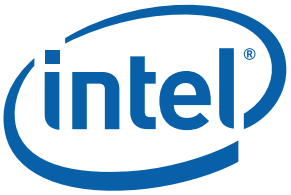


Intel® 7500 Chipset

Specification Update

October 2011

Notice: The Intel® 7500 Chipset 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.

Reference Number:323168-009

Contents

Revision History

Revision Number	Description	Date
001	• Initial release of the document	March 2010
002	• Updated Erratum 3 and added spec change 1 and spec clarification 2	April 2010
003	• Added Erratum 58	May 2010
004	• Added Spec Clarification 3	June 2010
005	• Added new errata 59 and 60, added specification change 2	September 2010
006	• Added specification clarifications 4 and 5. Added specification changes 2	November 2010
007	• Added specification clarifications 6 through 9	April 2011
008	• Added specification clarification 10 • Added erratum 61	September 2011
009	• Added specification clarification 11	October 2011

Preface

This document is an update to the specifications contained in the Affected Documents table below. This document is a compilation of device errata and documentation changes, 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 may also contain information that was not previously published.

Affected Documents

Document Title	Document Number/ Location
Intel® 7500 Chipset Datasheet	322827

Nomenclature

Errata are design defects or errors. These may cause the Intel® 7500 Chipset's 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.

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.

S-Spec Number is a five-digit code used to identify products. Products are differentiated by their unique characteristics, for example, core speed, L2 cache size, package type, etc. as described in the processor identification information table. Read all notes associated with each S-Spec number.

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).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 Table of Change

The following tables indicate the errata, specification changes, specification clarifications, or documentation changes which apply to the Intel 7500 chipset product. 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 use the following notations:

Codes Used in Summary Table

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 the 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.
Non-Si	This issue is not due to a erratum in the Intel 7500 chipset.

Row

|	Change bar on outside margins indicates this erratum is either new or new or modified from the previous version of the document.

Errata Summary (Sheet 1 of 3)

Number	Stepping	Segment		Status	Erratum
	B2	EX	MC		
1	X	X	X	No Fix	CPURST bit does not get cleared by hardware
2	X	X	X	No Fix	VTX-RCV-DETECT pulse too large during receiver detection
3	X	X	X	No Fix	PCIe* Surprise Link Down status may not be flagged during training failure or surprise removal
4	X	X	X	No Fix	PCIe Gen2 differential peak-to-peak transmit voltage swing is too low
5	X	X	X	No Fix	Extended Error Detect Mask Registers of all PCIe root ports mask error logging by default
6	X	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
7	X	X	X	No Fix	PCIe Header of a malformed TLP is not logged
8	X	X	X	No Fix	Intel VT-d does not support the draining of compatibility-format interrupts
9	X	X	X	No Fix	Hardware applies HPA_LIMIT to upstream memory request when Intel VT-d is disabled
10	X	X	X	No Fix	PCIe PMCSR Power State fields allow writing D1 and D2
11	X	X	X	No Fix	PCIe Gen2 Tx Return loss fails spec
12	X	X	X	No Fix	Persistent Jtag error reported at MIERRST register
13	X	X	X	No Fix	Interoperability issue of some PCIe Gen1 cards with Gen2 Devices
14		X	X	Non-Si	System reset is getting turned into a power cycle
15	X	X	X	No Fix	Intel VT-d: Memory read request with AT=11b results in malformed TLP
16	X	X	X	Doc	Memory writes to a certain address range are considered advisory non-fatal
17	X	X	X	No Fix	Transactions to addresses above TOCM are not setting the Master Abort
18	X	X	X	No Fix	Bandwidth very low for write traffic with noSnoop attribute set
19	X	X	X	No Fix	Modifying RTIDs causes system hangs
20	X	X	X	No Fix	System Hang due to VLW
21	X	X	X	No Fix	Intel® QuickPath Interconnect (Intel® QPI) Queue/Table overflow or underflow error observed
22	X	X	X	No Fix	Intel VT-d translated write transactions are blocked but not recorded
23	X	X	X	No Fix	Intel QuickPath Interconnect initialization abort failures logged during power-on resets
24	X	X	X	No Fix	ERRSID not logging ReqID for Inbound ERR_* messages
25	X	X	X	No Fix	Subsystem Vendor ID (SVID) and the Subsystem Device ID (SDID) of device 15 are not implemented
26	X	X	X	No Fix	Intel QuickPath Interconnect links do not link and train to the full width
27	X	X	X	No Fix	Link training failure due to multiple resets
28	X	X	X	No Fix	Setting bits 24 and 25 of the MISCCTRLSTS does not result in expected behavior
29	X	X	X	No Fix	Some PCIe inbound messages not ignored as expected

Errata Summary (Sheet 2 of 3)

Number	Stepping	Segment		Status	Erratum
	B2	EX	MC		
30	X	X	X	No Fix	Timeout values much larger than specified
31	X		X	No Fix	Spurious D8 error may be reported
32	X	X	X	No Fix	Failure during operation at PCI Express L1 power management state
33	X	X	X	No Fix	ACS Violation is not treated as Advisory when severity is set to Non-Fatal
34	X	X	X	No Fix	Snoop requests may not be possible when addressing above 2^41
35	X	X	X	No Fix	D5 and D7 errors will not have the Intel QPI headers logged
36	X	X	X	No Fix	Locked read time out bit not being set
37	X	X	X	No Fix	Reply Timer Timeout Errors logged when enabling ASPM L0s
38			X	No Fix	Link and PHY Layer Reset Can Cause CRC Errors With 16 Bit CRC
39	X	X	X	No Fix	Data Mismatch on Inbound MemWrs after MSI with payload greater than 1 DWORD payload
40	X	X	X	No Fix	MSI with greater than 1DWord payload is not logged in XPUNCERRSTS bit 8 as expected
41	X	X	X	No Fix	PCIe link degrades and surprise link down (SLD) on PCIe
42	X	X	X	No Fix	Error not logged due to a corrupt STP symbol
43	X	X	X	No Fix	Intel QuickPath Interconnect errors can occur on inband resets
44	X	X	X	No Fix	System may hang in a multi-IOH Platform with Intel® QuickData Technology enabled DMA traffic, PCIe inbound writes and remote non-posted peer-to-peer reads
45	X	X	X	No Fix	Potential Link Width degradation on Intel QuickPath Interconnect
46	X	X	X	No Fix	Intel QuickPath Interconnect Error Status D3 is observed
47	X	X		No Fix	Unpredictable PCI behavior accessing non-existent memory space
48	X	X	X	No Fix	Bandwidth changed status errors being escalated to Global RAS
49	X	X	X	No Fix	Intel® Trusted Execution Technology (Intel® TXT) writes may not complete as expected
50	X	X	X	No Fix	Intel VT-d: Address Remapping error when DMA/interrupt remapping is active
51	X	X	X	No Fix	Source ID for errors internally detected by PCIe root port 3 (ports 3-6) is not logged as expected
52	X	X	X	No Fix	Header Log information may not be captured correctly when accessed via JTAG
53	X	X	X	No Fix	Forwarded Clock Lane Detection status may not be indicated accurately
54	X	X	X	No Fix	ESI link cannot go into L1 state on the Intel® Itanium® Processor 9300 Series-Based Platform and Intel® Xeon® Processor 7500 Series-based Platform
55	X	X		No Fix	Intel QuickPath Interconnect interleave selection modes 0x5 and 0x6 do not allow NodeID port allocation for either port
56	X	X	X	No Fix	PCIe ASPM L1 can cause link degradation and speed change
57	X	X	X	No Fix	Links may not be able to recover after entering Live Error Recovery (LER) mode
58	X	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

Errata Summary (Sheet 3 of 3)

Number	Stepping	Segment		Status	Erratum
	B2	EX	MC		
59	X	X	X	No Fix	PCIe Slot Status Register Command Completed bit not always updated on any configuration write to the Slot Control Register
60	X	X	X	No Fix	Level-triggered multicast device interrupts may cause an MCA timeout
61	X	X	X	No Fix	The ESI upstream link between the ICH10 and IOH may intermittently fail to train on some IOH devices

Specification Changes

No.	Specification Changes
1	Intel QuickPath Interconnect Protocol Control Completion Retry Timeout Interval specification change
2	Intel QuickPath Interconnect Protocol APIC Source Address Decode Extended Logical Mode Interleave functionality specification change

Specification Clarifications

No.	Specification Clarifications
1	Memory Address Decode clarification
2	Clarification to the QPIPMADDATA Register, interleave select bit field [3:1]
3	Protocol status clarification to the Intel QPI Protocol Status Register Fields
4	Intel QuickPath Interconnect Protocol Power Control clarification to the State Level description
5	Intel QuickPath Interconnect Device Control register, clarification of the extended tag field functionality
6	Intel QuickPath Interconnect Interrupt Control (QPIPINTRC) Legacy Signal Mask behavior
7	PCI Express Link Status Register clarification (bit 12, RWO instead of RO)
8	IOH Hardware Support for Intel QPI Hot-Plug scratch pad and semaphore register count (17 instead of 16 registers)
9	Section 20.6.6.1 SR[0:3]: Scratch Pad Register 0-3 (Sticky) incorrectly shows SR[4:7] where it should be SR[0:3]
10	Table 8-6 MSI Address Format when Remapping is Enabled
11	Section 20.11.3.9 QPI[1:0]EP_SR: Electrical Parameter Select Register

Documentation Changes

No.	Documentation Changes
1	None

Identification Information

Component Identification via Programming Interface

The Intel 7500 chipset IOH stepping can be identified by the following register contents:

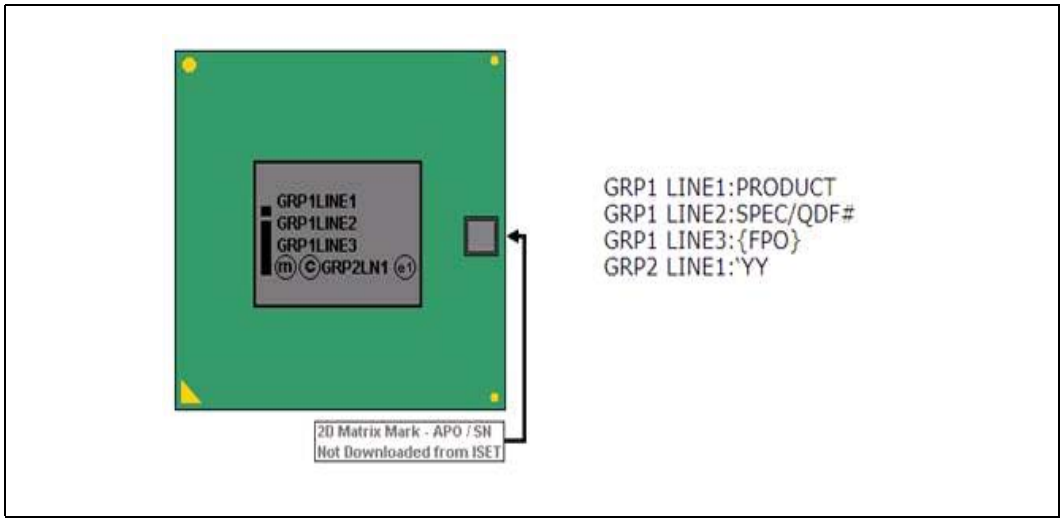
Stepping	Features	Vendor ID	Device ID	Revision Number
B2	Production	8086h	0011_0100_0000_0111b	0x22

Component Marking Information

the Intel 7500 chipset IOH stepping can be identified by the following component markings:

Stepping	⁄S-spec	Top Marking	Notes
B2	SLH3K	AC7500BXB	Production

Figure 1. Top Side Marking Example



Errata

1. CPURST bit does not get cleared by hardware

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

Implication: The Intel 7500 chipset IOH 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 Table of Changes.

2. VTX-RCV-DETECT pulse too large during receiver detection

Problem: The VTX-RCV-DETECT pulse has been measured as high as 700 mV.

Background:

VTX-RCV-DETECT - The amount of voltage change allowed during Receiver Detection. The maximum is 600 mV for both 2.5 GT/s and 5 GT/s.

Implication: This may overstress PCIe* agents.

Workaround: None at this time

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

3. PCIe* Surprise Link Down status may not be flagged during training failure or surprise removal

Problem: For PCIe* root ports (devices 1 - 10) not supporting hot-plug, Uncorrectable Error Status Register (offset 104h) bit [5] are not set when the connected endpoints are brought to a surprised link down state, with the PCI Express* Slot Control Register (offset A8h) Power Controller Control (bit [10]) left at default value.

Implication: Surprised link down state of the endpoints will not be detected, logged or reported by the PCIe root ports not supporting hot-plug slots.

Workaround: The logging of Surprise Down event caused by a non hot-plug removal event depends on condition of 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 (bit [10]) to 0, in addition to other proper configurations, to enable detection of Surprise Down condition of the associated endpoint.

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

4. PCIe Gen2 differential peak-to-peak transmit voltage swing is too low

Problem: The PCIe Gen2 transmit buffers are not generating a large enough peak-to-peak transmit voltage swing and may violate the PCIe Gen2 specification.

Implication: The PCIe Gen2 specification may be 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

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

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

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

Implication: All PCIe Advanced error status logging registers have a corresponding error detect mask register to control if error statuses 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 the 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 Table of Changes.

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

Problem: Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d) queue-based invalidation can be enabled only when queue-based invalidation is enabled on both isochronous and non-isochronous channels.

Implication: This is a violation of the Intel VT-d specification which allows each engine to be configured in either queue-based or register-based invalidation independently.

Workaround: None.

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

7. PCIe Header of a malformed TLP is not logged

Problem: A malformed Transaction Layer Packet (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 (Header Log register, Device:0-10, Function:0, Offset 11Ch).

Implication: The PCIe specifies that the header of the malformed TLP must be logged. Thus, the PCIe specification is violated.

Workaround: None.

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

8. Intel VT-d does not support the draining of compatibility-format interrupts

Problem: Intel 7500 chipset does not support the draining of compatibility-format interrupts which is a violation of the Intel VT-d specification. This issue only impacts software that use compatibility-format interrupts and then transition to remappable-format interrupts at run-time. Impact is limited to OS/VMMs that enable interrupt remapping outside of x2 APIC mode.

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

Workaround: Impacted software can use alternate methods to drain interrupts. Specifically, software can issue reads to any PCI/PCIe architectural registers without read side-effects at the interrupt source (endpoint device), and have the PCI ordering rules (read-completion push/drain) followed on any in-flight interrupts (including compatibility-mode interrupts) from that source.

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

9. 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. When Intel VT-d DMA translation is not enabled, hardware applies the HPA_LIMIT check to incoming DMA addresses. Note that the reset default value of HPA_LIMIT in hardware is 39 (that is, address bits 38:0 are valid).

Implication: The reset default value of HPA_LIMIT in hardware is address bits 38:0. Left at default value, incoming memory request with address greater than bit [38] will be aborted by the IOH.

Workaround: BIOS must set VTGENCTRL (Dev 20:Func 0:Offset 184h) Register bits[7:4] to 0Fh to allow address decode up to address bit[50].

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

10. 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: Given that the IOH does not support the D1 and D2 states, the IOH should not allow these values to be written. The IOH does not changes power state from D0 or D3hot when PMCSR bits 1:0 are written to D1 or D2, so there is no functional impact to the IOH when these states are written.

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

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

11. PCIe Gen2 Tx Return loss fails spec

Problem: The the differential and common mode transmit return loss fail to meet the PCIe Gen2 specification.

Implication: Some Gen2 PCIe agents may have increased inter-symbol interference, ISI, due to signal reflection at the driver of Intel 7500 chipset IOH.

Workaround: None

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

12. 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 IOH is configured to record errors related to the JTAG interface.

Workaround: None

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

13. Interoperability issue of some PCIe Gen1 cards with Gen2 Devices

Note: This is not an Intel 7500 chipset IOH issue, and the workaround documented below is to resolve the interoperability issue certain PCIe Gen1 cards have with Gen2 capable Intel® 5500/5520 chipset PCIe root ports.

Problem: During PCIe physical layer link initialization, Intel 5500/5520 chipset IOH Gen2 capable PCIe root ports will 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 downstream pre-2.0 PCIe Gen1 cards may have trouble dealing with these Gen1 reserved fields that are not reserved for PCIe Gen2.

Implication: These PCIe Gen1 cards may fail to properly train with Intel 5500/5520 chipset IOH.

Workaround: Below is the standard method to force Intel 5500/5520 chipset IOH PCIe root ports to Gen1 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 PCIe Gen1 cards fail to train with the standard method, BIOS should program the following registers to prevent Intel 5500/5520 chipset IOH PCIe root ports from advertising 2.0 Specification defined capabilities. There is one register control field for each Intel 5500/5520 chipset IOH 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 the default value of “1” to disable the corresponding root port from advertising upconfigure capability before doing Step 3 above.

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

14. System reset is getting turned into a power cycle

Problem: If a device behind the ICH9/ICH10 becomes non-responsive, the ESI link between the IOH and the ICH can run out of downstream credits. If the processor has a CATERR# due to the non-responsiveness of this device, some systems issue a reset so that the error data can be collected by BIOS. This reset causes a reset_warn to be issued from the ICH to the IOH. But since the IOH is out of credits it is unable to respond to the reset_warn. After 4 seconds the ICH times out waiting on the response to the reset_warn and issues a global reset (DC power cycle).

Implication: All error data about the CATERR# is lost due to the power cycle.

Workaround: If server management has knowledge of CATERR#, and has access to the reset line, processors, and chipset then the server management can watch for the CATERR#. If the CATERR# is asserted then collect the failure data followed by asserting the reset. The issue is ultimately considered a platform issue and will not be resolved via a silicon fix.

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

15. Intel VT-d: Memory read request with AT=11b results in malformed TLP

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

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

Workaround: None

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

16. 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: The corrected functionality is documented in Table 16-3 of the *Intel® 7500 Chipset Datasheet*.

17. Transactions to addresses above TOCM are not setting the Master Abort

Problem: Inbound Transactions to addresses above TOCM are master aborted, but BXB does 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 Table of Changes.

18. Bandwidth very low for write traffic with noSnoop attribute set

Problem: Upstream posted 100% 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 Table of Changes.

19. Modifying RTIDs causes system hangs

Problem: In some instances, modification of the QPI[1:0]PORB (Device:16, Function:1, Offset:B4h, B0h) bits 11:0 causes systems to hang.

Implication: The max request allocation pools for the RTIDs can not be modified once set.

Workaround: BIOS should set the QPI[1:0]PORB (Device:16, Function:1, Offset:B4h, B0h) bits 11:0 once after reset and should not make any modification to these bits thereafter.

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

20. System Hang due to VLW

Problem: Intel® QuickPath Interconnect (Intel® QPI) Virtual Legacy Wire (VLW) INIT routing is not enabled by default in the IOH preventing a response to a SHUTDOWN message in early boot. This is because the IOH Intel QuickPath Interconnect Response is blocked until BIOS programs broadcast node IDs and unmasks the INIT (NcMsgB-VLW) in the IOH.

Implication: If a catastrophic fault condition is encountered in early boot and a SHUTDOWN message is issued, the system hangs.

Workaround: The ICH watchdog timer can be used to reset the system. Once silicon is at PRQ quality, a healthy system will not exhibit this issue.

Status: This issue does not affect Intel Itanium processor 9300 series-based systems

21. Intel® QuickPath Interconnect (Intel® QPI) Queue/Table overflow or underflow error observed

Problem: Intel QuickPath Interconnect Queue/Table overflow or underflow error status (Dev 20:Func 2:Offset 84h.[27:26]) is observed.

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

Workaround: BIOS should program the following Intel 7500 chipset IOH registers to mask off the spurious sub-state status and allow the Intel QuickPath Interconnect 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 Table of Changes.

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

Problem: The Intel VT-d engine blocks Intel VT-d translates 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 Table of Changes.

23. Intel QuickPath Interconnect 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 a Intel QuickPath Interconnect link went through a successful retrain.

Workaround: Please contact your Intel field representative for the latest BIOS Specification Update for details.

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

24. 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 Table of Changes.

25. Subsystem Vendor ID (SVID) and the Subsystem Device ID (SDID) of device 15 are not implemented

Problem: The upper 8 bits of Subsystem Vendor ID (SVID) and the Subsystem Device ID (SDID) of device 15 are not implemented. Only the lower 8 bits of these registers can be written to, though the PCIe specification indicates that these are 16 bit registers.

Implication: When all 16 bits of the register are written and then read back only the first 8 bits are read correctly.

Workaround: None.

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

26. Intel QuickPath Interconnect links do not link and train to the full width

Problem: There may be some instances where the Intel QuickPath Interconnect links do not link and train to the full width when initializing operation at 4.8 GT/s.

Implication: All the lanes do not link and train as expected in full width mode when bringing up 4.8 GT/s operation.

Workaround: Firmware must set the ETPollingBitlock (bits 19:16) of register QPI[1:0]PH_PTV (device:13, function: 1-0, offset: 854h) to 0x1. The equivalent Intel® 7500 Scalable Memory Buffer register Px[n]_CSIPHPTV bits[19:16] should also be set to 1.

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

27. Link training failure due to multiple resets

Problem: When the Intel QuickPath Interconnect link is trained in slow mode and is followed by consecutive physical layer resets issued by the Tukwila processor, a link training failure may be observed after the thirty third (33) physical layer reset attempt.

Implication: Link training failure may be observed after multiple physical layer resets.

Workaround: None

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

28. 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 Completor Abort message may not be received.

Workaround: None

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

29. Some PCIe inbound messages not ignored as expected

Problem: Some PCIe inbound messages that should be ignored may be incorrectly reported as Unsupported Requests. Messages with codes between 0x40 and 0x48 (except 0x42 and 0x46) and 0x7f are logged as Unsupported Request.

Implication: These messages may not be ignored as expected.

Workaround: None

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

30. Timeout values much larger than specified

Problem: When using the bits 51:48 of the QPIPCTRL 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: Appropriate value programmed in BIOS.

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

31. Spurious D8 error may be reported

Problem: A spurious “D8: Protocol Layer Received Illegal packet field or Incorrect Target NodeID” error may be reported when EMP mode is enabled. This issue occurs when the DNID field of a NDR response is split across 2 flits. This issue does not result in any incorrect functional behavior.

Implication: A spurious “D8: Protocol Layer Received Illegal packet field or Incorrect Target NodeID” may be reported even when there are no related errors.

Workaround: The reporting of D8 errors can be disabled by setting bit 28 to a “1” in the following register: Device 13, Function 1-0, Offset: F94h.

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

32. Failure during operation at PCI Express L1 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 Table of Changes.

33. ACS Violation is not treated as Advisory when severity is set to Non-Fatal

Problem: If the severity of an Access Control Services (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 specification.

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 Table of Changes.

34. Snoop requests may not be possible when addressing above 2^41

Problem: Snoop requests to address range above 2^41 may not receive any responses.

Implication: Snoop requests accesses to memory ranges above 2^41 will not receive completions as expected.

Workaround: Set bit 0 to a “0” in the following register: Device :17,16, Function:0, Offset: 5ch.

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

35. D5 and D7 errors will not have the Intel QPI headers logged

Problem: When a D5 or D7 error occurs, the associated Intel QPI header will not be logged.

Implication: Header information when a D5 or D7 error occurs will not be available.

Workaround: D5 and D7 errors themselves are logged correctly in the error status registers.

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

36. Locked read time out bit not being set

Problem: The locked read timed out bit in the MISCCTRLSTS register (device :0, Function:0 Offset: 188h) may not be set as expected.

Implication: A locked read request that incurred a completion time-out on any of the PCI Express ports may not result in the lock read time out bit getting set.

Workaround: None.

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

37. Reply Timer Timeout Errors logged when enabling ASPM L0s

Problem: When enabling ASPM L0s Reply Timer Timeout errors may be logged in root ports of the IOH.

Implication: Reply Timer Timeout values may be observed when enabling L0s state on PCIe.

Workaround: Set bits [5:0] in the following registers Device: 14 Function: 0 Offset: 4C4h, Device: 13 Function: 3 Offset: 4C4h, Device: 13 Function: 1 Offset: 4C4h, Device: 13 Function: 0 Offset: 4C4h to a value of 9h.

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

38. Link and PHY Layer Reset Can Cause CRC Errors With 16 Bit CRC

Problem: A Link layer and Phy Layer reset can both cause multiple CRC errors when 16 bit rolling CRC is enabled in at least one direction in the link. This can possibly trigger ERROR# 301, “Rbox Intel QuickPath Interconnect CRC Error”. Please refer to the flow detailed in the *Intel® Itanium® Processor 9300 Series NDA Specification Update* when performing a link layer reset where the CRC mode is set or changed to 16 bit rolling CRC.

Implication: On Intel IPF based platforms, Intel QuickPath Interconnect CRC Errors can occur if 16 bit rolling CRC is enabled and a link layer or PHY layer reset occurs.

Workaround: Please refer to the erratum 8 of the *Intel® Itanium® Processor 9300 Series NDA Specification Update* for the details.

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

39. Data Mismatch on Inbound MemWrs 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: Bit 8 of XPUNCERRSTS register (Device:0-10, Function:0, Offset: 0x208) will have the indication of the error. Please refer to erratum 40 for more details when using in a x8 or x16 configuration

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

40. MSI with greater than 1DWord payload is not logged in XPUNCERRSTS bit 8 as expected

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: Read bit [8] of the XPUNCERRSTS register (Device:0-10, Function:0, Offset: 0x208) of all partner ports to see the status and escalate the error from partner port.

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

41. PCIe link degrades and surprise link down (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 Down will be reported as a fatal error by the IOH. Link degrades are non-fatal and can be checked in bit 0 of the Correctable Error Status register: XPCORERRSTS (Device:0-10, Function:0, Offset: 0x200h).

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. BIOS should implement the workaround following the steps below:

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
        end
        Else
                LNKCON[1:0] = 11b
else
        write Link Control Register (LNKCON, PCIE Offset 0x10)
                LNKCON[1:0] = 00b // Disable L0s and L1
end
Else
        LNKCON[1:0] = 01b // Enable only L0s
```

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 Table of Changes.

42. Error not logged due to a corrupt STP symbol

Problem: Intel 7500 chipset IOH does not log an error when STP (Start of Transaction Layer Packet) 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 Table of Changes.

43. Intel QuickPath Interconnect errors can occur on inband resets

Problem: On some inband resets (including physical layer resets), the IOH may take 1 or 2 extra cycles to determine that the Intel QuickPath Interconnect forwarded clock has stopped.

Implication: The IOH may see CRC errors, system hangs, or link level control errors.

Workaround: Please contact your Intel field representative for the latest BIOS Specification Update for details.

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

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

Problem: In a multi-IOH system, when at least two IOHs having 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 multi-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 Table of Changes.

45. Potential Link Width degradation on Intel QuickPath Interconnect

Problem: In some rare instances there could be a potential for link width reduction on Intel QuickPath Interconnect in situations where Phy Resets are being looped continuously.

Implication: Sub-optimal performance of the Intel QuickPath Interconnect due to link width reduction. Some of the other potential scenarios where this issue could be observed is during an Hot-Plug event, Dynamic Data Partitioning, System boot and CRC error escalation.

Workaround: In the event that link width reduction is observed, the following sequence can be used to restore the link to full-width operation.

- Read CSRs in Device: 13, Function: 1-0, Offset: 84Ch
- If either bits 27:24 or bits 19:16 of either CSR are not equal to 'Fh', then set to 1 bit 0 of the QPI[1:0]PH_CTR CSRs (Device: 13, Function: 1-0, Offset: 82Ch).

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

46. Intel QuickPath Interconnect Error Status D3 is observed

Problem: Intel QuickPath Interconnect 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 QuickPath Interconnect D3 error status to function normally: Set Dev 13:Func 0/1: Offset F8Ch (Dword) bit[14] to 1.

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

47. 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 Table of Changes.

48. 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 Table of Changes.

49. 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 when using Intel TXT.

Workaround: None.

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

50. 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 Table of Changes.

51. 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.

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 Table of Changes.

52. 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 non-fatal 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.

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

53. Forwarded Clock Lane Detection status may not be indicated accurately

Problem: Bit 31 of the QPI[1:0]PH_TDS register (Device: 13. Function: 1-0, Offset: 834h) is normally used to indicate detection of the forwarded clock. Due to this issue, this bit may not indicate the status of the forwarded clock accurately.

Implication: The status of the forwarded clock may not be determined by reading this bit.

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

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

54. ESI link cannot go into L1 state on the Intel® Itanium® Processor 9300 Series-Based Platform and Intel® Xeon® Processor 7500 Series-based Platform

Problem: The ESI link cannot go into the L1 state on the Intel Itanium processor 9300 series-based platform and Intel® Xeon® processor 7500 series-based platform.

Implication: As a result of this issue, the L1 power state cannot be engaged on the Intel Itanium processor 9300 series-based platform or Intel Xeon processor 7500 series-based platforms and results in a nominal increase to the published idle power number. There is expected to be no increase to any measured power values due to this issue.

Workaround: None Identified.

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

55. Intel QuickPath Interconnect interleave selection modes 0x5 and 0x6 do not allow NodeID port allocation for either port

Problem: Problem: When using Intel QuickPath Interconnect interleave selection modes 0x5 or 0x6, the port ID calculated for the final target node may not be consistent with the settings of the QPI[1:0]PORB register (Device: 16, Function: 1, Offset: B4h, B0h) and QPIRTBL (Device: 16, Function: 1, Offset: 44h) register.

Implication: Due to this issue, there are restrictions on how port allocations can be set up when using interleave modes 0x5 and mode 0x6.

Workaround: For a pair of NodeIDs that have only their ID bit (bit 1) toggled the mapping should be such that both NodeIDs are mapped to the same Intel QPI port.

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

56. PCIe ASPM L1 can cause link degradation and speed change

Problem: PCIe ASPM L1 can cause link degradation and speed change on Intel 7500 chipset root ports when additional power saving on PCIe Rx is enabled.

Implication: PCIe Gen2 x16 link can degrade to Gen1 x8 or x2 link.

Workaround: Intel 7500 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).

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

57. Links may not be able to recover after entering Live Error Recovery (LER) mode

Problem: When errors are detected by the PCIe port, the PCIe port goes into a Live Error Recovery mode (if enabled) and brings the associated link down. However, the links may not recover as expected after the relevant error status bits are cleared by software.

Implication: Live Error Recovery (LER) mode may not be used to bring the PCIe links back up after they have gone down due to errors being detected.

Workaround: To recover the link after entering LER mode, a hard reset to the entire Intel 7500 chipset should be issued followed by software clearing the affected error status registers.

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

58. 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 Table of Changes.

59. PCIe Slot Status Register Command Completed bit not always updated on any configuration write to the Slot Control Register

Problem: For PCIe root ports (devices 0 - 10) supporting hot-plug, the Slot Status Register (offset AAh) Command Completed (bit[4]) status is updated under the following condition: IOH will set Command Completed bit after delivering the new commands written in the Slot Controller register (offset A8h) to VPP. The IOH detects new commands written in Slot Control register by checking the change of value for Power Controller Control (bit[10]), Power Indicator Control (bits[9:8]), Attention Indicator Control (bits[7:6]), or Electromechanical Interlock Control (bit[11]) fields. Any other configuration writes to the Slot Control register without changing the values of these fields will not cause Command Completed bit to be set.

The PCIe Base Specification Revision 2.0 or later describes the "Slot Control Register" in section 7.8.10, as follows (Reference section 7.8.10, Slot Control Register, Offset 18h).

In hot-plug capable Downstream Ports, *a write to the Slot Control register must cause a hot-plug command to be generated* (see Section 6.7.3.2 for details on hot-plug commands). A write to the Slot Control register in a Downstream Port that is not hot-plug capable must not cause a hot-plug command to be executed.

The PCIe Spec intended that every write to the Slot Control Register is a command and expected a command complete status to abstract the VPP implementation specific nuances from the OS software. IOH PCIe Slot Control Register implementation is not fully conforming to the PCIe Specification in this respect.

Implication: Software checking on the Command Completed status after writing to the Slot Control register may time out.

Workaround: Software can read the Slot Control register and compare the existing and new values to determine if it should check the Command Completed status after writing to the Slot Control register.

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

60. Level-triggered multicast device interrupts may cause an MCA timeout

Problem: When multiple processors are the target of an I/O initiated legacy interrupt (INTx), generated by the IOH IOAPIC, one of the target threads can be handling the interrupt while any of the other target threads initiate an EOI for the same vector. Intel 7500 chipset Input/Output Hub's IOxAPIC can be kept busy by software handling the multicast interrupts in this way. As long as the actual handling thread has not deasserted the interrupt, Intel 7500 chipset Input/Output Hub will resend the same interrupt again whenever it sees the EOI (End Of Interrupt).

Implication: A CPU timeout MCA has been detected in a synthetic test environment when I/O initiated multicast interrupts (logical mode, level-triggered, fixed delivery mode interrupts with multiple bits set in the mask part of the logical APIC ID to specify more than one processor), generated from the IOH IOAPIC. Intel has not observed this erratum with any commercially available software.

Workaround: Do not program the IOxAPIC in the IOH to use logical mode, level-triggered, fixed delivery mode interrupts with multiple bits set in the mask part of the logical APIC ID. Either use physical mode instead of logical mode, set a single bit in the mask part of the logical APIC ID or use lowest priority delivery mode instead of fixed delivery mode. Intel is not aware of any commercially available operating system using logical mode, level-triggered, fixed delivery mode interrupts with multiple bits set in the mask part of the logical APIC ID.

Status: For the steppings affected, see the Summary Table 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 programed 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.

Specification Changes

1. Intel QuickPath Interconnect Protocol Control Completion Retry Timeout Interval specification change to the *Intel® 7500 Chipset Datasheet*

The following specification change to the QPIPCTRL Protocol Control Completion Retry Timeout Intervals, Device: 16, Function: 1, Offset: 4Ch, bits 51:48 will be updated in the next revision of the datasheet as follows:

Register: QPIPCTRL
Device: 16
Function: 1
Offset: 4Ch

Bit	Attr	Default	Description
51:48	RWL	0011	**Configuration Retry Timeout:** Applies only to PCI Express/ESI ports. Controls how long a configuration request is reissued (when enabled via the root control register) whenever a CRS response is received. Reissue applies to all configuration requests when CRS software visibility is disabled (via the root control register) and to all configuration requests except on configuration reads to Vendor/Device ID field at offset 0x0, when CRS software visibility is enabled. The timer that is controlled by this field starts when a configuration request is issued the very first time on PCI Express. When this timer expires and following that if either a CRS response is received for the configuration request or a completion time-out occurs on the configuration request, the request is aborted (that is, not reissued anymore) and a UR (/equivalent) response is returned. Note that a configuration request is not immediately aborted when this timer expires. Aborting a configuration request only happens when either a completion timeout condition is reached or when a CRS response is received with the retry timeout expired. 0000: 2.56 ms 0001: 40 ms 0010: 640 ms 0011: 10 s 0100: 160 s 0101: 320 s 0110:640 s 0111: 1280 s 1000: 2560 s 1001: 205 hrs 1010-1111: Reserved

2. Intel QuickPath Interconnect Protocol APIC Source Address Decode Extended Logical Mode Interleave functionality specification change to the *Intel® 7500 Chipset Datasheet*

The Extended Logical Mode Interleave functionality will be changed with the addition of Mode 0x02 as shown:

Register: QPIAPICSAD
Device: 16
Function: 1
Offset: 80h

Bit	Attr	Default	Description
3:1	RWLB	0	**Extended Logical Mode Interleave** Mode - Interleave - Local/remote Cluster ID 0x0 - APIC ID[18:16] - APIC ID[22:19] 0x1 - APIC ID[19:17] - APIC ID[23:20] 0x2 - APIC ID[20:18] - APIC ID[24:21] >0x2 - Reserved

Specification Clarifications

1. Memory Address Decode - *Intel® 7500 Chipset Datasheet*

The following clarification to the Memory Address Decoder Fields table will be made in the next revision of the specification.

Table 4-4 Memory Address Decode Fields

Field Name	Number of Bits	Description
Inter-leave Select	3	Determines how targets are interleaved across the range. Sys_Interleave value is set globally using the "QPIPINT: Intel QuickPath Interconnect Protocol Interleave Mask" register. Modes: 0x0 - Addr[8:6] 0x1 - Addr[8:7], Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x2 - Addr[9:8], Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x3 - Addr[8:6] XOR Addr[18:16] 0x4 - (Addr[8:7] XOR Addr[18:17]), Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x5 - Addr[8:6]: Similar to 0x0 encoding but the nodeid[1] will be overwritten by the Sys_Interleave bit (as calculated per the CSR setting), that is, final dnodeid = dnodeid[5:2], Sys_Interleave, dnodeid[0]. 0x6 - Addr[8:6] XOR Addr[18:16]: Similar to 0x3 encoding but the nodeid[1] will be overwritten by the Sys_Interleave bit (as calculated per the CSR setting), i.e. final dnodeid = dnodeid[5:2], Sys_Interleave, dnodeid[0] >0x6 - Reserved

2. The following clarification to the QPIPMADDATA Register, interleave select bit field [3:1], will be made in the next revision of the *Intel® 7500 Chipset Datasheet*

Register: QPIPMADDATA
Device: 16
Function: 1
Offset: 70h

Bit	Attr	Default	Description
3:1	RWL	0h	Interleave Select 0x0 - Addr[8:6] 0x1 - Addr[8:7], Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x2 - Addr[9:8], Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x3 - Addr[8:6] XOR Addr[18:16] 0x4 - (Addr[8:7] XOR Addr[18:17]), Sys_Interleave (Not intended for use with Intel Xeon processor 7500 series) 0x5 - Addr[8:6]: Similar to 0x0 encoding but the nodeid[1] will be overwritten by the Sys_Interleave bit (as calculated per the CSR setting), i.e. final dnodeid = dnodeid[5:2], Sys_Interleave, dnodeid[0]. 0x6 - Addr[8:6] XOR Addr[18:16]: Similar to 0x3 encoding but the nodeid[1] will be overwritten by the Sys_Interleave bit (as calculated per the CSR setting), i.e. final dnodeid = dnodeid[5:2], Sys_Interleave, dnodeid[0] >0x6 - Reserved

3. Protocol Status - *Intel® 7500 Chipset Datasheet*

The following clarification to the Intel QPI Protocol Status Register Fields (Register QPIPSTS, Device: 16, Function: 1, Offset: 54h) will be made in the next revision of the document.

Register: QPIPSTS
Device: 16
Function: 1
Offset: 54h

Bit	Attr	Default	Description
2	RO	0	**ORB non-lock_arb Not empty** This indicates that there are pending requests in the ORB with the exception of StopReq*/StartReq* messages from the lock arbiter. 1 - Pending ORB requests 0 - ORB Empty (except StopReq*/StartReq*)
1	RO	0	**ORB Empty** This indicates that there are pending requests in the ORB. 1 - Pending ORB requests 0 - ORB Empty

4. Intel QuickPath Interconnect Protocol Power Control clarification to the State Level description in the *Intel® 7500 Chipset Datasheet*

A change to the State_Level description (bits [15:0]) will be made in the next revision of the document as shown below.

Register: QPIPPOWCTRL
Device: 16
Function: 1
Offset: 64h

Bit	Attr	Default	Description
15:0	RW	0040h	State Level Sets the State_Level[15:0] in the CmpD response to a PMReq message. The value is priority encoded (similar to one-hot). Default is a state_level of 7. It is required that at least one bit be set in this field. state level: bit position 0: 1 1: 2 . 15: 16

5. Intel QuickPath Interconnect Device Control register, clarification of the extended tag field functionality, in the *Intel® 7500 Chipset Datasheet*

The PCI Express Device Control register extended tag field bit[8] functionality is clarified as shown below.

Register: DEVCTRL Device: 0-10 Function: 0 Offset: 98h			
Bit	**Attr**	**Default**	**Description**
8	RW	0h	Extended Tag Field Enable This bit enables the PCI Express port/ endpoint to use an 8-bit Tag field as a requester. Note that the IOH itself does not support the extended tag field as a requestor; nor, does it modify or check the extended tag bits [7:5] of a completion.[1]

1. The IOH does not process the Extended Tag bits [7:5] and forwards them unmodified. Since the IOH does not check the Extended Tag bits, a system crash could occur under the following conditions:
   The processor issues a Read transaction targeting an end-point
   The end point device malfunctions and corrupts the Extended Tag bits of the Read completion prior to assembling the PCI Express frame.
   The end point PCI Express port calculates correctly the CRC for Frame with the corrupted Tag
   The IOH will not detect the corrupted Tag and will forward the corrupted completion to the processor via QPI bus

6. Intel QuickPath Interconnect Interrupt Control (QPIPINTRC) clarification, in the *Intel® 7500 Chipset Datasheet*

Register: QPIPINTRC Device: 17 Function: 1 Offset: E0h			
Bit	**Attr**	**Default**	**Description**
15:8	RW	3Fh	**Legacy Signal Mask:** When set, the corresponding legacy wire from ICH is ignored by IOH and for FERR output, IOH does not assert FERR signal to ICH when masked. This mask also applies to internal IOH events generating NMI/SMI. 8: NMI 9: INIT 10: SMI 11: INTR 12: A20M 13: FERR 15-14: Reserved

7. Intel LNKSTS: PCI Express Link Status Register clarification, in the *Intel® 7500 Chipset Datasheet*

Register: LNKSTS Device: 0-10 Function: 0 Offset: A2h			
Bit	**Attr**	**Default**	**Description**
12	RWO	1	Slot Clock Configuration This bit indicates whether IOH receives clock from the same xtal that also provides clock to the device on the other end of the link. 1: indicates that same xtal provides clocks to devices on both ends of the link 0: indicates that different xtals provide clocks to devices on both ends of the link

8. Section 16.10.1.1 for IOH Hardware Support for Intel QPI Hot-Plug, item 6 in the *Intel® 7500 Chipset Datasheet*

Item 6 incorrectly states that there are 16 scratch pad and semaphore registers when there are actually 17.

9. Section 20.6.6.1 SR[0:3]: Scratch Pad Register 0-3 (Sticky)

The register description in the table below incorrectly shows SR[4:7] where it should be described as SR[0:3]. This will be corrected in the next revision of the *Intel® 7500 Chipset Datasheet*.

Register: SR[0:3]
Device: 20
Function: 1
Offset: 07Ch-088h

Bit	Attr	Default	Description
31:0	RWLBS	0h	**Scratch Pad – Sticky** Sticky scratch pad registers for firmware utilization

10. Table 8-6 MSI Address Format when Remapping is Enabled

The register description in the table below correctly defines the Interrupt Handle fields and adds the Interrupt Format field. This will be corrected in the next revision of the *Intel® 7500 Chipset Datasheet*.

Table 8-6. MSI Address Format when Remapping is Enabled

Bits	Description
31:20	FEEh
19:5	**Interrupt Handle[14:0]:** IOH looks up an interrupt remapping table in main memory using this field as an offset into the table
4	**Interrupt Format:** This field must have a value of 1b for Remappable format interrupts.
3	**Sub Handle Valid**: When IOH looks up the interrupt remapping table in main memory, and if this bit is set, IOH adds the bits 15:0 from interrupt data field to interrupt handle value (bit 19:4 above) to obtain the final offset into the remapping table. If this bit is clear, Interrupt Handle field directly becomes the offset into the remapping table.
2	**Interrupt Handle[15]:** IOH looks up an interrupt remapping table in main memory using this field as an offset into the table
1:0	00

11. Section 20.11.3.9 QPI[1:0]EP_SR: Electrical Parameter Select Register

The register description in the table below correctly defines the IEParamSel field. This will be corrected in the next revision of the *Intel® 7500 Chipset Datasheet*.

Register: QPI[1:0]EP_SR
Device: 13
Function: 1-0
Offset: 8A0

Bit	Attr	Default	Description
23:16	RWS	0h	**EParamSel:** Electrical parameter select [7:0] This selects the particular Electrical Parameter to be interfaced. 2h: VOC (Rx voltage offset) 3h: TDV (Tx Drive) 5h: TEQ (TX Equalization)

Documentation Changes

1. Thermal Sensor Control Register - *Intel® 7500 Chipset Datasheet*

The following additions to the TSCTRL Register (Device: 20, Function: 3, Offset: E8h) will be made in the next revision of the specification.

TSCTRL: On-Die Thermal Sensor Control Register addition of bit [0] TSEN: Thermal Sensor Enable.

Note for bits [9:8]: When SWTHROTTLE = 0 and TSDIS = 1, no throttling occurs, the thermal sensor does not assert the THERMALERT_N pin, and CTSTS.THRMALERT is not asserted.

Register: TSCTRL
Device: 20
Function: 3
Offset: E8h

Bit	Attr	Default	Description
0	RW	0	**TSEN: Thermal Sensor Enable** 0: Disable Thermal Sensor 1: Enable Thermal Sensor. When this bit is logic 1 a free-running counter is activated to produce a thermal measurement cycle.

§