CASE STUDY

Rockport fabric helps reveal the secrets of the cosmos

CLIENT: **Durham University**

CHALLENGE: Improve the performance and reliability of advanced workloads to better simulate and understand the universe.

SOLUTION: A Rockport fabric that delivers ultra-low latency and unprecedented monitoring capabilities.

RESULTS: 10%-20% better performance on some advanced workloads, insight to understand and improve how data flows through the fabric.

Overview

The Institute for Computational Cosmology (ICC) at Durham University in the UK is one of the leading research facilities studying the cosmos. Under the direction of Dr. Alastair Basden, technical manager of the COSMA High Performance Computing (HPC) Cluster at DiRAC/Durham University, researchers from around the world have access to powerful resources to investigate the complex interactions of matter, energy and space.

"We start with the Big Bang and we propagate the universe forward in time, really looking at things that we can put into our models of the universe," says Dr. Basden. "Understanding the science of things like dark energy, dark matter and black holes, we can then change these simulations in a way that allows them to match what we see with telescopes."

After being introduced to the Rockport scale fabric technology – a radical shift from 30-year-old spine and-leaf architectures – he decided to deploy Rockport on 224 nodes in Durham's COSMA7 cluster to explore exascale workloads and run codes to run on future exascale systems.

Congestion Kills Network Performance

Congestion in HPC environments has been one of the main culprits slowing workloads and, in some cases, preventing them from completing. This delays research and makes an expensive process even more costly.

The ICC's massively parallel performance- and data-intensive research, such as particle hydrodynamics and black hole collision modeling, requires tremendous computational power, where traditionally the network that interconnects the nodes becomes a limiting factor. Interconnect-related performance and congestion create bottlenecks when processing advanced computing workloads due to the spine-and-leaf architecture of traditional centralized switching. The result is slower and more unpredictable workload completion times, and underutilization of expensive compute and storage resources.

"We use a lot of networking," says Dr. Basden. "And we're very dependent on both network bandwidth and latency. If there's heavy use of the network, then a particular application can take 10%, maybe even 20% longer to complete, which means it's eating up research project allocations."

> *Based on the results and our first experience with Rockport's scalable architecture we were confident that Rockport is now a serious contender to improve our exascale modeling performance – all supported by the right economics.*

Dr. Alastair Basden
Technical manager,
COSMA High Performance
Computing Cluster at
DiRAC/Durham University

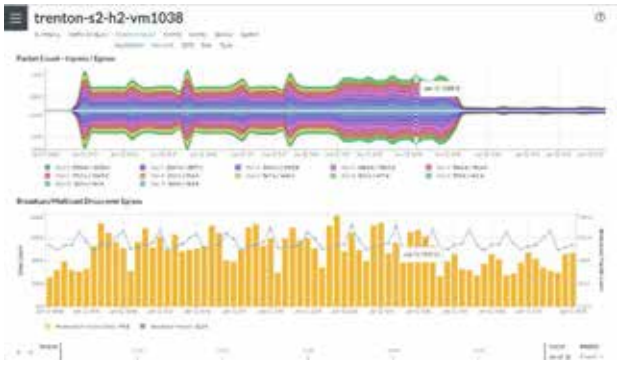
Scalable Fabric With Improved Visibility

The Rockport fabric distributes the network switching function to the COSMA7 endpoint devices (nodes), which directly interconnect to become the network fabric. The Rockport fabric eliminates the layers of switches along with their inherent congestion and latency. Rockport NC1225 Network Cards are installed in every node, and are directly connected to 12 other nodes using a passive optical patch panel configured in a six-dimensional torus.

The solution also includes the Autonomous Network Manager (ANM), which provides unparalleled insights into network performance, both in real time and historically.

"ANM allows us to see in real time what's going on in the network and where traffic is flowing," says Dr. Basden.

"Application developers can find out where congestion is occurring within their codes and then make changes, such as the order in which things are sent from each node, which can really help application performance."



The Rockport Autonomous Network Manager provides real-time visibility into network performance and traffic flows.

When combined with advanced congestion controls and other performance enhancements, the Rockport solution ensures compute and storage resources don't sit idle waiting for data, and researchers can count on predictable workload completion times.

Better Performance on Loaded Networks

Dr. Basden initially installed the Rockport fabric on the DINE cluster, where it was tested against a 100 GB InfiniBand EDR network. On a 16 node Rockport testbed, the results showed equal or better performance to InfiniBand, long considered the industry's network incumbent.

"When we used AREPO (moving-mesh hydrodynamics code for astrophysics) – which is very sensitive to network congestion – on the Rockport network and we introduced increasing levels of congestion, performance didn't suffer. This is unlike our current EDR InfiniBand, where AREPO really started to slow down as congestion increased," said Dr. Basden.

Based on those results, Durham decided to expand the Rockport deployment and install half of its COSMA7 cluster (224 nodes) with a Rockport fabric.

"At DiRAC we like to build our systems tailored for the jobs they're running," said Dr. Basden. "Where we find that particular workloads run better on Rockport, then we'll consider building larger systems with a Rockport fabric to achieve better performance with those workloads."

About Durham University

Durham University is a globally outstanding centre of teaching and research based in historic Durham City, UK.

We are a collegiate university committed to inspiring our people to do outstanding things at Durham and in the world. We conduct boundary-breaking research that improves lives globally and we are ranked as a world top 100 university with an international reputation in research and education (QS World University Rankings 2022).

We are a member of the Russell Group of leading research-intensive UK universities and we are consistently ranked as a top 10 university in national league tables (Times and Sunday Times Good University Guide, Guardian University Guide and The Complete University Guide).
For details see: durham.ac.uk/about

About DiRAC

DiRAC stands for Distributed Research Utilising Advanced Computing. The DiRAC High Performance Computing facility provides cutting-edge supercomputing resources for UK researchers working on world-leading scientific calculations across a wide range of areas, including particle physics, astrophysics, cosmology and nuclear physics. It comprises supercomputers at Cambridge, Durham, Leicester and Edinburgh, each designed to support specific types of calculations. DiRAC also provides access to a team of expert research software engineers to help researchers make the most efficient use of the available computing resources.
For details see: dirac.ac.uk

About ExCALIBUR

Exascale Computing ALgorithms & Infrastructures Benefiting UK Research (ExCALIBUR) is a UK research program that aims to deliver the next generation of high-performance simulation software for the highest-priority fields in UK research.

The Hardware and Enabling Software Programme (H&ES) provides grant funding for testbed projects that give UK researchers early access to novel hardware and software technologies that may play a part in future exascale systems and services. H&ES testbeds are generally small-scale systems that are provided on an experimental basis, and clearly delineated from formal service provision to avoid interference with production workloads.

Researchers can use the testbeds to validate the portability and performance of their codes, and to explore the potential of new paradigms such as workflow offload to data processing units or wafer scale compute engines. For details see: excalibur.ac.uk