PEX 8533

PCI Express Switch

Silicon Revisions and Errata List

CONFIDENTIAL PROPRIETARY INFORMATION

Version 1.7

March 2010

History:

Version 0.1, November 10, 2006

- Initial publication of the Errata list (1 item) based on Errata found during simulation and evaluation of early silicon samples

Version 0.2, December 20, 2006

- Added Erratum 2, 3, 4, & 5

Version 0.3, January 15, 2007

- Added Erratum 6

Version 0.4, February 27, 2007

- Edited Erratum 6
- Added Erratum 7

Version 0.5, March 1, 2007

- Edited Erratum 7

Version 0.5, March 8, 2007

- Edited Erratum 7

Version 0.6, June 27, 2007

- Added Erratum 8

Version 0.7, July 30, 2007

- Added Erratum 9, 10, 11, & 12

Version 1.0, September 5, 2007

- Added Caution 1

Version 1.1, October 1, 2007

- Added Erratum 13

Version 1.2, January 4, 2008

- Added Erratum 14

Version 1.3, March 7, 2008

- Added Erratum 15

Version 1.4, October 7, 2008

- Added Erratum 16 & 17

Version 1.5, November 10, 2008

- Edited Erratum 10

Version 1.6, February 3, 2009

- Added Erratum 18

Version 1.7, March 19, 2010

- Added Erratum 19

Errata Table

Errata in rev. AA silicon	Description	Risk Category
1.	ASPM Enabled on Both Sides of Link	Low
2.	Error Logging	Low
3.	Correctable Error Message Forwarding	Low
4.	Error Message Generation	Low
5.	Blinking Signal when Auto-Negotiating Down	Low
6.	Cut-Thru Cancellation with Relaxed Ordering	Low
7.	Possible Link Down-Train During Link Initialization	Medium
8.	No Unsupported Request Response for Non-Posted TLP Targeted to Non-existent Downstream Device After an Upstream Port Secondary Bus Reset	Low
9.	Weighted Round Robin Port Arbitration May Not Reflect Correct Weights	Low
10.	Dropped SKIP Ordered Set when Exiting L1 Power Management State	Low
11.	SDA and SCL I/O Pins can Load Down I^2C Bus when Chip is Not Powered	Low
12.	AC-JTAG Not Supported	Low
13.	No MSI Event Generated for Pending Events when MSI Enabled and Master is Enabled after Interrupt Event	Low
14.	I^2C Interface NAKs Last Expected Byte of Transaction	Low
15.	MSI Pending Bits Not Implemented	Low
16.	Sends “ACK” for Previous TLP when TLPs Replayed	Low
17.	Flow Control Frequency when Extended Synch Bit Set	Low
18.	Does Not Go to Detect State After Loopback Exit	Low
19.	Block Non-Posted TLP Can Block Posted and Completion TLPs if Cut-Thru is Enabled	Low

ERRATA LIST

Erratum #1 – ASPM Enabled on Both Sides of Link

Risk Category: Low **Silicon Revisions Affected: AA**

Description

When L0s Active State Power Management (ASPM) is enabled on one side of the link (i.e. either on the transmit side or the receive side, but not both), L0s operation happens as expected. But when L0s ASPM is enabled on both sides of the link (the transmit and receive side), the PEX 8533 initiates the process of going into recovery.

Workaround

Enabling bit "10" in Register "0x1E0" will allow ASPM to be enabled in the system. However, the PEX 8533 will not go into L0s on the transmit side.

Customer Impact

Customer will not be able to go into L0s state on the transmit side.

Erratum #2 – Error Logging

Risk Category: Low **Silicon Revisions Affected: AA**

Description

When the PEX 8533 receives "Fatal/Non-fatal error messages" from a downstream device, the "Received System Error Status bit" in the Secondary Status Register and the "Signaled System Error bit" in the PCI Command/Status Register (on the downstream and upstream ports) is not set. This error is also applicable when a downstream port of the device generates an internal error; the upstream port of the PEX 8533 should log status in both the aforementioned bits, but it does not. The correct behavior is for the "Received System Error Status bit" and the "Signaled System Error bit" to be set when error messages are received, given that "SERR# enable" in the Command Register is enabled.

Workaround

None.

Customer Impact

The "Received System Error Status bit" and "Signaled System Error bit" in the device cannot be relied upon by diagnostic software to trace the origin of the error. A scan of all the downstream devices is needed to trace the source of the error.

Erratum #3 – Correctable Error Message Forwarding

Risk Category: Low **Silicon Revisions Affected: AA**

Description

In the PEX 8533, forwarding of all types of error messages (fatal/nonfatal/correctable) is controlled by the "Bridge Control Register SERR# enable bit" and by the "Command Register SERR# enable bit." The correct behavior is for the forwarding of "Correctable error message" to only be controlled by the "Bridge Control Register SERR# enable bit," and not by the "Command Register SERR# enable bit."

Workaround

Set the "Command Register SERR# enable bit" in addition to the "Bridge Control Register SERR# enable bit."

Customer Impact

No "Correctable error messages" will be forwarded by the PEX 8533 unless the "Command Register SERR# enable bit" is set.

Erratum #4 – Error Message Generation

Risk Category: Low **Silicon Revisions Affected: AA**

Description

When a fatal or nonfatal Unsupported Request (UR) error occurs, and if "UR Reporting enable bit" is not set in the Device Control Register, then no error message will be generated regardless of the "SERR# enable bit" in the Command Register. The correct behavior is for an error message to be generated if either the "UR Reporting enable bit" or "SERR enable bit" in the Command Register is set, and the corresponding "Fatal or Non-fatal Reporting enable bit" is set.

Workaround

Set the "UR Reporting enable bit" to receive UR error messages.

Customer Impact

Setting only the "Command Register SERR# enable bit" does not ensure that the UR error messages will be sent.

Erratum #5 – Blinking Signal when Auto-Negotiating Down

Risk Category: Low **Silicon Revisions Affected: AA**

Description

When a port is up (meaning the link has trained successfully), the PEX_PORT_GOOD# signal is asserted. It was the design intent for the signal to blink if the negotiated link width is less than the configured link width (i.e. if a port is auto-negotiating down to a smaller link width). For example, if a port which is expected to come up as x16 comes up as x8 as a result of auto-negotiation, the signal for that port should assert and de-assert for equal intervals. This functionality does not work properly in the PEX 8533. Instead of blinking, the signal is always asserted when the link is up, and is always de-asserted when the link is down.

Workaround

None.

Customer Impact

The customer will be unable to visually determine if the link is auto-negotiating down or not. Instead, the customer will have to read the link width from the Link Status Register to determine the actual link width.

Erratum #6 – Cut-Thru Cancellation with Relaxed Ordering

Risk Category: Low **Silicon Revisions Affected: AA**

Description

There is a possibility of dropping a "Completion" type packet when all of the following conditions align:

1. Two or more PEX switches are being used in cascaded mode, or a single PEX 8533 is being used with End Points/Root Complexes that do not advertise infinite completion credits
2. Cut-Thru Cancellation Error (i.e. bad packet error, such as CRC, EDB, etc.) on a "Relaxed Ordered" (RO) Completion packet (with Data) while using Cut-Thru
3. Completions (RO & non-RO) blocked in the switch due to lack of credits
4. Mix of RO and non-RO Completions active in the queues

This Erratum does not affect systems with single stage PEX 8533s where devices connected to the switch advertise infinite credits for completions (as required by all Root Complexes not supporting Peer-to-Peer traffic and all Endpoints, per the PCIe specification).

Workarounds

Affected systems may implement *one* of the following workarounds:

1. Turn off Relaxed Ordering for the switch (0x664, bit 5)

a. RO can still be enabled in the system, switch will treat RO packets as normal completions without Relaxed Ordering

2. Do not allow a mix of RO and non-RO types of Completions be active in the Switch simultaneously

3. Disable Cut-Thru

Customer Impact

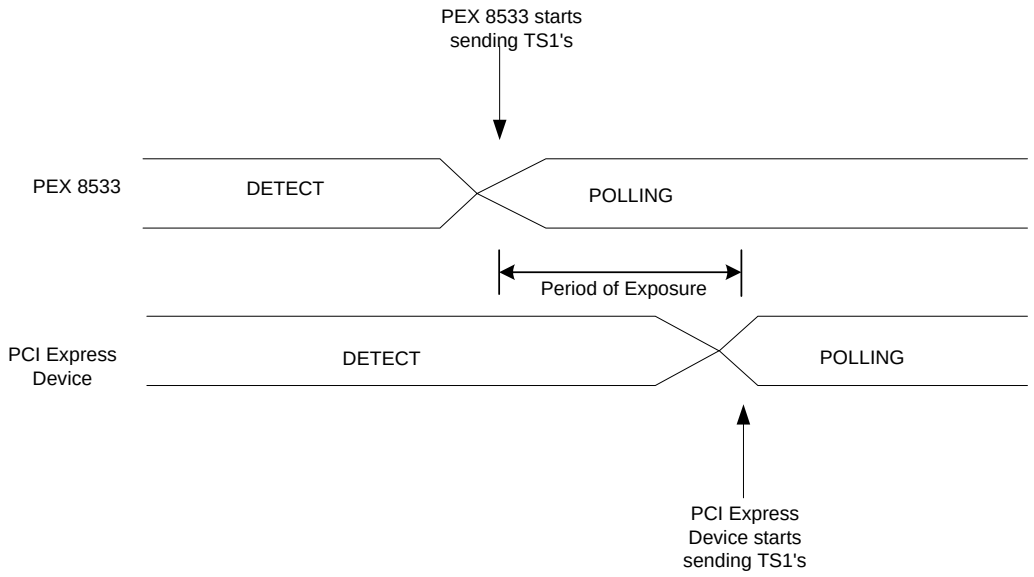
If the workaround is not implemented, when a completion is dropped, a completion timeout will occur at the original requestor. The event will not result in silent data corruption, and the software will need to take corrective action.

Erratum #7 – Possible Link Down-Train During Link Initialization

Risk Category: Medium Silicon Revisions Affected: AA

Description

When the PEX 8533 is connected to devices that start transmitting training sets to the PEX 8533 after the PEX 8533 has been transmitting training sets (PEX 8533 in Polling state), there is a period of exposure where the receivers in the PEX 8533 are in 'electrical idle' state while the transmitters in the PEX 8533 are sending valid data. During this period, the Elastic Buffer could be set in a state that prevents the lane from participating in the final link configuration.



Workaround

1. Skew the reset inputs such that the adjacent device starts transmitting ordered sets first. The Elastic Buffer will not enter the illegal state when the PEX 8533 detects valid data on its receiver lanes during its transition from Detect state to Polling

state. For slots where it is difficult to control polling entry by skewing reset, the downstream port could be held in link down state by setting the Port Disable bits (0x236 bits 0-2) through EEPROM and enabling the port prior to discovery.

Note: Port could be disabled on the downstream ports but cannot be done on the upstream port. For the upstream port if the "programmed" and "negotiated" link widths are the same, the switch enters polling ~13.5Msec after Reset, if the upstream device (e.g. Root Complex) starts transmitting valid data before this time period there is no exposure to this errata, empirical data on some chip-sets indicate RC enters polling (transmitting valid data) at ~12 Msec.

2. Re-start training sequence on link. When the adjacent device attempts to train at a link width less than its programmed width, the PLX port can be instructed to re-start its link training sequence by issuing a link disable closely followed by a link enable. The PEX 8533 link can be disabled/enabled via register 0x78 bit 4 of the appropriate port configuration space.

   Note: S/W can read the "Link Status" register of the switch and determine the negotiated width and read the downstream device "Link Capability" register and determine the configuration width and if they do not match can proceed to initiate the Link Disable/Enable sequence.

3. Use Different Reference clocks. The Elastic Buffer is able to recover from this condition in the Polling state when different reference clocks are used. The link will be reliable once it transitions to the Active state.

Customer Impact

If the workaround is not implemented, the PEX 8533 link may down-train to a smaller width when it enters the Polling state, thus affecting performance. During link up cycles where the switch comes up in the configured width, there is no run time exposure.

Erratum #8 – No Unsupported Request Response for Non-Posted TLP Targeted to Non-existent Downstream Device After an Upstream Port Secondary Bus Reset

Risk Category: Low **Silicon Revisions Affected: AA**

Description

If the PEX 8533 is strapped in a configuration where the downstream ports are vacant (no device connected) and a Secondary-Bus-Reset is applied to the upstream port of the switch (Reset on the virtual bus, upstream port register 3Ch[22]), subsequent Non-Posted Transaction Layer Packets (TLPs) targeted to the vacant downstream ports are incorrectly silently discarded. The PEX 8533 should respond to such requests with an Unsupported Request (UR). If a device is present on the downstream port of the switch then the

transactions on the corresponding down stream port are handled correctly. A Hot Reset received on the upstream port of the switch, or an upstream port link down condition (DL_Down), or a downstream port Hot Reset (Bridge Control 3C[22]) generated from the downstream port of the switch do not lead to this behavior.

Workaround

Rather than apply a Secondary Bus Reset to the upstream port, instead use a Hot Reset from the upstream device, or an upstream port DL_Down condition, or apply Secondary Bus Resets (3Ch[22]) from the downstream ports of the switch.

Customer Impact

The Device generating a Non-Posted TLP to a Non-Existent downstream device can experience a Completion Timeout, as the switch will not provide a Completion to the Non-Posted TLP.

Erratum #9 – Weighted Round Robin Port Arbitration May Not Reflect Correct Weights

Risk Category: Low **Silicon Revisions Affected: AA**

Description

The PEX 8533 has an ingress queue and an egress queue. Port arbitration moves the TLP from the ingress queue to the egress queue based on Round Robin or Weighted Round Robin arbitration schemes. The default arbitration scheme is Round Robin. When the Weighted Round Robin scheme is enabled on a particular port, it restricts the number of TLPs in the egress queue to provide finer port arbitration. An erratum in the restriction of the TLPs prevents the correct assignment of weights to the traffic flow.

Workaround

Use the device-specific source port weight arbitration scheme to assign traffic priority.

Customer Impact

The port arbitration scheme, as defined by the PCI Express Base Specification, may not assign the correct weights to the traffic flows.

Erratum #10 – Dropped SKIP Ordered Set when Exiting L1 Power Management State

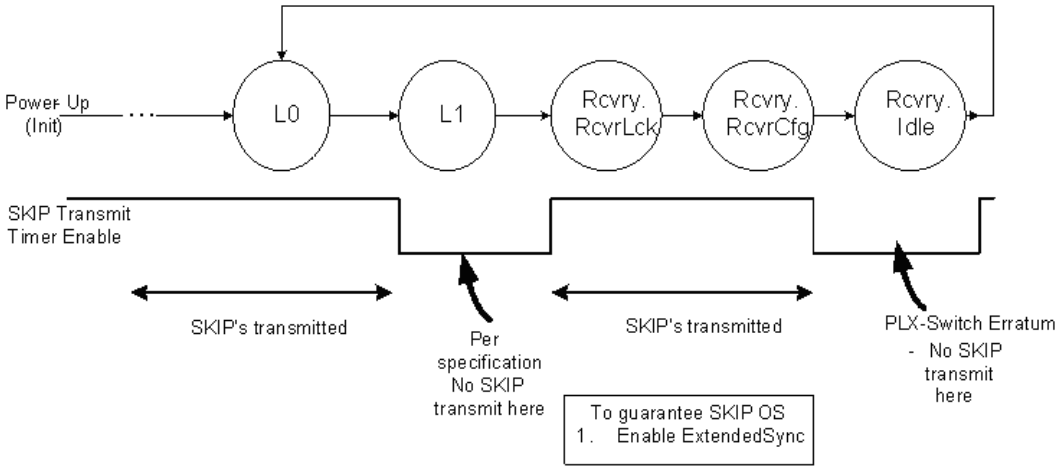
Risk Category: Low **Silicon Revisions Affected: AA**

Description

When exiting from the L1 Power Management (PM) state, there is a narrow window (recovery.idle state) during which no SKIPs are transmitted. The exit from L1 PM state steps through 3 sub-states – 1) rcvry.rcvrlock, 2) rcvry.rcvrconfig and 3) rcvry.idle. For the duration in which the link is in Recovery.idle state (waiting for receipt of 8 Idle

symbols) there will be no SKIP ordered sets transmitted. Note, SKIPs will be transmitted properly when entries into recovery sub-states are from another (L0 or L0s) state.

The Figure below illustrates the behavior.



Workaround

Disable the PEX 8533 upstream port's entry to L1 state for ASPM L1 [78h[1:0]] and PCI PM (device specific offset 1E0[7]]) mechanisms.

Customer Impact

PCI Express devices whose PHY can only deskew using SKIP ordered sets and need it to exit the recovery path shown above could fail L1 entry/exit and reach the L0 state >100ms later, taking the link down along the way. The workarounds listed above fix this, with Workaround 1 implying that the PEX 8533 will not realize any L1 power savings.

Erratum #11 – SDA and SCL I/O Pins can Load Down I^2C Bus when Chip is Not Powered

Risk Category: Low **Silicon Revisions Affected: AA**

Description

The I^2C specification requires that the open-drain SDA and SCL I/O pins in slave mode be floating and present no load to the wire-AND bus when power is switched off to the PEX 8533.

The PEX I/O implementation contains a clipping diode to the VDD33. When the power is off, this internal diode can turn on when the SDA or SCL line is asserted high by a master or any other slave device hanging on the bus. The resulting forward-bias current can lead to abnormal operation within the PEX 8533, resulting in system hang-up.

Workaround

At the board level, connect the SDA and SCL pins to the bus through a FET device controlled by the local PEX VDD33 supply rail. When the power is off, the two FET devices will isolate the PEX 8533 from the local I^2C bus. The SuperTex TN2501 is an example of a very good, low Vgs transistor.

Customer Impact

If the PEX 8533 is not powered up, the I^2C local bus may not function properly (result in system hang-up) without the two additional FET devices to isolate the PEX 8533 from the local I^2C bus.

Erratum #12 – AC-JTAG Not Supported

Risk Category: Low Silicon Revisions Affected: AA

Description

The SerDes receiver circuitry contains a circuit defect that causes the scan flip-flop to be falsely cleared just before it is to be scanned. The receiver fails to operate reliably at conditions that would normally be used for system level AC-JTAG operations. The conclusion is that the SerDes AC-JTAG model is not reliable for typical system environments and is no longer included in the BSDL file. DC-JTAG (1149.1) is unaffected.

Workaround

None. AC-JTAG has been removed from the BSDL file.

Customer Impact

Boundary-Scan instructions *EXTEST_PULSE* and *EXTEST_TRAIN* are not supported.

Erratum #13 – No MSI Event Generated for Pending Events when MSI Enabled and Master is Enabled after Interrupt Event

Risk Category: Low Silicon Revisions Affected: AA

Description

This erratum is applicable only in systems that use the MSI. It is only applicable when Legacy interrupts are disabled in the system and MSI is not yet enabled (MSI Enable bit in the MSI Control register is not set or the Bus Master Enable bit in the Command register is not set). In this case, if there is an event like a Hot Plug event which can cause an MSI event to be saved, when the MSI/Bus Master is later enabled, an MSI notification of the prior event is not sent. Also, the (optional) MSI Pending Bits status register is not implemented in the device.

Workaround

There are two options:

1) When the MSI enable bit is set by the host in the system, the host should scan all pending status bits that can generate an MSI event. The same is true for Bus master enable bit.

2) Keep INTx interrupt signaling mechanism enabled until MSI enable bit is set.

To work around the non-implemented MSI Pending Bits register issue, clear the Per-Vector Masking Capable bit in the Message Control register (48h[24]) by EEPROM or I^2C.

Customer Impact

If MSI (rather than Legacy INTx) is to be used (i.e. Hot Plug event notification), and MSI is enabled after Hot Plug event(s) occur (i.e. Attention Button pressed), the host will not be notified of the prior event when the host later enables MSI. Therefore, when the host enables MSI, the host should scan the interrupt status bits to determine whether a prior event occurred. Subsequent events will trigger an MSI write.

Erratum #14 – I^2C Interface NAKs Last Expected Byte of Transaction

Risk Category: Low Silicon Revisions Affected: AA

Description

This erratum applies to the I^2C interface of the device. When writing to the device, the I^2C interface always NAKs the last expected byte of the transaction. Instead, the device should ACK all bytes being written. If an external master tries to write more than the expected number of bytes, then the device should NAK those unexpected bytes.

Workaround

If the I^2C external master ignores the NAK on the last byte and does not retry the write of the last byte, the I^2C interface works as expected.

Customer Impact

After the I^2C interface NAKs the last expected byte of the transaction, if the I^2C master keeps retrying the I^2C write access, an infinite loop may be created.

Erratum #15 – MSI Pending Bits Not Implemented

Risk Category: Low Silicon Revisions Affected: AA

Description

The PEX 8533 implements per-vector masking capability bits for MSI. The per-vector masking capability structure requires that both the mask *and* the pending bits be implemented. However, the MSI pending bits are not implemented in the PEX 8533.

Workaround

Clear the per-vector masking capability bit in register 0x48 (bit 24) using EEPROM.

Customer Impact

Per-vector masking capability does not work properly in this device.

Erratum #16 – Sends "ACK" for Previous TLP when TLPs Replayed

Risk Category: Low Silicon Revisions Affected: AA

Description

If an external device replays TLPs inspite of receiving "ACK" DLLPs for TLPs that it had sent out, there is a case where the PEX 8533 could send an "ACK" with the sequence number for a previously "ACKed" TLP.

If the previously sent "ACK" got dropped due to link errors, there is no issue. However, if the device had latched the previously sent "ACK" yet it is replaying the TLP, it will see an ACK with a sequence number for which an ACK has been received before.

This issue can also be seen in case there is a particular sequence of traffic and there are link errors on the ACK that the PEX 8533 had sent out. If the device had sent another TLP whose ACK it had not received before it replays the TLP whose ACK got dropped, it could hit this issue where a bad DLLP error can cause a DLLP protocol error.

Workaround

Increase the replay timeout for the external device to avoid this corner case issue. The longer the replay timer, the lower the probability of hitting this issue.

Customer Impact

This erratum could lead to a DLL Protocol Error if the external device has a particular erratum of latching an ACK yet replaying the TLP, or if there are link errors for a particular traffic density.

Erratum #17 – Flow Control Frequency when Extended Synch Bit Set

Risk Category: Low Silicon Revisions Affected: AA

Description

When the extended sync bit is set on a particular port, flow control update timers and subsequently flow control updates seen on the PCI express bus on that port violate the PCI Express specified values in the PEX 8533.

The PCI Express specification requires flow control frequency with extended sync bit set to be 120us +50%, which calculates to a maximum of 180us gap between flow control update DLLPs.

In the PEX 8533, however, the flow control update timer is set to 256us.

Workaround

Do not set the extended synch bit.

Customer Impact

If the extended synch bit is set, the device flow control update timer violates the specification.

Erratum #18 – Does Not Go to Detect State After Loopback Exit

Risk Category: Low Silicon Revisions Affected: AA

Description

If the PEX 8533 is in the loopback active slave state and it sees an electrical idle ordered set, it is supposed to exit the loopback active state, go to loopback exit state, and after 2 milliseconds move on to detect state. After seeing an electrical idle ordered set in the loopback active state, the PEX 8533 does go to the loopback exit state, but it does not move on to detect state after 2 milliseconds. A fundamental reset is needed for the PEX 8533 to go back to detect state.

Workaround

Assert a fundamental reset to the device when going from loopback state to link up.

Customer Impact

If a PEX 8533 port is being used as a loopback slave, a fundamental reset will be required to bring the link up again.

Erratum #19 – Blocked Non-Posted TLP Can Block Posted and Completion TLPs if Cut-Thru is Enabled

Risk Category: Low **Silicon Revisions Affected: AA**

Description

If a Non-Posted TLP is blocked at the switch egress port due to lack of Non-Posted flow control credits from the connected device, the Non-Posted TLP remains queued at the switch egress port. If Cut-Thru is enabled, a Non-Posted TLP that is blocked (by insufficient credits at the switch egress port) can block pending Posted and Completion TLP(s) that are queued (within the switch) behind the blocked Non-Posted TLP.

Workaround

1. Disable Cut-Thru by clearing the Cut-Thru Enable bit (Port 0 register 1DCh[21] = 0). This will allow pending Posted and Completion TLPs to pass the stalled Non-Posted Request and prevent a possible cause of system deadlock.
2. Disable Completion Timeout in the requester to prevent Completion Timeout error, if a Completion TLP being blocked behind a Non-Posted TLP can be tolerated.
3. Ensure that any deficiency of Non-Posted flow control credits from the connected device is a transitory condition that is resolved before any resulting timeout error condition can occur

Customer Impact

If Cut-Thru is enabled, Posted or Completion TLP(s) can be blocked behind a Non-Posted TLP that is blocked at the switch egress port due to insufficient Non-Posted credits from the connected device. In some systems, such blockage could potentially lead to a deadlock condition.

CAUTIONS

Caution #1 – Switch Port Does Not Forward Completion TLPs in D3-Hot State

Risk Category: Low Silicon Revisions Affected: AA

Description

The PCI Express Base Specification r1.1 requires switch ports to forward Configuration TLPs and Messages while in D3-Hot State. Recently, the PCI SIG released an erratum to the PCI Express Base Specification 1.1 which describes the requirement for switch ports to forward Completion TLPs as well while in D3-Hot State.

Customer Impact

PCI Express devices expecting a Completion TLP while in D3-Hot state will experience a Completion Time-Out Error. However, many PCI Express devices do not support the forwarding of completion TLPs while in D3-Hot states since they were released prior to the PCI Express Base Specification r1.1 erratum.