NetApp®

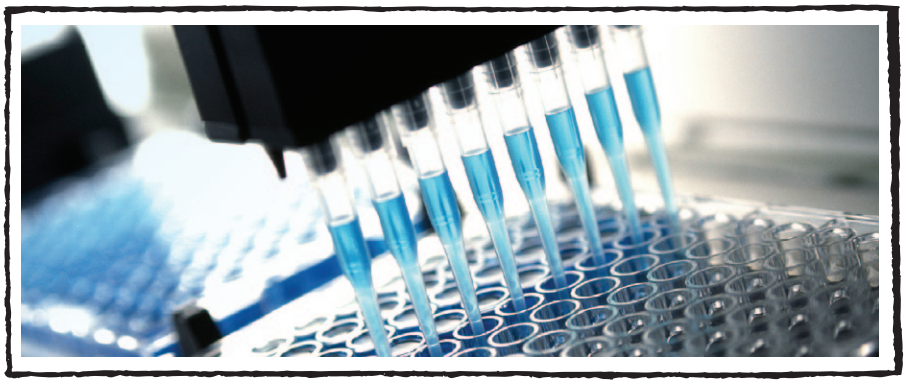
Success Story

JCVI Supports Groundbreaking Genomics Research with NetApp Storage

J. Craig Venter™ INSTITUTE

Another NetApp solution delivered by:

CDW®



KEY HIGHLIGHTS

Industry
Life Sciences

The Challenge
Support rapid growth in biological research data and provide the cost efficiencies vital to a non-profit research institute.

The Solution
Deploy a highly efficient tiered storage environment featuring NetApp® FAS and V-Series storage systems, with performance boosted by the NetApp Virtual Storage Tier with Flash Cache.

Benefits
- Consolidated and upgraded four storage controllers with a single FAS6080 cluster
- Saved tens of thousands of dollars with cost-efficient SATA disks
- Optimized EMC storage investments with NetApp V-Series storage efficiencies and high performance

Customer Profile

Founded in 2006, the J. Craig Venter Institute (JCVI) is a nonprofit research institute dedicated to the advancement of the science of genomics and the understanding of its implications for society. The organization communicates the results of its research to the scientific community, the public, and policymakers by publishing its findings in areas such as human genomic medicine, infectious disease, synthetic biology, bioinformatics, and environmental genomics. With locations in Rockville, Maryland, and San Diego, California, and more than 250,000 square feet of laboratory space, JCVI employs more than 300 scientists and staff.

The Challenge

Staggering data growth and the need for rock-solid storage

As pioneers in genomic research, JCVI scientists with expertise in human and evolutionary biology, genetics, high-throughput DNA sequencing, and other scientific disciplines generate and work with vast amounts of data every day. With JCVI data growing nearly 80% annually in recent years, the institute must have a high-performance computing environment and a reliable IT storage infrastructure to support its important scientific research.

"We are constantly generating research data and adding new project areas, while advances in applications that we depend on, such as the Illumina sequencing platform, are increasing data rates and putting new demands on our IT infrastructure," says Eddy Navarro, senior manager of IT at JCVI. "We need a flexible, scalable storage environment that can keep pace with our growth and that will provide the high performance we need—now and in the future."

Several years ago, after the merger of several institutes into one unified organization, JCVI found itself with multiple EMC and NetApp storage systems. JCVI needed an enhanced storage environment that could accommodate the growing influx of mission-critical data, such as genetic structure sequences generated by machines that analyze biological samples. Navarro was tasked with creating

"NetApp efficiency, manageability, and scalability play a critical role in our ability to handle steady growth and embark on new and important areas of biological scientific research, education, and policy making."

Eddy Navarro
Senior Manager, IT, JCVI

a robust network-attached storage (NAS) system that would enable his department to keep costs low, maximize the value of existing investments, and provide the high-availability infrastructure that scientists and other staff members require.

The Solution

An intelligent, high-performance storage environment

To meet its intensifying performance, availability, and capacity demands, JCVI replaced four NetApp FAS6070 storage controllers at its Maryland site with a single FAS6080 cluster. The NetApp storage system runs on Data ONTAP® 8.0 operating in 7-Mode, which supports 64-bit aggregates and gives JCVI enhanced flexibility and storage efficiency.

"Rather than increase our storage footprint to accommodate steady data growth, we benefit from fewer, more powerful controllers, thanks to the efficiency and manageability of the NetApp FAS6080 cluster," says Navarro. "Getting the most power possible from our IT resources is imperative to our budget and funding our ongoing research."

In the JCVI research environment, genomic sequencing machines such as Illumina and Roche systems take biological samples, sequence the genetic structure, and place that data into Oracle® and Sybase databases. Specialized applications analyze and store the data in NetApp NAS file shares on the NetApp FAS6080 cluster, accessed by a blade server farm using NFS to provide a high-performance computing grid. The NetApp FAS6080 cluster also supports the organization's MySQL environment.

Navarro upgraded a NetApp V3070 open storage controller to a V6240 open storage controller, which manages EMC arrays in Maryland that store 60TB of data for scientific project file shares. The FAS6080 and V6240 storage systems serve as primary storage for JCVI, while a FAS3240 storage system hosts various unstructured data and the organization's VMware® virtual desktop environment. Today the V3070 open storage controller supports JCVI's growing California facility.

The NetApp V-Series storage virtualization solution supports JCVI's tiered storage model, enabling the institute to move terabytes of data from Fibre Channel storage onto less expensive SATA disks, saving tens of thousands of dollars. The NetApp Virtual Storage Tier enabled by Flash Cache boosts performance of the V6240, FAS3240, and FAS6080 systems by increasing read rates and reducing latency without adding high-performance disk drives.

NetApp deduplication technology offers JCVI space efficiency savings of more than 60% for virtual desktop images. The Maryland facility uses NetApp SnapMirror® replication technology to migrate data between the V6240 and FAS3240 systems, for seamless management of NAS with NetApp Data ONTAP 8 operating in 7-Mode. NetApp FlexVol® technology offers additional efficiencies, enabling the IT staff to pool physical storage and to create and resize virtual volumes as application needs change.

The team leverages NetApp OnCommand™ Operations Manager to help monitor the NetApp arrays for capacity utilization and system performance management. The My AutoSupport™ Web-based support application also helps JCVI optimize utilization and performance by monitoring NetApp system performance and sending alerts to NetApp Services.

The IT team leverages NetApp Snapshot™ technology to create point-in-time copies of enterprise-wide data every 4 hours. The organization's second line of data defense is a weekly backup to disk, and third is a monthly backup to tape, which it sends off site.

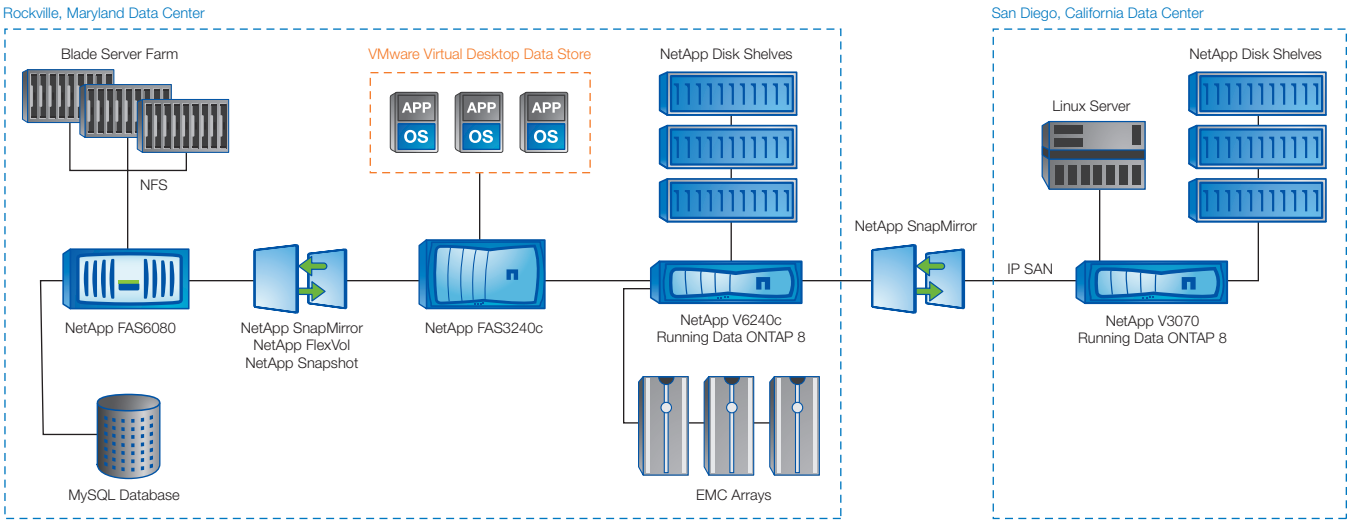


Figure 1) JCVI storage infrastructure.

Business Benefits

Supporting world-class, groundbreaking research

With its core focus on high-throughput genomic sequencing, the organization also depends on high-throughput storage processing. The NetApp Virtual Storage Tier increases I/O throughput, supporting JCVI scientists by speeding access to scientific data.

"Our day-to-day research work depends on the performance of our storage platform and availability of critical data," says Navarro. "The NetApp Virtual Storage Tier with Flash Cache was fundamental to our decision to upgrade to the V6240 system. Flash Cache allows us to leverage solid-state technology in a way that prevents us from having to buy a complete shelf of disks to support our most important data." Flash Cache modules are integrated as a standard feature in the V6240 and other high-end NetApp storage systems.

The data produced by JCVI scientists leads to breakthroughs such as the successful synthesizing of a complete microorganism and publication of the first diploid human genome sequence, taking science one step closer to individualized genomic medicine. NetApp Snapshot technology provides ongoing protection of invaluable research data. Typically, the end users themselves can easily restore research data within minutes of inadvertent deletion by retrieving the lost data directly from the Snapshot copy, boosting the productivity of both scientists and the IT team.

Freeing up resources for vital projects

As a nonprofit organization with tight program budgets, as well as a research institute generating massive amounts of project data every day, JCVI depends on the superior efficiency and manageability offered by NetApp storage. Two IT staff members effectively manage the NetApp storage environment, which contains 400TB of data. The Data ONTAP 8.0 operating system, NetApp deduplication, and NetApp FlexVol technology assist IT in streamlining storage management and optimizing storage resources.

JCVI uses FlexVol volumes to help maximize the amount of data it can store. The IT team leverages Data ONTAP 8.0 to create 64-bit aggregates, which reduces the number of aggregates needed to manage a given amount of data. "The 64-bit aggregates give us some breathing room where storage management and capacity planning are concerned," says Navarro. "Data ONTAP 8.0 enables giant pools of data and prevents us from having to move data between volumes, which allows us to focus on more important projects than shuffling data around."

My AutoSupport enables JCVI to monitor system uptime and storage efficiency. "My AutoSupport makes our lives easier," says Navarro. "We don't have to spend time opening up a support case and gathering relevant data. By collecting and analyzing data from NetApp storage systems, My AutoSupport gives us insight into the storage efficiencies we are achieving and provides early warning of a potential issue."

Positioned for continued growth and new opportunities

Beyond leveraging the storage efficiencies offered by NetApp, JCVI is in the process of building a highly efficient, sustainable, carbon-neutral laboratory facility on the campus of the University of California at San Diego. The 45,000-square-foot building will support 125 scientists and other staff members to further the institute's goals in genomic research and policy. In the meantime, JCVI continues to focus on analyzing E. coli and influenza, microbial oceanic life, biologically driven sources of energy, and other areas of pioneering science. Armed with the flexible, scalable NetApp storage platform, JCVI is well positioned to explore new areas of research as opportunities arise.

"Rather than increase our storage footprint to accommodate steady data growth, we benefit from fewer, more powerful controllers, thanks to the efficiency and manageability of the NetApp FAS6080 cluster."

Eddy Navarro
Senior Manager, IT, JCVI

"When we submit proposals to government agencies or other organizations, we emphasize that we can scale up our IT environment quickly to meet the scope and requirements of the proposed research project because of solutions we employ such as NetApp storage," says Navarro. "NetApp efficiency, manageability, and scalability play a critical role in our ability to handle steady growth and embark on new and important areas of biological scientific research, education, and policy making."

SOLUTION COMPONENTS

NetApp Products
FAS6080 and FAS3240 storage systems
V6240 and V3070 open storage controllers
Data ONTAP 8.0 operating in 7-Mode
SnapMirror replication technology
Snapshot technology
SnapRestore® data recovery software
FlexVol technology
OnCommand Operations Manager

Protocols
NFS
CIFS
iSCSI

Third-Party Products
Oracle 11*g* database
Sybase database
Illumina sequencing platform
Roche sequencing platform

Partner
CDW
www.cdw.com

NetApp

www.netapp.com

NetApp creates innovative storage and data management solutions that deliver outstanding cost efficiency and accelerate business breakthroughs. Discover our passion for helping companies around the world go further, faster at *www.netapp.com*.

Go further, faster®