Performance Characteristics of IBM® InfoSphere® Information Server 8.7 Running on VMware® vSphere® 5.0 on Intel® Xeon® Processors

Executive Summary
Virtualization has become one of the hottest technologies in recent years. With the continuous growth of processing power and the increasing number of processor cores in servers, virtualization has become very attractive for many companies. As more and more IBM® InfoSphere® Information Server customers begin to adopt virtualization technology, understanding the performance characteristics of InfoSphere Information Server running in virtualized environments has become critical. IBM, VMware, and Intel have collaborated to study the runtime performance of IBM® InfoSphere® DataStage® 8.7 running in a virtual machine on VMware® vSphere® 5.0 on a server featuring the Intel® Xeon® processor E7 family. This study demonstrated that InfoSphere DataStage runs well in a virtualized environment; InfoSphere DataStage was able to scale as the number of virtual CPUs (vCPUs) was increased. Moreover, under normal conditions, throughput was only 2 percent to 10 percent lower for DataStage running in a virtualized environment.

About the Authors
Samuel Wong is a member of the IBM InfoSphere InfoSphere Information Server performance team at the IBM Silicon Valley Lab. He graduated from University of Toronto and he has 13 years of software development experience with IBM. His email is samwong@us.ibm.com.

Jantz Tran is a Software Performance Engineer for Intel. He has been analyzing and optimizing enterprise software on Intel server platforms for 10 years. Jantz has a BSCE from Texas A&M University. His email is jantz.c.tran@intel.com.

Pin Lp Lv is a Software Performance Engineer from IBM. Pin has worked for IBM since 2006. He worked as a software tester for IBM WebSphere Product Center Team and RFID Team from September 2006 to March 2009, and joined IBM InfoSphere Information Server Performance Team in April 2009. Pin has a Master of Science degree in Computer Science from University of West Scotland. His email is pinlv@cn.ibm.com.
Introduction
Intel, VMware, and IBM performance teams partnered for the first comprehensive industry study on executing data integration workloads in a virtualized environment. Virtualization has grown to be an important tool in many enterprises for facilitating server provisioning, easing application manageability, and increasing workload availability. However, enterprises also worry about how processing-intensive and input/output-intensive workloads like data integration perform in virtual environments. To demonstrate the viability of virtualizing data integration and to help businesses realize the benefits of running these workloads on virtual machines, Intel, VMware, and IBM sought through this study to guide adoption of IBM InfoSphere DataStage in VMware environments on Intel processors and provide best-practice, tuning, and configuration information.

Virtualization Architecture
VMware® vSphere® 5.0 offers an exceptional platform for system virtualization to host very demanding workloads such as data integration. Features such as VMware® vMotion™ enable moving workloads from physical server to physical server with no downtime, simplifying common operations like hardware maintenance without impacting operations. VMware vSphere High Availability (HA) and VMware vSphere Fault Tolerance (FT) further enhance enterprise operations by providing administrators with additional options for application availability to reduce unplanned downtime.

Processors
The Intel Xeon processor E7 family is particularly well suited for running InfoSphere DataStage in a virtualized environment. Processors in this family can scale up to 10 cores per processor, have up to two times the input/output (I/O) capacity compared to Intel Xeon processors from other product families, and support up to 512 GB of memory per processor. The Intel Xeon processor E7 family’s advanced reliability, availability, and serviceability (RAS) features, such as Machine Check Architecture (MCA) Recovery, also improve the ability of VMware vSphere to recover from hardware memory errors. Moreover, the Intel Xeon processor E7 family hosts a number of technologies designed for virtualization. These technologies include Intel® Virtualization Technology (Intel® VT-x), Intel® Virtualization Technology for Directed I/O (Intel® VT-d), and Intel® Virtualization Technology for Connectivity (Intel® VT-c), which enhance virtualization performance and the robustness of the processor, memory, I/O, and networking. VMware natively supports many of these features, transparently providing these benefits to end users.

Test Objectives
This study of InfoSphere Information Server running on VMware on Intel Xeon processors focused on InfoSphere DataStage runtime performance with the following objectives:
• Understanding the amount of performance difference, under normal conditions, for DataStage jobs running in a virtualized environment using VMware vSphere compared to a native environment
• Testing file store access configurations such as Raw Device Mapping (RDM) versus Virtual Machine File System (VMFS)
• Gauging the effects on DataStage performance of overloading (over-committing) the host server with too many discrete virtual machines
• Judging the effectiveness of VMware recommendations for CPU reservations

The study used two workloads, one CPU intensive and the other I/O intensive. The CPU-intensive workload was a DataStage job that was part of a data integration scenario. The principal components of
this CPU-intensive job included lookup stages and transformer stages. The job first read the source from a sequential file and then processed the data in a transformer stage. Next, the data flowed through several lookup stages to be matched against reference tables. Finally, the data was processed by a second transformer stage. Since this job was designed to be representative of a CPU-intensive workload, the final output link was replaced with a Copy Stage that performed no operation. Though reading the source data file necessarily generated some disk I/O calls, their numbers were negligible and did not affect the CPU utilization. By design, CPU utilization was the bottleneck for this job.

Similar to the job used for the CPU-intensive workload, the I/O-intensive workload was also a DataStage job (and part of the same data-integration scenario). Beyond the lookup and transformer stages shared with the CPU-intensive job, the key difference came with the addition of a sort stage. In the course of sorting data, sort stages write data to a temporary sorting area (the scratch space area). I/O thus becomes the bottleneck for such operations. (See “Performance Guidelines for IBM InfoSphere DataStage Jobs Containing Sort Operations on Intel® Xeon® Servers” for technical details on sort operations.1, 2)

### Test Configuration and Methodology

The test used an IBM® System x3850 X5 for the server. It was a four-socket system powered by Intel Xeon E7-8870 processors with a total of 40 physical cores. Network-attached storage was provided by an IBM® System Storage® DS5300 system. The server was set up to dual boot between VMware® vSphere® 5.0 and Red Hat Enterprise Linux® (RHEL®) 6.1, thus ensuring that the same hardware was used to compare native and virtualized results. Moreover, the amount of memory and the number of CPUs available to the native environment were controlled to match the virtual machine configuration as much as possible. The test used the “mem=” kernel parameter\(^3\) to control the amount of memory. To control the number of CPU cores in the native environment, the test used the “CPU HotPlug Support in Linux Kernel” method. CPU cores were activated in groups of eight per socket. For example, in a test case using eight CPUs, all eight cores came from one socket. In contrast, for a test case using 16 CPUs, eight cores came from one socket while the other eight cores came from another socket. For the BIOS settings, the server was configured as “optimized for performance” and all of the virtualization-related options were enabled. Intel® Hyper-Threading Technology (Intel® HT Technology) was disabled to simplify the comparison between the two environments.

All virtual-machine files for the test resided on the storage area network (SAN). The operating system and InfoSphere DataStage were installed on a 35 GB virtual disk and each of the virtual machines running the InfoSphere DataStage workloads had an additional 200 GB virtual disk for source data, dataset, and scratch space. The virtual disks were thick provisioned and eager zeroed. VMware Paravirtual Small Computer System Interface (PVSCSI) was used for the SCSI controller type. The virtual network adapter for the virtual machines used VMXNET3 and was connected to a single physical 10 gigabit Ethernet (GbE) network adapter through a virtual switch. For the engine-cluster test scenario, a private network was configured for the cluster; virtual machines in the cluster were configured with an additional virtual adapter and connected to an additional virtual switch that was not connected to a physical network adapter.

In setting up InfoSphere DataStage, all of the components were installed on one physical server in the native environment and one virtual machine in the virtualized environment. This included the IBM® WebSphere® Application Server (WAS), XMeta repository, and DataStage engine. The test used the bundled version of WAS and IBM® DB2® Enterprise Server Edition 9.7 that came with InfoSphere Information Server 8.7 GA. The DataStage jobs were executed directly on the server using Orchestrate Shell Script (OSH) and the nmon\(^5\) system monitor was used to capture the performance counters for the guest operating system.

### Table 1. Hardware configuration in the IBM® System x3850 X5 server test environment

<table>
<thead>
<tr>
<th>Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>4x Intel® Xeon® processor E7-8870(^4) 2.4 GHz (10 cores per processor)</td>
</tr>
<tr>
<td>Memory</td>
<td>512 GB</td>
</tr>
<tr>
<td>Network</td>
<td>10 GbE</td>
</tr>
<tr>
<td>Storage</td>
<td>IBM® System Storage® DS5300(^6)</td>
</tr>
</tbody>
</table>

### Table 2. IBM® System Storage® DS5300 storage configuration details for this study

<table>
<thead>
<tr>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibre Channel Adapter</td>
<td>8 Gb</td>
</tr>
<tr>
<td>Number of logical unit numbers (LUNs) used</td>
<td>7</td>
</tr>
<tr>
<td>LUN size</td>
<td>200 GB</td>
</tr>
<tr>
<td>RAID level</td>
<td>0</td>
</tr>
<tr>
<td>Number of physical disks per LUN</td>
<td>8</td>
</tr>
<tr>
<td>Physical disks rpm</td>
<td>15,000 rpm</td>
</tr>
</tbody>
</table>

### Table 3. Software configuration in the test environment

<table>
<thead>
<tr>
<th>Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMware® vSphere®</td>
<td>5.0</td>
</tr>
<tr>
<td>Guest operating system</td>
<td>RHEL® 6.1</td>
</tr>
<tr>
<td>Native operating system</td>
<td>RHEL 6.1</td>
</tr>
<tr>
<td>IBM® InfoSphere® Information Server</td>
<td>8.7GA</td>
</tr>
</tbody>
</table>
Test Results

IBM® InfoSphere® DataStage® Runtime: CPU-Intensive Workload

Test results for the CPU-intensive workload running on both physical and virtual servers using 4, 8, 16, and 32 CPU cores are shown in Figure 1. The magenta and blue lines represent the average CPU utilizations during the runtime execution of the CPU-intensive job in native and virtual environments, respectively. CPU utilization was captured by nmon at five-second intervals and average CPU utilization was calculated when the system reached a steady state. Both ramp-up and ramp-down times were excluded for all test runs. The high CPU utilization recorded confirmed that the workload is CPU intensive across all test configurations.

The bar graph in Figure 1 presents the throughput of the job (magenta native and blue virtualized). Intuitively, throughput in both environments increased as the number of CPU cores increased. Moreover, performance in the two environments varied only between 2 and 10 percent. This confirmed that DataStage runtime on a virtualized environment can scale in a fashion similar to a native environment.

Performance Characteristics of IBM® InfoSphere® Information Server 8.7 Running on VMware® vSphere® 5.0 on Intel® Xeon® Processors
DataStage Runtime: I/O-Intensive Workload

In order to make the workload stress the disk I/O, we ran the test with 32 CPU cores and 4 DataStage nodes in both the virtual and native environments. This provided the workload with enough processing power to drive the disk I/O to higher levels. The test used 1, 2, and 4 LUNs in the SAN for the scratch space area. All LUNs were 200 GB with RAID-0 and were connected to the system via an 8 Gb Fibre Channel. The throughput results for this workload are presented in Figure 2; the relative throughputs for the workload run in the native and virtual environments are represented in the magenta and blue bar graphs, respectively. The overhead was 10 percent for all three LUN configurations. The specific I/O per second (IOPS) and I/O bandwidth figures captured for each test case are shown in Table 4.

RDM Versus VMFS

VMware vSphere 5.0 supports two methods of provisioning storage for access to data on its data stores: Raw Device Mapping (RDM) and Virtual Machine File System (VMFS). RDM accesses the physical disk directly while VMFS abstracts the file system. While VMFS can provide numerous advantages, because it introduces an additional layer to storage access, its use can potentially impact system performance. In order to evaluate the relative overhead of using VMFS versus RDM, we used a DataStage I/O-intensive workload to ensure that the storage system would be the bottleneck.

Table 4. IOPS and I/O bandwidth captured by nmon

<table>
<thead>
<tr>
<th>Test case</th>
<th>IOPS (average)</th>
<th>Disk read (average)</th>
<th>Disk write (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native—1 LUN</td>
<td>8,000</td>
<td>107 MB/s</td>
<td>121 MB/s</td>
</tr>
<tr>
<td>Native—2 LUN</td>
<td>6,900</td>
<td>126 MB/s</td>
<td>158 MB/s</td>
</tr>
<tr>
<td>Native—4 LUN</td>
<td>8,000</td>
<td>151 MB/s</td>
<td>210 MB/s</td>
</tr>
<tr>
<td>Virtual—1 LUN</td>
<td>14,000</td>
<td>89 MB/s</td>
<td>105 MB/s</td>
</tr>
<tr>
<td>Virtual—2 LUN</td>
<td>12,000</td>
<td>109 MB/s</td>
<td>145 MB/s</td>
</tr>
<tr>
<td>Virtual—4 LUN</td>
<td>11,000</td>
<td>149 MB/s</td>
<td>177 MB/s</td>
</tr>
</tbody>
</table>

Specifically, we configured the System Storage DS5300 system so the DataStage scratch space area used the ext3 file system in the guest operating system for the LUNs configured both with RDM and VMFS; when configured for VMFS, the LUNs were eager-zeroed thick disks provisioned to the maximum size. The average IOPS measured 23,000 for the I/O-intensive workload with LUNs configured both with VMFS and RDM; the throughput overhead of VMFS was measured to be less than 2 percent (as shown in Figure 3). Given that the amount of overhead was so small, our testing indicates that virtualized InfoSphere DataStage workloads can take advantage of the benefits and features that come with using VMFS for storage provisioning.

Single Large Virtual Machine Versus a Cluster of Smaller Virtual Machines

Virtualization confronts users with myriad configuration choices. Particularly when setting up the DataStage engine, users must decide whether to use a single large virtual machine (scaling up) or distribute the workload among a cluster of smaller virtual machines (scaling out).

To assess this question of configuration, we evaluated DataStage running on a different number of virtual machines while holding the total number of vCPUs constant. The first test case compared DataStage configured with 1 virtual machine with 8 vCPUs against a configuration with 2 virtual machines with 4 vCPUs each. The second test case compared a configuration with 1 virtual machine with 16 vCPUs against a configuration with 2 virtual machines with 8 vCPUs each and against a configuration with 4 virtual machines with 4 vCPUs each. The results are shown in Figure 4 and Figure 5. In both test cases, the configuration with DataStage engine running in a single virtual machine had higher throughput.
However, the root cause of this is currently unknown and we need to conduct further research. For this particular configuration, the better choice for performance was to run DataStage on a single virtual machine instead of on a cluster.

Impact of Over-committing CPU Resources

Lastly we assessed the potential performance impact to DataStage runtime while overcommitting the host server. A powerful driving force for virtualization is server consolidation to more efficiently use computational resources. Taken too far in terms of host server memory or CPU capacity, however, this process can impact the performance of virtualized workloads. We documented ways of minimizing these performance hits, specifically by reserving needed host-server resources.

To assess how DataStage performs as the virtual system becomes over-committed, we created ten virtual machines for the IBM System x3850 X5 server with 8 vCPUs each (on a total of 40 physical cores). We ran one InfoSphere DataStage workload on one of the virtual machines and different workloads on the other nine virtual machines. Each of the virtual machines was run close to 100 percent CPU utilization. We began the test with only the DataStage workload running and measured the DataStage throughputs as we subsequently started additional virtual machines. At five virtual machines running, the host server processors were 100 percent committed (that is, 40 CPUs serving 40 vCPUs). The results of the test can be seen in Figure 6. The y-axis shows the throughput and the x-axis shows host-processor commitment. Each virtual
machine committed roughly 20 percent of the CPU available to the host (the first data point at 20 percent processor commitment has only the DataStage virtual machine running). DataStage throughput remained roughly constant until the CPU was 80 percent committed, after which there was a 34 percent drop in throughput when the system was 100 percent committed. The throughput continued to drop as the host processors were further committed.

Based on discussions with VMware engineers, we determined that the large drop in performance was due to the 8 vCPUs not mapping neatly onto the 10-core, Non-Uniform Memory Access (NUMA)-node design of the Intel Xeon processor E7 family’s microarchitecture. At five virtual machines running, one of the server’s sockets has to run two virtual machines while the other three sockets run only one virtual machine each. This results in idle processor cores, which impacts performance. One workaround would be to use a 5 vCPUs configuration instead of an 8 vCPUs configuration (as 5 vCPUs virtual machines would map well to the 10-core NUMA nodes). Another workaround would be to turn off NUMA scheduling via either BIOS or VMware vSphere, which would allow all of the CPU cores to be used, though at the expense of memory performance. The test results (and the workarounds required to mitigate them) underscore the importance of understanding the NUMA architecture of the host system and configuring the virtual machines with the number of vCPUs accordingly.

If DataStage runtime is not performing as expected because the system is over-committed, the best option would be to increase the host system capacity or to reallocate the virtual machine to another host that has the needed capacity. If these options are not feasible, one can consider using the "reserve CPU" feature in VMware vSphere (Figure 7) for the performance-critical virtual machine. However, this can potentially impact the performance of the other non-reserved virtual machines running on the same host. By using this setting, a particular virtual machine will be guaranteed processing resources thus allocated.
When that virtual machine is shutdown, the CPU resources will be released back to the pool and will be available for other virtual machines to consume. The same over-commitment test was repeated with the "reserve CPU" option set to maximum for the DataStage guest. The result, Figure 8, shows that the performance impact to DataStage was minimal as the system became over-committed.

Performance Diagnostic

Overcommitted Host System

In situations where DataStage runtime is not performing as expected in a virtualized environment, performance data from the hypervisor level is an additional venue for investigation. It is important to understand if the system at the host level was over-committed. Traditionally, a DataStage administrator or user would use tools such as systools or nmon to capture the performance counters. However, when tools are used for monitoring within the guest operating system of a virtual machine, there is no easy way to tell if the host system is over-committed. Neither can monitoring CPU utilization at guest-level tell if the host system is over-committed. Figure 9 and Figure 10 show that the CPU utilizations were similar whether the host system was over-committed or not. In both cases, the same DataStage workload was executed. A user would only be able to tell if the host server is over-committed if the workload became much slower. Nor would the "steal time" performance counter from sar or vmstat for Linux guest operating systems be able to tell if the host server were over-committed.

In order to understand if the system is over-committed, performance counters at the hypervisor level should be monitored. Typical DataStage users or administrators might not have access to hypervisor-level counters; hypervisor access should be provided by hypervisor administrators for this purpose.
VMware provides some tools for gathering hypervisor-level performance stats. The first one is esxtop, a console tool available from the VMware ESXi shell. Figure 11 shows an example of esxtop output from an over-committed server. The other tool is VMware vCenter, which has a graphical interface. If the system is overcommitted, physical CPU utilization should be very high. If the %RDY and %CSTP values in the above tools are high, it is likely that the system is overcommitted.

Conclusions and Recommendations

InfoSphere Information Server runtime performed very well in a virtualized environment hosted on a platform powered by the Intel Xeon processor E7 family and using VMware vSphere 5.0. Running virtualized, CPU-intensive workloads, we observed only 2 percent to 10 percent throughput overhead compared to running the same workloads in a native environment. With I/O-intensive workloads, we consistently observed 10 percent throughput overhead. With regards to the two methods of provisioning data stores that are supported by VMware vSphere, we observed only a minimal performance difference between VMFS and RDM, meaning that organizations can reap the benefits and features that come with VMFS for provisioning storage for these virtualized workloads. In the particular configuration we tested, we observed performance overhead when the InfoSphere DataStage engine was set up on a cluster of smaller virtual machines rather than a single large virtual machine with the same capacity. However, we have not yet discovered the root cause of this overhead; until we follow up this investigation, we can only report that for the tested configuration using a single large virtual machine yielded better performance. Finally, we unsurprisingly confirmed that overcommitting host server resources can have a negative impact on DataStage runtime performance. If necessary, using CPU reservation for the virtual machine running DataStage can be a solution if moving other virtual machines to other host servers is not an option. However, these recommendations should not detract from the findings of this research that enterprises can successfully run processor- and I/O-intensive workloads like InfoSphere Information Server and still take advantage of the benefits of server virtualization.

References and End Notes

Performance Best Practices for VMware vSphere 5.0 (www.vmware.com/pdf/Perf_Best_Practices_vSphere5.0.pdf)


Performance Characteristics of IBM® InfoSphere® Information Server 8.7
Running on VMware® vSphere® 5.0 on Intel® Xeon® Processors

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other resources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, go to: http://www.intel.com/technology/performance/resources/benchmark_limitations.htm.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALES AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Inter compiler options, including the instruction sets and specific microprocessors they implicate, please refer to the "Intel® Compiler User and Reference Guides" under "Compiler Options." Many library routines that are part of Intel® compiler products are more highly optimized for Intel microprocessors than for other microprocessors. While the compilers and libraries in Intel® compiler products offer optimizations for both Intel and Intel-compatible microprocessors, depending on the options you select, your code and other factors, you likely will get extra performance on Intel microprocessors. Intel® compilers associated libraries and associated development tools may or may not include options that optimize for instruction sets that are available in both Intel® and non-Intel microprocessors (for example SIMD instruction sets), but do not optimize equally for non-Intel microprocessors. In addition, certain compiler options for Intel compilers, including some that are not specific to Intel micro-architecture, are reserved for Intel microprocessors. For a detailed description of Intel compiler options, including the instruction sets and specific microprocessors they implicate, please refer to the "Intel® Compiler User and Reference Guides" under "Compiler Options." Many library routines that are part of Intel® compiler products are more highly optimized for Intel microprocessors than for other microprocessors. While the compilers and libraries in Intel® compiler products offer optimizations for both Intel and Intel-compatible microprocessors, depending on the options you select, your code and other factors, you likely will get extra performance on Intel microprocessors. Intel® compilers associated libraries and associated development tools may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include Intel® Streaming SIMD Extensions 2 (Intel® SSE2), Intel® Streaming SIMD Extensions 3 (Intel® SSE3), and Supplemental Streaming SIMD Extensions 3 (Intel® SSSE3) instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors.

While Intel believes our compilers and libraries are excellent choices to assist in obtaining the best performance on Intel® and non-Intel microprocessors, Intel recommends that you evaluate other compilers and libraries to determine which best meet your requirements. We hope to win your business by striving to offer the best performance of any compiler or library; please let us know if you find we do not.

Copyright © 2013, Intel Corporation. All rights reserved. Intel, the Intel logo, Xeon, and the Xeon badge are trademarks of Intel Corporation in the U.S. and other countries.

* Other names and brands may be claimed as the property of others.

Printed in USA 0513/TA/FR/W/PDF Please Recycle 328952-001US