A POWER6 Virtualization Performance Study Using Integrated Virtual Ethernet

October 24, 2007

Andrew Theurer
Elisabeth Stahl

IBM Systems and Technology Group

Executive Overview

Data latency has long been a top concern **for corporations where** less than a millisecond can **determine** whether revenue is lost or **made.** The introduction of the IBM POWER6™ processor-based server brought to market several advances in server design, performance, and function. One of these key enhancements is the capability to virtualize Ethernet within a server. This new technology offers significant low-latency advantages in virtualized solutions which contribute to a business's bottom line.

This paper will highlight data latency's application to business results, discuss POWER6 virtualization and the Integrated Virtual Ethernet [1], present a virtualization performance study, and conclude with further recommendations.

Virtualization Performance – A Competitive Differentiator

Network latency is the time delay experienced when data is sent from one point to another. Many industries depend on split-second response times and high throughput between many different agents. This need for speed is growing in industry sectors such as the financial sector; media companies involved in video postproduction, digital animation, broadcasting, and the Web; oil and gas producers; big retail chains; and research institutions. These industries are finding that input and output bottlenecks are proliferating and that rapid access to data is increasingly a competitive differentiator.

Studies have shown that a 1-millisecond advantage in financial trading applications can be worth $100 million a year to a major brokerage firm [2]. The fastest systems, running from traders' desks to exchange data centers, can execute transactions in a few milliseconds. The physical distance between two servers processing a transaction can slow down how fast these transactions happen. To overcome this data latency, many high-frequency algorithmic traders are moving their systems as close to the Wall Street exchanges as possible [3].

Network latency, the lag experienced when data is moved from one place to another, robs firms of their ability to leverage investments and react quickly to market changes. Some organizations currently use physical proximity to overcome the technical barriers of data latency. As seen in the next section of the paper, with IBM’s POWER6 virtualization, virtual technologies can be used to mitigate the need for physical dependencies and increase performance.

The POWER6 Host Ethernet Adapter

The recent introduction of the POWER6 processor-based servers brought to market several advances in server design, performance, and function. One of these enhancements was the capability to virtualize Ethernet within a server using the Integrated Virtual Ethernet (IVE) adapter.

IVE is the integration of several technologies, including the Host Ethernet Adapter (HEA), advanced software, and updates to the POWER™ Hypervisor that provide integrated high-speed Ethernet adapter ports with hardware assisted virtualization capabilities. IVE is a standard set of features offered on every IBM System p™ POWER6 processor-based server.

IVE functionality provides:

- Either two 10 Gbps Ethernet ports or four 1 Gbps ports or two 1 Gbps ports
- External network connectivity for LPARs using dedicated ports without the need of a Virtual I/O Server
- Industry standard hardware acceleration, loaded with flexible configuration possibilities
- The speed and performance of the GX+ bus, faster than PCI Express x16

IVE consists of a physical Ethernet adapter that is connected directly to the GX+ bus of a POWER6 processor-based server, providing IVE with the high throughput and low latency of a bus imbedded in the I/O controller. IVE also includes special hardware features that provide logical Ethernet adapters. These adapters can communicate directly to logical partitions (LPARs), reducing the interaction with the POWER Hypervisor (PHYP).

Prior to IVE, virtual Ethernet provided a connection between LPARs. The use of a shared Ethernet adapter (SEA) and the Virtual I/O Server allowed connection to an external network. The IVE replaces the need for both the virtual Ethernet and the SEA. The IVE functionality eliminates the need to move packets (using virtual Ethernet) between partitions and then through an SEA to an Ethernet port; LPARs can share IVE ports with improved performance.

IVE extends the virtualization capabilities of a POWER6 processor-based system by providing:

- Internal network communication between logical partitions reducing hypervisor intervention using virtualized ports.
- Multiple Core Scaling that benefits highly parallel workloads.
- External network physical connection without the Virtual I/O server. The Hypervisor and SEA interaction is eliminated for data flow.

The IVE design provides a great improvement of latency for short packets. Messaging applications such as distributed databases require low-latency communication for synchronization and short transactions. The methods used to achieve low latency include:

- GX+ bus attachment.
- Immediate data in descriptors (reduced memory access).
- Direct user space per-connection queuing (OS bypass).
- Designed for up to three-times the throughput improvement over current 10 Gbps solutions.
- Provides additional acceleration functions in order to reduce host code path length. These include header/data split to help with zero-copy stacks.
- Provides I/O virtualization support so that all partitions of the system can natively take advantage of the features.
- Allows one 10 Gbps port to replace up to 10 dedicated PCI 1 Gbps adapters in a partitioned system.

The Host Ethernet Adapter (HEA) is the major hardware component of the IVE, and it is included in the P5IOC2 chip. HEA also includes all the logical ports and the virtual layer 2 switches and connects to the physical port.

IBM's HEA provides direct, protected access to ethernet adapters to many partitions at the same time. Direct access eliminates the need to use software based ethernet bridging in either the hypervisor or service partition This enhancement greatly reduces the latency to send and receive network packets. HEA is also engineered to be "close" to the system processors, providing very low hardware latency by integrating the adapter on IBM's high-performance GX bus. IBM's HEA can provide very low TCP/IP latency while providing secured, shared access to many partitions.

.

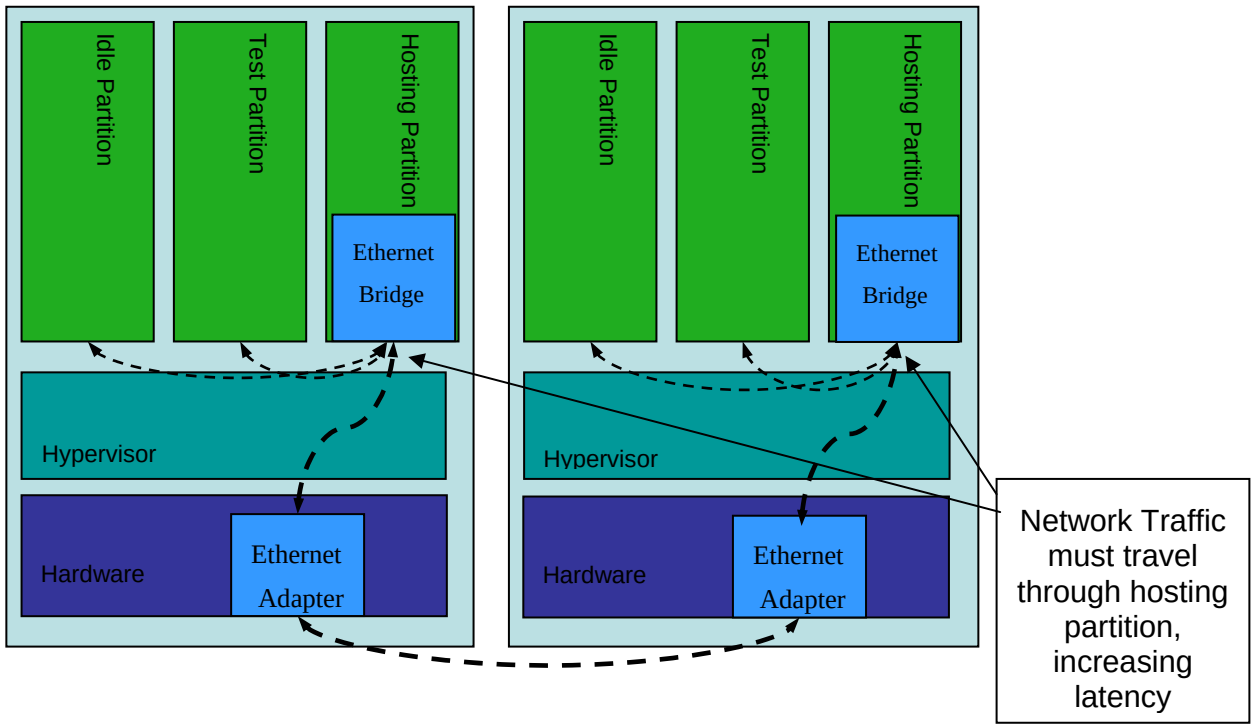
The POWER6 Virtualization Study

The POWER6 virtualization study described in this paper compared two virtualization solutions: (1) a pair of IBM System p™ 570 servers, using four POWER6 processor cores, Integrated Virtual Ethernet, and Advanced POWER Virtualization (2) A pair of 4-core x86 servers, using Intel® 5160 processors, Broadcom Gigabit Ethernet, and Xen hypervisor provided by RedHat Enterprise Linux® version 5 update 1. Each pair of systems was connected directly via Ethernet adapters in order to conduct network latency tests.

All systems had three partitions configured. The first partition was used for hosted I/O and management. On the p570 servers, SUSE Linux SLES10 SP1 was used, and on the x86 systems, Red Hat Enterprise Linux 5 update 1 was used. The second partitions were used to run the Netperf test, while the third partitions were active, but idle. On all systems, the second and third partitions used Red Hat Enterprise Linux 5 update 1.

On the x86 systems, the network adapter was hosted in the first partition and shared via software Ethernet bridge and virtual Ethernet drivers (see Figure 1). On the POWER6 processor-based systems, no Ethernet hosting is necessary, and all partitions had direct access to the same Ethernet device (see Figure 2). On all systems, the first partition was used to host the disk device and shared to their respective partitions with virtual disk drivers.

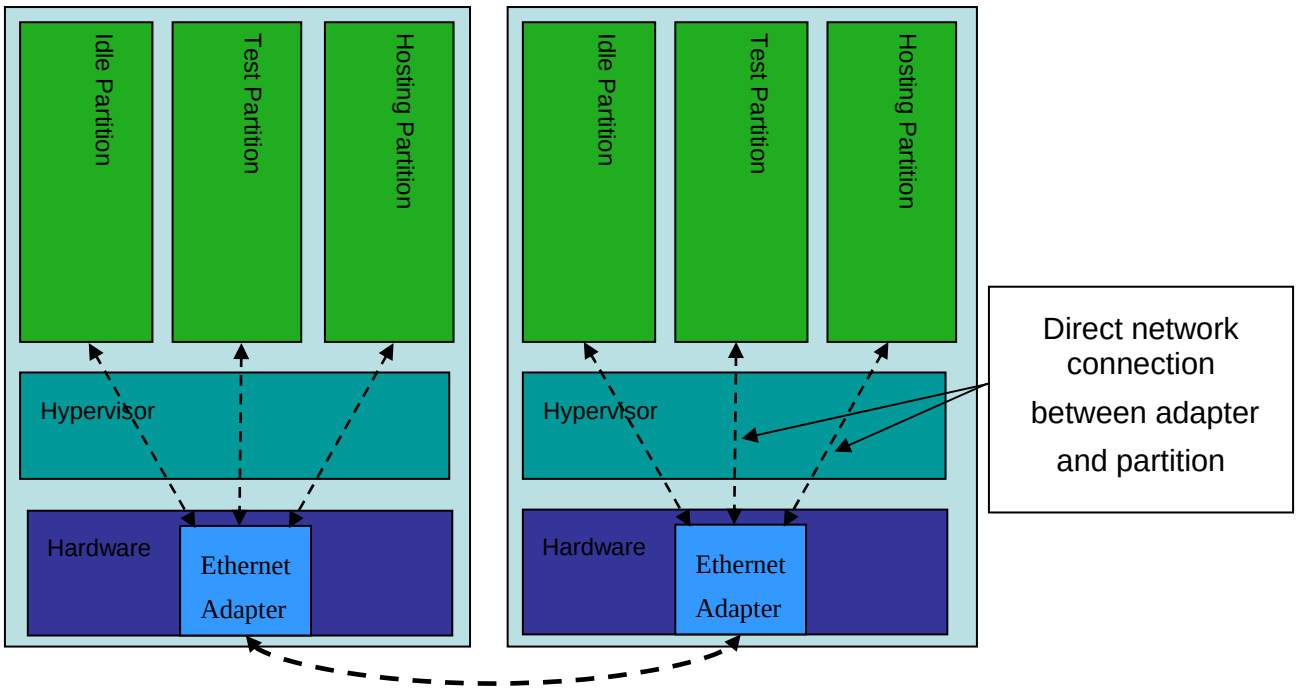
E



X86 Systems with four 4 cores and Shared Ethernet

Figure 1

In order to measure the impact of network latency, the request-response benchmark was used over TCP/IP protocol from the Netperf benchmark suite [4]. The benchmark consisted of a client and server program, and the benchmark was configured such that the client program was installed on Server A and the server program on Server B. For the request-response test, the client sends a request to the server, and the server responds with a message of N bytes. After the client receives the response, it will repeat with another request for a fixed number of transactions. The less time it takes to send a request and receive a response, the more transactions the system can perform per second. Therefore, the system with more transactions per second will have a better (lower) latency.



p570 Systems with four POWER6 cores and Integrated Virtual Ethernet

Figure 2

In all test cases, the POWER6 virtualization solution with Integrated Virtual Ethernet showed a significant advantage. The test was conducted with several response message sizes and, in all cases, the POWER6 processor-based solution outperformed the x86 Xen solution (see Figure 3) with as much as 6.8 times the number of transactions per second.

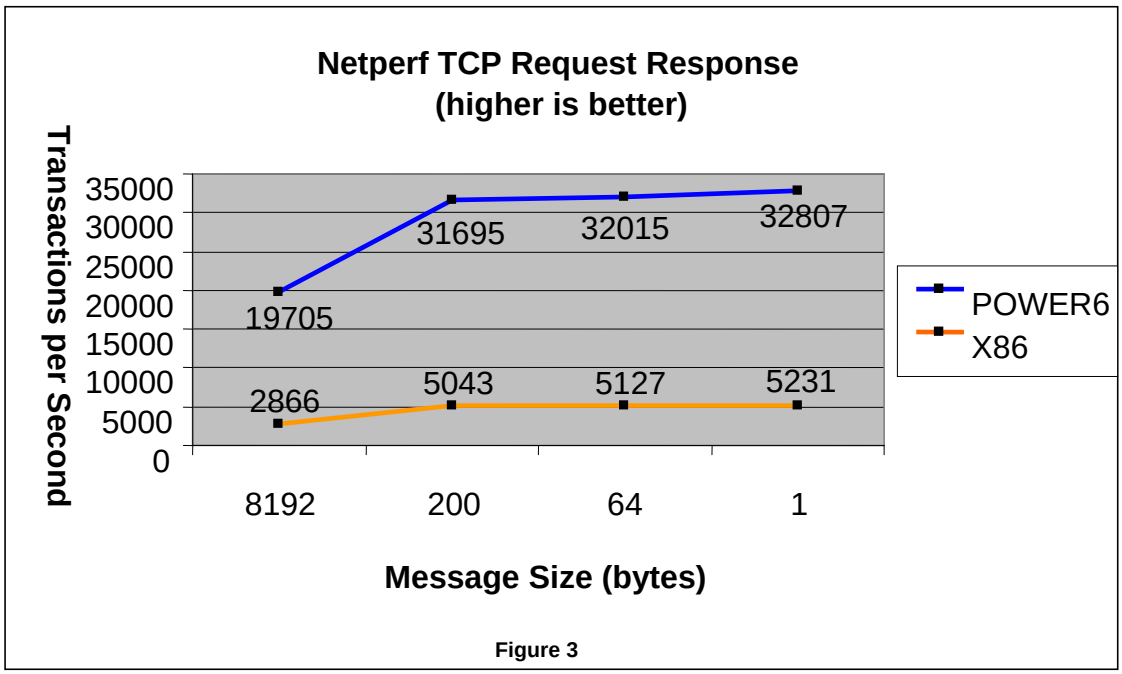

Netperf TCP Request Response
(higher is better)
Transactions per Second
35000
30000
25000
20000
15000
10000
5000
0
19705
31695
32015
32807
2866
5043
5127
5231
8192
200
64
1
Message Size (bytes)
POWER6
X86

Figure 3

Conclusion

Corporations in varied industries rely on quick access to data as a competitive differentiator. IBM's POWER6 virtual technologies reduce data latency while increasing performance.

The virtualization study outlined in this paper has shown that the POWER6 processor-based server's Host Ethernet Adapter produced a significant advantage in performance over other vendor technologies. IBM's HEA can provide very low latency while providing secured, shared access to many partitions.

References

[1] Integrated Virtual Ethernet Adapter Technical Overview and Introduction
http://www.redbooks.ibm.com/redpapers/pdfs/redp4340.pdf

[2] Wall Street's Quest to Process Data At The Speed Of Light
http://www.informationweek.com/story/showArticle.jhtml?articleID=199200297

[3] Five Ways to Avoid Data Latency
http://www.wallstreetandtech.com/showArticle.jhtml?articleID=60404306

[4] Netperf Benchmark
http://www.netperf.org

IBM Corporation
Systems and Technology Group
Route 100
Somers, New York 10589

Produced in the United States of America
October 2007

PSW03034-USEN-00