

# Intel<sup>®</sup> 5520 and Intel<sup>®</sup> 5500 Chipset

## Specification Update

---

*February 2015*

**Notice:** The Intel<sup>®</sup> 5520 and Intel<sup>®</sup> 5500 Chipset may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are documented in this Specification Update.



You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

Intel technologies may require enabled hardware, specific software, or services activation. Check with your system manufacturer or retailer.

No computer system can be absolutely secure. Intel does not assume any liability for lost or stolen data or systems or any damages resulting from such losses.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Copies of documents which have an order number and are referenced in this document, or other Intel literature may be obtained by calling 1-800-548-4725 or by visiting Intel's website at <http://www.intel.com/design/literature.htm>.

No computer system can provide absolute security under all conditions. Intel® Trusted Execution Technology (Intel® TXT) requires a computer system with Intel® Virtualization Technology, an Intel TXT-enabled processor, chipset, BIOS, Authenticated Code Modules and an Intel TXT-compatible measured launched environment (MLE). The MLE could consist of a virtual machine monitor, an OS or an application. In addition, Intel TXT requires the system to contain a TPM v1.2, as defined by the Trusted Computing Group and specific software for some uses. For more information, see <http://www.intel.com/technology/security/>.

Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain computer system software enabled for it. Functionality, performance or other benefits will vary depending on hardware and software configurations and may require a BIOS update. Software applications may not be compatible with all operating systems. Please check with your application vendor.

I<sup>2</sup>C is a two-wire communications bus/protocol developed by Philips. SMBus is a subset of the I<sup>2</sup>C bus/protocol and was developed by Intel. Implementations of the I<sup>2</sup>C bus/protocol may require licenses from various entities, including Philips Electronics N.V. and North American Philips Corporation.

Intel, the Intel logo, and Xeon are trademarks of Intel Corporation in the U. S. and/or other countries.

\*Other names and brands may be claimed as the property of others.

Copyright © 2015, Intel Corporation. All Rights Reserved.



## Contents

---

<b>Revision History</b> .....	4
<b>Preface</b> .....	5
<b>Summary Tables of Changes</b> .....	7
<b>Identification Information</b> .....	11
<b>Errata</b> .....	13
<b>Technology Technology Technology Technology Specification Changes</b> .....	30
<b>Specification Clarifications</b> .....	31
<b>Documentation Changes</b> .....	36



# Revision History

---

Date	Revision	Description
March 2009	-001	Initial Release
April 2009	-002	Added Errata 45 Updated Errata 23
May 2009	-003	Removed Errata 1 Added Errata 46,47,48,49,50
June 2009	-004	Added Erratum 51 Added Specification Clarification 1,2
July 2009	-005	Added Erratum 52, 53 Updated Erratum 12
August 2009	-006	Added Erratum 54 Added Specification Update on IOH/ICH Non-Posted Peer-to-Peer Support
December 2009	-014	Added Erratum 55 56 57 58 Updated Erratum 37and 46 Added C-2 Stepping to Summary Table of Changes Added C-2 Stepping Component Identification and Marking Information
April 2010	-008	Updated Erratum 46 Removed Erratum 30 Added Erratum 59
May 2010	-009	Added Erratum 60
August 2010	-010	Updated Erratum 42
November 2010	-011	Updated Errata Summary Table
November 2010	-012	Updated Erratum 36
January 2011	-013	Added Doc Change 1
September 2011	-014	Added Added Erratum 61 and 62
February 2015	-015	Added Added Erratum 63 - 65



# Preface

---

This document is an update to the specifications contained in the [Affected Documents](#) table below. This document is a compilation of device and documentation errata, specification clarifications and changes. It is intended for hardware system manufacturers and software developers of applications, operating systems, or tools.

Information types defined in [Nomenclature](#) are consolidated into the specification update and are no longer published in other documents.

This document is intended to errata for the Intel® 5520 and Intel® 5500 Chipset. Customers can use this data to know and understand the Intel® 5520 and Intel® 5500 Chipset issues that Intel has identified.

## Affected Documents

Document Title	Document Number/Location
<i>Intel® 5520 and Intel® 5500 Chipset Datasheet</i>	321328-001

## Related Documents

Document Title	Document Number/Location
<i>Intel® Virtualization Technology for Directed I/O Architecture Specification</i>	D51397-004 <a href="http://download.intel.com/technology/computing/vptech/Intel(R)_VT_for_Direct_IO.pdf">http://download.intel.com/technology/computing/vptech/Intel(R)_VT_for_Direct_IO.pdf</a>

## Nomenclature

**Errata** are design defects or errors. These may cause the Intel® 5520 and Intel® 5500 Chipsets behavior to deviate from published specifications. Hardware and software designed to be used with any given stepping must assume that all errata documented for that stepping are present on all devices.

**S-Spec Number** is a five-digit code used to identify products. Products are differentiated by their unique characteristics, for example, number of PCIe\* lanes. Read all notes associated with each S-Spec number.

**Specification Changes** are modifications to the current published specifications. These changes will be incorporated in any new release of the specification.

**Specification Clarifications** describe a specification in greater detail or further highlight a specification's impact to a complex design situation. These clarifications will be incorporated in any new release of the specification.

**Documentation Changes** include typos, errors, or omissions from the current published specifications. These will be incorporated in any new release of the specification.

**Note:** Errata remain in the Specification Update throughout the product's lifecycle, or until a particular stepping is no longer commercially available. Under these circumstances,



errata removed from the Specification Update are archived and available upon request. Specification changes, specification clarifications and documentation changes are removed from the Specification Update when the appropriate changes are made to the appropriate product specification or user documentation (datasheets, manuals, and so forth).



# Summary Tables of Changes

---

The following tables indicate the errata, specification changes, specification clarifications, or documentation changes which apply to the Intel® 5520 and Intel® 5500 Chipset. Intel may fix some of the errata in a future stepping of the component, and account for the other outstanding issues through documentation or specification changes as noted. These tables uses the following notations:

## Codes Used in Summary Tables

### Stepping

X:	Erratum exists in the stepping indicated. Specification Change or Clarification that applies to this stepping.
(No mark) or (Blank box):	This erratum is fixed in listed stepping or specification change does not apply to listed stepping.

### Status

Doc:	Document change or update will be implemented.
Plan Fix:	This erratum may be fixed in a future stepping of the product.
Fixed:	This erratum has been previously fixed.
No Fix:	There are no plans to fix this erratum.
Under Investigation:	All issues detailed in this section are still under investigation. Issues may or may not be root caused.

### Row

Change bar to left of table row indicates this erratum is either new or modified from the previous version of the document.



## Errata (Sheet 1 of 2)

Number	Steppings		Status	Erratum
	B3	C2		
1				Removed - Not an Errata
2	X	X	No Fix	CPURST bit does not get cleared by hardware
3	X	X	No Fix	VTX-RCV-DETECT pulse too large during receiver detection.
4	X		Fixed	PCIe*2 differential peak-to-peak transmit voltage swing is too low
5	X	X	No Fix	Surprise down error status not being flagged
6	X	X	No Fix	Extended Error Detect Mask Registers of all PCIe root ports mask error logging by default.
7	X	X	No Fix	Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d) queue-based invalidation is enabled only when enabled on both channels
8	X	X	No Fix	PCIe Header of a malformed Transaction Layer Packet (TLP) is not logged
9	X	X	No Fix	Intel VT-d does not support the draining of compatibility-format interrupts
10	X	X	No Fix	Hardware applies HPA_LIMIT to upstream memory request when Intel VT-d is disabled
11	X	X	No Fix	PCIe PMCSR Power State fields allow writing D1 and D2
12	X	X	No Fix	PCIe2 Tx Return loss fails spec
13	X	X	No Fix	Persistent JTAG error reported at MIERRST register
14	X	X	No Fix	Interoperability issue of issue of some PCIe1 cards with PCIe2 Devices
15	X	X	No Fix	Intel® QuickPath Interconnect (Intel® QPI) fails to retrain to L0 slow mode after 32 in-band resets
16	X	X	No Fix	Intel VT-d: Memory read request with Address Type equal to 11b results in malformed Transaction Layer Packet (TLP)
17	X	X	No Fix	PCIe Inbound Message that should be ignored is treated as Unsupported Request
18	X	X	Doc	Memory writes to a certain address range are considered advisory non-fatal
19	X	X	No Fix	Transactions to addresses above TOCM are not setting the Master Abort
20	X	X	Doc	Remote IOH P2P transaction not working on Dual-IOH Proxy mode.
21	X	X	No Fix	Setting bits 24 and 25 of the MISCCTRLSTS does not result in expected behavior
22	X	X	No Fix	Timeout values much larger than specified
23	X	X	No Fix	Failure during operation at PCIe power management state
24	X	X	No Fix	Bandwidth very low for write traffic with noSnoop attribute set
25	X	X	No Fix	Intel QPI Queue/Table overflow or underflow error observed.
26			Non-Si	Some PCIe2 endpoints will not complete the training in PCIe2 mode.
27	X	X	No Fix	Access Control Services (ACS) Violation is not treated as Advisory when severity is set to Non-Fatal
28	X	X	No Fix	Intel QPI L1 state is taking greater than 15 us from L1 exit to L0 state
29	X	X	No Fix	Dual-IOH: Intel QPI Protocol Status D7 Error Bit set
30				Removed - not an erratum
31	X	X	No Fix	Intel QPI Link Training failures in L0 when L0r enabled
32	X	X	No Fix	Intel VT-d translated write transactions are blocked but not recorded
33	X	X	No Fix	Intel QPI initialization abort failures logged during power-on resets
34	X	X	No Fix	ERRSID not logging ReqID for Inbound ERR_* messages
35	X	X	No Fix	Intel QPI errors can occur on inband resets



## Errata (Sheet 2 of 2)

Number	Steppings		Status	Erratum
	B3	C2		
36	X	X	No Fix	Link degrades and surprise link downs (SLD) on PCIe
37	X	X	No Fix	Intel QPI CRC errors experienced during L0s entry could cause system hangs
38	X	X	No Fix	Intel QPI D2 or B2 error on L1 exit
39	X		Fixed	DCA can block progress of other transactions in Dual-IOH, dual socket systems
40	X	X	No Fix	Data Mismatch on Inbound MemWrts after MSI with payload greater than 1 DWORD payload
41	X	X	No Fix	MSI with greater than 1DWord payload is not logged in XPUNCERRSTS bit 8
42	X		Fixed	PFP Flag (due to Intel VT-d ISOCH fault) in Fault Status Register not being cleared
43	X	X	No Fix	Error not logged due to a corrupt STP symbol
44	X	X	No Fix	PCIe ASPM L1 can cause link degrade and speed change.
45	X	X	No Fix	System may hang in a Dual-IOH Platform with Intel® QuickData Technology enabled DMA traffic, PCIe inbound writes and remote non-posted peer-to-peer reads
46	X		Fixed	Lost interrupts when MSI used
47	X		Fixed	Intel VT-d: Receiving two identical interrupt requests in back to back cycles may corrupt attributes of remapped interrupt, or hang subsequent interrupt-remap-cache invalidation command.
48	X	X	No Fix	Intel QPI Error Status D3 is observed
49	X	X	No Fix	Unpredictable PCI behavior accessing non-existent memory space
50	X	X	No Fix	Bandwidth changed status errors being escalated to Global RAS
51	X		Fixed	Intel® Trusted Execution Technology (Intel® TXT) writes may not complete as expected
52	X	X	No Fix	Intel VT-d: Address Remapping error when DMA/interrupt remapping is active
53	X		Fixed	Intel VT-d: In-flight remap-able interrupts not drained on interrupt invalidation command
54	X	X	No Fix	Source ID for errors internally detected by PCIe root port 3(Ports 3-6) is not logged as expected
55		X	No Fix	Device 0, Function 0's Revision ID (RID) is not reset to the Stepping Revision ID (SRID) by a CORERST_N Assertion
56	X	X	No Fix	Header Log information may not be captured correctly when accessed via JTAG
57	X	X	No Fix	Forwarded Clock Lane Detection status may not be indicated accurately
58	X	X	No Fix	ESI link cannot go to L1 state on the Intel® 5500 Platform
59	X		Fixed	Deadlock can occur in Dual IOH Platforms
60	X	X	No Fix	In-flight DMA requests received during the implicit DMA draining window when enabling Intel VT-d hardware may result in a spurious DMA fault
61	X	X	No Fix	The ESI upstream link between the ICH10 and IOH may intermittently fail to train on some IOH devices
62	X	X	No Fix	A Modification to the Multiple Message Enable field does not affect the AER Interrupt Message Number field
63	X	X	No Fix	PCIe* Hot-Plug Slot Status Register May Not Indicate Command Completed
64	X	X	No Fix	Intel® VT-d Memory Check Error on an Intel® QuickData Technology Channel May Cause All Other Channels to Master Abort
65	X	X	No Fix	PCIe* Header of a Malformed TLP is Logged Incorrectly



## Specification Changes

Number	SPECIFICATION CHANGES
1.	

## Specification Clarifications

Number	SPECIFICATION CLARIFICATIONS
1.	Update to Registers
2.	Updates for SYRE
3.	Update on IOH/ICH Non-Posted Peer-to-Peer Support
4.	Correction to the CRID Reset Functionality
5.	Updates to Resource conflicts on IOH PCIe

## Documentation Changes

Number	DOCUMENTATION CHANGES
1.	QPI L0s feature is not supported on Intel® 5520 and Intel® 5500 Chipset



# Identification Information

## Component Identification via Programming Interface

The Intel® 5520 and Intel® 5500 Chipset stepping can be identified by the following register contents:

Stepping	Features	Vendor ID <sup>1</sup>	Device ID <sup>2</sup>	Revision Number <sup>3</sup>
B-3	36D	8086h	3406h	13h
B-3	24D	8086h	3403h	13h
C-2	36D	8086h	3406h	22h
C-2	24D	8086h	3403h	22h

**Notes:**

1. The Vendor ID corresponds to bits 15:0 of the Vendor ID Register located at offset 00 - 01h in the PCI function 0 configuration space.
2. The Device ID corresponds to bits 15:0 of the Device ID Register located at offset 02 - 03h in the PCI function 0 configuration space.
3. The Revision Number corresponds to bits 7:0 of the Revision ID Register located at offset 08h in the PCI function 0 configuration space.

## Component Marking Information

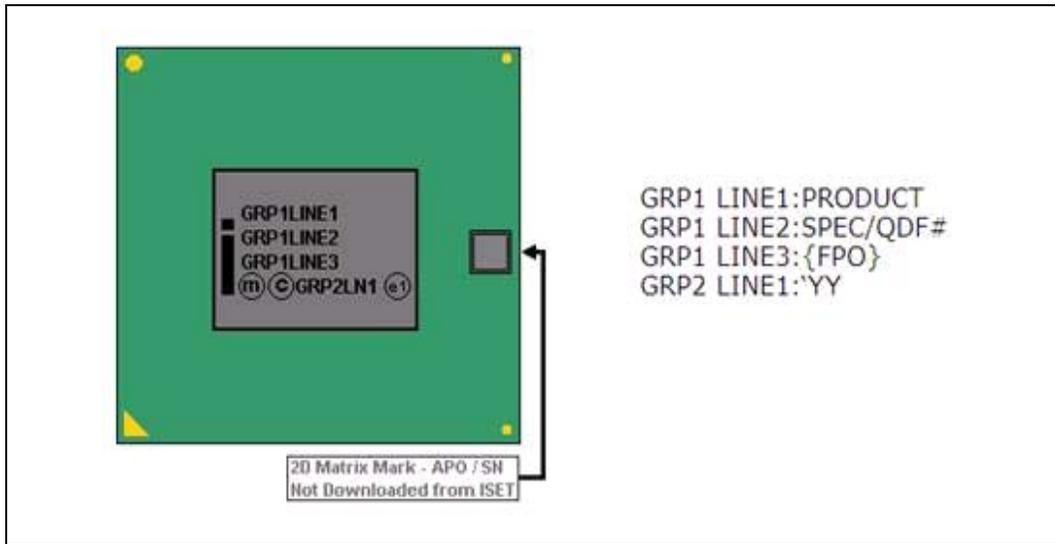
The Intel® 5520 and Intel® 5500 Chipset stepping can be identified by the following component markings:

Stepping	S-Spec	Top Marking	Notes
B-3	SLGMU	AC5520 SLGMU 901037	2, 4
B-3	SLGMT	AC5500 SLGMT 901036	2, 3s
C-2	SLH3P	AC5520 SLH3P 904729	2, 4, 5
C-2	SLH3N	AC5500 SLH3N 904728	2, 3, 5

**Notes:**

1. This is an Engineering Sample
2. This is a Production part
3. This part has 24 PCIe\* lanes
4. This part has 36 PCIe lanes
5. This part supports Intel® Trusted Execution Technology (Intel® TXT)

Figure 1. Top-Side Marking Example





# Errata

---

## 1. Removed.

## 2. CPURST bit does not get cleared by hardware

**Problem:** SYRE (Device:20, Function:2, Offset:0CCh) CPURST bit (bit 10) does not get cleared by hardware.

**Implication:** The Intel 5520 and Intel 5500 chipsets will not assert RESETO\_N when CPURST is set if it has been set previously.

**Workaround:** The CPURST bit must be cleared prior to setting it.

**Status:** For the steppings affected, see the Summary Tables of Changes..

## 3. V<sub>TX-RCV-DETECT</sub> pulse too large during receiver detection.

**Problem:** VTX-RCV-DETECT - The amount of voltage change allowed during Receiver Detection. The maximum specified for VTX-RCV-DETECT is 600 mV for both 2.5 GT/s and 5 GT/s. The VTX-RCV-DETECT pulse has been measured as high as 700 mV.

**Implication:** This may overstress PCIe\* agents.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

## 4. PCIe\*2 differential peak-to-peak transmit voltage swing is too low

**Problem:** The PCIe\*2 transmit buffers are not generating a large enough peak-to-peak transmit voltage swing, thus violating the PCIe2 specification.

**Implication:** The PCIe2 specification is violated.

**Workaround:** BIOS changes are required as part of the fix as outlined below:

Dev 13:Func 3

1. Set Offset 330h [15:12] to 5, [11:8] to 3

Dev 14:Func 0

1. Set Offset 330h [15:12] to 5, [11:8] to 3

Dev 13:Func 0

1. Set Offset 1A0h [23:16] to 5, [7:0] to FFh
2. Set Offset 11Ch [28:25] to n, (n = 0)
3. For each n setting in step 2, set Offset 1B4h [11:8] to 8, [7:4] to 6, [3:0] to 5

Dev 13:Func 1

1. Set Offset 1A0h [23:16] to 5, [7:0] to FFh
2. Set Offset 11Ch [28:25] to n, (n = 0, 1)
3. For each n setting in step 2, set Offset 1B4h [11:8] to 8, [7:4] to 6, [3:0] to 5

Dev 13:Func 3

1. Set Offset 1A0h [23:16] to 5, [7:0] to FFh
2. Set Offset 11Ch [28:25] to n, (n = 0, 1, 2, 3)
3. For each n setting in step 2, set Offset 1B4h [11:8] to 8, [7:4] to 6, [3:0] to 5

Dev 14:Func 0

1. Set Offset 1A0h [23:16] to 5, [7:0] to FFh



2. Set Offset 11Ch [28:25] to n, (n = 0, 1, 2, 3)
3. For each n setting in step 2, set Offset 1B4h [11:8] to 8, [7:4] to 6, [3:0] to 5

**Note:** The above configuration sequence needs to be repeated in the S3 state resume path. For Dual-IOH configuration, the above configuration sequence applies to both IOHs.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 5. Surprise down error status not being flagged

**Problem:** The Surprise down error status bit (Device:0-10, Function:0, Offset:104h, bit 5) is not being flagged during training failure or surprise removal.

**Implication:** The system will not be able to reliably detect if a PCIe card has not trained properly or been inadvertently removed.

**Workaround:** The logging of a Surprise Down event caused by a non hot plug removal event depends on the state of the Power Controller Control bit (offset A8h bit[10]). For slots that don't support hot plug, BIOS must clear the corresponding device's Power Controller Control bit control (bit [10]) to 0 in SLTCON register, to enable detection of Surprise Down condition of the associated endpoint.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 6. Extended Error Detect Mask Registers of all PCIe root ports mask error logging by default.

**Problem:** The mask fields of all PCIe root ports (devices 1 - 10) Extended PCIe Error Detect Mask Registers are set to 1 by default.

**Implication:** All PCIe Advanced error status logging registers have a corresponding error detect mask register to control if the respective error status will be logged. Below are the error status/error detect mask pairs:

Uncorrectable Error Status (offset 104h) and Detect Status Mask (offset 218h).

Correctable Error Status (offset 110h) and Detect Status Mask (offset 21Ch).

Root Port Error Status (offset 130h) and Detect Status Mask (offset 220h).

XP Correctable Error Status (offset 200h) and Detect Mask (offset 228h).

XP Uncorrectable Error Status (offset 208h) and Detect Mask (offset 224h).

With the default values of these error detect mask registers, no PCIe advanced errors will be logged and reported.

**Workaround:** BIOS must clear registers 218h, 21Ch, 220h, 228h, and 224h to zero in order to log and report corresponding errors.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 7. Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d) queue-based invalidation is enabled only when enabled on both channels

**Problem:** The isochronous and non-isochronous channels can both operate either in Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d) queue-based invalidation or in Intel VT-d register based invalidation mode.

**Implication:** This is a violation of the specification which allows both the isochronous and non-isochronous channels to be configured in either Intel VT-d queue-based or Intel VT-d register-based invalidation independently.

**Workaround:** None



**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **8. PCIe Header of a malformed Transaction Layer Packet (TLP) is not logged**

**Problem:** A TLP is logged in the UNCERRSTS register (Device:0-10, Function:0, Offset:0x104h), but the header of the malformed TLP is not logged in the HDRLOG register (Device:0-10, Function:0, Offset 11Ch).

**Implication:** This is a violation of the PCIe specification, that requires the header of the malformed TLP to be logged in the HDRLOG register.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **9. Intel VT-d does not support the draining of compatibility-format interrupts**

**Problem:** The Intel VT-d does not support the draining of compatibility-format interrupts which is a violation of the Intel VT-d specification.

**Implication:** The Draining of compatibility-format interrupts is used when software wants to convert an interrupt source from compatibility-format to remappable-format.

**Workaround:** For the software that uses compatibility-format interrupts, use an alternative approach to drain interrupt. Specifically, software issues reads to any PCI/PCIe architectural registers without read side-effects at the interrupt source (endpoint device), and have the PCI ordering rules of read-completion push/drain for any in-flight interrupts (including compatibility-mode interrupts) from that source.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **10. Hardware applies HPA\_LIMIT to upstream memory request when Intel VT-d is disabled**

**Problem:** When Intel VT-d DMA translation is enabled, hardware correctly applies the GPA\_LIMIT check to the incoming DMA address and HPA\_LIMIT to Intel VT-d page-walk addresses. Note that the reset default value of HPA\_LIMIT in hardware is 39 (that is, address bits 38:0 are valid).

**Implication:** When Intel VT-d DMA translation is not enabled, hardware should ignore the GPA\_LIMIT and HPA\_LIMIT so that hardware does not apply the HPA\_LIMIT check to the incoming DMA addresses.

**Workaround:** The BIOS must program the appropriate HPA\_LIMIT in VTGENCTRL register always, irrespective of the BIOS setup option to expose/disable Intel VT-d.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **11. PCIe PMCSR Power State fields allow writing D1 and D2**

**Problem:** The PMCSR (Devices: 0 to 10, Function: 0, Offset: E4h) bits 1:0 allow states D1 and D2 to be written.

**Implication:** The Intel® 5520 and Intel® 5500 Chipset do not support the D1 and D2 states. The IOH does not change power states from D0/ D3hot when PMCSR bits 1:0 are written to D1 or D2, thus there is no functional impact to the IOH when these states are written. Do not write states D1 and D2 to the PMCSR bits 1:0.

**Workaround:** Do not write states D1 and D2 to the PMCSR bits 1:0.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **12. PCIe2 Tx Return loss fails spec**

**Problem:** The differential and common mode transmit return loss fail to meet the PCIe2 specification.



**Implication:** Some PCIe2 agents may have increased inter-symbol interference, ISI, due to signal reflection at the driver of Intel 5520 and Intel 5500 chipsets. PCIe1 return loss will meet specifications.

**Workaround:** None.

**Status:** For the steppings affected, see the Summary Tables of Changes.

### 13. Persistent JTAG error reported at MIERRST register

**Problem:** The MIERRST register (Device: 20, Function: 2, Offset: 380h) bit 2 continuously gets set if the MIERRCTL register (Device: 20, Function: 2, Offset: 384h) bit 2 is set regardless of whether an error truly exists on the JTAG interface or not. Further, the MIERRCNT register (Device: 20, Function: 2, Offset 3C0h) will indicate an overflow via bit 7 as errors are continuously registered.

**Implication:** The MIERRST and MIERRCNT registers are not dependable sources of information if the Intel 5520 and Intel 5500 chipsets are configured to record errors related to the JTAG interface.

**Workaround:** None.

**Status:** For the steppings affected, see the Summary Tables of Changes.

### 14. Interoperability issue of issue of some PCIe1 cards with PCIe2 Devices

**Note:** This is not an Intel 5520 and Intel 5500 chipsets issue, and the workaround documented below is to resolve the interoperability issue certain PCIe1 cards have with PCIe2 capable Intel 5520 and Intel 5500 chipsets PCIe root ports.

**Problem:** During PCIe physical layer link initialization, Intel 5520 and Intel 5500 chipsets PCIe2 capable PCIe root ports advertise 2.0 Specification defined capabilities such as 5.0 GT/s data rate support and link upconfigure in the Training Sequence (TS) Ordered Sets. Some legacy downstream PCIe1 cards may not correctly interpret these fields that are reserved in PCIe1, but not in PCIe2.

**Implication:** These PCIe1 cards may fail to properly train with Intel 5520 and Intel 5500 chipsets.

**Workaround:** Below is the standard method to force Intel 5520 and Intel 5500 chipsets PCIe root ports to PCIe1 operation:

- 1. Set LNKCON2 (Dev 1-10: Func 0: Offset C0h) bits [3:0] to 0001b.
- 2. Clear LNKCON2 (Dev 1-10: Func 0: Offset C0h) bits [6] to 0b
- 3. Set LNKCON (Dev 1-10: Func 0: Offset A0h) bit [5] to 1b to retrain the link.

If some PCIe1 cards fail to train with the standard method, BIOS should program the following registers to prevent Intel 5520 and Intel 5500 chipsets PCIe root ports from advertising 2.0 Specification defined capabilities. There is one register control field for each Intel 5520 and Intel 5500 chipsets PCIe root port.

- Dev 13: Funcs 6 - 1: Offset 4B4h bit[23] for root ports 6 - 1.
- Dev 14: Funcs 3 - 0: Offset 4B4h bit[23] for root ports 10 - 7.

Clear this bit to disable the corresponding root port from advertising upconfigure capability before doing Step 3 above.

**Status:** For the steppings affected, see the Summary Tables of Changes.

### 15. Intel® QuickPath Interconnect (Intel® QPI) fails to retrain to L0 slow mode after 32 in-band resets

**Problem:** After 32 Intel® QuickPath Interconnect (Intel® QPI) inband resets from the CPU to an IOH port are executed when in slow mode the next inband reset will fail as the Intel QPI will not retrain to L0 slow mode.



**Implication:** The Intel QPI link will fail when operating in slow mode after 32 Intel QPI inband resets.  
**Workaround:** Bit 5 of device 13, functions 1-0, offset B4Ch must be set every time before an inband reset from the CPU to an IOH port.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **16. Intel VT-d: Memory read request with Address Type equal to 11b results in malformed Transaction Layer Packet (TLP)**

**Problem:** When in Intel VT-d mode, memory read requests with Address Type 11b result in a malformed Transaction Layer Packet.

**Implication:** This is a violation of the PCIe Address Translation Services Specification, Version 1.0. When in Intel VT-d mode, memory read requests with Address Type 11b should be completed with an Unsupported Request.

**Workaround:** None.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **17. PCIe Inbound Message that should be ignored is treated as Unsupported Request**

**Problem:** PCIe Inbound Messages with Message Codes between 0x40 and 0x48 (with the exception for x42 and 0x46) cause the root port to log errors as Unsupported Request when it should have ignored and discarded them.

**Implication:** Unsupported Request Error will be logged under the circumstances outlined above. End Point devices are strongly encouraged by PCIe spec 2.0 not to send those messages with message codes between 0x40 and 0x48.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **18. Memory writes to a certain address range are considered advisory non-fatal**

**Problem:** Accesses above address range  $2^{51}$  (TOCM) are required to be Master Aborted and the error logged in UNCERRSTS (Device:0-10, Function:0, Offset:104h). When severity of master-abort (UR) is set to non-fatal, memory write accesses should be considered normal non-fatal rather than advisory non-fatal. There is a range of address from  $2^{51}$  to  $2^{52}-1$  (0x8\_0000\_0000\_0000 to 0xF\_FFFF\_FFFF\_FFFF) for which memory write accesses are logged as advisory non-fatal. In this case, CORERRSTS (Device:0-10, Function:0, Offset:110h) bit 13 is set. Note that this issue does not arise if UR severity is set to Fatal.

**Implication:** The status of transactions occurring as described above will not be correctly reflected.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **19. Transactions to addresses above TOCM are not setting the Master Abort**

**Problem:** Inbound Transactions to addresses above TOCM are master aborted, but Intel 5520 and Intel 5500 chipsets do not set the C4 bit in the IOHERRST (Device:20, Function:2, Offset:300h) register.

**Implication:** Such cases of master aborts will not have status recorded.

**Workaround:** When Intel VT is enabled, there is no workaround. When Intel VT is disabled, program the HPA\_LIMT to the maximum value.

**Status:** For the steppings affected, see the Summary Tables of Changes.



## 20. Remote IOH P2P transaction not working on Dual-IOH Proxy mode.

**Problem:** P2P transactions between the two Intel® 5520 and Intel® 5500 Chipsets in Dual-IOH Proxy mode is not successful.

**Implication:** This causes a system hang.

**Workaround:** Dual-IOH proxy mode requires SAD registers to decode the destination of the P2P transaction regardless of EN/EP sprofile. BIOS needs to select SAD mode even for UP profile by setting QPIPCTRL (Dev 16: Func 1: Offset 4Ch, QWord) bit [0] to 1.

Note: Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code.

**Status:** For the steppings affected, see the Summary Tables of Changes.

## 21. Setting bits 24 and 25 of the MISCCTRLSTS does not result in expected behavior

**Problem:** Setting the bits 24 and 25 in the MISCCTRLSTS (Device: 0-10; Function: 0, Offset: 188h) register does not result in the peer-to-peer memory read/write transactions being disabled as expected.

**Implication:** Memory read/write transactions may still go through to the memory and the expected Completer Abort message may not be received.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

## 22. Timeout values much larger than specified

**Problem:** When using the bits 51:48 of the CSIPCTRL register (Device: 16, Function: 1, Offset 0x4C) to set the configuration retry timeout values, it has been observed that the actual timeout values may be much longer than what is specified for each of the settings.

**Implication:** The expected timeout signal for transactions exceeding the time limit may not be observed.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

## 23. Failure during operation at PCIe power management state

**Problem:** There may be intermittent failures observed when exiting from the L1 power management state back to the L0 power management state.

**Implication:** Due to this issue, successful transition from the L1 state back to the L0 state may not be achieved.

**Workaround:** Set the bit 26 in the following registers: Device 14, Function 3-0, offset 390h and Device 13, Function 6-0, Offset 390h to 1. Also set bit 15 in the following registers: Device 14, Function 0, Offset 31Ch and Device 13, Function 3, Offset 31Ch to a 0

**Status:** For the steppings affected, see the Summary Tables of Changes.

## 24. Bandwidth very low for write traffic with noSnoop attribute set

**Problem:** 100% upstream posted write traffic (PCIe to memory) with the noSnoop attribute set is achieving lower than expected bandwidth. In comparison, the same test with snooped traffic achieves much higher bandwidths.

**Implication:** Affected systems will achieve lower than expected performance.

**Workaround:** The PERFCTRLSTS (Device:1-10, Function:0, Offset:180h) register, bits 3 and 2 should be set to 0.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**25. Intel QPI Queue/Table overflow or underflow error observed.**

**Problem:** Intel QPI Queue/Table overflow or underflow error status (Device: 20, Function: 2, Offset: 84h. bits[27:26]) is observed.

**Implication:** This spurious error status will cause false error reporting if enabled.

**Workaround:** BIOS should program the following Intel 5520 and Intel 5500 Chipsets registers to mask off the spurious sub-state status and allow the Intel QPI Queue/Table overflow or underflow error status to function. Set Dev 13:Func 0/1: Offset F90h (Dword) bit 14 to 1.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**26. Some PCIe2 endpoints will not complete the training in PCIe2 mode.**

**Problem:** Some PCIe2 endpoints will not complete training as expected when operating in PCIe2 mode.

**Implication:** Affected PCIe2 endpoints will be unusable when operating in PCIe2 mode without the workaround.

**Workaround:** Please refer to the *Intel® 5520 Chipset and Intel® 5500 Chipset BIOS Specification Update*, Ref ID# 368193.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**27. Access Control Services (ACS) Violation is not treated as Advisory when severity is set to Non-Fatal**

**Problem:** An ACS Violation is defined as non-fatal and the completer sends a completion with CA completion status then this case must be handled as an Advisory Non-Fatal Error as described in the PCI Express\* Spec.

**Implication:** The ACS violations are treated as Non-Fatal errors instead of Advisory Non-Fatal errors.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

**28. Intel QPI L1 state is taking greater than 15 us from L1 exit to L0 state**

**Problem:** The IOH Intel QPI interface is taking longer than 15 us to transition from the L1 state to the L0 state.

**Implication:** As a result IOH L1 Exit latency needs to be reduced to prevent USB underun or potential audio glitches.

**Workaround:** Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**29. Dual-IOH: Intel QPI Protocol Status D7 Error Bit set**

**Problem:** The QPIP[1:0]ERRST (Device:20, Function:2, Offset:2B0h, 230h) register, D7 Error bit (bit 10) is being set in Dual-IOH systems in the IOH-to-CPU Intel QPI port during normal operation.

**Implication:** All indication is that the error is spurious as affected platforms have run with no functional issues observed.

**Workaround:** Mask the D7 bit by clearing bit 10 of the QPIP[1:0]ERRCTL, Device:20, Function:2, Offset:2B4h, 234h) register in Dual-IOH platforms

**Status:** For the steppings affected, see the Summary Tables of Changes.

**30. Removed - Peer-to-Peer (P2P) reads and throttling.**

This erratum was removed because allowed throttling is either 0 or 50%, thus erratum is not exposed.

**31. Intel QPI Link Training failures in L0 when L0r enabled**

**Problem:** Intel QPI Link Training failures seen when L0r is enabled. The failures observed include repetitive CRC errors resulting in a Link Layer Retry (LLR) Phy Reinit or LLR Abort.

**Implication:** Certain systems may fail to boot when L0r is enabled.

**Workaround:** Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**32. Intel VT-d translated write transactions are blocked but not recorded**

**Problem:** The Intel VT-d engine blocks Intel VT-d Translated write transactions to the interrupt address range (0xFEExxxxx) but does not record the error. Requests outside of the interrupt address range are not affected.

**Implication:** For Intel VT-d Translated write transactions to the interrupt address range (0xFEExxxxx), the transaction is blocked, but evidence of the blocked transaction will not be available in the error registers.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

**33. Intel QPI initialization abort failures logged during power-on resets**

**Problem:** During power-on reset, the IOH may log Intel QPI initialization abort (D2) failures via the QPI[1:0]ERRST (Device:20, Function:2, Offset:280h, 200h) register, bit 4.

**Implication:** This has been seen on a small number of parts. If the power-on reset is executed and D2 is logged, then this indicates that an Intel QPI link went through a successful retrain.

**Workaround:** Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**34. ERRSID not logging ReqID for Inbound ERR\_\* messages**

**Problem:** Error Source Identification Register (Device:0-10, Function:0, Offset:134h) does not log Requester ID for Inbound ERR\_\* messages

**Implication:** While internally generated error messages in the IOH will have their Requester ID logged correctly in this register, incoming ERR\_\* messages' Requester ID will not be.

**Workaround:** Software needs to read downstream devices' error logs to identify the source of the error.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**35. Intel QPI errors can occur on inband resets**

**Problem:** On some inband resets (includes L1 and physical layer resets), the IOH may take 1 or 2 extra cycles to determine that the Intel QPI forwarded clock has stopped.

**Implication:** The IOH may see CRC errors, system hangs, or link level control errors on a small percentage of L1 Entries.

**Workaround:** Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code.

**Status:** For the steppings affected, see the Summary Tables of Changes.



### 36. **Link degrades and surprise link downs (SLD) on PCIe**

**Problem:** PCIe link degrades and surprise link downs seen on PCIe ports 1 and 5 with Rx L0s enabled.

**Implication:** Surprise link downs will be reported as a fatal error, but the link degrades are non-fatal and can be checked in the XP Correctable Error Status register (Device:0-10, Function:0, Offset:200h) bit 0.

**Workaround:** BIOS must disable the devices on the other end of the PCIe link directly connected the IOH PCIe root ports from generating L0s requests. IOH root port Tx L0s and L1 are not affected and should still be enabled. BIOS should implement the following workaround:

1. Find the device directly connected to the IOH root port (the device residing on the secondary buses of the root port). Steps 2 to 4 are applied to this device.
2. Read offset 06h bit[4] to check for Capability List Enable. It must be set to 1 if the device has capability list.
3. Scan through capability list chain, starting from offset 34h, to find the PCIe Capability Identifier with ID of 10h (the offset is referred as PCIe block below).
4. Follow the pseudo code below:

Read Link Capabilities Register (LNKCAP, PCIe offset 0x0c)

if LNKCAP[11:10] = 11b // Support both L0s and L1

write Link Control Reg (PCIe Offset 0x10)

LNKCON[1:0] = 10b // Disable L0s, Enable only L1

else

write Link Control Register (LNKCON, PCIe Offset 0x10)

LNKCON[1:0] = 00b // Disable L0s and L1

5. Apply steps 1 to 4 for each IOH PCIe root port.

6. Modify ACPI FADT table to set byte offset 109 (decimal, IAPC\_BOOT\_ARCH) bit[4] (PCIe ASPM Controls) to 1 to prevent OS from tampering with ASPM values.

**Status:** For the steppings affected, see the Summary Tables of Changes.

### 37. **Intel QPI CRC errors experienced during L0s entry could cause system hangs**

**Problem:** If CRC errors are experienced as the IOH Intel QPI enters the L0s state, the IOH may subsequently hang.

**Implication:** The system may hang if the problem scenario occurs.

**Workaround:** Disable the IOH Intel QPI L0s state. Note that L0s state must be disabled at the CPUs for the CPU-to-CPU Intel QPI link.

*Note: QPI L0s feature is not supported on Intel® 5520 and Intel® 5500 Chipset*

**Status:** For the steppings affected, see the Summary Tables of Changes.



### 38. Intel QPI D2 or B2 error on L1 exit

- Problem:** When exiting Intel QPI L1 state or executing an Intel QPI physical layer reset, a rare race condition can occur between when the clock buffers become stable and initial data is sent from the transmitter. If data arrives before the clock buffers are stable, the training sequence is corrupted.
- Implication:** A D2 or B2 error (Intel® QPI Physical Layer Init Failure) in conjunction with B0 errors is occasionally logged on Intel 5520 and Intel 5500 chipsets after Intel QPI L1 exit or after an Intel QPI physical layer reset. The link will automatically retrain without exposure to the issue. The D2/B2 error and the B0 error can be safely cleared without impact to the system.
- Workaround:** Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code. This changes the severity of a D2 error from uncorrectable to correctable.
- Status:** For the steppings affected, see the Summary Tables of Changes.

### 39. DCA can block progress of other transactions in Dual-IOH, dual socket systems

- Problem:** Direct Cache Access transactions can block progress of other transactions in Dual-IOH, dual socket systems.
- Implication:** The Dual-IOH, dual socket system will hang if the problem scenario occurs.
- Workaround:** On Dual-IOH, dual socket systems, disable DCA on both of the IOH parts. Note that the workaround is not required for Dual-IOH, single socket systems.
- Status:** For the steppings affected, see the Summary Tables of Changes.

### 40. Data Mismatch on Inbound MemWr's after MSI with payload greater than 1 DWORD payload

- Problem:** If for some reason an MSI is incorrectly sent with a payload greater than 1 DWORD then data mismatches may result on subsequent inbound memory writes.
- Implication:** This issue may lead to data corruption.
- Workaround:** When this error occurs, it will be logged at XPUNCERRSTS (Dev 0 - 10:Func 0:Offset 208h) Register bit[8]. The error can be logged at any one of the partner ports. Partner ports are defined as ports in the bifurcation. For example, partner ports for port 3 x16 are ports 3, 4, 5, and 6. Partner ports for port 9 x8 are port 9, and 10. BIOS must make sure that XPUNCERRMSK (Dev 0 - 10:Func 0:Offset 20Ch) Register bit[8] is 0 (default value) for all partner ports to escalate the error. BIOS may need to check XPUNCERRSTS Register bit[8] of all partner ports to identify the error source.
- Status:** For the steppings affected, see the Summary Tables of Changes.

### 41. MSI with greater than 1DWord payload is not logged in XPUNCERRSTS bit 8

- Problem:** When ports are used in a x8 or x16 configuration, bit 8 of the XPUNCERRSTS register (Device:0-10, Function:0, Offset: 0x208) on the port that received the MSI, may not log that an MSI write with a payload greater than 1 DWORD has been received.
- Implication:** The expected indication is not available on the port that received on the MSI.
- Workaround:** When this error occurs, it will be logged at XPUNCERRSTS (Dev 0 - 10:Func 0:Offset 208h) Register bit[8]. The error can be logged at any one of the partner ports. Partner ports are defined as ports in the bifurcation. For example, partner ports for port 3 x16 are ports 3, 4, 5, and 6. Partner ports for port 9 x8 are port 9, and 10. BIOS must make sure that XPUNCERRMSK (Dev 0 - 10:Func 0:Offset 20Ch) Register bit[8] is 0 (default value) for all partner ports to escalate the error. BIOS may need to check XPUNCERRSTS Register bit[8] of all partner ports to identify the error source.
- Status:** For the steppings affected, see the Summary Tables of Changes.



#### 42. **PFP Flag (due to Intel VT-d ISOCH fault) in Fault Status Register not being cleared**

**Problem:** When a Intel VT-d Fault is detected in the ISOCH engine, subsequently setting bit 127 [F fld] of the Fault Record Register (Register:FLTREC[7:0], Addr: MMIO, BAR: VTBAR, Offset:[170h:100h], [1170:1100h]) fails to result in the clearing of the PFP (Primary Fault Pending) status bit in the Fault Status Register (Register: FLTSTS, Addr: MMIO, BAR: VTBAR, Offset:34h, 1034h). In the current design, once the PFP bit is set, it remains set and is not cleared by hardware even after the fault is cleared by software in the Fault Record Register.

**Implication:** Does not impact normal functionality. Only the first fault can be recorded and reported to software. The additional faults will be ignored.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 43. **Error not logged due to a corrupt STP symbol**

**Problem:** Intel 5520 and Intel 5500 Chipsets do not log an error when STP (Start of Transaction Layer) symbol is corrupted. However NAK (Negative Acknowledge) is sent back to alert the sender that this TLP (Transaction Layer Packet) needs to be resent. After a properly framed TLP is resent, the IOH processes the TLP as normal.

**Implication:** Because the error is not logged, the sender cannot determine if the NAK was sent due to the IOH detecting a receiver error or a bad TLP.

**Workaround:** None.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 44. **PCIe ASPM L1 can cause link degrade and speed change.**

**Problem:** PCIe ASPM L1 can cause link degrade and speed change on Intel 5520 and Intel 5500 chipsets root ports when additional power saving on PCIe Rx is enabled.

**Implication:** PCIe2 x16 link can degraded to Gen1 x8 or x2 link.

**Workaround:** Intel 5520 and Intel 5500 chipsets additional power saving on PCIe Rx must be disabled. BIOS must make sure that Dev 13:Func 3 and Dev 14:Func 0 Offset 31Ch.[15] = 0 (power-on default). These bits might have been set as a result of following Chapter 14 of the BIOS Specification Update ("Additional PCIe Power Savings in Power Management States") Note: The *Intel® Xeon® 5500 Platform Design Guide* (document #364117) will be updated to reflect an increase in idle power from 8 W to 8.8 W given the workaround above.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 45. **System may hang in a Dual-IOH Platform with Intel® QuickData Technology enabled DMA traffic, PCIe inbound writes and remote non-posted peer-to-peer reads**

**Problem:** In a Dual-IOH system, on both IOHs - with remote Non-posted Peer-to-Peer reads, PCIe inbound writes to memory and Intel® QuickData Technology enabled DMA traffic, it is likely the traffic may block one of the transactions progress.

**Implication:** A system hang may occur.

**Workaround:** Disable remote Non-posted Peer-to-Peer reads on the Dual-IOH system on one IOH. Alternatively, disable Intel QuickData Technology on one IOH in a multi-IOH system.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### 46. **Lost interrupts when MSI used**

**Problem:** In a system where Message Signaled Interrupt (MSI) is used to forward PCIe interrupts to the processor, and Uncorrectable (UC) Errors are classified as a mix of either fatal or



non-fatal errors, if there are multiple UC Errors occurring in a certain window, a non-related local interrupt waiting to get serviced may get lost.

**Implication:** This may result in an interrupt getting lost before it is serviced. Please note this issue does not affect any interrupts that are delivered via the ERR# pins or the Global Error Reporting Mechanism.

**Workaround:** BIOS must program the IOH root ports (dev 1-10: Func 0, for both IOHs in the DIOH mode) PCIe Uncorrectable Error Severity (offset 10Ch) Register to report all PCIe uncorrectable errors with the same severity (fatal is strongly recommended) Since PCIe AER registers are architectural, additional step is required to prevent OS from overriding the values set by BIOS.

WHEA and ACPI 4.0spec APEI have defined the Hardware Error Source Table (HEST) to describe a system's hardware error sources to the OS. This table can contain multiple error structures. PCIe Root Port AER Structure is designated as type 6. BIOS must implement WHEA/APEI support and set "Uncorrectable Error Severity" field (offset 32 decimal, DWord) to the same value as programmed into the PCIe Uncorrectable Error Severity (offset 10Ch) register by the BIOS thus ensuring that all uncorrectable errors are reported with the same severity.

For non WHEA/APEI capable but PCIe aware Oses, the WHEA/AEPI ACPI objects will be ignored by the workaround above. For certain OS which may change the PCIe AER registers, pls contact the OS vendor for patch in the OS.

Non PCIe aware Oses will not change the PCIe AER registers and are covered by the workaround above.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **47. Intel VT-d: Receiving two identical interrupt requests in back to back cycles may corrupt attributes of remapped interrupt, or hang subsequent interrupt-remap-cache invalidation command.**

**Problem:** If Intel VT-d interrupt-remapping hardware receives two identical back to back interrupt requests the attributes of remapped interrupt returned may be corrupt. This interrupt sequence may hang the system if the software executes a subsequent interrupt-remap-cache invalidation command.

**Implication:** This scenario may lead to unpredictable external interrupt behavior or a system hang.

**Workaround:** BIOS to set bit 25 in (Dev/Func: 14/4, Offset: 168h). De-feature interrupt remapping capability by clearing Interrupt-Remap (IR) RWO field (bit 3) in Extended Capability Register and clearing INTR\_REMAP bit (bit 0) in Flags field of ACPI DMAR tables used to report VT-d hardware capability to software.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **48. Intel QPI Error Status D3 is observed**

**Problem:** Intel QPI D3 error status (Dev 20:Func 2:Offset 280h/200h.[11]) is observed due to false sub-state illegal link reset detection

**Implication:** This spurious error status may cause false error reporting if enabled in system debug scenarios

**Workaround:** BIOS should program the following IOH registers to mask off the spurious sub-state status and allow the Intel QPI D3 error status to function normally: Set Dev 13:Func 0/1: Offset F8Ch (Dword) bit[14] to 1.**Note: this workaround will be implemented in future release of Intel QPI RC.**

**Status:** For the steppings affected, see the Summary Tables of Changes.

**49. Unpredictable PCI behavior accessing non-existent memory space**

**Problem:** Locked instructions whose memory reference is split across cache line boundaries and are aborted on PCI behind ICH may cause subsequent PCI writes to be noticeably unpredictable.

**Implication:** Aborted split lock accesses to non-existent PCI memory space behind ICH may cause PCI devices to become inoperable until a platform reset. Intel has not observed this issue with commercially available software and has only observed this in a synthetic test environment.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Tables of Changes.

**50. Bandwidth changed status errors being escalated to Global RAS**

**Problem:** It has been observed that setting the bit 0 in the XPCORERRMSK register (Device: 0-10, Function: 0, Offset: 204h) only prevents bandwidth changed status errors from being escalated to Global error registers for Ports 0, 3 and 7.

**Implication:** Setting the bit 0 in the XPCORERRMSK register (Device: 0-10, Function: 0, Offset: 204h) does not prevent the bandwidth changed status errors from being escalated to Global error registers for Ports 1,2,4,5,6,8,9 and 10.

**Workaround:** The bandwidth status changed error is a correctable error and under normal operation is permitted to be escalated to the global error register.

**Status:** For the steppings affected, see the Summary Tables of Changes.

**51. Intel® Trusted Execution Technology (Intel® TXT) writes may not complete as expected**

**Problem:** When Intel® Trusted Execution Technology (Intel® TXT) is enabled, in a system with back to back Configuration Retry transactions and certain Intel TXT write requests, the Intel TXT write request may not complete as expected.

**Implication:** This may result in incorrect system behavior. *Please note that Intel TXT is not supported on S-Spec SLGMU/SLGMT.*

**Workaround:** None Identified. *BIOS must not enable Intel TXT in processors via IA32\_FEATURE\_CONTROL MSR for S-Spec SLGMU/SLGMT. See the Chapter 6 of Intel® 64 and IA-32 Architectures Software Developer's Manual Volume 2B for details regarding this MSR.*

**Status:** For the steppings affected, see the Summary Tables of Changes.

**52. Intel VT-d: Address Remapping error when DMA/interrupt remapping is active**

**Problem:** With Intel VT-d enabled, when software updates the root table pointer or interrupt remapping table pointer while DMA/interrupt remapping is active, it is possible that the address used to access the page-table structure for DMA requests or interrupt remapping could be corrupted and cause an address remapping error.

**Implication:** Software cannot update the root table pointer or interrupt remapping table pointer while DMA/interrupt remapping is active.

**Workaround:** None Identified

**Status:** For the steppings affected, see the Summary Tables of Changes.

**53. Intel VT-d: In-flight remap-able interrupts not drained on interrupt invalidation command**

**Problem:** The Intel VT-d hardware is not draining in-flight remapped interrupts when an interrupt invalidation command occurs. The software expectation is that when it performs an invalidation, all in-flight interrupts are received and acknowledged by the processor.



**Implication:** As a result of this issue, software may lose interrupts or receive spurious interrupts.

**Workaround:** BIOS to set bit 25 in (Dev/Func: 14/4, Offset: 168h). Disable interrupt remapping capability by clearing Interrupt-Remap (IR) RWO field (bit 3) in Extended Capability Register and clearing INTR\_REMAP bit (bit 0) in Flags field of ACPI DMAR tables used to report VT-d hardware capability to software.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **54. Source ID for errors internally detected by PCIe root port 3 (Ports 3-6) is not logged as expected**

**Problem:** The Source ID for errors internally detected by PCIe root port 3 (ports 3-6) is not logged in the ERRSID register (Device: 0-10, Function: 0, Offset: 134h) as expected. However the error message details are logged correctly as specified by the PCIe Specification in the root port 3 RPERRSTS register (Device 0-10, Function: 0, Offset 130h) as expected. The root port 3 ERRSID register does not log the internally detected errors as expected.

**Implication:** The root port 3 (ports 3-6) ERRSID register does not log the internally detected errors as expected.

**Workaround:** Other methods to determine if root port 3 (ports 3-6) detected an error: Reading the UNCERRSTS (Device: 0-10 Function: 0 Offset: 104h) and CORERRSTS (Device: 0-10 Function: 0 Offset: 110h) registers will allow software to determine if root port 3 internally detected an error as well as the type of error.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **55. Device 0, Function 0's Revision ID (RID) is not reset to the Stepping Revision ID (SRID) by a CORERST\_N Assertion**

**Problem:** If the RID is set by BIOS to report the Compatible Revision ID (CRID), the RID reported by Device 0, Function 0 is not restored back to the SRID by a "hard reset" (CORERST\_N assertion) without a "power-good" reset (COREPWRGOOD not de-asserted). Furthermore, re-writing the "key" to set the CRID in all devices is ignored (Device 1 through Device 22 do not change to the CRID).

**Implication:** Platforms that do not use the CRID feature are not affected. For platforms that utilize the CRID functionality, BIOS will enable the CRID before handing off to boot the Operating System. When the Operating System is running, and if a hard reset without cycling is issued (I/O port CF9h write = '6', or other source), the ID on Device 0: Function 0, does not change back to the SRID. Device 1 through Device 22 will report the SRID in the RID field as expected. Since the CRID is locked, a subsequent write of the "key" to set the CRID is ignored. Upon rebooting, the Operating System will see a change in the RID (from the CRID to the SRID) for Device 1 through Device 22. Operating System images that expect the RID to report the CRID may re-enumerate the PCI bus, may display informational messages (new hardware found) and reload the associated driver. An extended boot time (varies by OS) and a reboot request may be generated (requiring user intervention) in order to re-configure the OS for the new RID value.

**Workaround:** For platforms that support the CRID functionality, BIOS can issue an IOH-only powergood reset when a hard reset without cycling is detected. A power-good reset to the IOH will reset the RID to the SRID in all devices and allow the BIOS to re-enable the CRID functionality. Please note that a power-good reset will cause all "sticky" registers in the IOH to be reset to their default values. Logic supporting an IOH-only power-good.reset is documented in the "Intel® Xeon® 5500 Platform Design Guide (PDG)" Section 5.2.7: IOH CSIFREQSEL[1:0] Guidelines. Note: Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code. For more BIO implementation detail refer to the IOH BIOS Spec Update. Depending on the BIOS and platform implementation the following should be considered:



- a. If the BIOS supports error logging, then before the BIOS issues the IOH-only power-good reset it should log any errors present in the IOH (these will be cleared by the IOH-only power-good reset). Error logging is BIOS and platform implementation dependant. If the platform implements runtime error logging, the probability of a pending error is low and this step may be ignored.
- b. If the BIOS uses any "sticky" registers for storage of proprietary values, the values will be cleared by the IOH-only power-good reset. An alternative nonvolatile storage location should be considered.
- c. The BIOS should write the CRID "key" to Device 0, Function 0 to restore the desired CRID functionality before handing off to the Operating System.
- d. The additional resets and BIOS execution may add additional time to the normal hard reset execution time.
- e. Effects on platform manageability (which is platform implementation dependant) should also be evaluated. It is not expected that this workaround will affect manageability functionality.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **56. Header Log information may not be captured correctly when accessed via JTAG**

**Problem:** Header Log information in the IOHNFERRHD (Device: 20, Function: 2, Offset: 324h) and IOHFFERRHD (Device: 20, Function: 2, Offset: 30Ch) registers for the first nonfatal error and first fatal error respectively may get corrupted if the original request is received from the JTAG port.

**Implication:** Header Log information may not be accurate if the IOH aborts the packet for a request from JTAG port. The actual errors themselves will be logged accurately in the relevant error status registers.

**Workaround:** None Identified

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **57. Forwarded Clock Lane Detection status may not be indicated accurately**

**Implication:** Bit 31 of the QPI[1:0]PH\_TDS register (Device: 13, Function: 1-0, Offset: 834h) may not indicate the status of the forwarded clock accurately. The status of the forwarded clock may not be accurately reflected by the QPI[1:0]PH\_TDS register due to this issue.

**Workaround:** Bit 24 of the following register (Device: 13, Function: 1-0, Offset: 850h) provides accurate indication on the status of the forwarded clock. A value of "1" for Bit 24 indicates that the forwarded clock has been detected.

**Status:** For the steppings affected, see the Summary Tables of Changes.

#### **58. ESI link cannot go to L1 state on the Intel® 5500 Platform**

**Problem:** ESI link cannot go to L1 state on the Intel® 5500 Platform, which prevents the Coarse-Grain Dynamic Clock Gating (CGCG) from being engaged.

**Implication:** Since the Intel 5520 Chipset and Intel 5500 chipsets cannot engage CGCG, all platform power that has been measured has been with CGCG disengaged, regardless of whether the CGCG feature is enabled/disabled in BIOS. There is no change in the measured Power Consumption on the Intel 5500 Platform as a result of this Erratum. However, the specified Idle Power for the Intel 5520 chipset and Intel 5500 chipset has increased. Please refer to the *Intel® 5520 and Intel® 5500 Chipsets Thermal/Mechanical Design Guidelines* for the updated Idle Power Numbers on the Intel 5520 chipset and Intel 5500 chipset.

**Workaround:** None Identified.

**Status:** For the steppings affected, see the Summary Tables of Changes.



### 59. **Deadlock can occur in Dual IOH Platforms**

**Problem:** Dual IOH Platform with PCIe to Remote Memory Write Ordering Conflict May Cause deadlock to Occur

**Implication:** A deadlock could occur

**Workaround:** To avoid deadlock without enabling B3 deadlock breaker, an alternate method was devised: "csipooldfx0[31:24] = (n-2)/2" should be set, where n is number of RTID allocated on QPI link connected to CPU. This workaround reduces impact to remote upstream write performance. Note: Please see your Intel representative for the latest Intel® QuickPath Interconnect Initialization Reference Code (RC1.05 or ≥1.85 are both OK to use)

**Status:** For the steppings affected, see the Summary Tables of Changes.

### 60. **In-flight DMA requests received during the implicit DMA draining window when enabling Intel VT-d hardware may result in a spurious DMA fault**

**Problem:** In-flight DMA requests, during the 2 cycle window for DMA draining when enabling Intel VT-d hardware (GCMD\_REG.TE = 1), may result in a spurious DMA fault.

**Implication:** BIOS features, such as legacy keyboard emulation, can result in in-flight DMA requests (such as from a USB controller) when the OS/VMM is booting and enabling Intel VT-d. If such DMA requests happen to arrive during the 2 cycle implicit DMA draining window when enabling Intel VT-d, they may result in a DMA fault irrespective of the programming of Intel VT-d translation structures, which could result in a system hang.

**Workaround:** BIOS should set the Disable Address Drain bit (Bit 25) at Device Eh, Function 4h, Offset 168h. Software that depends on implicit address draining when Intel VT-d is enabled should perform explicit IOTLB invalidation after enabling Intel VT-d to drain in-flight DMA requests.

**Status:** For the steppings affected, see the [Summary Tables of Changes](#).

### 61. **The ESI upstream link between the ICH10 and IOH may intermittently fail to train on some IOH devices**

**Problem:** On the upstream link between the ICH10 and the IOH, by default, the ICH10 uses a "detect" mechanism to determine if the IOH is present on the ESI link. On some IOH components the ESI termination resistors are too weak to support a consistent detect process by the ICH10. This causes the IOH to intermittently not be detected by the ICH10. The ICH detect failure results in the failure of the IOH ESI training and the system will not boot. Cold temperatures can aggravate this issue. The IOH may be detected and the system will boot on a subsequent RESET or Power Cycle.

**Implication:** Due to this issue the system may occasionally not boot. If the system does boot, the system will operate normally until the next RESET or Power Cycle.

**Workaround:** The following workarounds are available:

- If a SPI flash is present in the design, and the ME is utilized, version SPS\_01.01.02.007.0 (or later) of the ME Firmware will workaround this erratum.
- If a SPI flash is present in the design, but ME Firmware is not utilized, then the SPI Flash should be programmed to set FISBA + Offset 000h: ICHSTRP0 - ICH Softstrap 0 Bits [2:1] to "01". This selects the DMI Force Detect[1:0] to "Force DMI to x4 Link".
- If a SPI flash is not present in the design, weak 100K pulldown terminators can be placed on the IOH ESIR{P}[3:0] inputs.

**Status:** For the steppings affected, see the Summary Table of Changes.



## 62. A Modification to the Multiple Message Enable field does not affect the AER Interrupt Message Number field

**Problem:** The Advanced Error Interrupt Message Number field (RPERRSTS Devices 0-3; Functions 0-3; Offset 178H; bits[31:27]) should be updated when the number of messages allocated to the root port is changed via modifying the Multiple Message Enable field (MSIMSGCTL Device 3; Function 0; Offset 62H; bits[6:4]). However, when the number of messages allocated to the root port in the Multiple Message Enable field is modified, a subsequent change in the interrupt number assigned for AER (Advanced Error Reporting) errors in the Advanced Error Interrupt Message Number field is not observed.

**Implication:** Software can allocate only one MSI (Message Signaled Interrupt) to the root port.

**Workaround:** None

**Status:** For the steppings affected, see the Summary Table of Changes.

## 63. PCIe\* Hot-Plug Slot Status Register May Not Indicate Command Completed

**Problem:** The PCIe Base Specification requires a write to the Slot Control register (Offset A8H) to generate a hot-plug command when the downstream port is hot-plug capable. Due to this erratum, a hot-plug command is generated only when one or more of the Slot Control register bits [11:6] are changed.

**Implication:** Writes to the Slot Control register that leave bits [11:6] unchanged will not generate a hot-plug command and will therefore not generate a command completed event. Software that expects a command completed event may not behave as expected.

**Workaround:** It is possible for software to implement a one-second timeout in lieu of receiving a command completed event

**Status:** For the steppings affected, see the Summary Table of Changes.

## 64. Intel® VT-d Memory Check Error on an Intel® QuickData Technology Channel May Cause All Other Channels to Master Abort

**Problem:** An Intel QuickData Technology DMA access to Intel® VT-d protected memory that results in a protected memory check error may cause master abort completions on all other Intel QuickData Technology DMA channels.

**Implication:** Due to this erratum, an error during Intel QuickData Technology DMA access to an Intel® VT-d protected memory address may cause a master abort on other Intel QuickData Technology DMA channels.

**Workaround:** None identified.

**Status:** For the steppings affected, see the Summary Table of Changes.

## 65. PCIe\* Header of a Malformed TLP is Logged Incorrectly

**Problem:** If a PCIe port receives a malformed TLP (Transaction Layer Packet), an error is logged in the UNCERRSTS register (Device 0; Function 0; Offset 14CH and Device 2-3; Function 0-3; Offset 14CH). Due to this erratum, the header of the malformed TLP is logged incorrectly in the HDRLOG register (Device 0; Function 0; Offset 164H and Device 2-3;

**Implication:** The PCIe header of a malformed TLP is not logged correctly.

**Workaround:** None identified.

**Status:** For the steppings affected, see the Summary Table of Changes.



# Technology Technology Technology Technology Specification Changes

---

The Specification Changes listed in this section apply to the following documents:

1. *Intel® 5520 and Intel® 5500 Chipset Datasheet*, Revision -001 (Document Number 321328)

All Specification Changes will be incorporated into a future version of the appropriate Intel 5520 and Intel 5500 Chipsets documentation.

***There are no Specification Clarifications for this revision.***

1.



# Specification Clarifications

The Specification Clarifications listed in this section apply to the following documents:

1. *Intel® 5520 and Intel® 5500 Chipsets Datasheet, Revision -001* (Document Number 321328)

All Specification Changes will be incorporated into a future version of the appropriate Intel 5520 and Intel 5500 Chipsets documentation.

## 1. Updates to CTOCTRL, DEVCON2, DEVCAP2

### 19.13.8 CTOCTRL: Completion Time-Out Control Register

Register: CTOCTRL Device: 0-10 Function: 0 Offset: 1E0h			
Bit	Attr	Default	Description
31:10	RV	0	Reserved
9:8	RW	00	XP-to-PCIe time-out select within 2s to 6.5s range: When OS selects a time-out range of 2s to 6.5s for XP (that affect NP tx issued to the PCIe/DMI) using the root port's DEVCTRL2 register, this field selects the sub-range within that larger range, for additional controllability. 00: 2s 01: 4s 10: 6.5s 11: Reserved Note: These values can deviate +/- 10%
7:6	RV	00	Reserved
5	RO	0	Reserved
4:0	<b>RV</b>	0	Reserved

### 19.4.3.17 DEVCON2: PCI Express\* Device Control 2 Register

The PCI Express\* Device Control register controls PCI Express specific capabilities parameters associated with the device.

Device: 20 Function: 0-2 Offset: 68h			
Bit	Attr	Default	Description
15:6	RO	0h	Reserved
5	RO	0	<b>Alternative RID Interpretation (ARI) Enable</b> When set to 1b, ARI is enabled in Root Port.
4	RO	0	<b>Completion Timeout Disable</b> When set to 1b, this bit disables the Completion Timeout Mechanism for all NP tx that IOH issues on the PCIe/DMI link and in the case of CBDMA, for all NP tx that DMA issues upstream. When set to 0b, completion timeout is enabled. Software can change this field while there is active traffic in the root port.



<b>Device:</b> 20 <b>Function:</b> 0-2 <b>Offset:</b> 68h			
Bit	Attr	Default	Description
3:0	RO	0000b	<p><b>Completion Timeout Value on NP Tx that IOH Issues on PCIe/DMI:</b>  <b>In devices that support Completion Timeout Programmability, this field allows system software to modify the Completion Timeout range. The following encodings and corresponding timeout ranges are defined:</b></p> <p>0000b: 2ms          0001b: Reserved (IOH aliases to 0000b)          0010b: Reserved (IOH aliases to 0000b)          0101b: 4ms          0110b: 10ms          1001b: 40ms          1010b: 210ms          1101b: 800ms          1110b: 2s to 6.5s</p> <p>Note: These values can deviate +/-10%</p> <p>When the OS selects the 2s - 6.5s range, CTOCTRL further controls the timeout value within that range. For all other ranges selected by OS, the timeout value within that range is fixed in the IOH hardware.</p> <p>Software can change this field while there is active traffic in the root port.</p> <p>This value is also used to control PME_TO_ACK timeout. This field sets the timeout value for receiving the PME_TO_ACK message after a PME_TURN_OFF message has been transmitted. The PME_TO_ACK timeout has meaning only if Bit 6 of MISCCTRLSTS is set to 1b.</p>

### 19.4.3.16 DEVCAP2: PCI Express Device Capabilities Register 2

<b>Register:</b> DEVCAP2 <b>Device:</b> 0-10 <b>Function:</b> 0 <b>Offset:</b> B4h			
Bit	Attr	Default	Description
31:6	RO	0h	Reserved
5	RO	1	Alternative RID Interpretation (ARI) Capable: This bit is set to 1b indicating Root Port supports this capability
4	RO	1	Completion Timeout Disable Supported: IOH supports disabling Completion Timeout



<b>Register:</b> DEVCAP2 <b>Device:</b> 0-10 <b>Function:</b> 0 <b>Offset:</b> B4h			
Bit	Attr	Default	Description
3:0	RO	1110b	<p><b>Completion Time-Out Values Supported</b> – This field indicates device support for the optional Completion Time-Out programmability mechanism. This mechanism allows system software to modify the Completion Time-Out range. Bits are one-hot encoded and set according to the table below to show time-out value ranges supported. A device that supports the optional capability of Completion Time-Out Programmability must set at least two bits.</p> <p>Four time value ranges are supported:                      Range A: 50 us to 10ms                      Range B: 10ms to 250ms                      Range C: 250ms to 4 s                      Range D: 4s to 64s</p> <p>Bits are set according to the table below to show timeout values supported.                      0000b: Completion Timeout programming not supported. Value is fixed by implementation in the range 50us to 50 ms                      0001b: Range A                      0010b: Range B                      0011b: Range A &amp; Range B                      0110b: Range B &amp; Range C                      0111b: Range A,B &amp; C                      1111b: Range A,B,C &amp; D                      All other values are Reserved</p> <p>IOH supports time-out values up to 2.5ms to 6s. The values specified above are required by PCIe 2.1 Specification.</p>

## 2. Updates to SYRE

### 19.6.9.17 SYRE: System Reset

<b>Register:</b> SYRE <b>Device:</b> 20 <b>Function:</b> 2 <b>Offset:</b> 0CCh			
Bit	Attr	Default	Description
31:15	RV	0	Reserved
14	RV	0	<p><b>S5</b>                      1 - Translate ESI.GO_S3 to QPI.SpcPMReq (S5)                      0 - Forward ESI.GO_S3 to QPI.SpcPMReq(S3)</p>
13:12	RWSLB	0	Reserved
11	RW	0	<p><b>RSTMSK</b>                      0 - the IOH will perform the appropriate internal handshakes on RST_N signal transitions to progress through the hard reset.                      1 - IOH ignores RST_N, unaffected by the RST_N assertion</p>
10	RW	0	<p><b>CPURESET</b>                      1 - IOH asserts RESETO_N                      The IOH clears this bit when the CPURESET timer elapses.</p>
9:1	RV	0	Reserved



<b>Register:</b> SYRE <b>Device:</b> 20 <b>Function:</b> 2 <b>Offset:</b> 0CCh			
Bit	Attr	Default	Description
0	RWS	1	<b>Enable CPU BIST</b> 1- Enable CPU BIST 0- Disable CPU BIST This bit controls whether or not BIST is run in the CPU on reset. It's value will correspond to the BIST value in the POC exchanged from IOH on Intel QuickPath Interconnect. This value will only make a difference in CPU's that observe POC (like Xeon). By default BIST is disabled. If BIST is desired, then after this bit is set the CPU must be reset to cause the CPU to capture the new value.

### 3. Non-Posted Peer-to-Peer Support between IOH and ICH.

This Specification Clarification applies to the following document:

*Intel® 5520 and Intel® 5500 Chipset Datasheet*

- Product Feature Section -  
 Currently published: Full peer-to-peer support between PCI Express interfaces.  
 Clarification: IOH to ICH Non-Posted peer-to-peer transactions are not supported.

### 4. Correction to the CRID Reset Functionality

#### 19.3.3 Conceptual Description:

Following a power-on reset (COREPWRGOOD de-asserted) or hard reset (CORERST\_N asserted), the SRID value may be read from the RID register at offset 08h of all devices and functions in the IOH chipset, which reflects the actual product stepping. To select the CRID value, BIOS/configuration software writes a 32-bit key value of 00000069h to Bus 0, Device 0, Function 0 (ESI port) of the IOH's RID register at offset 08h. Through a comparator, the written value is matched with a key value of "00000069h". The comparator output is flopped and controls the selection of either CRID or RID. Subsequent reads to RID register at offset 08h will return CRID if the comparator flop is set. Otherwise it will always return the SRID when the comparator flop is reset. The internal RID comparator flop in the ESI port (Bus 0 device 0 Function 0) is a "write-once" register and gets locked after the first write to offset 08h.

### 5. Updates to Resource conflicts on IOH PCIe

#### 5.10.3 System Resource Conflicts

Resource conflicts are generally not a problem in the Intel QuickPath Interconnect because of the independent nature of the message classes. Two cases are explicitly stated here for how resources are managed to prevent resource problems. The first basic Intel QuickPath Interconnect rule is that completions are absorbed at the source, unconditional on any other message classes. This generally requires preallocation of completion resources before a request is sent. See Section 5.11.1 for more details on message class ordering details. The resource sharing creates input transaction dependency on output resources availability (IN/OUT Dependency). This dependency is allowed by the PCIe spec only for Root Complex (which IOH is). The IN/OUT dependency applies to all Peer-to-Peer inbound transactions via all the IOH PCIe ports (PE[0-10]). The IOH PCIe implementation groups the resources in two pools:



pool0 : IOU0(PE3-PE6) and IOU2(DMI, PE1-PE2)

pool1 : IOU1(PE7-PE10), ME, and MISC internal agents, (for example, IOAPIC, Hot-Plug, JTAG, Intel VT-d table walk master, broadcast msi, vlv, power management, messages from ICH that need to be broadcast to the CPUs, messages from CPU such as EOI and GPE).

When inbound Peer-to-Peer transaction reaches the top of the queue it can make forward progress only if all the outbound buffers of pool1 have free space (i.e. IOU1, ME, MISC)

### 5.10.3.1 Peer-to-Peer Legacy Interrupts- INTx

All inbound INTx transactions from any PCIe port (PE0-10) target the MISC unit in pool1 regardless of final destination and are subject to the IN/OUT dependency on free outbound buffers in Pool1. A stalled INTx at the inbound buffer (due to lack of free outbound buffer) will block all subsequent inbound posted transactions until space is freed in the outbound buffers.

Normally the concern will be with IOU1 (PE7-PE10) outbound buffer getting full due to heavy traffic load such as in DMA operations. The outbound free buffer requirement is regardless of the IOU1 ports configuration. For example if IOU1 is configured as 2X8, Inbound Peer-to-Peer transaction on one port will be stalled by a full outbound buffer of the other port.

When the IOH PCIe link is connected to an End-Device (directly or via bridge or a switch) the delay is temporary and removed when an outbound transaction is processed. (Deadlock may occur if the End-Device or Switch is none compliant and implemented with an IN/OUT dependency).

However, in system topologies that connect two IOH RC ports via a NTB (such as failover and mirroring applications common in storage systems) system designers need to pay special attention to this IN/OUT dependency at both ends of the link. Under certain heavy transaction loads and sufficiently high density of inbound INTx transaction- the blockage can develop at both systems leading to a deadlocked link. In these situations it is recommended to replace the INTx transactions with none broadcasting MSI transactions. The development of blockages that lead to deadlock is a dynamic process that is dependent on the traffic loads and resources of the entire link. Deadlock can develop from a steady-state heavy traffic loads or from momentary spikes in traffic load and Peer-to-Peer transactions.

### 5.10.3.2 Peer-to-Peer Request Redirect - ACS

When Multiple Root complex ports are connected via NTB, and ACS redirection is implemented in any flow control class (P, NP, C) a deadlock can also occur due to the IN/OUT dependencies in the link. System designers need to consider these dependencies in order to avoid deadlocks.



# Documentation Changes

---

The Documentation Changes listed in this section apply to the following documents:

1. *Intel® 5520 and Intel® 5500 Chipset Datasheet*, Revision -001 (Document Number 321328)

All Specification Changes will be incorporated into a future version of the appropriate Intel 5520 and Intel 5500 Chipsets documentation.

1. **Intel QPI L0s feature is not supported on the Intel® 5520 and Intel® 5500 Chipset.**

§