A Concurrent Real-Time White Paper

2881 Gateway Drive
Pompano Beach, FL 33069
(954) 974-1700
www. real-time.ccur.com

PCI Express Expansion Limitations

By: Bob Hartley
Consulting Engineer

Phone: 954-973-5460
bob.hartley@ccur.com

February 2013

PCI Express has a great increase in bandwidth over legacy PCI and the specifications imply that it is easy to expand into large I/O configurations. This capability is very appealing to our HIL simulation customers. However In our experience, we have found that the I/O board and motherboard vendors are patching legacy product firmware and BIOS to market PCI-E equipment. They are not designing hardware that takes into account the resources that PCI-E requires and unfortunately customers wind up with non-functional systems when attempting to build anything but the smallest PCI-E configuration.

Several of our customers have purchased iHawk systems with the intent of integrating their own I/O. Some of them find that they cannot get the desired configuration working and the system will either hang in the BIOS after power-on or if successfully booted, will not provide kernel resources for all of the I/O. These problems can generally not be solved by Concurrent and require the customer to re-evaluate their I/O configuration.

As an OEM and integrator, we have had extensive experience with PCI-Express expansion I/O systems. We have run into situations where we could not get the desired configuration to function and the board vendors are of little or no help. As of now, any request for a system with PCI-Express expansion requires a configuration review by Concurrent Special Systems to evaluate our risk in accepting the order. In most cases we have to steer the customer towards a different configuration than originally desired. Let me explain what we (and any other OEM/integrator) are up against:

PCI-E to PCI expansion works well, PCI-E to PCI-E is very problematic for several reasons.

A) The chipsets used on all current motherboards, restrict the BIOS from accessing beyond to 64K of *IOSPACE*. The motherboards also reserve a portion of this 64K to on-board devices. The more on-board devices, the less *IOSPACE* available for external devices. A four socket system reserves more on-board I/O resources than a two socket.

B) Each PCI-E bridge with *IOSPACE* requirements downstream will reserve a minimum of 4K of I/O space. A motherboard will theoretically not support more than 16 bridges.

C) Most PCI-E cards available today are not designed as native PCI-E cards and use PCI-E to PCI bridges to access the PCI-E interface. Most of these boards also unnecessarily use the legacy *IOSPACE* for register access where *MEMSPACE* is the preferred method. This normally is not a problem in PCI because the cards reserve as little as 256 bytes of I/O space. However, when used behind a PCI-E to PCI bridge, the 256 bytes translates into reserving a minimum of 4K bytes of I/O space per card.

D) In addition to the 64K *IOSPACE* limit, there appears to be bus enumeration limits in the motherboard BIOS as well. The specification is for 255 busses but it appears that the BIOS engineers limit the number of bus probes to improve the time to boot. The vendors have not anticipated the way that PCI-E uses bus resources, where each PCI-Slot and bridge appears as a bus. The number of busses gets to be very high in an expansion system, especially with non-native PCI-E cards included.

E) The motherboard vendors do not attempt to meet the PCI-E specifications and test their boards using the on-board slots only. They do not know what the I/O limits are for a specific motherboard. Our experience so far is that vendors like Tyan and Supermicro purchase AMI or Phoenix BIOS and a have limited knowledge about the internals. They perform some relatively minor changes to support the basic functions on a specific motherboard chipset, which is good enough for 99.9% of their customers. Other vendors such as Dell, IBM, and Intel produce their own BIOS. We have not been successful in engaging these vendors to support large I/O systems. Some vendors simply do not support any expansion beyond the on-board slots.

F) There have been discussions within the industry that these BIOS problems will not exist in the next generation base-board management firmware known as “UEFI”. To this end, we have experimented with an Intel UEFI motherboard and have found the 64K *IOSPACE* limit still exists.

We have built a PCI-Express I/O evaluation test bed that we are using to qualify motherboards for use with PCI-E expansion. Currently we have qualified only a few motherboards that are feasible for use with large expansion systems. The list of qualified motherboards is not published because there are several factors to consider that involve the type of expansion chassis and the design of the I/O cards to be installed. For instance; the largest mixed PCI-E and PCI configuration we have been successful configuring using a dual socket motherboard contains 36 I/O cards in four PCI/PCI-X expansion chassis. This same motherboard will support only 11 bridged non-native PCI-E cards in an 18-slot PCI-E expansion.

The good news is that there is industry demand to support large GPU (CUDA) computing clusters that require expansion systems to contain a number of PCI-Express GPU modules. IBM and DELL currently have high end server systems available that support large GPU configurations. However, the GPU boards are modern designs with a native PCI-Express interface which has no *IOSPACE* requirements. The high end server systems, while targeting the GPU market by supporting a large number of native PCI-Express devices are still limited to supporting a relatively small number of non-native PCI-Express cards.

About Concurrent Real-Time

Concurrent (NASDAQ:CCUR) is a global leader in multi-screen video delivery, media data management and real-time computing solutions. Concurrent Real-Time is the industry's foremost provider of high-performance, real-time computer systems and software solutions. The company's core competencies include hardware-in-the-loop and man-in-the-loop simulation, high-speed data acquisition, visual imaging and low-latency transactions for a wide range of markets, including aerospace and defense, automotive, energy, transportation and finance. Concurrent Corporation is headquartered in Atlanta with offices in North America, Europe and Asia. For more information, please visit Concurrent Real-Time at www.real-time.ccur.com.