analytics**—Collaborating with Harvard University and the Allen Institute for Brain Science, Bridges analyzed about 35 terabytes of data to reconstruct the "wiring diagram" of part of a mouse brain's visual center. "The work identified nerve cells that respond to specific visual elements, and how these cells pass along their signals. The finding is a major step in reconstructing brain connections in a way that helps scientists understand how the millions of nerve cells in the brain communicate and work together."

Bridges' Large Memory Nodes Supporting De Novo Assemblies

The most memory-intensive work in genomics is assembly, and now de novo transcriptome assemblies from the Galaxy platform are using Bridges' 12 TB large memory nodes for rapid assembly of massive RNA sequence data. Assembly requires loading the entire sequencer data output into memory and then running code that finds the matches among the millions of pieces of the transcriptome. "This type of work is ideal for Bridges' large memory nodes," commented Nystrom.

Solution Summary

Bridges is helping make new discoveries possible, by putting supercomputing in the hands of users who are not HPC programmers, while also delivering supercomputing resources to traditional HPC scientists. PSC's system was architected to enable a wider variety of research across different compute domains (traditional HPC, big data, visualization, and machine learning/artificial intelligence) by a broader range of users and scientists. Built on HPE Superdome, ProLiant, and Apollo servers, Intel OPA interconnects the system into a powerful solution for science and research.

Where to Get More Information

Learn more about Bridges at www.psc.edu/resources/computing/bridges.

Learn more about Intel Omni-Path Architecture at www.intel.com/omnipath.

Solution Ingredients

- Four HPE Integrity Superdome X servers with 12 TB of memory—Intel Xeon processor E7-8880 v3 and v4
- 42 HPE ProLiant DL580 servers with 3 TB of memory—Intel Xeon processor E7-4800/8800/v4/v3
- 848 HPE Apollo 2000 servers with 128 GB of memory—Intel Xeon processor E5-2695v3
- Intel® Omni-Path Host Fabric Adapters
- Intel® Omni-Path Edge Switches (48 ports)

1 https://www.psc.edu/news-publications/2437-bridges-brain-reconstruction-win-hpcwire-awards