Next-Generation Sequencing (NGS) technologies are rapidly transforming the capabilities of biotech, healthcare, pharmaceutical, and life sciences organizations. However, the tremendous demands of NGS workloads present a new set of business challenges and place immense strain on computational resources. In this white paper, the results of two recent benchmark tests will be used to help companies select High Performance Computing (HPC) solutions that will accelerate the performance of sequencing workloads, and demonstrate the computing efficiencies that can be achieved with HPE ProLiant Gen9 servers, Intel® Xeon® Processor E5 v4 family, Intel® Omni-Path Architecture (Intel OPA), and Intel Enterprise Edition for Lustre® software (Intel EE for Lustre*).

Introduction

Genomic sequencing has progressed rapidly over the last couple of decades to the point where a whole human genome can be sequenced in under a day for about $1,245 USD.¹ The proliferation of next-generation sequencing (NGS) is producing volumes of genetic data which pose tremendous business challenges of accelerating discovery and development, ensuring regulatory compliance, speeding time-to-solution, and reducing costs. From a computing standpoint, NGS also places high demands on processing speed, throughput, storage, and memory.

To help alleviate this strain on computational resources, sequencing data is typically transferred from instruments to more efficient HPC systems that use shared file systems for processing large volumes of genomes and exomes. Scientists depend on standardized workflows, such as variant analysis, to transform the data over a series of steps which include aligning, variant calling, and annotation. The individual applications that compose the workflow range from computationally-efficient parallel programs to comparatively slow Java or Python scripts.

The performance study

Alignment, which maps sequences against a large reference genome, is the first processing step in an NGS pipeline. This step reduces the amount of data by an order of magnitude, with subsequent steps decreasing storage requirements even further. While read mapping is a critically important step, it is often the most computationally-intensive portion of the pipeline.

The performance of the alignment step is enabled with the bwakit, a self-consistent installation-free package of scripts and precompiled binaries that provides an end-to-end solution for read mapping. The run-bwamem perl script in this package creates and prints mapping command lines for a workflow which uses the other programs or data in bwakit.

In this study, the majority of the compute time is spent running BWA-MEM², which provides the basic mapping functionality. The other programs in bwakit, while not compute-intensive, generate proper human reference genomes and take advantage of ALT contigs (if present) to improve read mapping and perform HLA typing for high-coverage human data.

Test hardware

Three HPE server models are used; the configuration details are provided below in table 1.

Different families of Intel Xeon processors dictate the configuration options for the servers.

- The Intel Xeon Processor E5 v2 family requires the use of HPE ProLiant Gen8 servers as it does not support DDR4 memory chips or Omni-Path Architecture interconnects.
- The Intel Xeon Processor E5 v3 family requires the use of HPE ProLiant Gen9 servers with DDR4 memory chips running no faster than 2133 MHz as it does not support 2400 MHz DDR4 memory chips.

---

¹ Wetterstrand KA. (2016). DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). genome.gov/sequencingcostdata
• The Intel Xeon Processor E5 v4 family requires the use of HPE ProLiant Gen9 servers. It supports up to 2400 MHz DDR4 memory.
• The Intel Xeon Processor E5 v3 and v4 families support both InfiniBand and Omni-Path Architecture interconnects.

<table>
<thead>
<tr>
<th>HPE server model</th>
<th>HPE ProLiant SL230s Gen8 SE</th>
<th>HPE ProLiant XL230a Gen9</th>
<th>HPE ProLiant XL170r Gen9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor family</td>
<td>Intel Xeon Processor E5 v2</td>
<td>Intel Xeon Processor E5 v3</td>
<td>Intel Xeon Processor E5 v4</td>
</tr>
<tr>
<td>Processor speed</td>
<td>2.4 GHz</td>
<td>2.3 GHz</td>
<td>2.2 GHz</td>
</tr>
<tr>
<td>Processor model</td>
<td>E5-2695 v2</td>
<td>E5-2698 v3</td>
<td>E5-2698 v4</td>
</tr>
<tr>
<td>Processor/Node</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cores/Node</td>
<td>24</td>
<td>32</td>
<td>40</td>
</tr>
<tr>
<td>Memory</td>
<td>256 GB</td>
<td>128 GB</td>
<td>128 GB</td>
</tr>
<tr>
<td></td>
<td>1666 MHz</td>
<td>2133 MHz</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>DDR3 2 DPC</td>
<td>DDR4 2 DPC</td>
<td>DDR4 2 DPC</td>
<td>DDR4 2 DPC</td>
</tr>
<tr>
<td>L3 Cache</td>
<td>30 MB shared between 12 cores</td>
<td>60 MB shared between 16 cores</td>
<td>50 MB shared between 20 cores</td>
</tr>
<tr>
<td>Interconnect</td>
<td>FDR InfiniBand</td>
<td>FDR InfiniBand</td>
<td>Intel Omni-Path Architecture</td>
</tr>
<tr>
<td>Operating system</td>
<td>RHEL 6.6</td>
<td>RHEL 6.6</td>
<td>RHEL 7.2</td>
</tr>
</tbody>
</table>

The HPE ProLiant XL170r Gen9 cluster described in table 1 is used in every performance study described in this paper. If the server is modified from its original configuration, then the modifications will be fully described in each section.

**Test cases**

**Benchmark Test 1: GRCh38+ALT**

Run-bwamem is used to map 10,000,009 101bp sequences of synthetic data against hs38a, a human reference genome in FastA format that consists of GRCh38 (including chromosomes, unplaced, and unlocalized contigs) and EBV plus ALT contigs. The BWT index for hs38a is calculated in advance using the BWT-SW algorithm that is capable of handling the whole human genome.

Run-bwamem prints the following command line for mapping:

```
cat r0-3000000.fq \
| ./bwa mem -p -t{NTHREAD} hs38a.fa - 2> out.log.bwamem \
| ./k8 ./bwa-postalt.js hs38a.fa.alt \
| ./samtools view -1 - > out.aln.bam;
```

**Note**

NTHREAD is the amount of parallelism to use on the system.
The processing steps are as follows:

1. The sequence data is read into memory and sent to BWA
2. The basic mapping functionality is performed with the BWA-MEM algorithm designed for processing Illumina sequence reads up to 100bp
3. The k8 interpreter for Java scripts is used to run bwa-postalt.js, which is a Java script to post-process alignments to ALT contigs/decosy/HLA gene
4. SAMtools\(^3\) \(^4\) is used to convert the data from SAM to BAM format

Following these processing steps yields the results shown in figures 1 and 2:

![Figure 1 and Figure 2](image-url)

Figures 1 and 2 show CPU usage while running Benchmark Test 1 on a 40 core HPE ProLiant XL170r Gen9 Server. In figure 1, BWA begins slowly as the sequencing data is read, then spikes as it uses all of the available cores for the majority of the run time. Figure 2, which shows only the bwa-postalt and SAMtools processes, demonstrates that when BWA finishes, the bwa-postalt and SAMtools processes continue running serially and simultaneously.

The presence of ALT contigs in the reference genome causes the use of the serial bwa-postalt to post-process the BWA results.

Benchmark Test 1, which is marked by periods of both heavy and light computation, is representative of the CPU usage in a typical NGS pipeline.

---


**Benchmark Test 2: Homo_sapiens_assembly38**

Run-bwamem is used to map ERR009171, consisting of 18,120,339 read pairs which come from the 1000 Genomes Project Pilot 2, against Homo_sapiens_assembly38, a human reference genome. The BWT index for Homo_sapiens_assembly38 is calculated in advance using the BWT-SW algorithm that is capable of handling the whole human genome.

Run-bwamem prints the following command line for mapping:

```
./seqtk mergepe ERR009171_1.filt.fastq ERR009171_2.filt.fastq  \
| ./bwa mem -p -t{ NTHREAD} Homo_sapiens_assembly38.fasta - 2> out.log.bwamem  \
| ./samtools view -1 - > out.aln.bam;
```

**Note**

NTHREAD is the amount of parallelism to use on the system

The processing steps are as follows:

1. `seqtk mergepe` reads both read pairs, interleaves the information, and sends it to BWA
2. The basic mapping functionality is performed with the BWA-MEM algorithm designed for processing Illumina sequence reads up to 100bp
3. `SAMtools` is used to convert the data from SAM to BAM format

Following these processing steps yields the results in figures 3 and 4:

![Figure 3](image1.png)  ![Figure 4](image2.png)

**Figure 3 and Figure 4.** CPU usage while running Benchmark Test 2 on the 40 core HPE ProLiant XL170r Gen9 Server described in table 1. The y-axis is percentage of CPU usage and the x-axis is time in seconds. The maximum percent CPU usage on the 40 core system is 4,000 100% for each core. Figure 3 (left) shows usage for all processes, including BWA, seqtk, and SAMtools. Figure 4 (right) shows CPU usage for the seqtk and SAMtools processes only

Figures 3 and 4 show CPU usage while running Benchmark Test 2 on a 40 core HPE ProLiant XL170r Gen9 Server. These figures show a CPU usage profile that differs dramatically from figures 1 and 2. In this benchmark test, `seqtk` takes a few seconds to read and merge an initial section of data. BWA dominates computer usage in figure 3, consuming all of the available cores for most of the run time. The BWA profile dips when waiting for data from `seqtk` or sending data to `SAMtools`. In figure 4, the `seqtk` and `SAMtools` applications are serial, therefore they never use more than 100% of the CPU time. In this example, they compete with BWA for CPU time.
The results of these two benchmark tests demonstrate that CPU usage is dominated by BWA. BWA is considered the most compute-intensive portion of the NGS pipeline, often consuming all the compute resources of a compute node.

**Performance**

**Serial performance**

It is important to examine single core performance as most NGS pipelines include many serial, minimally parallel subtasks. Serial performance is strongly dependent on the processor and memory, with other factors that influence performance typically associated with heavy usage. Processor frequency is a significant feature of single core performance, as higher frequency processors generally offer better performance.

Though often overlooked, the processor family goes hand-in-hand with computational speed. In 2007, Intel introduced the Tick-Tock model to new processor family releases. A “tick” represents a shrinking of the process technology in even-numbered families, resulting in more cores per chip. Sometimes, like in the release of the current v4 processors, new instructions are introduced. A “tock” is a change to the microarchitecture, which results in faster instructions.

Processor families designate the speed of memory chips; the newer the family, the faster the memory DIMMs.

Figure 5 shows the relative performance of the Intel Xeon Processor E5 v2, v3, and v4 families of chips during Benchmark Tests 1 and 2.

![Figure 5](attachment:image.png)

**Figure 5.** Relative performance corrected for frequency of both benchmark tests on servers with the Intel Xeon E5 v2, v3, and v4 processors

In figure 5, the actual application performance is scaled by the processor frequency speed to compensate for frequency differences; thus, the performance reflects the effects of differences in processor families and memory speeds. Benchmark Test 1 shows a performance improvement of approximately 12% with each family upgrade, while Benchmark Test 2 shows a performance improvement of approximately 8% with each family upgrade.

Figure 5 shows that the Intel Xeon Processors E5 v4 family—the latest generation of E5 processors—delivers the best results in terms of serial performance.

**More cores versus faster clock**

The selection of processors for an NGS computing solution must be strongly correlated to performance. Intel offers a wide selection of processors within the Xeon E5 v4 family of processors, ranging from processors with many slower cores to processors with fewer but faster cores. The HPE ProLiant XL170r Gen9 system is modified to incorporate three of these v4 processors, which are described in table 2.
The HPE ProLiant XL170r Gen9 system described in Table 1 is modified with three Intel Xeon E5 v4 processors.

<table>
<thead>
<tr>
<th>Processor model</th>
<th>E5-2690 v4</th>
<th>E5-2698 v4</th>
<th>E5-2699 v4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor family</td>
<td>Intel Xeon Processor E5 v4</td>
<td>Intel Xeon Processor E5 v4</td>
<td>Intel Xeon Processor E5 v4</td>
</tr>
<tr>
<td>Processor speed</td>
<td>2.6 GHz</td>
<td>2.2 GHz</td>
<td>2.2 GHz</td>
</tr>
<tr>
<td>Cores/Node</td>
<td>28</td>
<td>40</td>
<td>44</td>
</tr>
<tr>
<td>L3 Cache</td>
<td>35 MB shared between 14 cores</td>
<td>50 MB shared between 20 cores</td>
<td>55 MB shared between 22 cores</td>
</tr>
</tbody>
</table>

The relative performance of the three v4 processor models (E5-2690, E5-2698, and E5-2699) on the HPE ProLiant XL170r Gen9 system is shown in figure 6.

![Figure 6](image-url)

Benchmark Test 1, which spends approximately half the time using fewer than 2 cores, appears to be virtually insensitive to the processor choice. The benefit from the higher core count of the E5-2699 v4 processor is similar to that of the faster, fewer cores of the E5-2690 v4 processor. Benchmark Test 2 is more compute-intensive, and therefore better able to take advantage of parallelism, making the 22 core E5-2699 v4 the clear choice for best processor in terms of performance.

This example demonstrates that processors with higher-count, slower cores can be a better compute solution for NGS pipelines than processors with fewer, faster cores. The E5-2699 v4 processor with twenty-two 2.2 GHz cores delivers higher levels of performance than the E5-2690 v4 processor with fourteen 2.6 GHz cores.

**Storage performance**

Computing centers that process NGS data are under intense pressure to upgrade to greater amounts of efficient storage. In a typical NGS workflow, data is taken from the sequencer and placed on a shared file system, and compute nodes attached to that shared file system are used to process the data. Growth in the number of instruments used by the sequencing applications typically leads to an exponential expansion in the file systems size.

To show the degree to which efficient storage affects time-to-solution, the benchmarks are run on the HPE ProLiant XL170r Gen9 system using the NFS and Intel Enterprise Edition for Lustre® V3.0 shared storage systems described in Table 3.
Stress on a shared file system typically requires demand from multiple nodes. To measure the effect of demand, the benchmark tests are run repeatedly on different numbers of nodes, with the start times on each node staggered. The assumption is that sequencing at NGS centers often uses the same reference genome, so a typical run is able to leverage memory for the reference. The file system is stressed for reads of sequences and writes. A single 40-way parallel job is run on each compute node.

Performance is measured for only jobs using all nodes. The average performance is used to measure the number of jobs that will run on a compute node in one day.

Figure 7 shows the measured number of Benchmark Test 1 jobs that can be run on a single compute node in one day as a function of the number of compute nodes using the shared storage.

---

### Table 3. The storage systems attached to the HPE ProLiant XL170r Gen9 system described in table 1

<table>
<thead>
<tr>
<th>Storage system</th>
<th>NFS</th>
<th>Intel EE for Lustre* V3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage server</td>
<td>HPE ProLiant DL560 Gen8</td>
<td>HPE ProLiant Apollo 4520</td>
</tr>
<tr>
<td>Number of servers</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Processor family</td>
<td>Intel Xeon Processor E5 v2</td>
<td>Intel Xeon Processor E5 v4</td>
</tr>
<tr>
<td>Processor speed</td>
<td>2.4 GHz</td>
<td>2.6 GHz</td>
</tr>
<tr>
<td>Processor model</td>
<td>E5-4650 v2</td>
<td>E5-2690 v4</td>
</tr>
<tr>
<td>Processor/Node</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cores/Node</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Memory/Node</td>
<td>256 GB 1600 MHz DDR3 2 DPC</td>
<td>128 GB 2400 MHz DDR4</td>
</tr>
<tr>
<td>L3 Cache</td>
<td>25 MB shared between 10 cores</td>
<td>35 MB shared between 14 cores</td>
</tr>
<tr>
<td>Interconnect</td>
<td>10G Ethernet</td>
<td>Omni-Path Architecture</td>
</tr>
<tr>
<td>Storage</td>
<td>2.2TB raid 5 set of 5 600GB SAS drives</td>
<td>23 4TB 12G 7.2K SAS disks per node; JBOD containing 140 4TB 12G 7.2K SAS disks</td>
</tr>
</tbody>
</table>

---

**Figure 7.** The y-axis is the number of Benchmark Test 1 jobs that can be run on a compute node per day. The x-axis is the number of compute nodes.
The Intel EE for Lustre* software file system shows no degradation in performance up to 32 compute nodes. The NFS file system performance begins to decline starting at 8 compute nodes. At 32 compute nodes, the degradation in performance for the NFS file system is about 60%.

Figures 8 and 9 show a performance profile on a single compute node for Benchmark Test 1 when running 32 compute nodes with the NFS file system. The Intel EE for Lustre* software file system is unstressed by the multi-node workload; the Benchmark 1 performance profile on the heavily loaded Intel EE for Lustre* file system is the same as on a single compute node.

**Figure 8 and Figure 9.** CPU usage while running Benchmark Test 1 using the NFS file system on the 32 forty core HPE ProLiant XL170r Gen9 computers described in table 1. The y-axis is percentage of CPU usage and the x-axis is time in seconds. The maximum percent CPU usage on the 40 core system is 4,000, 100% for each core. Figure 8 (left) shows usage for all processes, including BWA, bwa-postalt, and SAMtools. Figure 9 (right) shows CPU usage for the bwa-postalt and SAMtools process only.

Figures 10 and 11 show a performance profile on a single compute node for Benchmark Test 1 when running 32 compute nodes with both the NFS and Intel EE for Lustre* file system.

**Figure 10 and Figure 11.** CPU usage while running Benchmark Test 1 using both the NFS and Lustre* file systems on the 32 forty core HPE ProLiant XL170r Gen9 computers described in table 1. The y-axis is percentage of CPU usage and the x-axis is time in seconds. The maximum percent CPU usage on the 40 core system is 4,000, 100% for each core. Figure 10 (left) shows usage for BWA alone. Figure 11 (right) shows CPU usage for the bwa-postalt and SAMtools process only.
The dependence of BWA on the heavily utilized file systems is shown in figure 10. The read of the sequences on the overloaded NFS file system takes 20 seconds longer than the Intel EE for Lustre* file system, causing the delay of the start of BWA. In addition, figure 11 shows that the slow NFS file system is responsible for a degradation in the bwa-postalt CPU usage.

The measured number of Benchmark Test 2 jobs that can be run on a single compute node as a function of the number of compute nodes using the shared storage is shown in figure 12.

The Intel EE for Lustre* software file system shows no degradation in performance up to 32 compute nodes. The NFS file system performance begins to decline starting at 4 compute nodes. At 32 compute nodes, the degradation in performance is significant, about 85%.

![Figure 12. The y-axis is number of benchmark test 2 jobs that can be run on a compute node per day. The x-axis is the number of compute nodes.](image)

Figures 13 and 14 show the CPU usage for Benchmark Test 2 run on 32 nodes with overloaded NFS storage. The performance profile for Benchmark 2 on a single compute node is the same as when 32 compute nodes stress the Intel EE for Lustre* software file system.

![Figure 13 and Figure 14. CPU usage while running Benchmark Test 2 using the NFS shared file system on the 32 forty core HPE ProLiant XL170r Gen9 computers described in table 1. The y-axis is percentage of CPU usage and the x-axis is time in seconds. The maximum percent CPU usage on the 40 core system is 4,000, 100% for each core. Figure 13 (left) shows usage for all processes, including BWA, seqtk, and SAMtools. Figure 14 (right) shows only the seqtk and SAMtools processes.](image)
Figure 15 and Figure 16. CPU usage while running Benchmark Test 1 using both the NFS and Lustre* file systems on the 32 forty core HPE ProLiant XL170r Gen9 computers described in table 1. The y-axis is percentage of CPU usage and the x-axis is time in seconds. The maximum percent CPU usage on the 40 core system is 4,000, 100% for each core. Figure 15 (left) shows usage for BWA alone. Figure 16 (right) shows CPU usage for the bwa-postalt and SAMTools process only.

Figures 15 and 16 show the performance profiles on the heavily utilized NFS and Intel EE for Lustre* software shared file systems. The differences are dramatic. Figures 13 and 14, illustrate that the jobs on the overloaded NFS file system play a waiting game with spikes of activity. Seqtk waits for data from the NFS file system, BWA waits for seqtk to deliver the merged data, and SAMTools spends its time waiting for data from BWA.

These examples prove that using the efficient Intel EE for Lustre* software V3.0 shared file system with Intel Omni-Path Architecture card provides a significant improvement in time-to-solution at computing centers that run multiple NGS workflows. Intel EE for Lustre* software delivers the fast, massively scalable storage required to accelerate the performance of NGS workloads, while Intel OPA establishes a next-generation fabric optimized for density, low latency, and high bandwidth.

Conclusions

The results of these two benchmark tests confirm that CPU usage in NGS workloads is dominated by compute-intensive BWA. At such a high CPU usage, the Intel Xeon E5 v4 family of processors deliver maximum performance and efficiency for a wide range of workloads. Specifically, the E5-2699 v4 processor with twenty-two 2.2 GHz cores delivered the best performance, indicating that processors with higher-count, slower cores are a better compute solution for NGS pipelines than processors with fewer, faster cores. Furthermore, computing centers that often run multiple NGS workflows can reduce time-to-solution by utilizing an Intel EE for Lustre* software V3.0 shared file system run on an Intel Omni-Path Architecture card.

Next-generation sequencing requires HPC solutions and storage infrastructures that deliver unmatched levels of performance, scalability, and affordability for supporting today’s most data-intensive workloads. A compute solution comprised of HPE ProLiant Gen9 servers, Intel Xeon E5 v4 processors, Intel EE for Lustre* software V3.0 shared file systems can significantly reduce the processing time for today’s large genomic datasets and accelerate discovery and development cycle times.

Hewlett Packard Enterprise and Intel have extensive experience working with industry-leading pharmaceutical companies and independent software providers to deliver solutions that offer high performance, flexible capacity, secured availability, and simplified management. A broad choice of standards-based solutions, proven in many of the most demanding research and production environments in the world, will ensure businesses will find the optimal solution to fit their needs and budget.
Learn more at
hpe.com/info/hpc-bigdata-industrysolutions