



Life Sciences Computational Workflows

Delivering the performance that matches today's variable and demanding workloads

SOLUTION HIGHLIGHTS

- Data searches, routine analysis and managing output of new lab equipment create complex and mixed workloads on storage infrastructure
- Next-generation sequencers are generating unprecedented volumes of data
- Without special care, this data can stress high performance computational workflows
- What's needed is a robust and scalable network data storage and intelligent management solution

Introduction

The rapidly growing amounts of new data being created in Life Sciences highlight challenges when selecting and configuring a storage infrastructure that can support a productive workflow. It's not enough to base a decision on single benchmark numbers, what's needed is a deeper understanding of the overall environment — an understanding that only can come from actual application testing. This testing was conducted by BlueArc and Illumina to examine how BlueArc network storage and data management solutions can be used to ensure optimal next-generation sequencing computational workflow throughputs, while enabling optimal handling of the mixed workloads that are routinely encountered in life sciences organizations today.

The Rush to Discovery

A new generation of laboratory sequencing instruments is driving unprecedented changes in the life sciences. Specifically, next-generation sequencers (NGS) such as the solutions provided by Illumina perform their work faster than before, thus helping to accelerate research. They produce a much richer set of data per sequence experiment offering researchers more insight into the biological systems they are studying. And NGS instruments now cost much less to operate, on a per sequenced base basis, than their predecessors.

The faster run times and lower operational costs allow life sciences organizations to conduct many more sequencing experiments in a shorter period of time. At the same time, each of these faster and more frequent runs produces several orders of magnitude more data than was the case with previous generation systems. This means that special care must be taken to handle this data explosion so as to not disrupt the very data-intensive computational workflows, which are commonly used in life sciences organizations today.

What's needed is a shared network data storage and management solution that easily scales to accommodate the growing data from NGS equipment, provides the performance to feed high-throughput computational workflows, and addresses long-term data management cost of ownership issues by automating data migration to suitable storage tiers over time.

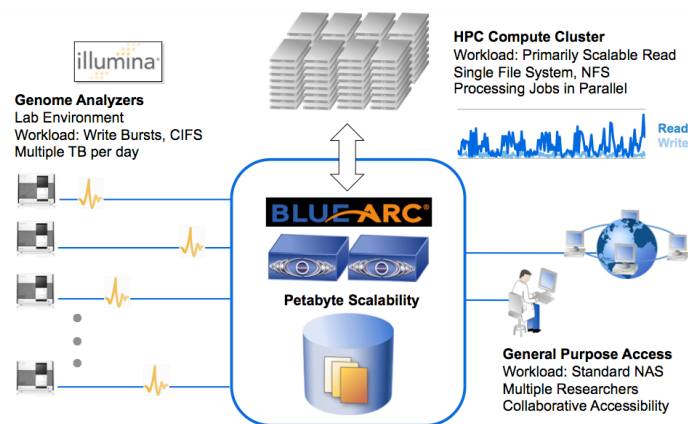
Illumina HiSeq 2000



User Environment

The user environment below demonstrates the three competing uses of any shared network storage infrastructure in a production life sciences solution. The genome analyzers are the point of data acquisition and generate the initial wave of results to be stored and analyzed. The compute cluster performs much of the analysis of the data, typically in the form of parallel batch jobs that share resource between multiple departments and research users. The third element of the processing is the ability to have transparent access to the data and results by individual users that also may participate in collaborative research.

These three important uses of the shared network storage infrastructure compete for performance, capacity and utilization of the system. The combined workload is mixed and unpredictable in nature. As the research focus or the application mix change, so to do the performance requirements. It's important that the shared storage solution can grow with the complexity and handle a broad spectrum of workloads in order to simplify and consolidate management of the data assets.



Application example: Illumina

When trying to match a suitable storage solution with an NGS environment, it is important to keep in mind that you cannot rely on narrow benchmarks. Real-world application workloads are very complex, and overall analysis throughput can vary greatly over time depending on something as minor as which version of an analysis program is used. Additionally, most organizations have multiple applications that produce mixed workloads, so optimizing for one is not practical. What's needed is a system designed to handle mixed workloads – one that can meet the demands of today's demanding, data-intensive workflows.

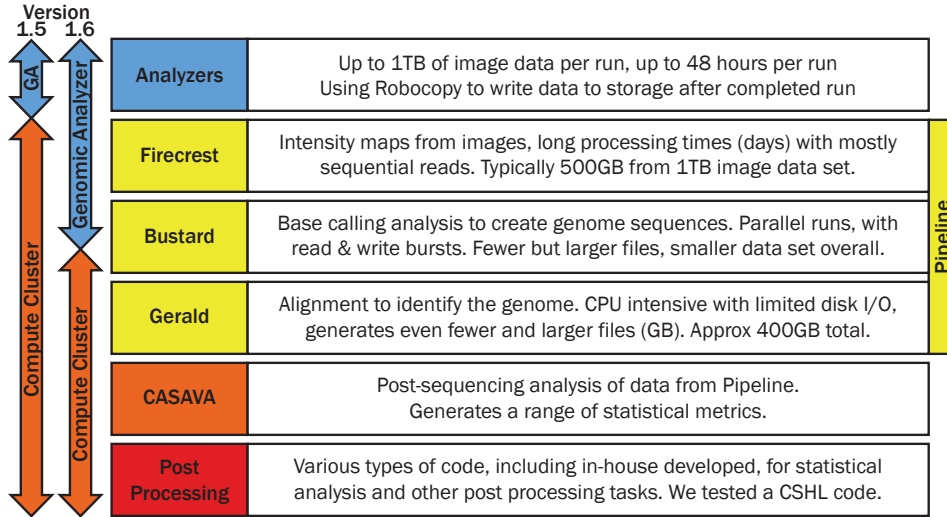
With these issues in mind, BlueArc and Illumina teamed to examine the challenges, nuances, and potential benefits when integrating and optimizing next-gen sequencing systems, storage/data management solutions, and high-throughput workflows. (For information about the IT infrastructure used in this testing, see the sidebar: Test Setup.)

There are numerous software elements in a genomic analysis workflow. Elements can include software to manage the individual genome analyzers (Sequence Control Software or SCS, in the case of Illumina), imaging analysis software, base calling and alignment software, and post-processing statistical analysis routines.

From an IT perspective, the version of the genomic analysis software (as noted above) can be significant. The latest version was used in this test — Illumina Genome Analyzer OffLine Basecaller (OLB) software, version 1.6. This software carries out the preliminary image analysis of the sequencer's output data using Firecrest and base calling analysis of the data to create genomic sequence reads using Bustard.

With Illumina's Real Time Analysis Software (RTA) version 1.6, this part of the processing work is done on a SCS workstation. In contrast, OLB 1.6 relies on the compute cluster to carry out this task. This is accomplished by submitting multiple batch processing jobs in parallel to the compute cluster.

Illumina Workflow Tasks



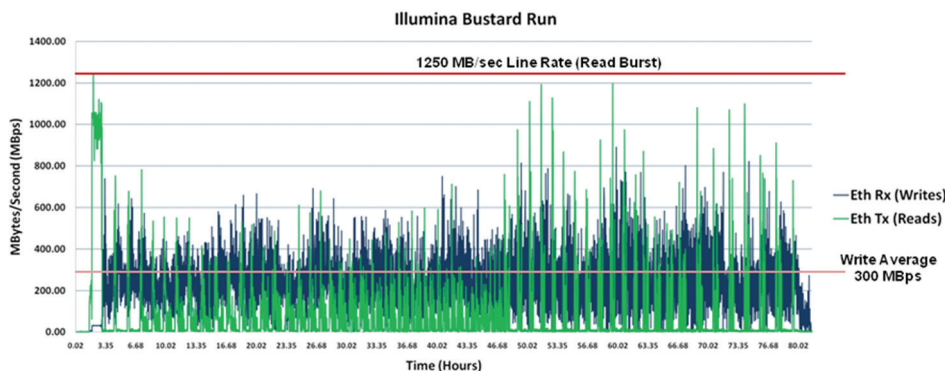
This illustrates the difficulty in evaluating overall performance. The image analysis operations done in the OLB version 1.6 software produce a predominately sequential read workload. When RTA version 1.6 software is used, this step is offloaded to the SCS processor, thus lowering the dependence on a storage system's raw sequential read performance on the total analysis system throughput and instead requiring a more balanced performance profile.

Once the newly acquired data set on the genome analyzer is ready, it's transferred as a stream of writes to the shared network storage system. It's common to have multiple genome analyzers running in parallel producing an aggregate workload that scales with the number of instruments. This workload competes with the downstream processing steps for the shared resources.

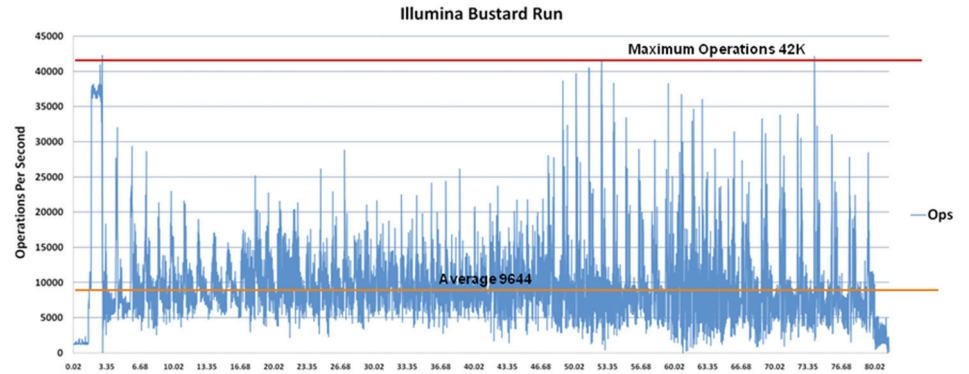
A variety of tests were conducted to illustrate the storage system performance requirements needed to accommodate the workloads of the various analysis software routines.

For example, a Bustard analysis was done using two lanes worth of data amounting to a 800 GB dataset. Running multiple parallel jobs, the test ran for three-plus days and exhibited a workload pattern dominated by bursty reads. Specifically, the test produced 8.8 TB of data and resulted in about 300 MB/sec of average sustained write and 200 MB/sec average sustained read, with bursts of 500 MB/sec write and 1,000 MB/sec read. IOPS for the test averaged 9,644 IOPS, with a maximum of 42,000 IOPS.

Illumina Bustard Run (Throughput)



Illumina Bustard Run (IOPS)



In another test, Gerald analysis (which includes an alignment analysis to identify the locations of genomic reads within a genome) was done running a maximum of 16 parallel jobs. This operation is very CPU intensive and produced about 400 GB of data. The run generated over 200 MB/sec read bursts and between 100 to 150 MB/sec write bursts. The average read was practically zero between the bursts; the average write was about 50 MB/sec. The IOPS typically ran well below 1,000 IOPS on average and peaked at 5,000 IOPS.

Test Setup

- The testing reported in this solution brief was done on a rented compute cluster from R Systems.
- It included 120 nodes with 480 cores total. Each node was a Dell 1950 system with 3.0 MHz Dual Core Intel Woodcrest CPU, four cores per node, 8 GB memory, 1 GbE NIC, and 73 GB of local storage. The nodes ran standard Red Hat RHEL 5.2. Sun Grid Engine 6.2 was used to manage the batch compute jobs.
- BlueArc systems used included three Titan 3210 systems and two Mercury 100 systems. All had two 10 GbE interfaces. Additional lower-tier storage was provided using an array of 144 SAS 300 GB drives and 144 NLSAS drives. The storage was SAN-connected to allow the Titan or Mercury cluster to attach to any of the arrays as required.
- A Force10 Networks E600 switch was used to connect all of the elements together.

Meeting varied workload demands

These two commonly used analysis operations required to derive genomic sequences from an analyzer's raw data have vastly different storage read/write performance and IOPS requirements.

Again, this reinforces the notion that the storage solution that is chosen must have a range of performance capabilities to satisfy the varied computational workloads used in life sciences organizations today.

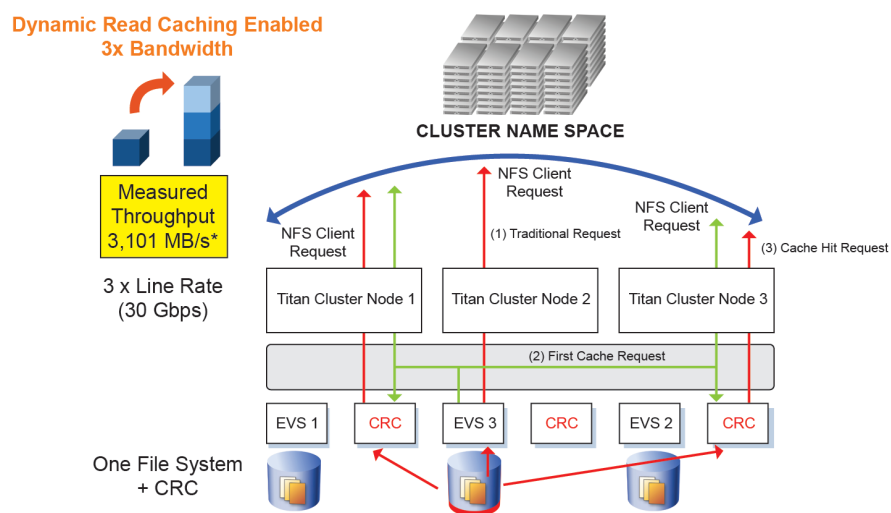
This was particularly evident in another test. Quite often, once sequencing data has been processed using algorithms such as Firecrest, Bustard, and Gerald, additional analysis must be done. In fact, depending on an organization's focus, there are numerous post-processing applications that can be applied and they are a significant part of the overall analysis work. Researchers develop their own code to perform more specialized analysis specific to the mission of their research. The starting point for this analysis is the data produced by the standard Illumina analysis software.

In this testing, a frequency counting statistical analysis application from a BlueArc customer with multiple Illumina NGS instruments in production was used on a relatively small (40 GB) dataset. The files in this dataset ranged from a half to 6 MB. For the test, 400 CPU cores in the compute cluster were used. Traffic of 1,050 MB/sec maximum was pinned to one node and a single 10 GbE link. The run generated very heavy read-oriented traffic.

This initial test highlighted a common concern in many life sciences computational workflow environments. In particular, in such settings, the storage system cannot achieve more than the line rate out of any single file system.

To overcome this limitation, BlueArc used a feature of its systems' software called Dynamic Read Caching. This is essentially a dynamic caching mechanism applied to BlueArc network storage. When used with a cluster of BlueArc network storage clusters, each network storage node maintains its own dynamic cache, but is aware of the files accessed by all the other servers in the cluster. Copies of so called "hot" files from anywhere in the cluster therefore make their way to every cache on every BlueArc server. This can result in dramatic aggregate read performance improvements since every server can respond to any read request for a given set of hot files.

In the test, using Dynamic Read Caching on three BlueArc Titan network storage servers produced an amazing rate of 3,068 MB/sec maximum traffic, a close to linear speedup in this case. The test setup was limited to three servers, but the Dynamic Read Caching can be used with more servers and also with BlueArc's midrange network storage server Mercury for smaller installations. This performance enhancement can be quite useful in many situations. Applications with read-intensive workload profiles and a need to stage data in an optimized workflow process can leverage read caching as a way to scale performance when and how they need it.



This can help significantly improve the performance of many life sciences high performance computational analysis workflows. In particular, wherever storage systems are hitting hard limitations with performance or scalable and sustainable client/server access, dynamic read caching can help achieve new levels of optimization and speed time to results, all while retaining older data not used in a specific workflow calculation on lower cost storage.

Take-Aways

This testing was done to better understand the demands being placed on shared storage systems by next-generation sequencing systems and their associated analysis routines and computational workflows.

The new sequencers generate much more data than their predecessors. Simply storing the data on the highest performance storage for long periods does not make sense.

BlueArc provides the mix of high performance storage in its Titan network storage systems and cost-effective Mercury network storage systems for longer-term storage, both with a range of disk technology options. BlueArc solutions are designed to work in today's heterogeneous IT infrastructures found on most organizations today. Features like BlueArc Dynamic Read Caching can significantly improve throughputs for read-intensive applications.

Additionally, the BlueArc's SiliconFS file system and portfolio of intelligent data management tools offer simplified volume management and easy-scalability to match the growth of lab data in life sciences organizations. Its data management solutions help IT managers and administrators more simply manage and migrate data over its lifetime.

For more information

Please visit www.bluearc.com

BlueArc Corporation

Corporate Headquarters
50 Rio Robles Drive
San Jose, CA 95134
t 408 576 6600
f 408 576 6601
www.bluearc.com

BlueArc UK Ltd.

European Headquarters
Queensgate House
Cookham Road
Bracknell RG12 1RB, United Kingdom
t +44 (0) 1344 408 200
f +44 (0) 1344 408 202

