Big data technologies for ultra-high-speed data transfer in life sciences

Using Aspera's high-speed data transfer technology to achieve 10 Gbps WAN speeds on the Intel® Xeon® Processor E5-2600 family in virtualized environments

aspera
an IBM® company

Contents:

Executive summary

A major challenge in high-performance cloud computing is moving big data in and out of the back-end data center. While high-performance servers and hardware are already deployable inside the data center, and WAN bandwidth can be provisioned beyond multi-gigabits per second (Gbps), existing transport technology lacks the capability of fully using the end-to-end capacity provided by the underlying hardware platform, particularly over the wide area.

Improvements in widely available compute, storage, and network components are rapidly improving the ability of commercial, academic, and government organizations to effectively handle big data. By integrating 10-gigabit Ethernet (10 GbE) into its mainstream servers with built-in technologies that enhance input/output (I/O) throughput for network and storage traffic, Intel has demonstrated I/O improvements by up to 2.3 times[1, 2], reduced network latency by up to 30 percent[2, 3], and improved PCIe support[4] for increased bandwidth by as much as two times[5] per server.

IBM Aspera develops high-speed data transfer technologies that provide speed, efficiency, and bandwidth control over any file size, transfer distance, network condition, and storage location (i.e., on-premises or cloud). Aspera's FASP®* transport technology has no theoretical throughput limit and can only be constrained by the available network bandwidth and the hardware resources at both ends of the transfers.

This white paper describes the results of a joint effort between Aspera and Intel to investigate ultra-high-speed (10 Gbps and beyond) data transfer solutions built on Aspera's FASP transport technology and Intel's hardware platform.

The investigation shows the performance advantage for Aspera transfers on Intel hardware in virtualized computing environments typical of cloud computing platforms. The high-end Intel platform tested is powered by the Intel® Xeon® processor E5-2600 family, which supports Intel® Data Direct I/O Technology (Intel® DDIO). The test system also had built-in support for single-root I/O virtualization (SR-IOV), a technology that allows virtual machine platforms to bypass the hypervisor to directly access resources on the physical network interface for exceptional I/O performance. While the Aspera transfers can easily achieve 10 Gbps speed with a single stream using jumbo frames, the emphasis of this investigation is to demonstrate maximum possible speed using regular packet sizes (1492B) to show what is possible on commodity Internet connections.

The testing showed dramatic throughput improvement for Aspera transfers on the Intel® platform with Intel DDIO and built-in support for SR-IOV (over 300 percent) versus a baseline system with no support for Intel DDIO and SR-IOV. It also showed that the Aspera transfer speed performance on the high-end Intel Xeon processor-based system is about the same for both physical and virtualized computing environments. Furthermore, both the LAN and WAN transfer speeds had similar results, confirming the Aspera FASP distance-neutral transfer performance.

Challenges in life sciences

The field of life sciences is faced with a series of big data challenges as the emphasis shifts from raw sequencing performance to mapping, assembly and analytics. The need to transmit terabytes of genomic information between sites worldwide is both essential and daunting, including:

- Collaboration with research and clinical collaborators worldwide to establish statistically significant patient cohorts and use expertise across different institutions.
- Reference genomes used to assemble sequences, conduct quality control, identify and annotate variants, and perform genome-wide analysis studies (GWAS).
- Cloud-based analytics to address critical shortages in bioinformatics expertise and burst capacity for high-performance computing (HPC) clusters.
- Data management and resource utilization across departments in shared research HPC cluster environments, analytics clusters, storage archives, and external collaborators.
- Medical genomics extends data management considerations from research to clinical collaborators, CLIA labs, hospitals, and clinics.

Most institutions still rely on shipping physical disks due to inherent problems with commodity 1 Gb Ethernet (GbE) networks and TCP inefficiencies. The transition from 1 GbE to 10 GbE and beyond has been unusually slow in life sciences, likely due to an overemphasis on shared compute resources out of context from the broader usage and system architecture requirements. Data centers in other segments have been quick to adopt 10 GbE and unified networking due to impressive cost savings, performance, and manageability considerations. Adopting a balanced compute model – where investments in processor capacity are matched with investments in network and storage – yields significant performance gains while reducing data center footprint and power and cooling costs. Improved server density and shared resource utilization drive the need for virtualization. While I/O optimization historically has addressed jumbo packet transmissions on physical infrastructure, a more realistic test is that of regular packets, comparing physical and virtualized environments over both LAN/WAN traffic conditions.

Investigation set-up

The test had two basic set-ups:

1. One to investigate transfers over the local area network (LAN)
2. One to investigate transfers over the wide area network (WAN)

Figure 1 shows the first set-up, for the LAN test, with two high-performance Intel Xeon processor-based servers connected back-to-back via a 10Gbps SFP+ Ethernet cable.

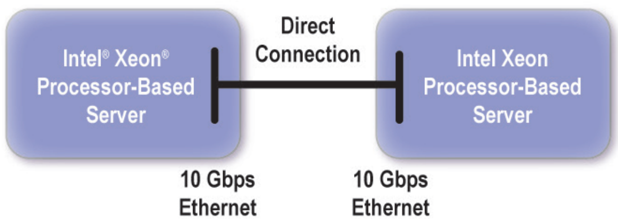

Figure 1: Local-area network set-up

Each of the Intel Xeon processor-based servers uses the Intel DDIO and non-uniform memory access (NUMA) technologies to deliver exceptional I/O throughput. Each system is installed with Red Hat Enterprise Linux* 6.3 OS with loadable kernel-based virtual machine (KVM) modules. IBM® Aspera® Enterprise Server 3.1 was also installed on both systems to test the data transfer throughput between the two servers.

The hardware platform included:

- Intel Xeon processor E5-2600 family (3.10 GHz, 16 cores with Intel® Hyper-Threading Technology [Intel® HT Technology])
- NUMA and Intel DDIO enabled as default
- 32 GB DDR3 main memory
- Intel® Solid-State Drive (Intel® SSD) 910 series, 800 GB RAID0 each system
- Intel Corporation 82599EB 10-Gigabit SFI/SFP+ Network Connection

The software platform included:

- Red Hat Enterprise Linux 6.3 (KVM disabled by default)
- Aspera Enterprise Server 3.1.1.66573
- Aspera Performance Automation Suite*

The set-up for the test transfers over the WAN is shown in *Figure 2*. In this set-up, the two Intel Xeon processor-based servers were connected with the 10 Gbps Netropy* WAN emulator via two sets of 10 Gbps SFP+ Ethernet cable to emulate typical WAN latency and packet loss conditions. Except for the network set-up, the other hardware and software configurations were the same as for the LAN test.

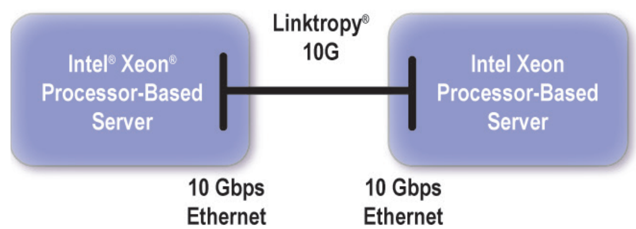

Figure 2: Wide-area network set-up

To demonstrate transfer performance over virtualized platforms, the test used KVM, a well-known virtualization environment supporting SR-IOV, on top of the standard Linux distribution, and installed Aspera Enterprise Server on a number of virtual machines hosted by KVM. Specific software deployed on top of Intel's high-performance platform for the test involving transfers on virtualized platforms include:

Hardware (BIOS) configuration (to enable virtualization and SR-IOV):

- Intel VT-d enabled
- Intel Virtualization Technology enabled

Virtualization configuration:

- Host OS: Red Hat Enterprise Linux 6.3 with KVM enabled
- Guest OS: Red Hat Enterprise Linux 6.3 (up to four-guest OS)

Intel® Integrated I/O

Intel® Integrated I/O manages data traffic through:

- The integrated I/O controller, which significantly reduces latency
- Integrated PCIe* 3.0 support with more lanes than the previous generations for greater flexibility in connecting peripheral devices to the processor, which can double bandwidth
- Support for a growing ecosystem of PCIe 3.0 add-in cards
- Intel® Data Direct I/O Technology (Intel® DDIO)

Intel Data Direct I/O Technology (Intel DDIO)

Intel DDIO is a key component of Intel Integrated I/O that increases performance by allowing Intel Ethernet controllers and server adapters to talk directly with cache and maximize throughput. Traditional I/O transfer must first be moved and stored in main memory before it can go to cache for processing. Then, once processing is complete, the data has to reverse its course. With Intel DDIO, Intel has rearchitected the processor and dedicated a portion of cache to I/O so that data transfers directly to cache and bypasses main memory. This places less demand on main memory to deliver greater bandwidth scalability, lower power utilization, and reduced latency.

Intel® Virtualization Technology (Intel® VT) for Connectivity (Intel® VT-c)

Intel® Virtualization Technology (Intel® VT) for Connectivity (Intel® VT-c) is a collection of platform-level I/O virtualization technologies and initiatives that enable lower CPU utilization, reduced system latency, and improved networking and I/O throughput. Intel VT-c optimizes virtualized systems with a multifaceted approach to I/O virtualization:

- **Virtual Machine Device Queues** (VMDq) improve traffic management within the server, helping to enable better I/O performance from large data flows while decreasing the processing burden on the software-based virtual machine monitor (VMM).
- **PCI-SIG* SR-IOV** provides near-native performance by providing dedicated I/O to virtual machines and completely bypassing the software virtual switch in the hypervisor. It also improves data isolation among virtual machines and provides flexibility and mobility by facilitating live virtual machine migration.

Test details

While Aspera transfers can easily achieve 10 Gbps speed with a single stream using jumbo frames, the emphasis of this test was to demonstrate maximum possible speed using regular packet sizes (1492B) to show what is possible on commodity Internet connections.

The team tested data transfer speed for three cases using regular packet sizes. In all of the tests, unless mentioned specifically, each FASP session was used to transfer a single large file (i.e., larger than typical CPU and file cache sizes, to eliminate caching benefits). To minimize the overhead other than I/O, encryption was turned off and reading and writing from and to the hard drive were disabled.

Case I: Benchmarking test

The initial test benchmarked the raw performance of Aspera transfers on the Intel Xeon processor E5-2600 family equipped with DDIO. The test also compared the performance of the Aspera transfers on an Intel Xeon processor E5-400 family-based server that did not support DDIO. The intent was to demonstrate the potential benefit from transfers on machines with DDIO.

Case II: LAN transfers over virtual platforms

This test focused on data transfer over a 10 Gbps LAN connection between high-end Intel Xeon processor-based systems with built-in support for SR-IOV. The Intel Xeon processor E5-2600 family-based system includes an Ethernet driver with built-in support for SRIOV, which allows virtual machine platforms to bypass the hypervisor to directly access resources on the physical Ethernet interface. The SR-IOV support enabled the team to test the throughput users can expect from the Intel platform in virtualized environments, using SR-IOV to speed transfers over the LAN.

Case III: WAN transfers over virtual platforms

This test focused on data transfer over a 10 Gbps WAN connection between Intel Xeon processor E5-2600 family-based systems with built-in support for SR-IOV. The test investigated how the performance benefit in Case II could extend to the wide-area network (WAN) where latency and packet losses are common. The team used the configuration in *Figure 2*, with two Intel Xeon processor-based servers connected via a Netropy* WAN emulator. The Netropy WAN emulator allowed the team to try different combinations of network latency and packet loss conditions, for speeds up to 10 Gbps.

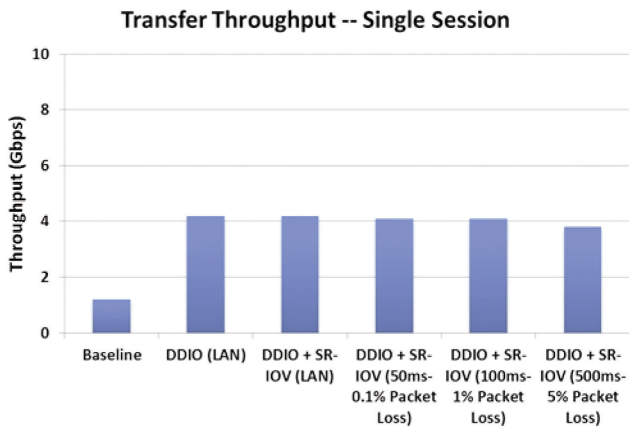

Figure 3: Aspera transfers on physical and virtualized platforms for both LAN and WAN scenarios

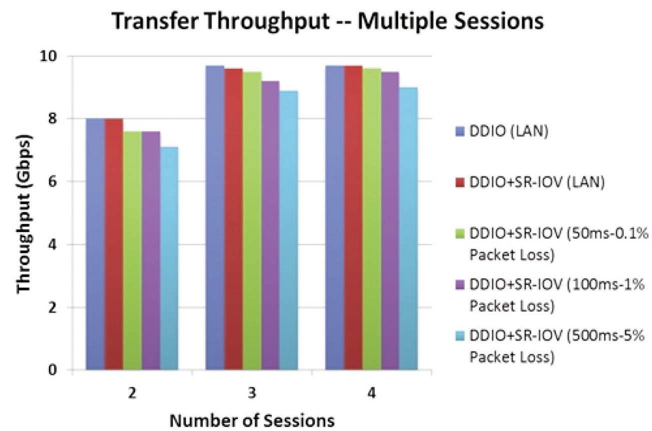

Figure 4: Aspera transfers on physical and virtualized platforms for both LAN and WAN scenarios when running in multi-threaded mode

The team tested 10 Gbps Aspera transfers over the following emulated WAN conditions (high latency and typical packet loss ratios):

- 50ms, 0.1 percent packet loss (coast-to-coast)
- 100ms, 1 percent packet loss (intercontinental)
- 500ms, 5 percent packet loss (poor connection, global distance, and/or wireless media)

Results

Figure 3 shows the results of the test for transferring a single stream using normal packet sizes over the LAN and WAN and on both physical hardware and virtualized platforms.

As noted, the baseline machine is an Intel Xeon processor E5-400 family-based system with no built-in support for DDIO. The Intel Xeon processor E5-2600 family-based system with built-in support for DDIO (bare metal LAN) achieved a transfer rate over 300 percent better than the baseline system. *Figure 3* also shows that the Aspera transfer performance on the Intel Xeon processor E5-2600 family-based system with DDIO support is about the same for transfers running on both physical and virtualized platforms. Moreover, the performance over the virtual platforms is about the same for both transfers over the LAN and over the WAN under different network conditions, even when the network conditions degrade significantly (500 milliseconds latency and 5 percent packet loss). The slight reduction in throughput when the network degrades is primarily due to injection of packet losses.

Figure 4 shows the results of test transfers using multiple streams to achieve the maximum possible end-to-end throughput. The test compared the maximum possible performance achievable using two, three, or four streams when FASP is running in multi-threaded mode. Again, the LAN performance was about the same as the WAN performance, even when the packet loss rate is as high as 5 percent and the network latency is 500 milliseconds. *Figure 4* also shows that Aspera only needs three to four streams to fully utilize the available 10 Gbps bandwidth. The slight reduction in throughput when the network degrades is primarily due to injection of packet losses.

Conclusions

Aspera transfers can easily achieve 10 Gbps with a single stream using jumbo frames. However, the investigation by Intel and Aspera focused on demonstrating the maximum possible speed using regular packet sizes (1492B). The intent is to show what performance users can expect for transfers over commodity Internet connections. Aspera transfers on an Intel Xeon processor-based platform equipped with DDIO and built-in support for SR-IOV offer over 300 percent throughput improvement over a baseline system with no support for DDIO and SR-IOV.

The Aspera transfer speed performance on the Intel Xeon processor E5-2600 family-based system is about the same for both physical and virtualized computing environments. Also, both the LAN and WAN transfer speeds had similar results, confirming a key attribute of Aspera FASP. In both the LAN and WAN use case, Aspera needed only three to four streams when running in multi-threaded mode to fully use the available 10 Gbps bandwidth.

Future work

Future investigation will identify the bottleneck components along the end-to-end transfer path that limited the effective throughput for a single transfer versus the theoretical throughput of 10 Gbps. This future investigation will also optimize the transfer for a single data stream and may involve integration of Intel® Data Plane Development Kit (Intel® DPDK), which allows multiple applications to directly and concurrently control the physical NIC interface, therefore bypassing the kernel networking stack. The Intel DPDK technology reduces the number of memory copies needed to send and receive a packet. As a result, Intel expects to be able to boost data transfer speed for a hardware configuration that supports DPDK.

Aspera FASP transfer technology Aspera's patented FASP transfer technology is software that eliminates the fundamental shortcomings of conventional, TCP-based file transfer technologies such as FTP and HTTP. It can achieve speeds up to hundreds of times faster than FTP/HTTP and provide a predictable and reliable delivery time regardless of file size, transfer distance, or network conditions, including transfers over satellite, wireless, and inherently long distance and unreliable international links. FASP also provides visibility into bandwidth utilization and control over transfer rates and bandwidth sharing with other network traffic. Built-in security, including secure endpoint authentication, on-the-fly data encryption, and integrity verification.

To learn more, visit www.asperasoft.com/technology/transport/fasp.

About Aspera, an IBM Company

Aspera, an IBM company, is the creator of next-generation transport technologies that move the world's data at maximum speed regardless of file size, transfer distance and network conditions. Based on its patented, Emmy® award-winning FASP® protocol, Aspera software fully utilizes existing infrastructures to deliver the fastest, most predictable file-transfer experience. Aspera's core technology delivers unprecedented control over bandwidth, complete security and uncompromising reliability. Organizations across a variety of industries on six continents rely on Aspera software for the business-critical transport of their digital assets.

For more information

For more information on IBM Aspera solutions, please visit ibm.com/software/aspera and follow us on Twitter @asperasoft.

1 Ethernet consolidation source: Intel 10 GbE ROI Calculator. This ROI calculator is a cost comparison for a highly virtualized solution, using multiple 1 GbE connections, versus a dual-port 10 GbE implementation: event-management online.de/LAD/calculator.

2 Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests such as SYSmark* and MobileMark* are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

3 Intel internal measurements of average time for an I/O device read to local system memory under idle conditions comparing Intel Xeon processor E5-2600 family (230 ns) vs. Intel Xeon processor 5500 series (340 ns). Baseline Configuration: Green City system with two Intel Xeon processor E5520 (2.26 GHz, 4 C), 12 GB memory @ 1333, C-States Disabled, Turbo Disabled, SMT Disabled. New Configuration: Meridian system with two Intel Xeon processor E5-2665 (2.4 GHz, 8 C), 32 GB memory @1600 MHz, C-States Enabled, Turbo Enabled. The measurements were taken with a LeCroy* PCIe* protocol analyzer using Intel internal Rubicon (PCIe 2.0) and Florin (PCIe 3.0) test cards running under Windows* 2008 R2 w/SP1.

4 Intel IT. *Upgrading Data Center Network Architecture to 10 Gigabit Ethernet* (January 2011). intel.com/content/dam/www/public/us/en/documents/best-practices/inte-it-datacenter-efficiency-upgrading-data-center-network-architecture-to-10-gigabit-ethernet-practices.pdf.

5 Eight GT/s and 128 b/130 b encoding in PCIe 3.0 specification enables double the interconnect bandwidth over the PCIe 2.0 specification. Source: pcisig.com/news_room/November_18_2010_Press_Release.

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
February 2015

ZZW03369-USEN-00