Blueprint

Solution Brief

The Dell Difference

Dell's comprehensive high performance computing and data management solutions, domain expertise and technical know-how help genomic research organizations in healthcare, government, universities and industrial labs accelerate their research and engineering to:

- Concentrate resources on R&D, not complex infrastructure
- Optimize genomic pipelines for quicker results
- Overcome obstacles to mainstream product viability
- Identify treatments in clinically relevant timeframes
- Enable cost-effective bioinformatics centers
- Maintain compliance and protect confidential data using secure, in house resources

Dell HPC System for Life Sciences

Achieving high performance in high-volume life sciences research

Amazing work is being done today in the healthcare and life sciences. With the advent of genome sequencing researchers now have a critical gateway to understanding the underlying molecular pathways for disease. And for some diseases, every hour closer to discovery and result can mean the difference between life and death.

However, many factors are hindering discovery and causing inefficiencies. Genomic processing requires immense computational power and storage. Many researchers, lacking adequate offsite resources, are faced with the challenge of designing and deploying their own infrastructure with little IT expertise or support. Add to this complex integration and performance tuning of the environment—which can take months and result low system usage and delays in key projects.

A complete, end-to-end genomic processing system

The Dell HPC System for Genomics is designed to meet the needs of genomic research organizations, enabling bioinformatics centers to deliver cost-effective results and identify treatments in clinically relevant timeframes while maintaining compliance and protecting confidential data.

This solution provides a flexible, high performance computing environment architected to deliver high-throughput and fast turnaround of genomic workflows in a diverse range of fields including drug design, cancer research, agriculture, biofuels and forensics. Sequenced data is written from Next Generation Sequencing (NGS) machines directly into the system's computational scratch space for processing; output data and user files are network-accessed by researchers for further investigation.

- Available to be ordered, deployed, and delivered as a factory integrated, single rack system with software deployment, integration and training on-site
- Flexible design that scales from 4 to 40 nodes per system with an optional large memory node providing up to 1.5TB of memory with 10 GbE or FDR Infiniband fabrics
- Includes up to 480TB of Lustre storage and up to 240TB of NFS storage with a CIFS gateway server to import data from the sequencer
- Software integration using Bright Cluster Manager 7.1, Red Hat Enterprise Linux 6.6 and Lab7 BioBuild with pre-built open-source binaries for common bioinformatics tools
- Single point of contact for HPC system support
- Achieve fast results with maximum efficiency
  - Up to 163 genomes per day with 10X data; and up to 54 genomes per day with 50X data[i]
  - A complete genome analysis in less than 8 hours
  - As low as 2kWh/Genome with 10X data samples[ii]

Enable bioinformatics innovation by maximizing productivity

Dell's HPC System for Genomics is designed to help you achieve timely results and insights faster, through quick deployment, simple management, and high

performance, backed by Dell's dedicated solutions, support and services teams. As part of the new Dell HPC Systems Portfolio, it combines the flexibility of a custom solution with the simplicity, reliability and value of a preconfigured, factory-built product.

Why choose Dell?

Dell provides HPC systems for all areas that need better, faster solutions to bigger problems, with extra focus on the key markets that require HPC solutions now: research, federal, energy, life, manufacturing and finance. Dell can offer HPC solutions that are designed, configured and scaled for big simulation and big analytics. Dell's CTO group, HPC Engineering group, and new Datacenter Scalable Solutions (DSS) efforts, plus partnerships in showcase projects with leading HPC centers, will drive innovations that go into future HPC systems: Delivering market-ready HPC Systems Portfolio, custom architected systems with partners, and further integration of HPC, big data and cloud.

Dell helps you choose the products and services that best fit your needs and deliver on speed-to-value:

- Integrate into existing environments without disruption
- Maximize choice and evolve at your own pace without lock-in or forklift upgrades
- Leverage plug and play building blocks to grow capacity, capability as needed

Accelerate discovery and innovation through collaboration

Through insightful partnership and collaboration, Dell is committed to help our life sciences and research customers dedicate more time to science and engineering, enabling new discoveries, novel innovations, and timely trials and treatments.

To further accelerate research in clinically relevant timeframes, the Dell HPC System for Genomics can be integrated with other Dell healthcare solutions such as the Dell Cloud Clinical Archive (DCCA), which provides an end-to-end managed archival service for medical images and associated patient information across multiple imaging applications.

Services can be provided by Dell or by Dell partners who have expertise in high performance computing and industry verticals, and who can supplement their services with additional services and support from Dell.

Get more value out of HPC with Dell

We help researchers, lab directors, clinicians, and program managers:

- Analyze volumes of FASTQ, BAM, VCF files with pre-configured NGS toolsets and latest pipelines
- Manage petabytes worth of local omics data into insights faster through tested and validated architecture for life science uses
- Enable better patient outcomes by processing high volume production pipelines at record speed
- Manage compliance by securing and scaling replicable pipelines on a healthcare standards-based compliant architecture

We provide IT managers:

- Simplified design, lightning fast deployment
- Rapid order and deployment options
- Flexible, and open architecture based on engineering-tested and validated reference architectures
- Reduced risks, controlled cost — Dell engineering tested, validated, and characterized.
- Optimized operations – Pre-tuned for specific application sets with single point of support from Dell
- Simplified management using Dell iDRAC with Lifecycle Controller to update, monitor and maintain PowerEdge servers
- Single point of contact for support

[i] Performance metrics above based on August 2015 internal Dell HPC lab benchmark testing (SW: bwa, GATK, bcbio-nextgen GRCh37 (Genome Reference Consortium Human build 37) Input reads: 212M, ~10X coverage; 813M, ~50X), confirming 1) 163 genomes per day throughput while using 10X data on an HPC System for Genomics configuration across a total of 1,092 compute cores 2) confirming 4.4X improvement in terms of genomes/day than GDAP v1.0. 3) the complete analysis of a genome in 11.45 hours across 1,092 compute cores 4) total energy consumption of 154.5 kWh when running 10x coverage whole genome human sequencing data from the Illumina platinum genomes project, named ERR091571_1.fastq.gz and ERR091571_2.fastq.gz, resulting in approximately 2 kWh/genome when using 10X data. 5) Confirmed sustained 29.6 Tflops, 87% of a theoretical maximum of 34 Tflops, based on August 2015 internal Dell HPC lab benchmark testing. Actual performance will vary based on configuration, usage and manufacturing variability.

[ii] The Dell HPC System for Genomics can process genomes with an energy efficiency rate of approximately 2 Kilowatt-hours per genome processed. Results based on additional internal Dell HPC lab benchmark testing completed in August 2015, confirming energy consumption of 154.5 kilowatt hours (kWh) when running 78 concurrent genome analyses when using 10X data while consuming ~2kWh per genome (~3.75X lower energy compared to GDAP v1). To obtain the power and energy consumption at a whole rack level, a Fluke 1735 Three-Phase Power Logger was used. Actual performance will vary based on configuration, usage and manufacturing variability.

Contact your Dell representative or visit www.dell.com/hpc to take the next steps to accelerate your genomics research results