CASE STUDY

High Performance Computing
Intel® Omni-Path Architecture

University of Colorado Boulder's RMACC Summit Supercomputer Broadens Subatomic to Galactic Horizons

The value of an Intel hardware and software stack pays off for scientific discovery

Summit

- An Intel® Scalable System Framework system utilizing Intel Xeon and Intel Xeon Phi processors and Intel OPA
- Dell EMC PowerEdge* C6320 Server nodes
- 2.5X faster than previous supercomputer
- Serves wide variety of science and engineering
- Joint effort of the University of Colorado Boulder and Colorado State University

University of Colorado Boulder's (CU Boulder) Research Computing Group, part of the Rocky Mountain Advanced Computing Consortium (RMACC), had been running on their Janus supercomputing cluster since late 2010. In 2016, they installed their next-generation system named RMACC Summit, which is supported by CU Boulder and Colorado State University (CSU), to enhance ongoing research and discovery in a variety of areas. RMACC Summit is built on Dell EMC PowerEdge* C6320 Server nodes with Intel® Xeon® processor E5 and E7 family with Intel® Xeon Phi™ processor nodes planned in 2017, all interconnected by Intel® Omni-Path Architecture (Intel® OPA) fabric. RMACC Summit is expected to be nearly 2.5X faster with only one-third the node count and almost half the cores.

Challenge

Janus was a 1,368-node, 16,416-core cluster based on Intel Xeon processor 5000 sequence (formerly codenamed Westmere), delivering 184 TFLOPs. While Janus had contributed to much research, as the university approached 2016, Janus was five years old, new technologies were emerging, and "support costs would rise as warranties began to expire", said Jonathon Anderson, part of CU Boulder's Research Computing Group.

"With the newest generation of interconnect and processors, the amount of bang you get for your space and electrical costs is just so much greater with newer equipment," added Peter Ruprecht, Senior HPC Analyst in the Research Computing Group." Thus, CU Boulder began the process of acquiring a replacement for the aging system. RMACC Summit began installation in the Summer of 2016.

According to Ruprecht, CU Boulder has many dozens of ongoing research projects to support. "Traditionally, the majority have come out of the hard sciences and engineering. Physics, astrophysics, aerospace, and material science are what you might consider the largest of our users. But we nearly support every department on campus." RMACC Summit has to support this wide variety of users with a host of different needs, coding capabilities, and programming models. Some scientists want to use the power of High Performance Computing (HPC) without having to be computer scientists. Others are domain scientists, interested in computing and writing their own applications. And there are computer scientists developing algorithms. "It's a wide range. We deal with everything from undergrads to grads to post docs and faculty," commented Ruprecht.

Solution

CU Boulder chose Intel® architecture and technologies as the foundation for RMACC Summit, including Intel OPA for the next-generation fabric. "When writing an NSF proposal, it's good to have things that are relatively new, like the Intel fabric" said Anderson, "So, we wanted to be a test bed for other sites that might be interested in deploying Intel OPA. We thought there might be interesting experiences that could come out of that."

"Performance for the system was a key metric," according to Mark West of Intel. "When we talked with the University, we emphasized that with Intel OPA fabric instead of an alternative, they'd spend less on the network without sacrificing the fabric's throughput or latency. A single Intel OPA switch contains 48 ports, compared to 36 ports in other solutions. That means fewer switches delivering high bandwidth across more servers. They could direct more of their budget to additional compute nodes to improve overall system performance."

This aligned with Anderson's position on technology. "I'm a kind of standards-oriented person, so I like the idea of going down the Intel HPC path as far as we could, or as wide as we could, with a full-stack solution as much as possible in a single cluster."

RMACC Summit was purchased and began installation in mid-2016. Dell EMC and Intel worked together to bring up the cluster and network. It was installed as a 395-node, 9600-core system built on Intel Xeon processor E5-2680 v3 and Intel Xeon processor E7-4830 v3. Another 20 nodes with Intel Xeon Phi processor 7200 series were recently added. Eleven GPU nodes run Intel Xeon processor E5-2680 v3. The entire system runs on a 100 Gbps Intel OPA fabric.

"The cluster is mostly built on Dell EMC PowerEdge C6320 Rack Servers," said Anderson. "What we have today is a denser version of our previous cluster, but with more heterogeneity, with GPUs and high-mem nodes." According to Ruprecht, RMACC Summit's aggregate measured performance (Rmax) with the new nodes is 375 TFLOPs. The 20 Intel Xeon Phi processor-based nodes alone deliver 34.5 TFLOPs. RMACC Summit has been rolling out into production over the first part of 2017, with early users already performing ground-breaking research.

Result

CU Boulder is pleased with their new supercomputer. "The synthetics have all been good," said Anderson, "and the first impressions of our early users have all been good." The CU Boulder team ran RMACC Summit through acceptance testing using the Ohio State University (OSU) benchmark suite: paired bandwidth tests, paired latency tests, and all-to-all across the various scales of the fabric. "From the beginning of testing we hit all the benchmarks we expected."

"The admin and management tools that are part of the Intel OPA software stack have been really good," commented Ruprecht. "As system-side people, that's an important factor for us as well. Not only the hardware works as expected and designed, but that we can manage it easily as well."

"From an applications perspective, users who have ported their codes onto it haven't had any real issues if they had existing MPI codes," said Anderson.

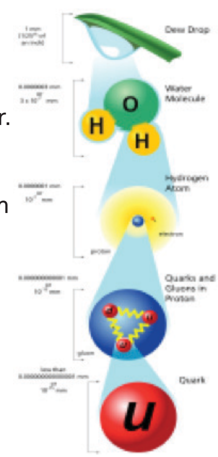
"And, when using the Intel version of MPI implementation, it has been completely seamless for users to take advantage of the fabric," added Ruprecht.

The users they are talking about are researchers like Professor Anna Hasenfratz, a theoretical physicist.

Theories in BSM Particle Physics

"I'm studying Beyond Standard Model prototype systems, specifically in high-energy particle physics" said Dr. Hasenfratz. "My colleagues and I are exploring possibilities of what makes up the Higgs-boson particle, which was observed in 2012 using the Large Hadron Collider (LHC)." Beyond Standard Model (BSM) theories explore the unknown aspects of particle physics that cannot be explained by the well-known Standard Model of physics. "I'm working on theories of a model, which is, I believe, quite an attractive possible BSM description. And the best way to understand the properties of these models is through computer simulations."

Her colleagues are Peter Boyle and Oliver Witzel, researchers in the School of Physics and Astronomy at the University of Edinburgh. Professor Hasenfratz is leveraging a code base called Grid[1] that Dr. Boyle co-developed at the Intel® Parallel Computing Center (IPCC) in Edinburgh. Grid is a library written for lattice quantum chromodynamics (QCD) calculations used in particle computer simulations. Lattice QCD principles are used to model the interactions of quarks and gluons, the sub-atomic particles in hadrons; protons and neutrons are hadrons. "Lattice QCD principles are applicable to BSM simulation, allowing us to explore BSM theories," said Dr. Hasenfratz.



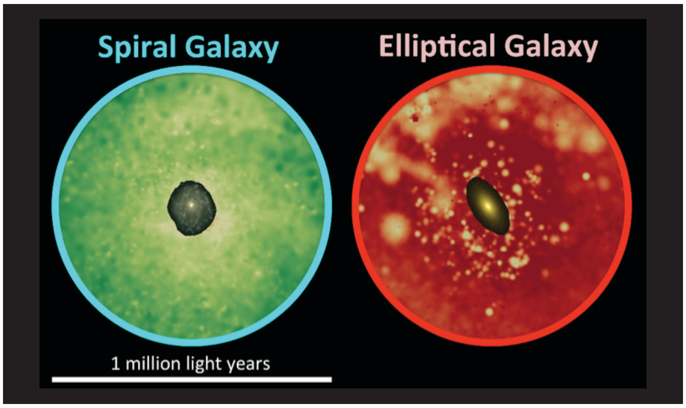
"Grid started out as a project to update an existing lattice QCD code base to optimize for the Knights Corner version of Intel Xeon Phi processor," commented Dr. Boyle. "But when we began to explore it more deeply, and considered other interesting programming techniques used in game development, we decided to rewrite the entire code in C++11, which took critical portions down from 100,000 lines of code to just 200 and gave Grid flexibility the previous code base didn't have." Now optimized on the Intel Xeon Phi processor 7200 series, single-node benchmarks reveal the code runs nearly 3X faster than on the previous Intel Xeon Phi processor and on both Intel Xeon processors E5 v3 and v4 families.

"We were lucky to have Summit come on line and be given early access to it," added Professor Hasenfratz. "We have been running Grid production code on it, and we are happy with the results so far. We expect important insights soon."

Simulating Entire Galaxies

Going from the sub-atomic to the galactic, Dr. Benjamin Oppenheimer of CU Boulder uses RMACC Summit to simulate entire galaxies to study the distribution of gases in and

1 https://arxiv.org/abs/1512.03487



around a galaxy. “When you look at something like a spiral galaxy, like our Milky Way, you just see the tip of the iceberg, the stellar components with its grand-design spiral arms,” said Dr. Oppenheimer. “There is a lot of gas that is much more massive than the galactic disk. But it’s very diffuse, low density, and can’t be optically observed. So, we use different instruments, like the Cosmic Origin Spectrograph (COS) on Hubble, to measure absorption by the gas in the ultraviolet spectrum of this circumgalactic medium to measure its distribution. I simulate this distribution over time using supercomputers, like Summit.”

Dr. Oppenheimer’s simulations become part of creating realistic galactic simulations to help astrophysicists understand the physics of galaxies of different types—evolving from their early beginnings, their star formations, hydrodynamics, gravity, growing of black holes and other astrophysics. But one of his key objectives is to create “zoom” simulations that show more detail.

“Galaxies are massive objects with billions of solar masses across hundreds of thousands of light years of the galactic disk. But I’m simulating the observed gases across one hundred times that distance. When you run a simulation, you’re not resolving individual stars, but an integrated population of stars. My simulations resolve the stellar population at eight times better resolution than the current suite of runs produced in 2014. On Summit, we aim to increase the resolution by another eight times for a total of 64 times better resolution.”

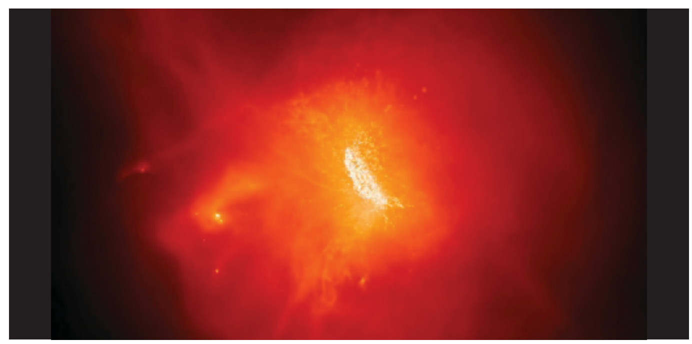
Other Active and Planned Research

Another research team is using RMACC Summit to analyze satellite data, looking at earth record data in a wide variety of different wavelengths from different kinds of satellites. That data needs to be combined into products that are more useful to the scientific community. This kind of processing requires a fast storage system and a lot of compute nodes.

RMACC Summit is a collaborative project of CU Boulder, Colorado State University (CSU), and the Rocky Mountain Advanced Computing Consortium (RMACC). The expansion to 468 nodes was the result of ‘condo’ purchases by faculty at CU Boulder, CSU and University of Colorado Denver. While CU Boulder and CSU researchers are currently the main users, it will soon be made available to students and faculty across the six-state Rocky Mountain Region.

Solution Summary

The RMACC Summit cluster, supported by CU Boulder and CSU, is built on Dell EMC PowerEdge C6320 Server nodes with Intel architecture and technologies, following Intel’s model for HPC called the Intel® Scalable System Framework (Intel® SSF). It combines Intel Xeon processors, Intel Xeon Phi processors, Intel OPA high-performance fabric, and software to create a holistic infrastructure designed for performance and scalability. RMACC Summit is early in its production cycle, but it is already being used for discovering new theories in particle physics, for advanced galactic simulations, and high-speed data processing.



Where to Get More Information

Learn more about RMACC Summit at https://rc.colorado.edu/resources/compute/summit.

Learn more about Intel Omni-Path Architecture at http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-architecture-fabric-overview.html.

Learn more about Intel Xeon Phi processor at http://www.intel.com/content/www/us/en/products/processors/xeon-phi.html.

Learn more about Dell EMC PowerEdge* C6320 Servers at http://www.dell.com/en-us/work/shop/povw poweredge-c6320

Solution Ingredients

- 395-node, 9600-core cluster
- Dell EMC PowerEdge C6320 Server nodes
- Intel® Xeon® processor E5-2680 v3
- Intel® Xeon® processor E7-4830 v3
- Intel® Xeon Phi™ processor 7200 series
- Intel® Omni-Path Architecture fabric