



PCI-Express 3.0 technology and optimizations on HP Z820 Workstation

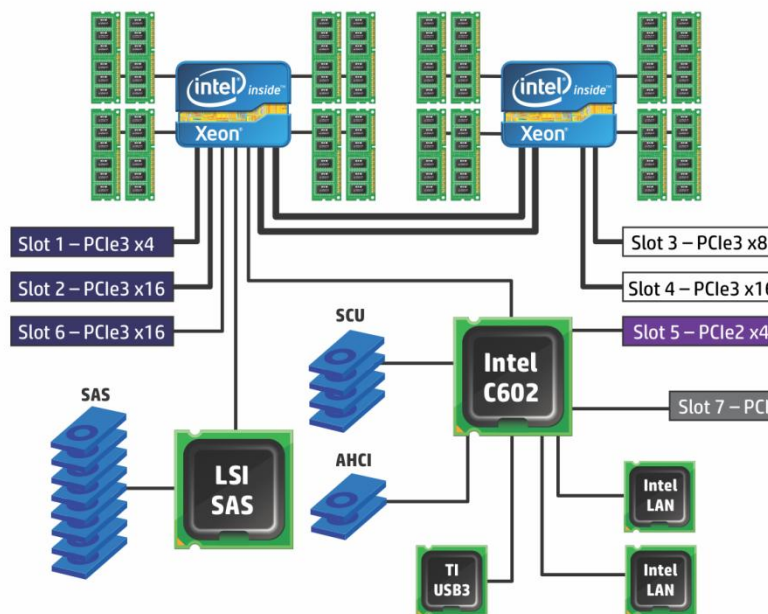
Introduction

The HP Z820 Workstation introduces the new PCI-Express 3.0 technology, as well as I/O slot improvements, and higher performance over a variety of conditions. This paper covers some of the key technological aspects of PCI-Express 3.0 and provides guidance on optimizing the performance of your system when using PCI-Express cards.

Integrated PCI-Express 3.0

The HP Z820 uses the Intel® Xeon® processor E5-2600 series, with integrated PCI-Express 3.0 controllers delivering a peak bandwidth of 16GB/s per direction, i.e. a doubling over the HP Z800. PCI-Express 3.0 is backward compatible with 1.0 and 2.0, and slots will train to the highest common speed. PCI Express 3.0 slots will initialize at 1.0 and then transition to 3.0 through a training sequence that involves four adaptive training phases. When HP Z820 was introduced, the industry had validated PCI Express 3.0 with a limited number of 3.0 devices available in the market. Therefore, it is recommended to carefully evaluate and validate PCI-Express 3.0 devices that are not available or supported from HP.

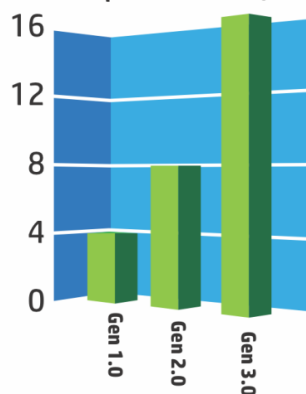
Figure 1: HP Z820 Workstation Block Diagram



PCI-Express I/O Slot Improvements

The HP Z820 provides a total of seven high-performance Graphics and I/O slots, including support for up to three PCIe 3.0 graphics cards in PCIe 3.0 x16 slots. With the standard 850W power supply, certain system configurations can support up to three cards totaling 160W. With the optional 1125W supply, certain configurations can support up to two 300W, or three 225W, cards. An additional bulkhead allows for an eighth mechanical-only IO card/cable (e.g. SDI, eSATA, mini SAS 4x). In single processor configurations, slot 3 and 4 are not available.

Figure 2: x16 Peak Bandwidth per Direction—GB/s



Higher PCI-Express Performance

The integration of PCIe 3.0 controllers within the processor combined with DMA caching in the CPU, an integrated 4-channel memory controller, PCIe 3.0 speeds, and a dual QPI processor interconnect at up to 8GT/s, resulted in dramatic improvements in I/O bandwidth, remote bandwidth, and latency. Unlike previous workstations based on Intel NUMA architectures, bandwidth performance on the HP Z820 is mostly equivalent for local versus remote memory accesses. Furthermore, improvements in

PCI Express credit allocation improve latency and bandwidth across various traffic conditions, including small TLP payload sizes (i.e. 16 or 32B).

Recipe for Optimizing PCI-Express I/O Performance

For high I/O bandwidth applications, the choice of slot loading, CPU, and memory configuration can be optimized to ensure maximum bandwidth available. Applications and cards sensitive to I/O latency may benefit as well from some of the tips below.

Recommended Configuration Steps

1. Place the graphics cards first, following the slot order listed in the Recommended column of Table 1.
2. Place I/O cards next, from highest bandwidth to lowest, following the slot order listed in the Recommended column of Table 1. This is the optimal load order for most applications.
3. If the onboard SAS controller is not used and there is an I/O card in slot 1, then disable the SAS controller (BIOS setup menu -> Security -> Device Security -> SAS Controller = Device hidden).
4. If PCIe slot 1 is not used and the onboard SAS controller is used, then disable PCIe slot 1 (BIOS setup menu -> Security -> Device Security -> Slot 1 = Device hidden). *This option is only available for BIOS revision J2.01 and later.*
5. If PCIe 2.0 I/O cards fail to train at full Gen 2 speeds (5 Gbps) in Gen3 slots, then try slot 5 which only trains up to PCIe Gen2.
6. Some applications may perform better using one of the optional load orders. For dual CPU systems try load Option A. For single CPU systems try load Option B.
7. Additional I/O bandwidth refinements may be possible. If necessary, refer to the tips below.

Table 1. Z820 I/O Slot Recommended Load Order

Slot #	PCI-Express Lane Width	I/O Card Load Order			Peak Bandwidth (GB/s)	CPU/ C602 chipset
		Recommended	Option A	Option B		
Slot 1	x4 (Gen3)	5 th	5 th	3 rd	4.0	1 st
Slot 2	x16 (Gen3)	1 st	1 st	1 st	16.0	1 st
Slot 3 (dual CPU only)	x8 (Gen3)	4 th	4 th	-	8.0	2 nd
Slot 4 (dual CPU only)	x16 (Gen3)	3 rd	2 nd	-	16.0	2 nd
Slot 5	x4 (Gen2)	6 th	6 th	4 th	2.0	C602
Slot 6	x16 (Gen3)	2 nd	3 rd	2 nd	16.0	1 st
Slot 7	PCI	-	-	-	0.125	C602

Additional Tips

- For applications doing direct bus Peer-to-Peer transfers between cards, load the corresponding cards in slots located behind the same CPU. For instance, load cards in slots 1, 2 and 6, or in slots 3 and 4.
- For very high bandwidth applications in dual CPU systems, select CPU models with the highest QPI frequency (8GT/s).
- Make sure all I/O cards are loaded in slots that have a PCI-Express Lane Width at least as wide as the card (see Table 1).
- For predictable latencies, try disabling NUMA (Non-Uniform Memory Access) mode (BIOS setup menu -> Advanced -> Bus Options -> NUMA = Disabled).
- For cards that are latency sensitive, load these cards in CPU slots.
- Ensure Idle Power Savings BIOS setting is set to Normal (BIOS setup menu -> Power -> OS Power Management -> Idle Power Savings = Normal).
- Use the latest BIOS version available on hp.com.
- Check for updates in the latest performance optimization white papers (link below).

Additional Resources

hp.com/go/whitepapers

hp.com/support/Z820_manuals

© Copyright 2012 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Xeon are trademarks of Intel Corporation in the U.S. and other countries. All other trademarks are the property of their respective owner.

