



PCI-SIG ENGINEERING CHANGE NOTICE

TITLE:	Posted Request Acceptance Limit
DATE:	February 21, 2004
AFFECTED DOCUMENT:	PCI Express Base Specification 1.0a
SPONSOR:	Dan Neal; QuestTech Consulting, Inc., 512-630-8440

Part I

1. Summary of the Functional Changes

To help reduce or avoid excessive bus traffic congestion, Conventional PCI and PCI-X require that devices never back-pressure (repeatedly terminate with retry) memory write requests longer than specified limits. This can be accomplished through device design, a restricted usage model implemented by drivers, or a combination. See Conventional PCI 2.3, Section 3.5.3 and PCI-X 2.0a, Section 2.13.

This ECN adds similar requirements for Posted Requests (Memory Writes and Messages) in PCI Express, based on a newly defined Posted Request Acceptance Limit. Without such a limit, a PCI Express device may back-pressure Posted Requests for excessive periods of time and cause PCI Express fabric congestion, which may degrade overall system performance, force data overruns or underruns in other devices, or even cause system crashes.

2. Benefits as a Result of the Changes

For many devices, places requirements on device drivers to avoid them issuing more requests than the device can handle concurrently, thus avoiding degradation of system performance and/or avoiding causing system crashes.

Similar requirements exist in PCI and PCI-X to prevent data underruns. For example, prior to this requirement in PCI and PCI-X, excessive posted writes to a device from its device driver through a PCI host bridge, that were being held off by the targeted device, prevented a USB device from getting sufficient data to maintain the desired continuous sound effect from a set of speakers.

3. Assessment of the Impact

Requires PCI Express device drivers to meet the requirements specified.

4. Analysis of the Hardware Implications

None.

5. Analysis of the Software Implications

Requires device drivers to issue requests no faster than devices can process them.

Part II

Detailed Description of the change

Add the following text to Section 2.3.1, page 80:

2.3.1. Request Handling Rules

...

- ❑ If the Request is a type which requires a Completion to be returned, generate a Completion according to the rules for Completion Formation (see Section 2.2.9)
 - The Completion Status is determined by the result of handling the Request
- ❑ Under normal operating conditions, PCI Express Endpoints and Legacy Endpoints must never delay the acceptance of a Posted Request for more than 10 μ s, which is called the Posted Request Acceptance Limit. The device must either (a) be designed to process received Posted Requests and return associated Flow Control credits within the necessary time limit, or (b) rely on a restricted programming model to ensure that a Posted Request is never sent to the device either by software or by other devices while the device is unable to accept a new Posted Request within the necessary time limit.
 - The following are not considered normal operating conditions under which the Posted Request Acceptance Limit applies:
 - ◆ The period immediately following a Fundamental Reset (see Section 6.6).
 - ◆ TLP retransmissions or Link retraining.
 - ◆ One or more dropped FCPs.
 - ◆ The device being in a diagnostic mode.
 - ◆ The device being in a device-specific mode that's not intended for normal use.
 - The following are considered normal operating conditions, but any delays they cause do not count against the Posted Request Acceptance Limit:
 - ◆ Upstream TLP traffic delaying upstream FCPs.
 - ◆ The link coming out of a low-power state.
 - ◆ Arbitration with traffic on other VCs.
 - Though not a requirement, it is strongly recommended that Root Complex Integrated Endpoints also honor the Posted Request Acceptance Limit.
 - If the device supports being a target for I/O writes, which are Non-Posted Requests, it is strongly recommended that each associated Completion be returned within the same time limit as for Posted Request acceptance, although this is not a requirement.



IMPLEMENTATION NOTE

Restricted Programming Model for Meeting the Posted Request Acceptance Limit

Some hardware designs may not be able to process every Posted Request within the required acceptance time limit. An example is writing to a command queue where commands can take longer than the acceptance time limit to complete. Subsequent writes to such a device when it is currently processing a previous write could experience acceptance delays that exceed the limit. Such devices may rely on a restricted programming model, where the device driver limits the rate of memory writes to the device, the driver polls the device to determine buffer availability before issuing the write transaction, or the driver implements some other software-based flow control mechanism.