

Intel[®] X58 Express Chipset

Specification Update

— For Intel[®] X58 Express Chipset-based High-End Desktop (HEDT) Platforms Only

January 2009

Notice: The Intel[®] X58 Express Chipset IOH 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.

Document Number: 320839-003



INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. Intel products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications.

Intel may make changes to specifications and product descriptions at any time, without notice.

Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them.

The Intel® X58 Express 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.

‡Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain platform 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.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

I²C is a two-wire communications bus/protocol developed by Philips. SMBus is a subset of the I²C bus/protocol and was developed by Intel. Implementations of the I²C bus/protocol may require licenses from various entities, including Philips Electronics N.V. and North American Philips Corporation.

Intel and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

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

Copyright © 2008 - 2009, Intel Corporation. All Rights Reserved.



Contents

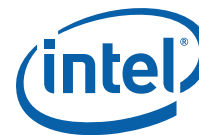
Preface	5
Summary Tables of Changes	6
Identification Information	9
Errata	10
Specification Changes	23
Specification Clarifications	24
Documentation Changes	25



Revision History

Revision	Description	Date
-001	Initial Release	November 2008
-002	Added errata 38-45.	December 2008
-003	Added errata 46-49.	January 2009

§ §



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 corrections, 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® X58 Express Chipset Datasheet	320838-001

Nomenclature

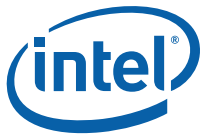
Errata are design defects or errors. These may cause the Intel X58 Express Chipset IOH's behavior to deviate from published specifications. Hardware and software designed to be used with any given stepping must assume that all errata documented for that stepping are present on all devices.

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.

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, etc.).



Summary Tables of Changes

The following tables indicate the errata, specification changes, specification clarifications, or documentation changes which apply to the IOH 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 Tables

Stepping

- X: An 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.

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	Errata
	B2		
1	X	No Fix	PCIe Link bandwidth permanently set
2	X	No Fix	CPURST bit does not get cleared by hardware
3	X	No Fix	VTX-RCV-DETECT pulse too large during receiver detection.
4	X	No Fix	Testlo[20] must be pulled high.
5	X	No Fix	Surprise down error status not being flagged
6	X	No Fix	Extended Error Detect Mask Registers of all PCIe root ports mask error logging by default.
7	X	No Fix	Long XOR chain is broken
8	X	No Fix	PCI-E Header of a malformed TLP is not logged
9	X	No Fix	PCI-E PMCSR Power State fields allow writing D1 and D2
10	X	No Fix	DEVCON2[3:0] and CTOCTRL not accurately set the completion timeout value.
11	X	No Fix	PCIe Gen2 Tx Return loss fails spec
12	X	No Fix	Persistent Jtag error reported at MIERRST register
13	X	No Fix	Interop issue of some PCIe Gen1 cards with Gen2 Devices
14	X	No Fix	CSI fails to retrain to L0 slow mode after 32 in-band resets
15	X	No Fix	Boot-dependent QPI CRC errors
16	X	No Fix	PCIe Inbound Msg that should be ignored is treated as Unsupported Request
17	X	Doc	Memory writes to a certain address range are considered advisory non-fatal
18	X	No Fix	PCIe lane-to-lane performance variation
19	X	No Fix	Transactions to addresses above TOCM are not setting the Master Abort
20	X	No Fix	Setting bits 24 and 25 of the MISCCTRLSTS does not result in expected behavior
21	X	No Fix	Timeout values much larger than specified
22	X	No Fix	Failure during operation at PCI-Express L1 power management state
23	X	No Fix	Bandwidth very low for write traffic with noSnoop attribute set
24	X	No Fix	System Hang due to VLW
25	X	No Fix	QPI Queue/Table overflow or underflow error observed.
26	X	No Fix	ACS Violation is not treated as Advisory when severity is set to Non-Fatal
27	X	No Fix	MIERRCNT does not properly count persistent SMBus retry failures and MINNERRST doesn't log the errors
28	X	No Fix	EOI to the I/OxAPIC can be blocked
29	X	No Fix	GTIME upper 32 bits can not be read or written.
30	X	No Fix	IOH May Falsely Assert THERMTRIP_N Signal After a Reset Event
31	X	No Fix	VT-d queue-based invalidation is enabled only when enabled on both channels
32	X	No Fix	VTd does not support the draining of compatibility-format interrupts
33	X	No Fix	Hardware applies HPA_LIMIT to upstream memory request when VT-d is disabled
34	X	No Fix	VT-d: Memory read request with AT=11 results in malformed TLP
35	X	No Fix	VT-d: IOTLB Domain-page-selective invalidation not working correctly.
36	X	No Fix	VT-d translated write transactions are blocked but not recorded.
37	X	No Fix	The IOH ignores Snoop Behavior Bit in VT-d Page Table.
38	X	No Fix	QPI L1 state is taking greater than 15 us from L1 exit to L0 state



Errata (Sheet 2 of 2)

Number	Steppings	Status	Errata
	B2		
39	X	No Fix	Intel® QuickPath Interconnect (QPI) L0s and L1 Power Management Link State Fails
40	X	No Fix	PCI Express* 2.0 L0s Link Recovery Fails
41	X	No Fix	QPI Link Training failures in L0 when L0r enabled
42	X	No Fix	VT-d reports an Invalidation Queue Error when the Queue size is 7.
43	X	No Fix	ERRSID not logging ReqID for Inbound ERR_* messages
44	X	No Fix	QPI PhyResets causing QPI Errors to be flagged
45	X	No Fix	QPI initialization abort failures logged during power-on resets
46	X	No Fix	QPI errors can occur on inband resets.
47	X	No Fix	Data Mismatch on Inbound MemWrS after MSI with payload greater than 1 DWORD payload
48	X	No Fix	MSI with greater than 1DWord payload is not logged in XPUNCERRSTS bit 8
49	X	No Fix	PFQ Flag (due to VT-d ISOCH fault) in Fault Status Register not being cleared

Specification Changes

Number	SPECIFICATION CHANGES
	There are no specification changes in this revision of the specification update

Specification Clarifications

No.	SPECIFICATION CLARIFICATIONS
	There are no specification clarification in this revision of the specification update

Documentation Changes

No.	DOCUMENTATION CHANGES
1	Documentation update to the Intel® X58 Express Chipset – External Design Specification (EDS) – Rev. 2.0 #367244



Identification Information

Component Identification via Programming Interface

The Intel® X58 Express Chipset (IOH) stepping can be identified by the following register contents:

Stepping	Features	Vendor ID	Device ID	Revision Number
B2	36S	8086h	3405h	12h

Component Marking Information

The Intel® X58 Express Chipset (IOH) production stepping can be identified by the following component markings:

Stepping	S-Spec	Top Marking	Notes
B2	SLGBT	AC82X58	Production Units

§



Errata

1. PCIE Link bandwidth permanently set

Problem: The Link Autonomous Bandwidth Status and Current Link Speed are bits are set prior to reaching L0 after the Link has been brought down and then back up. These bits maintain their values through the Link down event. The bits should be reset during the Link down event.

Background

As part of the Link speed negotiation process the PCIE component enters the Recovery.RcvrCfg state from L0 at 2.5GT/s and then moves on to Recovery.Speed to negotiate the Link bandwidth.

Recovery.Speed

The Transmitter enters Electrical Idle and stays there until the Receiver Lanes have entered Electrical Idle, and then additionally remains there for at least 800 ns on a successful speed negotiation (i.e., `successful_speed_negotiation = 1b`) or at least 6 μ s on an unsuccessful speed negotiation (i.e., `successful_speed_negotiation = 0b`), but stays there no longer than an additional 1 ms. The frequency of operation is changed to the new data rate only after the 35 Receiver Lanes have entered Electrical Idle.

LNKSTS: PCI Express Link Status Register - The PCI Express Link Status register provides information on the status of the PCI Express Link such as negotiated width, training etc.

Link Autonomous Bandwidth Status (LNKSTS bit 15) - This bit is set to 1b by hardware to indicate that hardware has autonomously changed link speed or width, without the port transitioning through DL_Down status, for reasons other than to attempt to correct unreliable link operation. IOH sets this bit when it receives eight consecutive TS1 or TS2 ordered sets with the Autonomous Change bit set. Note that if the status bit is set by hardware in the same clock software clears the status bit, the status bit should remain set and if MSI is enabled, the hardware should trigger a new MSI.

Link Bandwidth Management Status (LNKSTS bit 14) - This bit is set to 1b by hardware to indicate that either of the following has occurred without the port transitioning through DL_Down status: a) A link retraining initiated by a write of 1b to the Retrain Link bit has completed b) Hardware has autonomously changed link speed or width to attempt to correct unreliable link operation Note that if the status bit is set by hardware in the same clock software clears the status bit, the status bit should remain set and if MSI is enabled, the hardware should trigger a new MSI.

Current Link Speed (LNKSTS bits 3:0) - This field indicates the negotiated Link speed of the given PCI Express Link.

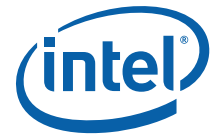
0001- 2.5 Gbps
0010 - 5 Gbps
Others - Reserved

The value in this field is not defined and could show any value, when the link is not up. Software determines if the link is up or not by reading bit 13 of this register.

Implication: Hot Plug operations involving cards capable of different bandwidths should not be performed as the Link bandwidth will not get adjusted accordingly.

Workaround: None

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



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

Problem: SYRE: CPURST bit does not get cleared by hardware.

Background

SYRE: System Reset - This register controls IOH reset behavior.

CPURESET (SYRE bit 10) - When this bit is set to "1", the IOH asserts RESET0_N. The IOH clears this bit when CPURESET timer elapses.

Implication: The IOH will not assert RESET0_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_{TX-RCV-DETECT} pulse too large during receiver detection.**

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

Background

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

Implication: This may overstress PCIE agents.

Workaround: None

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

4. **Testlo[20] must be pulled high.**

Problem: The Testlo[20] signal is inverted.

Implication: If the workaround is not implemented, the IOH will not correctly function.

Workaround: The Testlo[20] pin must be pulled to the VCCEPW rail via a 10K ohm $\pm 1\%$ resistor.

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 is not being flagged during training failure or surprise removal.

Background

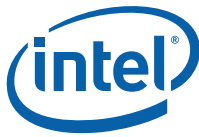
UNCSTS: Uncorrectable Error Status - This register identifies uncorrectable errors detected for PCI Express/DMI port.

Surprise Down Error Status (UNCSTS bit 5) - Flags the system when a training failure or surprise removal has occurred.

Implication: The system will not be able to reliably detect if a PCIE card has not trained properly or been inadvertently removed unless the workaround is used.

Workaround: Workaround: BIOS fix

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 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 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. Long XOR chain is broken

Problem: The long XOR chain used for high volume manufacturing test of the PCIe interface is not functional.

Implication: Tests dependent on using the long XOR chain for high volume manufacturing test will not function

Workaround: The individual XOR chains for each PCIe port are operable and can be used to cover those areas covered by the long XOR chain.

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

8. PCI-E 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 HDRLOG register (Header Log register, Device:0-10, Function:0, Offset 11Ch) do not log the header of the malformed TLP.

Implication: Given that the PCI-E specifies that the header of the malformed TLP must be logged, the PCI-E specification is violated.

Workaround: None

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

9. PCI-E 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 change 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 Tables of Changes.



10. DEVCON2[3:0] and CTCTRL not accurately set the completion timeout value.

Problem: Problem: The DEVCON2 (device: 0-10, function: 0, offset: B8h) bits 3:0 in conjunction with CTCTRL (device: 0-10, function: 0, offset: 1E0h) bits 9:8 does not accurately set the completion timeout value.

Implication: The duration of the PCIe timeout will differ from that programmed.

Workaround: See the documentation change for 3319486.

Status: For the steppings affected, see the Summary Tables 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 the IOH. Gen1 return loss will meet specifications.

Workaround: Following Tylersburg PDG PCIe Gen2 guidelines will result in no Gen2 performance or functional issues. Contact your Intel field representative for the latest BIOS information.

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



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

Note: Note: this is not an Intel X58 Express Chipset IOH issue, and the workaround documented below is to resolve the interoperability issue certain PCIe Gen1 cards have with Gen2 capable IntelX58 Express Chipset PCIe root ports.

Problem: During PCIe physical layer link initialization, the 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 X58 Express Chipset IOH.

Workaround: Below is the standard method to force Intel X58 Express 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 X58 Express Chipset IOH PCIe root ports from advertising 2.0 Specification defined capabilities. There is one register control field for each Intel X58 Express 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 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.

14. CSI fails to retrain to L0 slow mode after 32 in-band resets

Problem: After 32 CSI inband resets from the CPU to an IOH port are executed when in slow mode the next inband reset will fail as the CSI will not retrain to L0 slow mode.

Implication: The CSI link will fail when operating in slow mode after 32 CSI 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.

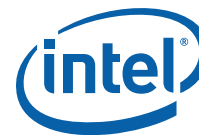
15. Boot-dependent QPI CRC errors

Problem: QPI CRC errors may be observed as a function of boot-dependent (probabilistic) phase alignment of internal clock domains

Implication: QPI port may fail to train or may train with sustained non-zero CRC error rate

Workaround: If CRC's are not user-tolerable, power cycle system.

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



16. **PCIE Inbound Msg 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 PCI-E spec 2.0 not to send those messages.

Workaround: None

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

17. **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 error is to be logged in UNCERRSTS (Device:0-10, Function:0, Offset:104h). The accesses are aborted and logged; however, when severity of master-abort (UR) is set to non-fatal, memory write accesses are not to be considered advisory non-fatal, rather, they should be considered normal 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.

18. **PCIE lane-to-lane performance variation**

Problem: The lane-to-lane performance is inconsistent on the PCIE interface. Some lanes source/sink more current than others leading to analog variations.

Implication: Current compensation is compromised insomuch as not all of the lanes can be counted on to perform consistently. Optimizing the current compensation for one lane will not lead to optimized analog behavior in a neighboring lane. Ultimately, the overall analog behavior is limited by the lowest performing lane, thus overall channel analog behavior is compromised.

Workaround: Under Investigation.

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

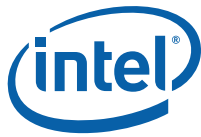
19. **Transactions to addresses above TOCM are not setting the Master Abort**

Problem: Inbound Transactions to addresses above TOCM are master aborted, but TBG 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. When VT is enabled, there is no workaround.

When VT is disabled, program the HPA_LIMT to the maximum value.

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



20. **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 Change.

21. **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 timelimit may not be observed.

Workaround: None

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

22. **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 1.

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

23. **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.

24. **System Hang due to VLW**

Problem: 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 QPI Response is blocked until BIOS programs broadcast node IDs and unmask the INIT (NcMsgB-VLW) in the IOH.

Implication: If a catastrophic fault condition is encountered in early boot and a SHUTDOWN message 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: For the steppings affected, see the Summary Tables of Changes.

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

Problem: QPI 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 X58 Express Chipset IOH registers to mask off the spurious sub-state status and allow the 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. 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 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.

27. MIERRCNT does not properly count persistent SMBus retry failures and MINNERRST doesn't log the errors

Problem: The first time a SMBus invalid access is performed the error is logged on the MIERRST register and the MIERRCNT register counts the first error as expected. However, subsequent SMBUS retry failures are not captured by the MIERRCNT register and also the MINNERRST register does not log the error as expected.

Implication: Persistent SMBus retry failures are not logged as expected.

Workaround: None.

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

28. EOI to the I/OxAPIC can be blocked

Problem: The End of Interrupt (EOI) message targeted to the I/OxAPIC will be blocked in the case that an INTx from a PCIe device and an MSI are pending inside the Intel X58 Express Chipset .

Implication: The I/OxAPIC will continue to treat a level triggered interrupt as asserted as the EOI message is never received by the I/OxAPIC to clear the interrupt.

Workaround: Avoid using the I/OxAPIC in the Intel X58 Express Chipset and use only the I/OxAPIC in the ICH component.

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

29. GTIME upper 32 bits can not be read or written.

Problem: The GTIME register (Device: 20, Function: 2, Offset 0x1D4) always returns 0 and writes 0 to the upper 32 bits of data. The register does count correctly.

Implication: The GTIME upper 32 bits can not be read or written accurately. The GFFERRTIME (Device:20, Function:2, Offset:1E0h) and GNFFERRTIME (Device:20, Function:2, Offset:1F0h) registers do get logged correctly given that the register does count correctly.

Workaround: None

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



30. IOH May Falsely Assert THERMTRIP_N Signal After a Reset Event

Problem: When Intel® X58 Express Chipset based platforms undergo a reset event, such as a front panel reset or any warm reset, the THERMTRIP_N pin on the IOH may falsely assert, resulting in unexpected system behavior (i.e.- possible platform shutdown, IOH initiated platform throttling, etc. depending on implementation).

Implication: Unexpected system behavior. **Note:** Platforms that have THERMTRIP_N unconnected or connected to a pull-up resistor are NOT impacted.

Workaround: None.

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

31. VT-d queue-based invalidation is enabled only when enabled on both channels

Problem: VT-d queue-based invalidation is enabled only when queue-based invalidation is enabled on both isochronous and non-isochronous channels.

Implication: This is a violation of the 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 Tables of Changes.

32. VTd does not support the draining of compatibility-format interrupts

Problem: The Vt-d does not support the draining of compatibility-format interrupts which is a violation of the 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: Under investigation

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

33. Hardware applies HPA_LIMIT to upstream memory request when VT-d is disabled

Problem: When VT-d DMA translation is enabled, hardware correctly applies the GPA_LIMIT check to the incoming DMA address and HPA_LIMIT to VT-d page-walk addresses. When 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 (i.e. address bits 38:0 are valid).

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

Workaround: The easiest workaround for this hardware bug is to have BIOS program the appropriate HPA_LIMIT in VTGENCTRL register always, irrespective of the BIOS setup option to expose/disable VT-d. Contact your Intel field representative for the latest BIOS information.

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

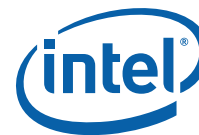
34. VT-d: Memory read request with AT=11 results in malformed TLP

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

Implication: When in 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 Tables of Changes.



35. VT-d: IOTLB Domain-page-selective invalidation not working correctly.

Problem: When VT-d is doing the IOTLB Domain-page-selective invalidation for leaf only (IH=1) while Vt-d is still in the middle of pagewalk, Vt-d needs to invalidate all entries in the cache that are still waiting for the pagewalk data from memory to complete a walk. Because of a logic bug, those entries did not get invalidated and are still waiting to receive the pagewalk data. This can never happen because the data could be thrown away by the invaldaiton command.

Implication: The pagewalk never completes and no address will be given back to the IOH and this could hang the system.

Workaround: Workaround is to not set the INVADDRREG (Device: MMIO, BAR: VTBAR), bit 6.

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

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

Problem: The VT-d engine blocks 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 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.

37. The IOH ignores Snoop Behavior Bit in VT-d Page Table.

Problem: When the snoop behavior control bit <11> of the VT-d leaf page table entry is set, all inbound untranslated transactions should be treated as snooped, irrespective of the NS (Non-Snoop) attribute in the request. However, the state of this bit appears to not have any affect on snoop behavior.

Implication: If any DMI agent sends non-snoop traffic and the page-table entry configures the request to force snoop behavior, the traffic will remain non-snoop traffic.

Workaround: The BIOS should force the IOH to disable the "snoop control" capability for the nonisochronous DMI VT-d unit. Contact your Intel field representative for the latest BIOS information.

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

38. QPI L1 state is taking greater than 15 us from L1 exit to L0 state

Problem: The IOH 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 contact your Intel field representative for the latest BIOS Specification Update for details.

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

39. Intel® QuickPath Interconnect (QPI) L0s and L1 Power Management Link State Fails

Problem: The Intel® QuickPath Interconnect (QPI) on the IOH-side of the Intel® X58 Express Chipset may exhibit some instability entering or exiting the link power management states L0s and L1.

Implication: System hang or a low rate of CRC errors on QPI link.



Workaround: BIOS workaround- Disable Intel® QPI power management states L0s and L1. CPU Package C-states will be disabled as well. Please contact your Intel field representative for the latest BIOS Specification Update for details.

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

40. PCI Express* 2.0 L0s Link Recovery Fails

Problem: When the Intel® X58 Express Chipset receives a request to enter L0s on the PCI Express* 2.0 interface and a subsequent link recovery is initiated the link training may fail. This issue does not affect the DMI interface on the Intel® X58 Express Chipset.

Implication: System hang.

Workaround: BIOS workaround- Disable the transmission of L0s requests on devices connected to the Intel® X58 Express Chipset PCI Express* interface. Please contact your Intel field representative for the latest BIOS Specification Update for details.

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

41. QPI Link Training failures in L0 when L0r enabled

Problem: 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: The fix is integrated into the QPI RC 0.90RC1 and above.

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

42. VT-d reports an Invalidation Queue Error when the Queue size is 7.

Problem: When the Queue Size is set to 7 in the Invalidation Queue Address register (Addr: MMIO, BAR: VTBAR, Offset:90h, 1090h), VT-d incorrectly reports Invalidation Queue Error in the Fault Status register (Addr: MMIO, BAR: VTBAR, Offset:34h, 1034h, bit 4) at the time of fetching a descriptor.

Implication: A fault event will be generated thus leading to an interrupt request if the Fault Event Control Register (Addr: MMIO, BAR: VTBAR, Offset:38h, 1038h) is programmed to respond to this error.

Workaround: Do not set the Queue Size to 7.

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

43. 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.

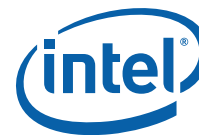
Status: For the steppings affected, see the Summary Tables of Changes.

44. QPI PhyResets causing QPI Errors to be flagged

Problem: When QPI PhyResets are executed, unexpected QPI errors may be logged. The link remains operational regardless of the logged errors. The following errors have been observed: B6 and DG. These errors will be flagged in the QPI[1:0]ERRST (Device:20, Function:2, Offset:280h, 200h) register.

Implication: This issue may manifest itself at all QPI speeds upon assertion of PhyReset, which should not occur during normal operation. Regardless of the issue, the link remains operational and system does not fail.

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



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

45. QPI initialization abort failures logged during power-on resets

Problem: During power-on reset, the IOH may log 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 at least one QPI lane hasn't trained correctly.

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

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

46. 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 CSI 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 contact your Intel field representative for the latest BIOS Specification Update for details.

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

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

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

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

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

49. PFP Flag (due to VT-d ISOCH fault) in Fault Status Register not being cleared

Problem: When a 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.



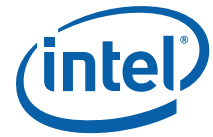
Specification Changes

No Specification Changes in this revision of the specification update.



Specification Clarifications

No Specification Clarifications in this revision of the specification update.



Documentation Changes

1. Documentation update to the *Intel® X58 Express Chipset – External Design Specification (EDS) – Rev. 2.0 #367244*. Chapter 14, page 203- Intel® Virtualization Technology. The next revision of the EDS will reflect the following specification changes to the documentation.
 - Remove line- *"Support for root entry, context entry, and default context."*
 - Remove line- *"Support for Intel VT-d read prefetching/snarfing, that is, translations within a cacheline are stored in an internal buffer for reuse for subsequent transactions."*
 - Remove line- *"Support for PCISIG endpoint caching (ATS)."*
 - Remove line- *"No support for Intel VT-d translation bypass address range (such usage models need to be resolved with VMM help in setting up the page tables correctly)."*
 - Remove line- *"Support for queue-based invalidation interface."*
 - Change *"No support for 1 or 2 level page walks for 1, 2, or 3 level walks for non-isoch remap engine"* to *"No support for 1 or 2 level page walks for isoch remap engine, and 1, 2, or 3 level walks for non-isoch remap engine"*