



# A Universe of Data

Durham University works with Dell EMC to build an expansive and scalable HPC storage environment for cosmological research.



Scientific Research

United Kingdom

## Business needs

Durham University needs an efficient and scalable HPC storage environment to support cosmological studies and other scientific investigations.

## Solutions at a glance

- PowerEdge R640 servers with Intel® Xeon® processors
- Ready Solutions for HPC Lustre Storage with ZFS
- PowerVault ME4024 and ME484 storage arrays
- PowerSwitch networking and Mellanox InfiniBand EDR interconnects

## Business results

- Delivering efficient storage for petabytes of data
- Reducing the data storage footprint
- Accelerating research with high performance storage
- Helping researchers unlock the secrets of the universe

The cluster's storage system can move data at I/O speeds up to

**200** GB/s



The cluster's Dell EMC PowerVault bulk storage has a usable capacity of

**2.6** PB



# Big data from a big universe

When you're in the business of running large-scale simulations of the Universe, you live in a world of really big data. That's the way it is at Durham University, a leading research institution in the fields of astrophysics and particle physics.

What is the Universe? What is it made of? What is dark matter? What is dark energy? These are the types of questions explored by distinguished researchers working with the University's Institute for Computational Cosmology (ICC). These cosmologists calculate and model structures of our universe using high-performance computing clusters hosted by the ICC.

Among those clusters is a new memory-intensive system spearheaded by the UK's Distributed Research utilizing Advanced Computing (DiRAC) organization and designed for cosmological calculations. Named COSMA 7, the cluster features 452 compute nodes and 220 terabytes of memory, making it one of the largest HPC memory facilities available to the research community.

## A storage environment built for huge simulations

The type of research carried out on the COSMA 7 cluster requires a robust storage environment built to handle — literally — astronomical amounts of data. To meet this need, the DiRAC resources team at Durham University chose a comprehensive Dell EMC storage solution.

This storage environment spans the flow of data in the COSMA 7 cluster, from scratch storage on the front end, to archival storage on the back end. The environment includes:

- [Dell EMC PowerEdge R640](#) rack servers with NVMe SSDs for scratch storage, or space dedicated to temporary storage
- [Dell EMC Ready Solutions for HPC Lustre Storage](#) with a ZFS (Z File System) backend for file system management
  - [Dell EMC PowerVault ME4024](#) storage arrays for SSD storage for file system metadata
  - [Dell EMC PowerVault ME484](#) dense JBOD expansion enclosure for object storage
- Spectra Logic tape library for archival storage

This end-to-end combination of storage products and technologies enables the COSMA 7 team to store data in an efficient and economical manner, while delivering a level of performance that helps keep researchers productive.

Alastair Basden, technical manager at Durham University, notes that the PowerEdge R640 rack servers with NVMe for scratch storage provide very fast I/O with a parallel file system that reads and writes data to disk very quickly.

“What that means is that our researchers can dump their data to disk very quickly, and read it from disk very quickly when they need to restart,” he says. “By reducing the length of time that it takes to write and restart files, we're improving the efficiency of the cluster. That means we can really concentrate on the core computation rather than waiting for files to be saved.”



## All the right stuff

Among other leading-edge features, the COSMA 7 supercomputer includes:

- A software-defined Lustre storage system that runs on Dell EMC PowerEdge servers with NVMe SSDs and Intel® processors
- A fast I/O storage area network built with Dell EMC PowerVault storage that provides 480 terabytes of capacity with a maximum I/O speed of about 200 gigabytes per second
- A high-speed network that includes Dell EMC PowerSwitch networking and Mellanox InfiniBand EDR interconnects

## Lustre, ZFS and PowerVault

In a groundbreaking feature, the COSMA 7 storage environment uses Dell EMC Ready Solutions for HPC Lustre Storage with a ZFS (Z File System) backend for file system management of data held in Dell EMC PowerVault storage arrays. ZFS delivers a rich array of features that can be beneficial for HPC storage. These features include snapshots, end-to-end data integrity, performance optimizations, software RAID and more.



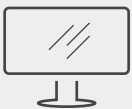
In another important gain, the density of the Dell EMC storage arrays allows the COSMA 7 team to pack a lot more data into the same footprint when compared to the previous-generation storage environment, Basden says.

“The old system was about 2.5 racks in size, and we replaced it with something that is about two-thirds of a rack,” Dr. Basden says. “So we have achieved significant space savings. More importantly, our power consumption was reduced from over 22kW to only 6kW, leading to massive savings in CO<sub>2</sub> production.”

## Unlocking the secrets of the universe

With its rich mix of technologies, COSMA 7 delivers the robust high performance compute and storage resources that scientists require when they are trying to unlock the secrets of a very big and complex universe. This work involves enormous amounts of data that must be processed at high speeds. A single simulation that runs in COSMA 7 can produce hundreds of terabytes of data.

“About 75 percent of the universe is made up of dark matter that we don’t understand,” Dr. Basden explains. “We know very little about it, aside from a few hypotheses and a few ideas about what it might be. By running these simulations, we are able to find out more about it. And, of course, when we do that, we begin to understand more about what the universe is made from. COSMA 7 opens all sorts of science that can be done, and that can actually have an impact over ideas that we can’t even conceive right now.”



[Learn more](#) about  
Dell EMC HPC



[Learn how](#) to unlock the  
value of data with AI



[Share this story](#)

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Intel, Intel Logo are trademarks of Intel Corporation in the U.S. and/or other countries. Other trademarks may be trademarks of their respective owners. This case study is for informational purposes only. The contents and positions of staff mentioned in this case study were accurate at the point of the interview conducted in January 2019. Dell and EMC make no warranties — express or implied — in this case study.

