Virtualization Performance Analysis
November 2010

Effect of SR-IOV Support in Red Hat KVM on Network Performance in Virtualized Environments

Steve Worley
System x Performance Analysis and Benchmarking
IBM Systems and Technology Group

Abstract

Single Root I/O Virtualization (SR-IOV) is a technology that has the potential to benefit virtualization network performance. SR-IOV defines extensions to the PCI Express (PCIe) specification suite to enable multiple guest operating systems to directly access subset portions of physical I/O resources. The analysis presented in this paper is based on the use of the Standard Performance Evaluation Corporation's new virtualization benchmark, which was used to analyze the benefit of SR-IOV. SPECvirt_sc2010 (SPECvirt) is SPEC's first benchmark that addresses the need to compare the performance of datacenter servers used in virtualized server consolidation. Red Hat Enterprise Linux® 5.5 Kernel-based Virtual Machine (KVM) was used for this analysis because it supports SR-IOV.

Introduction

As servers are virtualized and consolidated, the network has the potential to become a performance bottleneck. One of the solutions for this bottleneck is to offload the hypervisor emulation of I/O devices to specialized hardware and device drivers. One of these solutions is Single Root I/O Virtualization.

Single Root I/O Virtualization (SR-IOV) is a Peripheral Component Interconnect Special Interest Group (PCI-SIG) virtualization I/O specification. There are three PCI-SIG groupings of specifications for virtualization I/O:

- Address Translation Services (ATS): This specification provides a grouping of transactions for PCI Express components to exchange and use translated addresses in support of native I/O Virtualization.
- Single Root IOV: This specification provides native I/O Virtualization in existing PCI Express topologies where there is a single root complex (i.e., a single computer that supports virtualization technology).
- Multi-Root IOV: This specification builds on the Single Root IOV Specification to provide I/O Virtualization in topologies where multiple root complexes share a PCI Express hierarchy (i.e., a server blade enclosure that uses a PCI Express switch-based topology to connect server blades to PCI Express Devices).

Recently the Red Hat/KVM kernel began support of SR-IOV. SR-IOV support for the VMware® ESX hypervisor is not currently available.

It is useful to quantify the advantage of using SR-IOV on Red Hat/KVM. One way to do this is to use the SPECvirt_sc2010 benchmark released in 2010 by Standard Performance Evaluation Corporation (SPEC) as the workload.

The SPECvirt_sc2010 Workload

The workload used for this analysis was the SPECvirt_sc2010 virtualization benchmark. See www.spec.org for more information about the SPECvirt_sc2010 benchmark. The SPECvirt benchmark is the first benchmark produced by SPEC to address performance evaluation of datacenter servers used in virtualized server consolidation. SPECvirt measures the performance of all system components, including the hardware, virtualization platform, virtualized guest operating systems, and application software.

The benchmark uses several workloads representing applications that are commonly consolidated into a virtualized environment. Scaling is achieved by running additional sets of virtual machines, called “tiles,” until overall throughput reaches a peak. Each tile consists of six different virtual machines:

- Infraserver – Serves file downloads directory for web workload, runs web back end simulator
- Webserver – Runs Web workload
- Mailserver – Runs Internet Message Access Protocol (IMAP) workload
- Appserver – Runs application server workload
- Dbserver – Runs database as backend to application server workload
- Idleserver – Runs a poll workload

Because each tile represents six VMs, a system running 10 tiles would consist of 60 VMs.

Test Bed Configuration

The test bed used for this analysis consisted of an IBM System x®3850 X5 server with MAX5 memory expansion as the system under test (SUT). Only two sockets on the x3850 X5 were used. The system was configured with a total of 768GB of memory; 512GB was installed in the x3850 X5 and 256GB was installed in the MAX5 memory expansion unit. Less memory could have been used for this analysis and the results would have been the same. The SUT was configured with two 8-core Intel® Xeon® X7560 (2.26GHz) processors. The SUT was configured with a 4Gb QLogic QL2462 Fibre Channel host bus adapter that was connected to the HDD arrays through an IBM System Storage™ DS4800 dual-disk controller.

Red Hat Enterprise Linux 5.5 with the KVM was used as the hypervisor.

A detailed list of the configuration can be found as part of the IBM System x3690 X5 submission at http://www.spec.org/virt_sc2010/results/res2010q3/virt_sc2010-20100727-00014-perf.html. This submission used the same client configuration and the same software stack on the system under test as was used for this analysis.

Results

Results were obtained for the SPECvirt_sc2010 metric (score) and Quality of Service (QoS) as part of the SPECvirt harness reporting. The SPECvirt_sc2010 metric is a “supermetric” that is the arithmetic mean of the normalized submetrics for each workload (Mail server, Web server, and Application server). The Quality of Service requirements are also an aggregate of the three workloads for the mail server, Web server, and application server. For the Web server, for example, the QoS requirements are defined in terms of Time_Good and Time_Tolerable. Time_Good equals 3 seconds and Time_Tolerable equals 5 seconds. For each page, 95% of the page requests are expected to be returned within Time_Good, and 99% of the requests within Time_Toerable. The CPU utilization was obtained using the mpstat command while SPECvirt was running and is an average utilization of all cores.

Table 1 shows the results for 13 tiles with SR-IOV enabled, 10 tiles with and without SR-IOV enabled, and 8 tiles with and without SR-IOV enabled.

Configuration	Score	QoS	CPU Utilization All
13 Tile SR-IOV	1275	99.98	86.56
10 Tile SR-IOV	982.1	100	68.24
10 Tile no SR-IOV	892.6	85.23	82.48
8 Tile SR-IOV	785.7	100	55.04
8 Tile no SR-IOV	778	100	66.23

Table 1 – SR-IOV Data Results

From Table 1 the 13-tile run with SR-IOV enabled shows an average CPU utilization of all cores as 86.56% and a QoS of 99.98. When attempting to run with 13 tiles with SR-IOV disabled, the run could not be completed due to a CPU utilization bottleneck. In order to get a valid comparison with SR-IOV disabled, the number of tiles was dropped to 10. At 10 tiles QoS without SR-IOV was 85.23, which is still not valid for a compliant SPECvirt run. Reducing the number of tiles to 8 tiles allowed for good QoS when running with SR-IOV. The CPU utilization at 8 tiles jumped from 55.04% to 66.23% when SR-IOV was disabled.

Conclusion

The analysis presented in this paper shows that SR-IOV can be of great benefit when using the Red Hat Enterprise Linux 5.5 KVM. The results show a minimum of 30% increase in number of tiles supported when enabling SR-IOV. This increase in the number of tiles is due to decreased CPU utilization. Because SR-IOV offloads the network I/O processing from the hypervisor to specialized hardware and device drivers, the CPU cycles that would be used for processing of network I/O can be used to increase application performance. The effect on CPU utilization can be seen in the 8-tile runs. CPU utilization increases 20% normalized (11% real) at 8 tiles with SR-IOV disabled. The impact on QoS when running without SR-IOV appears to be magnified as the system approaches higher average CPU utilization. The result is a lower number of tiles that can be supported on the system under test while still passing QoS requirements.

IBM Systems and Technology Group
3039 Cornwallis Road
Research Triangle Park, NC 27709

Produced in the USA
November 2010