ISTITUTO MARIO NEGRI

RozoFS addresses the requirements for high-performance and high-availability data storage for Next Generation Sequencing

MARIO NEGRI
ISTITUTO DI RICERCHE FARMACOLOGICHE

The vast amount of data produced by modern sequencers at the Department of Oncology at IRCCS Istituto di Ricerche Farmacologiche Mario Negri creates an important data storage challenge. RozoFS was evaluated in real world scenarios and proven to deliver the performance and robustness required for Next Generation Sequencing applications.

HIGHLIGHTS

Industry:
Life sciences - Next generation sequencing

Challenges:

- The need for high-performance storage for next generation sequencing
- Resilience to hardware failures
- Scalability for large data sets

Solution:
RozoFS software-defined scale-out NAS deployed on four x86 servers

Results:

- Performance exceeding existing file systems in real-world cases, independent of file size
- Manages intense workloads without timeouts and stale file handles
- Scales to large number of nodes

THE CUSTOMER

The IRCCS Istituto di Ricerche Farmacologiche Mario Negri was the first private Italian foundation dedicated entirely to biomedical research. The Institute was established over 50 years ago when philanthropist Mario Negri met Silvio Garattini, a professor at the Milan University of Pharmacology. The Institute's main aim is to help defend human health and life. It has adhered to a policy not to patent its discoveries, preferring to make them freely available to scientists and patients. Today the Institute employs about 700 people that have produced over 14,000 scientific publications. It has three headquarters in Milan, Bergamo and Ranica.

THE CHALLENGES

Analysis of NGS (Next Generation Sequencing) data is a computationally demanding task requiring large amounts of CPU, memory and disk space. There is also a requirement for high-performance data storage systems, resilient to hardware failure, to be connected directly to the computing infrastructure (typically a multi-node cluster) to store large quantities of NGS data reliably. Traditional shared file systems such as NFS (Network File System) do not offer the performance, scalability or cache coherence required by modern NGS data analysis, so alternatives including GlusterFS, Ceph, and Lustre have been developed. However, there is a trade-off between data safety on replicated local storage and degradation of performance across distributed storage. Resilience to hardware failure is typically provided by RAID (Redundant Array of Independent Disks) and redundant storage nodes.

FILE SYSTEM OVERVIEW

RozoFS is a POSIX-compatible file system for distributed storage, offering both high performance and high availability of the data. It is based on the Mojette Transform, a discrete version of the Radon transform used in tomography, This method is used to divide the data into chunks and distribute them to the storage devices, which are part of the storage pool, guaranteeing data integrity and accessibility even if some servers are unavailable. At the same time, some form of data confidentiality is kept as the single chunks (projections) are not usable individually. At the implementation level, RozoFS uses a combination of nodes serving metadata (export servers) and nodes providing storage (storage servers). Client nodes access the data through a file system in userspace (FUSE). The Mario Negri team selected the free and open source version of RozoFS.

TESTING METHODOLOGY

To evaluate basic file system performance, a single system (BaseStation T80; Baseport, Calgary, Canada) was set up and benchmarked RozoFS against a traditional file system (ext4 configured in RAID6) using the “iozone” tool. Afterwards, a RozoFS cluster was set up composed of four storage nodes each carrying sixteen 8 TB disks and two metadata (export) servers replicated with DRBD, offering a total of 154 TB of storage. As a control, a traditional file system was set up with a 7 TB RAID10 partition exported via NFS v4.

Actual testing was done using a community-developed NGS pipeline, “bcbio” (version 0.9.9). The scripts included in the pipeline documentation were used to download the DREAM-TCGA synthetic data set 3 (limited to chromosome 6). The pipeline was configured to run a complete tumor-normal analysis, including alignment, recalibration, variant calling and annotation, and was run on a 15 node cluster running the Simple Linux Resource Manager (SLURM). The pipeline was run twice, once on the NFS export and once on the RozoFS storage pool, each using 40 cores. Data (read and write operations throughput) were collected for one hour with the “nfsiostat” and the “rozodiag” tools for NFS and RozoFS exports, respectively.

THE RESULTS

Compared to a traditional ext4 file system, RozoFS showed stable reading and writing speeds independently from the file size (Figure), while ext4 with RAID6, although faster in reading, suffered a severe performance loss writing larger files, noteworthy because most NGS pipelines produce large files.

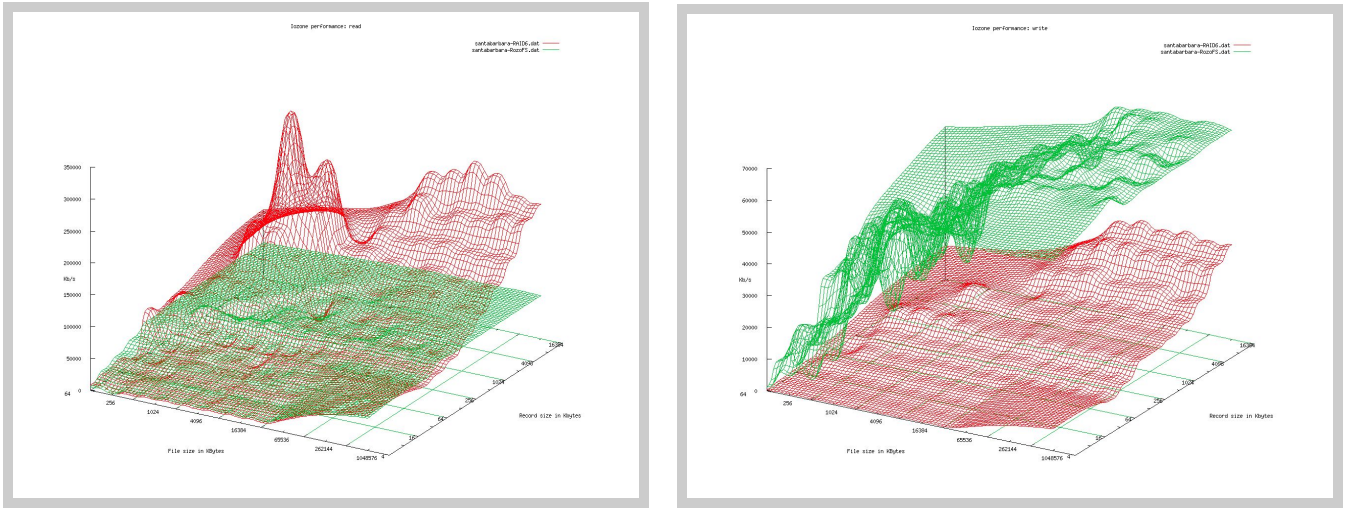

Figure - iozone benchmarks on read (left panel) and write (right) performance. Red, ext4; green, RozoFS

RozoFS was then compared against NFS in a real-world use case involving exome-level tumor-normal variant calling and annotation. Our preliminary data show that RozoFS has a good read performance under a typical NGS workload, and has a higher write performance on comparable hardware when compared to NFS. Lastly, we made a preliminary assessment on scalability, as previous reports indicate that NFS does not scale well with large sample numbers. A high-performance cluster (34 nodes) was used. It was set up a NFS share on one of the nodes and a RozoFS storage cluster spanning 20 nodes. We then executed a “bcbio” analysis on a 33 sample paired tumor-normal data set. Under high I/O load, NFS suffered from repeated timeouts and stale file handles that prevented the analysis from completing. On the other hand, RozoFS did not suffer from this problem.

“Our benchmarks of RozoFS showed great performance, eliminating the severe performance loss suffered by traditional file systems in key applications. Running traditional NFS file systems on a high-performance cluster spanning 20 nodes under high I/O load produced repeated timeouts and stale file handles that prevented the analysis from completing. RozoFS does not suffer from this problem making it a great fit for next generation sequencing workloads.”

Luca Beltrame
Senior Scientist
Istituto Mario Negri

CONCLUSION

The sheer size of NGS data produced by modern sequencers is an important challenge for data storage. Traditional “scale-up” approaches rapidly hit physical space constraints and do not provide enough performance and reliability with modern pipelines. “Scale-out” distributed file systems are much better suited in this regard, as shown by the number of available approaches, such as Ceph, GlusterFS or Lustre. The IRCCS investigation indicates that RozoFS is a viable distributed file system suited for use in NGS data analysis with added guarantees of performance, data integrity and scalability.

This research was initially published at the European Conference on Computational Biology (2016) under the title ‘Distributed file systems for storage and analysis of Next-Generation Sequencing data’ by Luca Beltrame, Anthony J. Travis, Luca Clivio, Sergio Marchini, and Maurizio d’Incalci

CONTACT US

To learn more about how Rozo products, services and solutions can help solve your business and IT challenges, contact us at info@rozosystems.com or visit us at www.rozosystems.com.