D∈LL Technologies

HPC Ready Architecture for Genomics

Sequence and assemble more genomes, faster

Customer Results

56

human genomes processed in 54 hours at 50x coverage[1]

~10 billion

reads per month in SPAdes de novo assembly[1]

90 billion

data points sequenced[2]

Advanced computing technologies have rapidly transformed genomics, giving researchers the power to work with vast amounts of data. For example, High Performance Computing (HPC) has been enabling genomic research for many years, providing the horsepower for faster and faster sequencing and assembly. However, as the science advances, the amount of data increases as well. This necessitates continuing innovation to fuel the breakthroughs that help build our understanding of the human genome.

In addition to HPC, other advanced computing technologies — such as data analytics and artificial intelligence (AI) — are becoming more widespread for their ability to advance the study of human genomes faster and with lower costs. While they have existed as separate technologies for many years, the three are converging as the powerful, scalable compute, networking and storage provided by HPC is being used to empower data analytics and AI for genomics use cases. This convergence is reshaping the ability of researchers to better understand and respond to the mysteries of human life.

But building an HPC system for genomics workloads can be challenging. For example, designing a system that provides the immense computational power and storage capacity required is a specialized task. Some organizations struggle to design their own HPC infrastructure while performing complex integration and performance tuning with little IT expertise or support.

Dell Technologies is dedicated to making HPC more accessible and invests in creating a portfolio of solutions designed to simplify the configuration, deployment and management of HPC systems. The building blocks for the Dell EMC HPC Ready Architecture for Genomics are optimized, tested and tuned for a variety of key use cases, such as next generation sequencing (NGS) and de novo assembly. They include the servers, storage, networking, software and services that have been proven in our labs and in customer deployments to meet workload requirements and customer outcomes. The modular building block approach provides a customized yet validated approach for deploying new clusters, or for scaling or upgrading existing environments.

Dell EMC HPC Ready Architectures

The Dell EMC HPC Ready Architecture for Genomics uses a flexible and modular approach to HPC system design where individual building blocks can be combined to build HPC systems that are optimized for specific workloads and use cases. These integrated, tested and tuned solutions leverage Dell EMC servers, networking and storage, with services available from consulting to financing. They include the hardware resources required for various forms of genomic data analysis while providing an optimal balance of compute density, energy efficiency and performance.

[1] Dell EMC HPC Ready Architecture for Genomics, July 2020.

[2] Dell Technologies Case Study, Giving Hope to Children with Rare Disorders.

Resources

- Get architecture guidance and performance testing Information at hpcatdell.com.
- Explore the Dell Technologies HPC & AI Innovation Lab.
- Join the Dell Technologies HPC Community at dellhpc.org.

Learn more

delltechnologies.com/hpc
delltechnologies.com/healthcare

The HPC Ready Architecture for Genomics outlines the system design and performance benchmarking results. The solution can process up to 56 genomes in 54 hours for whole human genome with 50X depth coverage with the medium configuration.

The HPC Ready Architecture for Genomics is designed to speed time to production, improve performance with a purpose-built solution, and scale more easily with modular building blocks. The document provides configuration guidance based on next generation sequencing (NGS) and de novo assembly applications and workloads.

The options below serve as a starting point for a solution that can be customized for your specific environment and needs.

Infrastructure nodes	Compute nodes	Storage	Networking	Software	Operating system
Master nodes: PowerEdge R440 Servers Login nodes: PowerEdge R640 Servers	DNA sequencing: PowerEdge C6400 enclosure with 4x PowerEdge C6420 Servers De novo assembly: PowerEdge R740xd Servers	• Ready Solutions for HPC BeeGFS Storage • Isilon Scale-Out or Hybrid Scale-Out NAS storage	PowerSwitch S3048-ON Ethernet Mellanox® InfiniBand™ QM8790 (HDR) or SB7890 (EDR)	Bright Cluster Manager® recommended BioBuilds tested and recommended	Red Hat® Enterprise Linux® or CentOS

You don't have to go it alone. Collaborate with experts in Dell Technologies Customer Solution Centers worldwide to chart your course. Dell Technologies Services — ranging from consulting and education to deployment and support — are available when and where you need them. Whether on-premises or on-demand, Dell Technologies flexible consumption solutions enable you to pay for technology as you need or use it.

DELLTechnologies