Caringo Swarm

STFC Scientific Computing Department Deploys Caringo Swarm to Streamline Management & Data Access

Benefits

- Equivalent performance to parallel file systems for S3 throughput
- NFS and S3 access to the same objects
- Easy management of thousands of tenants and billions of objects
- Simple internal and external file access and file sharing

From monitoring volcanoes and earthquakes to crop yield analysis, wildlife and insect migratory patterns, JASMIN is giving mankind unrivaled insight into our natural world. The JASMIN facility is a "super-data-cluster" that delivers infrastructure for data analysis funded by the Natural Environment Research Council (NERC), the UK Space Agency (UKSA), educational institutions and private industry. It is managed jointly by The Science and Technology Facilities Council (STFC) Scientific Computing Department (SCD) and the Centre for Environmental Data Analysis (CEDA), which is part of STFC Rutherford Appleton Laboratory (RAL) Space.

Datasets come from various projects and a mix of instruments, including 12 ESA satellites that stream 10 Terabytes of data daily. Many datasets, both model and observed data, are too big to be easily shipped around; JASMIN enables scientists to bring their processing to the data.

The JASMIN infrastructure provides compute and storage linked together by a high-bandwidth network in a unique topology with significant compute connected with much greater bandwidth to disk than is typical of a normal data center. JASMIN has a supercomputer's network and storage, but without quite as much compute. JASMIN provides four basic services: storage (including disk and tape), batch computing, hosted computing and cloud computing.

The infrastructure for JASMIN is designed, deployed and managed by SCD's Jonathan Churchill and his team at STFC RAL. STFC is a world-leading, multi-disciplinary science organization with the goal to deliver economic, societal, scientific and international benefits to the UK and its people—and more broadly, to the world. STFC supports an academic community of around 1,700 in particle physics, nuclear physics, and astronomy (including space science), who work at more than 50 universities and research institutes in the UK, Europe, Japan and the United States—including a rolling cohort of more than 900 PhD students. STFC's large-scale scientific facilities in the UK and Europe are used by more than 3,500 users each year, carrying out more than 2,000 experiments and generating around 900 publications. The first phase of JASMIN was funded in 2011 and deployed in 2012. Since then, hundreds of Petabytes of data have been processed and thousands of researchers have accessed the platform.

Due to increasing capacity needs and a growing researcher base, a shift in access from traditional file POSIX protocols to RESTful object interfaces became necessary. Churchill and his team started investigating solutions that would ease the migration path from file to object without losing performance back in 2016, leading to deployment of an object storage solution as part of the JASMIN Phase 4 upgrade in 2018.

First and foremost, because of the scale and nature of JASMIN, object storage had to meet specific performance metrics to be equivalent to the read access of parallel file systems. In 2018, benchmarking of most major object storage solutions commenced on JASMIN infrastructure.

The performance requirements defined by STFC using 2 Gigabyte objects with sequential access and erasure-coded data via the S3 and NFSv4 protocols were set by a minimum requirement and targets based around the parallel file system performance for similar number of HDD spindles. JASMIN data analyses work flows are mostly WORM (Write once read many), so the STFC was less concerned with matching the write performance of the parallel file system than the read performance.

S3 Aggregate Throughput

	Minimum Performance Requirements	Caringo Results	Parallel File System for Comparison (similar number of spindles)
S3 Read	21.5 GB/s	35.0 GB/s	37 Gbytes/sec
S3 Write	6.5 GB/s	12.5 GB/s	37 Gbytes/sec

The NFS requirement for an object storage platform was mostly based around functionality and not expected to meet parallel file system performance. However, the team was looking for a solution that would linearly scale in NFS performance by adding more NFS server instances.

NFS Single Instance Throughput

	Minimum Performance Requirements	Caringo Results
NFS Read	150 MB/s	349 MB/s
NFS Write	110 MB/s	392 MB/s

With performance criteria met, Churchill and team's evaluation moved to solving user management challenges in an expanding environment with limited staff. On a system-wide shared file system with thousands of users and hundreds of projects, users and groups need their own unique identifiers. Maintaining those, allocating them, and updating them is difficult. With object storage, the need for these identifiers can largely be removed or automated. Tenants inside each of the projects can manage storage allocations and quotas up and down without going to STFC to intervene on their behalf.

Another key function the team at STFC is streamlining with object storage is file sharing and data access. Before, some form of web server or grid FTP server needed to be put in front of files to share them. If you wanted to access a file behind the institutional firewall, it had a different file name and path than if you accessed it externally. When the data is stored as objects, users can access data using the same URL internally, using the significant 20-Tbit networking capability, as they do externally over the internet.

Object storage is also being used to enable researchers to quickly find data they need by leveraging metadata to store date and time, and splitting large files into individual objects. Examples of files being used are NetCDF and HDF5 files ranging in size from 2 GBs to 2 TBs that contain a sequential time and/or location series of data. Researchers have to read through the file in series or they have to jump from the header to where the information they want is stored. With object storage, researchers can store files in individual pieces/objects with a much smaller size, enabling them to go directly to the data needed without file access "skipping patterns."

The impact of the continuous evolution of the underlying infrastructure of JASMIN has produced quantifiable improvements. For instance, instead of being able to study the national distribution of hundreds of animal and insect species, researchers can now study thousands at a time. And, real-time monitoring of all volcanoes and earthquakes in the world is within the near future. With object storage, the team at STFC once again expanded the capabilities of JASMIN, enabling streamlined management, access and search—setting the stage for more rapid data analysis and discovery by researchers.