Understanding the Role of Dell EMC Isilon SmartConnect in Genomics Workloads

Kihoon Yoon. Dell EMC HPC Innovation Lab. November 2016

Coming together with EMC has opened many new opportunities for the Dell EMC HPC Team to develop high-performance computing and storage solutions for the Life Sciences. Our lab recently stood up a 'starter' 3 node Dell EMC Isilon X410 cluster. As a loyal user of the Isilon X210 in a previous role, I couldn't wait to start profiling genomics applications using the X410 with Dell EMC HPC System for Life Sciences.

Because our current Isilon X410 storage cluster is currently fixed at the 3 node minimum, we aren't set up yet to evaluate the scalability of the X410 with genomics workflows. We will tackle this work once our lab receives additional X nodes and the new the Isilon All-Flash node (formerly project Nitro).

In the meantime, I wanted to understand how the Isilon storage behaves relative to other storage solutions and decided to focus on the role of Isilon SmartConnect.

Through a single host name, SmartConnect enables client connection load balancing and dynamic network file system (NFS) failover and failback of client connections across storage nodes to provide optimal utilization of the Isilon cluster resources.

Without the need to install client-side drivers, administrators can easily manage a large and growing number of clients and ensure in the event of a system failure, in-flight reads and writes will successfully finish without failing.

Traditional storage systems with two-way failover typically sustain a minimum 50 percent degradation in performance when a storage head fails, as all clients must fail over to the remaining head. With Isilon SmartConnect, clients are evenly distributed across all remaining nodes in the cluster during failover, helping to ensure minimal performance impact.

To test this concept, I ran the GATK pipeline varying the number of samples and compute nodes without and with SmartConnect enabled on the Isilon storage cluster.

The configuration of our current lab environment and whole human genome sequencing data used for this evaluation are listed below.

Table 1 System configuration, software, and data

Dell EMC HPC System for Life Sciences	
Server	40 x PowerEdge C6320
Processor	2 x Intel Xeon E5-2697 v4. 18 cores per socket, 2.3 GHz
Memory	128 GB at 2400 MT/s
Interconnect	10GbE NIC and switch for accessing Isilon & Intel Omni-Path fabric
Software	

Operating System	Red Hat Enterprise Linux 7.2
BWA	0.7.2-r1039
Samtools	1.2.1
Sambamba	0.6.0
GATK	3.5
Benchmark Data	
ERR091571	10x Whole human genome sequencing data from Illumina HiSeq 2000, Total number of reads = 211,437,919

As noted earlier, the Isilon in our environment is currently set up in a minimum 3 node configuration. The current generation of Isilon is scalable up to 144 nodes. As you add additional Isilon nodes, the aggregate memory, throughput, and IOPS scale linearly. For a deep dive on Isilon and OneFS file system, see this technical white paper.

The Isilon storage cluster in our lab is summarized in Table 2. The Isilon storage is mounted on each compute nodes, up to 40 nodes, through NFS service (version 3) over a 10GbE network.

Table 2: Current Isilon Configuration

Dell EMC Isilon	
Server	3 x X410
Processor	2 x Intel(R) Xeon(R) CPU E5-2640 v2 @ 2.00GHz (1995.20-MHz K8-class CPU), 16 cores.
Memory	256 GB at 2400 MT/s
Back End Networking	2 x IB QDR links
Front-end Networking	2 x 1GbE ports and 2 x 10GbE SFP+ ports
Storage Capacity	4TB x 30 HDDs, 120TB (usable)
Software	
Operating System	OneFS 8.0
Isilon SmartConnect	Round Robin Mode

Table 3 summarizes all the tests we performed. To mimic a storage environment without proper load balancing, all the tests were performed *without* SmartConnect enabled except the concurrent 120 sample run.

Each pipeline (job) runs with one sample and uses 11 cores on a single compute node. A maximum three pipelines run concurrently on a single compute node. The tests were performed up to 40 nodes and 120 samples.

I included the detailed running times for each sub-step in BWA-GATK pipeline in Table 3. The running times for Aligning & Sorting and HaplotypeCaller steps are the bottlenecks in the pipeline, but the

Aligning & Sorting step is more sensitive to the number of samples. In this benchmark, GenotypeGVCFs is not a bottleneck since we used an identical sequence data for all the concurrent pipelines. In real data analysis, a large number of different samples are used, and GenotypeGVCFs becomes the major bottleneck.

Table 3 Test results for BWA-GATK pipeline without and with Isilon SmartConnect enabled.

Number of samples (Data Size)	3	15	30	60	90	120	*120 with SmartConnect
Number of compute nodes	1	5	10	20	30	40	40
Aligning & Sorting	3.48	3.54	3.64	4.21	4.94	5.54	4.69
Mark/Remove Duplicates	0.46	0.49	0.79	1.27	2.52	3.07	1.84
Generate Realigning Targets	0.19	0.18	0.19	0.19	0.18	0.20	0.18
Realign around InDel	2.22	2.20	2.24	2.27	2.26	2.27	2.29
Base Recalibration	1.17	1.18	1.19	1.18	1.16	1.13	1.17
HaplotypeCaller	4.13	4.35	4.39	4.34	4.31	4.32	4.29
GenotypeGVCFs	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Variant Recalibration	0.58	0.50	0.53	0.55	0.57	0.53	0.41
Apply Recalibration	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Total Running Time (Hrs)	12.3	12.5	13.0	14.1	16.0	17.1	15.0
Number of Genomes per Day	6	29	53	96	129	156	185

As shown in Table 3, after 30 samples on 10 compute nodes, the total running time of BWA-GATK began to increase. When the number of compute nodes doubled, the total run time continued to climb without SmartConnect enabled. Also starting at 30 samples, we saw jobs starting to fail presumably due to unbalanced client connections and inability to failover and failback those connections.

However, when we enabled SmartConnect using the default round-robin settings, we saw a significant improvement on the total run time and daily sample throughput.

As expected, SmartConnect, maximized performance by keeping client connections balanced across all three Isilon storage nodes. In the three X410 configuration with SmartConnect enabled, the 120 samples processed with 40 compute nodes showed a 14% speed-up and 19% increased daily sample throughput.

This test also suggests a starting point identifying the number client connections per Isilon node for this genomics workflow. In our case, adding one additional X410 to the Isilon storage cluster for each 15 additional compute nodes may be a reasonable place to start.

As we add additional Isilon nodes to our cluster, we will perform additional studies to refine recommendations for the number of client connections per Isilon node for this genomics workflow. We'll also take a deeper dive with the advanced SmartConnect load balancing options like CPU utilization, connection count, and network throughput. The Isilon SmartConnect White Paper provides a detailed summary and examples for each of these modes.

If you are using Bright Cluster Manger and need some tips to set it up with Isilon SmartConnect, read this [post](#).