Variant Calling (BWA-GATK) pipeline benchmark with Dell EMC Ready Bundle for HPC Life Sciences

13G/14G server performance comparisons with Dell EMC Isilon and Lustre Storage

Overview

Variant calling is a process by which we identify variants from sequence data. This process includes making a decision if there is single nucleotide polymorphisms (SNPs)[i], insertions and deletions (indels)[ii], and/or structural variants (SVs)[iii] at a given position in an individual genome or transcriptome. The main goal of identifying genomic variations is linking to human diseases. Although not all human diseases are associated with genetic variations, variant calling can provide a valuable guideline for geneticists working on a particular disease caused by genetic variations. BWA-GATK is one of the Next Generation Sequencing (NGS) computational tools that is designed to identify germline and somatic mutations from human NGS data. There are a handful of variant identification tools, and we understand that there is not a single tool performs perfectly (Pabinger et al., 2014). However, we chose GATK which is one of most popular tools as our benchmarking tool to demonstrate how well Dell EMC Ready Bundle for HPC Life Sciences can process complex and massive NGS workloads.

The purpose of this blog is to provide valuable performance information on the Intel® Xeon® Gold 6148 processor and the previous generation Xeon® E5-2697 v4 processor using a BWA-GATK pipeline benchmark on Dell EMC Isilon F800/H600 and Dell EMC Lustre Storage. The Xeon® Gold 6148 CPU features 20 physical cores or 40 logical cores when utilizing hyper threading. This processor is based on Intel's new micro-architecture codenamed "Skylake". Intel significantly increased the L2 cache per core from 256 KB on Broadwell to 1 MB on Skylake. The 6148 also touts 27.5 MB of L3 cache and a six channel DDR4 memory interface. The test cluster configurations are summarized in **Table 1**.

Table 1 Test Cluster Configurations

	Dell EMC PowerEdge C6420	Dell EMC PowerEdge C6320
CPU	2x Xeon® Gold 6148 20c 2.4GHz (Skylake)	2x Xeon® E5-2697 v4 18c 2.3GHz (Broadwell)
RAM	12x 16GB @2666 MHz	8x 16GB @2400 MHz
OS	RHEL 7.3	
Interconnect	Intel® Omni-Path	
BIOS System Profile	Performance Optimized	
Logical Processor	Disabled	
Virtualization Technology	Disabled	
BWA	0.7.15-r1140	
Sambamba	0.6.5	
Samtools	1.3.1	
GATK	3.6	

The test clusters and F800/H600 storage systems were connected via 4 x 100GbE links between two Dell Networking Z9100-ON switches. Each of the compute nodes was connected to the test cluster side Dell Networking Z9100-ON switch via single 10GbE. Four storage nodes in the Dell EMC Isilon F800/H600 were connected to the other switch via 8x 40GbE links. The configuration of the storage is listed in Table 2. For the Dell EMC Lustre Storage connection, four servers (a Metadata Server pair (MDS) and an Object Storage Server pair (OSS)) were connected to 13/14th Generation servers via Intel® Omni-Path. The detailed network topology is illustrated in **Figure 1**.
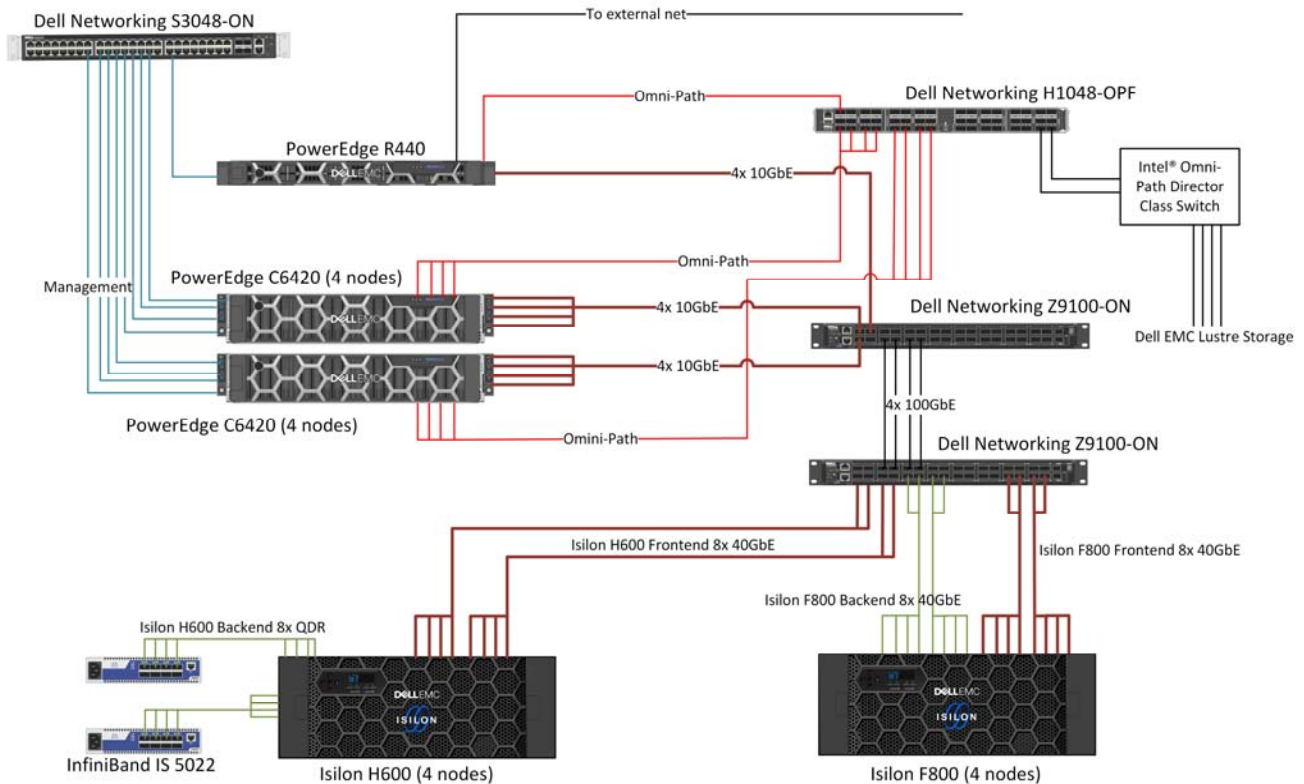

Figure 1 Illustration of Networking Topology: only 8 compute nodes are illustrated here for the simplicity, 64 nodes of 13G/14G servers used for the actual tests.

Table 2 Storage Configurations

	Dell EMC Isilon F800	Dell EMC Isilon H600	Dell EMC Lustre Storage
Number of nodes	4	4	2x Dell EMC PowerEdge R730[iv] as MDS 2x Dell EMC PowerEdge R730 as OSS
CPU per node	Intel® Xeon™ CPU E5-2697A v4 @2.60 GHz	Intel® Xeon™ CPU E5-2680 v4 @2.40GHz	2 x Intel® Xeon™ E5-2630V4 @ 2.20GHz
Memory per node	256GB	256GB	256GB
Storage Capacity	Total usable space: 166.8 TB, 41.7 TB per node	Total usable space: 126.8 TB, 31.7 TB per node	960TB Raw, 768TB (698 TiB) usable MDS Storage Array: 1 x Dell EMC PowerVault MD3420 (Total 24 - 2.5" 300GB 15K RPM SAS) OSS Storage Array: 4 x Dell EMC PowerVault MD3460 (Total 240 3.5" 4 TB 7.2K RPM NL SAS)
SSD L3 Cache	N/A	2.9 TB per node	N/A
Network	Front end network: 40GbE Back end network: 40GbE	Front end network: 40GbE Back end network: IB QDR	Front end network: Intel Omni-Path Internal Connections: 12Gbps SAS
OS	Isilon OneFS v8.1.0 DEV.0	Isilon OneFS v8.1.0.0 B_8_1_0_011	Red Hat Enterprise Linux 7.2 (3.10.0-327.el7.x86_64)

The test data was chosen from one of Illumina's Platinum Genomes. ERR091571 was processed with Illumina HiSeq 2000 submitted by Illumina and can be obtained from EMBL-EBI. The DNA identifier for this individual is NA12878. Although the description of the data from the linked website shows that this sample has a 30x depth of coverage, in reality it is closer to 10x coverage according to the number of reads counted.

Performance Evaluation

BWA-GATK Benchmark over Generations

Dell EMC PowerEdge C6320s and C6420s were configured as listed in **Table 1**. The tests performed here are designed to demonstrate performance at the server level, not for comparisons on individual components. At the same time, the tests were also designed to estimate the sizing information of Dell EMC Isilon F800/H600 and Dell EMC Lustre Storage. The data points in Figure 2 are calculated based on the total number of samples (X axis in the figure) that were processed concurrently. The number of genomes per day metric is obtained from total running time taken to process the total number of samples in a test. The smoothed curves are generated by using a polynomial spline with the piecewise polynomial degree of 3 generating B-spline basis matrix. The details of BWA-GATK pipeline information can be obtained from the Broad Institute web site.
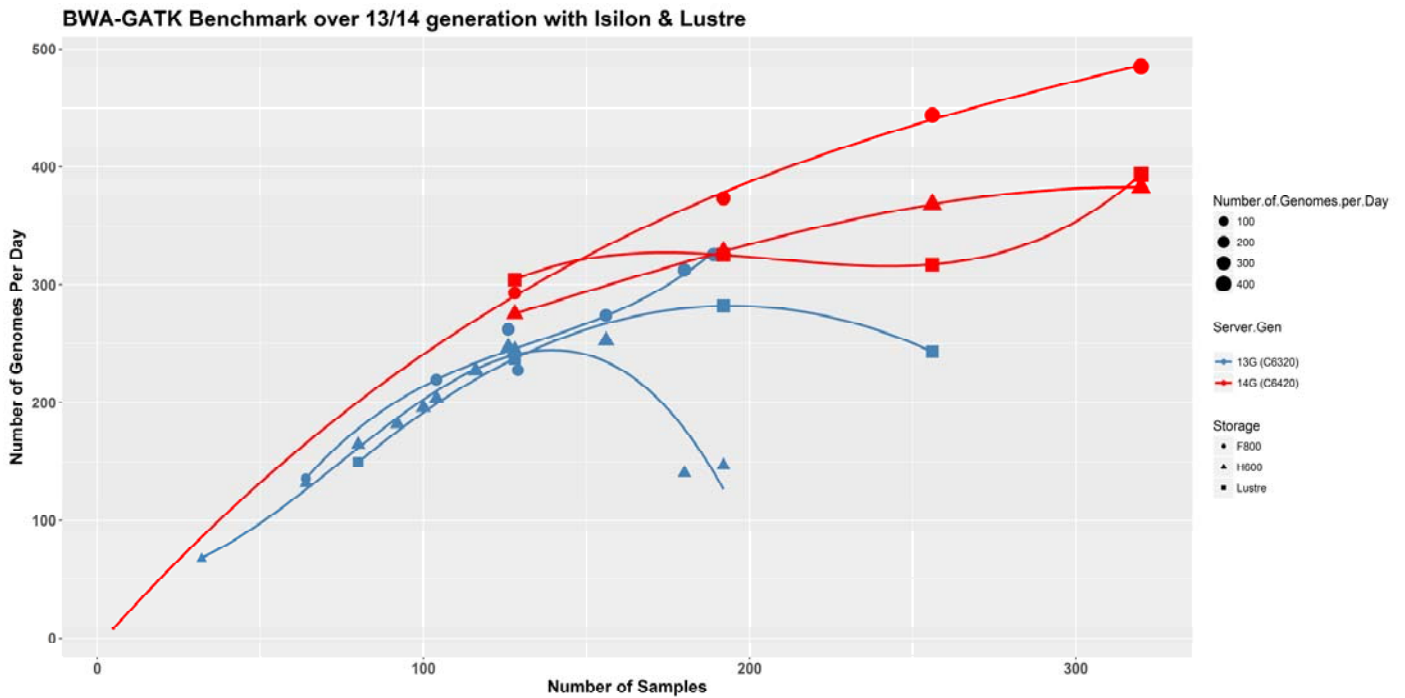

Figure 2 BWA-GATK Benchmark over 13/14 generation with Dell EMC Isilon and Dell EMC Lustre Storage: the number of compute nodes used for the tests are 64x C6420s and 63x C6320s (64x C6320s for testing H600). The number of samples per node was increased to get the desired total number of samples processed concurrently. For C6320 (13G), 3 samples per node was the maximum number of samples each node can process. 64, 104, and 126 test results for 13G system (blue) were with 2 samples per node while 129, 156, 180, 189 and 192 sample test results were obtained from 3 samples per node. For C6420 (14G), the tests were performed with maximum 5 samples per node. The plot for 14G was generated by processing 1, 2, 3, 4, and 5 samples per node. The number of samples per node is limited by the amount of memory in a system. 128 GB and 192 GB of RAM were used in 13G and 14G system, respectively as shown in **Table 1**. C6420s show a better scaling behavior than C6320s. 13G server with Broadwell CPUs seems to be more sensitive to the number of samples loaded onto system as shown from the results of 126 vs 129 sample tests on all the storages tested in this study.

The results with Dell EMC Isilon F800 that indicate C6320 with Broadwell/128GB RAM performs roughly 50 genomes per day less when 3 samples are processed per compute node and 30 genomes per day less when 2 samples are processed in each compute node compared to C6420. It is not clear if C6320's performance will drop again when more samples are added to each compute node; however, it is obvious that C6420 does not show this behavior when the number of samples is increased on each compute node. The results also allow estimating the maximum performance of Dell EMC Isilon F800. As the total number of genomes increases, the increment of the number of genomes per day metric is slow down. Unfortunately, we were not able to identify the exact number of C6420s that would saturate Dell EMC Isilon F800 with four nodes. However, it is safe to say that more than 64x C6420s will require additional Dell EMC Isilon F800/H600 nodes to maintain high performance with more than 320 10x whole human genome samples. Dell EMC Lustre Storage did not scale as well as Dell EMC Isilon F800/H600. However, we observed that some optimizations are necessary to make Dell EMC Lustre Storage perform better. For example, the aligning, sorting, and marking duplicates steps in the pipeline perform extremely well when the file system's stripe size was set to 2MB while other steps perform very poorly with 2MB stripe size. This suggests that Dell EMC Lustre Storage needs to be optimized further for these heterogeneous workloads in the pipeline. Since there is not any concrete configuration for the pipeline, we will further investigate the idea of using multiple tier file systems to cover different requirements in each step for both Dell EMC Isilon and Lustre Storage.

Dell EMC PowerEdge C6320 with Dell EMC Isilon H600 performance reached the maximum around 140 concurrent 10x human whole genomes. Running three 10x samples concurrently on a single node is not ideal. This limit appears to be on the compute node side, since H600 performance is much better with C6420s running a similar number of samples.

Conclusion

Dell EMC PowerEdge C6420 has at least a 12% performance gain compared to the previous generation. Each C6420 compute node with 192 GB RAM can process about seven 10x whole human genomes per day. This number could be increased if the C6420 compute node is configured with more memory. In addition to the improvement on the 14G server side, four Isilon F800 nodes in a 4U chassis can support 64x C6420s and 320 10x whole human genomes concurrently.

Resources

Internal web page

http://en.community.dell.com/techcenter/blueprints/blueprint_for_hpc/m/mediagallery/20442903

External web page

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3956068/

Contacts

Americas

Kihoon Yoon

Sr. Principal Systems Dev Eng

Kihoon.Yoon@dell.com

+1 512 728 4191

---

[i] SNPs (pronounced "snips") represent a difference in a single DNA.
[ii] Indel is the addition or removal of a small number of consecutive nucleotides in a DNA sequence.
[iii] Structural variation is genomic alterations that involves DNA segments larger than 1Killo base (Kb).
[iv] All Dell EMC PowerEdge R730s were configured identically.