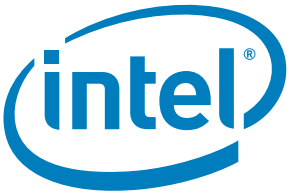


Intel® Ethernet Controller I350 Specification Update

LAN Access Division (LAD)

Revision: 2.04
January 2012

Legal

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.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL'S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

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 information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document 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.

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

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

Intel and Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

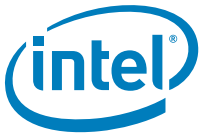
* Other names and brands may be claimed as the property of others.

Revisions

Date	Revision	Description
1/18/2011	1.0	Initial release.
2/17/2011	1.1	Brand string in title updated. Moved to Software Clarification section: • 1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB Specification Clarification updated: • 1. SerDes: AN_TIMEOUT Only Works When Link Partner Idle Specification Changes added or updated: • 1. Updates to PXE/iSCSI EEPROM Words • 2. Updates to Software Compatibility EEPROM Word 0x3 Errata updated (errata numbers impacted by this): • 4. SGMII: Counters Incorrectly Increment on Collision - Text updated. • 6. NC-SI: Hardware Arbitration Issues - Combined six hardware arbitration issues into one master entry. The issues combined are related and minor. • 9. TSYNC: Auxiliary Timestamp from SDP is Unreliable - Text updated.
4/15/2011	1.9	PRQ version (by convention revision 1.9). Specification Clarification added: • 2. SR-IOV and DMA Coalescing Cannot Be Set Together Errata added or updated: • 5. HOST<-->BMC Only Packets Not Counted as Host Sent/Received Packets (previously combined under another errata; broken out) • 10. Rx Packet May Be Dropped After Receiving 2 Bit Errors While in An Idle State (text re-written) • 11. SDPx_1 Remains an Input in D3 9. TSYNC: Auxiliary Timestamp from SDP is Unreliable • 12. NC-SI Enable Global Multicast Filter Command - No Differentiation of Neighbor Discovery Modes • 12. NC-SI Enable Global Multicast Filter Command - No Differentiation of Neighbor Discovery Modes • 13. NC-SI: Select Package Command Increments NC-SI Commands Received Counter for Ports 0 and 1 and Not for Ports 2 and 3 • 14. VLAN Tags Byte-Swapped in Loopback Packets • 15. DEV_OFF_N May Not Properly Disable the Device • 16. DMA Coalescing Reacts to MSI-X Vectors Instead of Interrupt Causes • 17. IEEE Std 802.3™-2008 Tx Distortion Marginality
5/6/2011	1.91	Specification Change added: • 3. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)
5/27/2011	1.92	Specification Clarification added: • 3. PCIe Completion Timeout Mechanism Compliance
6/21/2011	2.00	Device information: • Table 1-1 updated. Combines data in several of the older tables. Errata added or updated: • 17. IEEE Std 802.3™-2008 Tx Distortion Marginality

Date	Revision	Description
8/18/2011	2.01	Specification Change: • 4. Minimum Value for Flow Control Receive Threshold Low , Added. Software Clarification added: • 2. Serial Interfaces Programmed By Bit Banging
9/14/2011	2.02	Specification Clarification added: • 4. Malicious Driver Detection
12/14/2011	2.03	Errata added: • 18. Alternate MAC Address Port Order • 19. Neighbor Discovery Offload: Override Bit not Set in Neighbor Advertisement Packet
1/30/2012	2.04	Specification Clarifications added: • 5. Padding on Transmitted SCTP Packets • 6. MCTP/DMTF Standard Compliance Errata added: • 20. VFTA Register Access Might Not Be Performed

§ §



1. Introduction

This document applies to all the I350 variants. The information supplements data provided in the product datasheet.

1.1 Product Code and Device Identification

Product Code: I350

The following describes identifying markings on each device package.

Table 1-1. Product and Device Identification

MM#	Supplier/Mfg Part Number	Top Marking	Spec #	Status & Media	Description	Step	Pack Qty
915808	L2A3341A1-BH-DT	NHI350AM4	S LJ2Z	Tape & Reel	1 Gbs, 4-port, 17x17	A1	850/reel
915799	L2A3341A1-BH-DB	NHI350AM4	S LJ3Z	Tray		A1	90/tray
915809	L2A3341A1C-BH-DT	NHI350AM2	S LJ3R	Tape & Reel	1 Gbs 2-port, 17x17	A1	850/reel
915801	L2A3341A1C-BH-DB	NHI350AM2	S LJ3S	Tray		A1	90/tray
915810	L2A3341A1B-BH-DT	NHI350AS4	S LJ3T	Tape & Reel	1 Gbs 4-port, 17x17 SERDES	A1	850/reel
915802	L2A3341A1B-BH-DB	NHI350AS4	S LJ3U	Tray		A1	90/tray
915811	L2A3341A1-DT	NHI350BT2	S LJ3X	Tape & Reel	1 Gbs, 2-port, 25x25	A1	650/reel
915804	L2A3341A1-DB	NHI350BT2	S LJ3Y	Tray		A1	44/tray
913157	M-L2A3341A1-BH-DB	NHI350AM4	Q P3L	Engineering, Tray	1 Gbs, 4 port 17x17	A1	90/tray
913159	M-L2A3341A1C-BH-DB	NHI350AM2	Q P3R	Engineering, Tray	1 Gbs, 2-port 17x17	A1	90/tray
913193	M-L2A3341A1B-BH-DB	NHI350AS4	Q P5A	Engineering, Tray	1 Gbs, 4-port 17x17 SERDES	A1	90/tray
913194	M-L2A3341A1-DB	NHI350BT2	Q P5B	Engineering, Tray	1 Gbs, 2-port 25x25	A1	44/tray

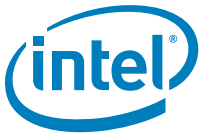
Table 1-3. Device ID

Device	Vendor ID	Device ID
I350 (EEPROM-Less)	0x8086	0x151F
I350 (SerDes (KX/BX))	0x8086	0x1523
I350 (SGMII)	0x8086	0x1524
I350 (Copper)	0x8086	0x1521
I350 (Fiber)	0x8086	0x1522

1.2 Marking Diagrams

Figure 1-1. Example Component With Identifying Marks

Line 1	“intel”
Line 2	Marketing Name
Line 3	Fab Lot Number “XXXXXXXXX” (Wafer Lot no. concatenated with Assembler vendor code)
Line 4	Assembly Date Code YYWW; Engineering samples have additional Intel data.
Line 5	Copyright line; includes two number date code and the Pb-free mark (e1)
Line 6	Country of Origin



1.3 Nomenclature Used In This Document

This document uses specific terms, codes, and abbreviations to describe changes, errata, sightings and/or clarifications that apply to silicon/steppings. See Table 1-4 for a description.

Table 1-4. Nomenclature

Name	Description
Specification Changes	Modifications to the current published specifications. These changes will be incorporated in the next release of the specifications.
Errata	Design defects or errors. Errata may cause device 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.
Sightings	Observed issues that are believed to be errata, but have not been completely confirmed or root caused. The intention of documenting sightings is to proactively inform users of behaviors or issues that have been observed. Sightings may evolve to errata or may be removed as non-issues after investigation completes.
Specification Clarifications	Greater detail or further highlights concerning a specification's impact to a complex design situation. These clarifications will be incorporated in the next release of the specifications.
Software Clarifications	Applies to Intel drivers, EEPROM loads.
Documentation Corrections	Errors, or omissions in current published specifications. These changes are incorporated in the next release of the applicable document and then dropped from the specupdate. You may also check for changes in the revision history of specific documents.
A1, B1, etc.	Stepping to which the status applies.
Doc	Document change or update that will be implemented.
Fix	This erratum is intended to be fixed in a future stepping of the component.
Fixed	This erratum has been previously fixed.
NoFix	There are no plans to fix this erratum.
EEPROM/NVM Fix	This indicates the Errata was in the EEPROM/NVM and is fixed in an updated version.
Eval	Plans to fix this erratum are under evaluation.
Red Change Bar/or Bold	This Item is either new or modified from the previous version of the document.

2. Changes and Updates

See Section 1.3 for an explanation of terms, codes, and abbreviations.

Table 2-1. Summary of Hardware Sightings, Clarifications, Changes, Errata, and Software Clarifications

Sightings	**Status**
None.	NA
Specification Clarifications	**Status**
1. SerDes: AN_TIMEOUT Only Works When Link Partner Idle	NA
2. SR-IOV and DMA Coalescing Cannot Be Set Together	NA
3. PCIe Completion Timeout Mechanism Compliance	NA
4. Malicious Driver Detection	NA
5. Padding on Transmitted SCTP Packets	NA
6. MCTP/DMTF Standard Compliance	NA
Specification Changes	**Status**
1. Updates to PXE/iSCSI EEPROM Words	NA
2. Updates to Software Compatibility EEPROM Word 0x3	NA
3. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)	NA
4. Minimum Value for Flow Control Receive Threshold Low	NA
Errata	**Status**
1. I2C Data Out Hold Time Violation	A1 NoFix
2. ECRC Capability Not Set in VF Configuration Space	A1 NoFix
3. RQDPC Register is RC (Read-Clear)	A1 NoFix
4. SGMII: Counters Incorrectly Increment on Collision	A1 NoFix
5. HOST<-->BMC Only Packets Not Counted as Host Sent/Received Packets	A1 NoFix
6. NC-SI: Hardware Arbitration Issues	A1 NoFix
7. Entering D3 State May Be Delayed	A1 NoFix
8. PCIe Gen2 Transmitter $V_{TX\text{-}AC\text{-}CM\text{-}PP}$ Max Value Violation	A1 NoFix
9. TSYNC: Auxiliary Timestamp from SDP is Unreliable	A1 Fixed
10. Rx Packet May Be Dropped After Receiving 2 Bit Errors While in An Idle State	A1 NoFix

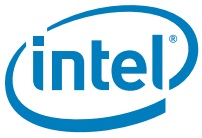


Table 2-1. Summary of Hardware Sightings, Clarifications, Changes, Errata, and Software Clarifications

11. SDPx_1 Remains an Input in D3	A1 NoFix
12. NC-SI Enable Global Multicast Filter Command - No Differentiation of Neighbor Discovery Modes	A1 NoFix
13. NC-SI: Select Package Command Increments NC-SI Commands Received Counter for Ports 0 and 1 and Not for Ports 2 and 3	A1 NoFix
14. VLAN Tags Byte-Swapped in Loopback Packets	A1 NoFix
15. DEV_OFF_N May Not Properly Disable the Device	A1 NoFix
16. DMA Coalescing Reacts to MSI-X Vectors Instead of Interrupt Causes	A1 NoFix
17. IEEE Std 802.3™-2008 Tx Distortion Marginality	A1 NoFix
18. Alternate MAC Address Port Order	A1 EEPROM Fix
19. Neighbor Discovery Offload: Override Bit not Set in Neighbor Advertisement Packet	A1 EEPROM Fix
20. VFTA Register Access Might Not Be Performed	A1 NoFix
Software Clarifications	**Status**
1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB	NA
2. Serial Interfaces Programmed By Bit Banging	NA

1.1 Sightings

None.

1.2 Specification Clarifications

1. SerDes: AN_TIMEOUT Only Works When Link Partner Idle

Clarification:

The auto-negotiation time out mechanism (PCS_LCTL.AN_TIMEOUT_EN) only works if the SerDes partner is sending idle code groups continuously for the duration of the time out period (the usual case).

If the partner is transmitting packets in auto-negotiation, time out will not occur since auto-negotiation is restarted at the beginning of each packet. If the partner has an application that indefinitely transmits data in spite of the lack of response, it is possible that link will not be established.

Workaround:

If this is a concern, the auto-negotiation time out mechanism may be considered unreliable and an additional software mechanism can be used to disable auto-negotiation (if sync is maintained without a link being established (PCS_LSTS.SYNC_OK=1 and PCS_LSTS.LINK_OK=0) for an extended period of time).

2. SR-IOV and DMA Coalescing Cannot Be Set Together

Clarification:

DMA coalescing assumes a specific mapping between interrupt causes and MSI-X vectors. This mapping doesn't match the distribution of MSI-x vectors to Virtual functions.

Workaround:

DMA coalescing and SR-IOV cannot be activated together.

3. PCIe Completion Timeout Mechanism Compliance

Clarification:

The I350 Completion Timeout Value[3:0] must be properly set by the system BIOS in the I350 PCIe Configuration Space Device Control 2 register (0xC8; W). Failure to do so can cause unexpected completion timeouts.

The I350 complies with the PCIe 2.0 specification for the completion timeout mechanism and programmable timeout values. The PCIe 2.0 specification provides programmable timeout ranges between 50us to 64s with a default time range of 50us - 50ms. The I350 defaults to a range of 16ms - 32ms.

Workaround:

The completion timeout value must be programmed correctly in PCIe configuration space (in Device Control 2 register); the value must be set above the expected maximum latency for completions in the system in which the I350 is installed. This ensures that the I350 receives the completions for the requests it sends out, avoiding a completion timeout scenario. Failure to properly set the completion timeout value can result in the device timing out prior to a completion returning.

4. Malicious Driver Detection

Clarification:

Section 7.8.3.8.3 of the Datasheet [search for “Interrupt on Misbehavior of VM (Malicious Driver Detection)”] describes the actions of the device when an error that might be caused by a malicious driver is detected. These actions are usually sufficient to keep the PF and the other VFs functioning, but there are situations where the transmit DMA function may hang as a result of malicious action by a VF driver.

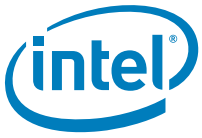
Workaround:

We recommend that the following actions be performed by the PF driver upon receipt of an MDDET interrupt to ensure that the PF and othe VFs are not blocked:

- Clear VFTE and VFRE bits for the offending VF.
- Set VTCTRL.RST of the offending VF.

The above actions were implemented in Intel igb driver v3.1.16

5. Padding on Transmitted SCTP Packets



Clarification:

When using the I350 to offload the CRC calculation for transmitted SCTP packets, software should not add Ethernet padding bytes to short packets(less than 64 bytes). Instead, the TCTL.PSP bit should be set so that the I350 pads the packets after performing the CRC calculation.

6. MCTP/DMTF Standard Compliance

Clarification:

The I350 MCTP protocol implementation is based on the DMTF DSP0236, DSP0237 and DSP0239 standards.

The I350 NC-SI over MCTP implementation is described in Chapter 10 of the datasheet.

1.3 Specification Changes

1. Updates to PXE/iSCSI EEPROM Words

Change:

Word 0x30, 34, 38, 3A are now defined as follows:

Bit(s)	Value	Port Status	CLP (Combo) Executes	iSCSI Boot Option ROM CTRL-D Menu	FCoE Boot Option ROM CTRL-D Menu
2:0	0	PXE	PXE	Displays port as PXE. Allows changing to Boot Disabled, iSCSI Primary or Secondary.	Displays port as PXE. Allows changing to Boot Disabled, FCoE Enabled.
	1	Boot Disabled	NONE	Displays port as Disabled. Allows changing to iSCSI Primary/ Secondary.	Displays port as Disabled. Allows changing to FCoE enabled.
	2	iSCSI Primary	iSCSI	Displays port as iSCSI Primary. Allows changing to Boot Disabled, iSCSI Secondary.	Displays port as iSCSI. Allows changing to Boot Disabled, FCoE Enabled.
	3	iSCSI Secondary	iSCSI	Displays port as iSCSI Secondary. Allows changing to Boot Disabled, iSCSI Primary.	Displays port as iSCSI. Allows changing to Boot Disabled, FCoE Enabled.
	4	FCoE	FCOE	Displays port as FCoE. Allows changing port to Boot Disabled, iSCSI Primary or Secondary.	Displays port as FCoE. Allows changing to Boot Disabled.
	5-7	Reserved	Same as Disabled	Same as Disabled.	Same as Disabled.
4:3	Same as before.				
5	Bit 5, formerly used to indicate iSCSI enable / disable, is no longer valid and is not checked by software.				
15:7	Same as before.				

2. Updates to Software Compatibility EEPROM Word 0x3

Change:

Bit	Name	Default	Description
1:0	Reserved	0	Reserved
2	Reserved	0	Reserved
3	Reserved	0	Reserved
4	Reserved	0	Reserved
7:5	Reserved	0	Reserved
8	Reserved	0	Reserved.
9	Client	0	Client/Not a Client NIC 0b = Server. 1b = Client. This bit is used by DMIX to verify the NIC is server or client. Server NIC or LOM required.
10	Reserved	0	Reserved
11	LOM	0	Indicates whether dedicated flash for the option ROM is attached to LAN silicon. Used by option ROM update applications, QV and DMIX. 0b = NIC (A dedicated flash is attached) 1b = LOM (No dedicated flash is attached)
12	Reserved	0	Reserved
14:13	Reserved	0	Reserved
15	Reserved	0	Reserved

3. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)

Driver software provides a method to identify an external port on a system through a command that causes the LEDs to blink. Based on the setting in word 0x4, the LED drivers should blink between STATE1 and STATE2 when a port identification command is issued.

When word 0x4 is equal to 0xFFFF or 0x0000, the blinking behavior reverts to a default.

Bit	Description
15:12	Control for LED 3 0000b or 1111b: Default LED Blinking operation is used. 0001b = Default in STATE1 + Default in STATE2. 0010b = Default in STATE1 + LED is ON in STATE2. 0011b = Default in STATE1 + LED is OFF in STATE2. 0100b = LED is ON in STATE1 + Default in STATE2. 0101b = LED is ON in STATE1 + LED is ON in STATE2. 0110b = LED is ON in STATE1 + LED is OFF in STATE2. 0111b = LED is OFF in STATE1 + Default in STATE2. 1000b = LED is OFF in STATE1 + LED is ON in STATE2. 1001b = LED is OFF in STATE1 + LED is OFF in STATE2. All other values are Reserved.
11:8	Control for LED 2 – same encoding as for LED 3.
7:4	Control for LED 1 – same encoding as for LED 3.
3:0	Control for LED 0 – same encoding as for LED 3.

4. Minimum Value for Flow Control Receive Threshold Low

Change:

If FCRTL0.XONE is 1, the minimum value allowed in FCRTL0.RTL is 3 (48 bytes).

1.4 Errata

1. I^2C Data Out Hold Time Violation

Problem:

The I350 should provide a data out hold time of 50 ns on the SFPx_I2C_DATA pins. The actual hold time is about 16 ns.

Implication:

Timing specification violation. There have been no reports of failures resulting from this timing. Note that the data input hold time required is zero, so the provided output hold time should be more than enough as long as the I^2C CLK and DATA signals are reasonably matched on the board.

Workaround:

None.

Status: A1 NoFix

2. ECRC Capability Not Set in VF Configuration Space

Problem:

According to SR-IOV v1.1 specification, the ECRC capability reporting bits (Advanced Error Capabilities and Control Register; bits 5 & 7) should be set in VF if the device supports ECRC generation and checking.

Although the I350 supports ECRC generation and checking, the capabilities are reported in the PF and not in the VF configuration space. In the VF configuration space, these bits are set to zero.

Implication:

ECRC capabilities should be read from PF. The device's behavior (e.g. ecrc generate and check) is correct.

Workaround:

None.

Status: A1 NoFix

3. RQDPC Register is RC (Read-Clear)

Problem:

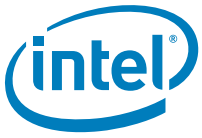
The RQDPC register should not be cleared by a read. It is RC for VF and PF.

Implication:

The RQDPC register may be cleared by a VF driver or a PF driver and thus can not be used by both the PF and the VF to gather receive-queue-drop statistics. Only one of the drivers may use it.

Workaround:

The VF driver may request the receive-queue-drop statistics from the PF.



Status: A1 NoFix

4. SGMII: Counters Incorrectly Increment on Collision

Problem:

In SGMII mode/half duplex, the statistics counters listed below incorrectly increment when a collision occurs:

Name	Definition	Location
RLEC	Length error counter	0X4040
CRCERRS	CRC error counter	0x4000
RFC	receive frame counter	0x40A8

Implication:

Error counters may not be accurate.

Workaround:

None.

Status: A1 NoFix

5. HOST<-->BMC Only Packets Not Counted as Host Sent/Received Packets

Description:

When OS2BMC is enabled, packets that do not reach the LAN are not counted as packets sent by the host (HGPTC register). Similarly, packets received from the MC are not counted as packets received by the host (RPTHC register)

Implication:

HGPTC and RPTHC counts are not accurate.

Workaround:

Add the O2BGPTC count to the HGPTC count to get the accurate number of packets sent by the host. Add the B2OGPRC count to the RPTHC count to get the accurate number of packets received by the host.

Status: A1 NoFix

6. NC-SI: Hardware Arbitration Issues

These are several minor issues:

1. Receiving a FLUSH Smaller then the I350 ID Causes Issue
2. Wait IDLE State Violation
3. Wait-Idle-to-Normal Operation Time Violation

4. Timeout Mechanism Stops Upon Receiving Pause Packets from MC
5. XON Not Sent When Expected

Description(s):

1. In the middle of normal operation, the device may get FLUSH commands with a smaller ID then the Device ID. The device should pass on the received FLUSH; but it sends it's own ID for ~2us and then passes on the lower ID FLUSHs.
2. While the device is in Wait Idle state, it sends IDLE on ARB_OUT even though it did not get anything from the ARB_IN pin.
3. The time from Received-Idle (while in Wait Idle state) until the device sends IDLE on ARB_OUT is 1.7us; the max time allowed (by spec) is T9 = 640ns.
4. Hardware arbitration timeout mechanism stops upon receiving pause packets from the MC. The timer stops counting until the pause indication drops.
5. If a token timeout occurs while the device waits to send an XON packet, the internal state machine is reset and the XON is never sent.

Implication(s):

1. No implication in actual operation. Eventually, the lower IDs pass and arbitration succeeds.
2. No effect on the functionality of HW-ARB. If the device doesn't receive valid op-codes from ARB-IN, this means that the HW-ARB loop is broken and the system should be moved to SW arbitration; so IDLE op-codes generated in such a situation are irrelevant.
3. The issue is not expected to cause problems because the timeout period is longer. Minor NC-SI compliance violation related to hardware arbitration.
4. Longer than expected timeout (no spec violation).
5. The MC is released by the XOFF timer expiration. Minor NC-SI compliance violation related to hardware arbitration.

Workaround(s):

None.

Status: A1 NoFix

7. Entering D3 State May Be Delayed

Problem:

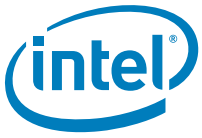
If one of the functions driver accesses the flash and this access is delayed, a request to enter another function (or this one) to D3 will be delayed until the flash transaction is done. In some extreme cases (internal FW is erasing a sector and a function tries to access the flash), this may delay the D3 entrance by hundreds of ms.

Implication:

Entering D3 state may be delayed.

Workaround:

None.



Status: A1 NoFix

8. PCIe Gen2 Transmitter $V_{TX-AC-CM-PP}$ Max Value Violation

Problem:

According to PCIe 2.1 spec the maximum value for $V_{TX-AC-CM-PP}$ (Tx AC common mode voltage (5.0 GT/s)) is 100 mV. In some cases, the I350 may violate this maximum value. The worst case expected value is 114 mV.

Implication:

Specification violation.

Workaround:

None.

Status: A1 NoFix

9. TSYNC: Auxiliary Timestamp from SDP is Unreliable

Problem:

The SDP inputs to the timestamp logic are not properly synchronized when using AUX1. As a result, the Auxiliary Timestamp Register values and the Auxiliary Timestamp Taken bits in TSAUXC are sometimes loaded incorrectly when using AUX1

Implication:

The auxiliary timestamp feature should be considered unreliable when using AUX1. Use AUX0 whenever possible.

Workaround:

If AUX1 is used (2 aux timestamps on the same port) in applications that use the auxiliary timestamp feature to synchronize to an external clock, it might be acceptable to drop some of the samples. For such applications, software can filter out many of the incorrect timestamp values by comparing them to an approximate expected timestamp and discarding unreasonable values.

Status: A1 Fixed

10. Rx Packet May Be Dropped After Receiving 2 Bit Errors While in An Idle State

Problem:

On SGMII Rx lines, when receiving false carrier indication following an ESD (end of stream delimiter) with no idle cycles (i.e. back to back), the valid packet may be dropped by MAC. According to IEEE802.3 (section 24.3.1, Figure 24-11 ‘Receive state diagram’) such packets should not be dropped.

In order to experience such failures, the device will need to observe Rx_bit[0] and Rx_bits[9:2]!=1111111 while in IDLE state. This means 2 bit errors in 2 symbols.

The probability for this is much lower than, for example, receiving a single bit error in middle of a valid stream (which will have the same impact – frame drop at the MAC).

Implication:

When performing the End of Stream Delimiter Test (UNH: 24.1.1) - the I350 was not observed to reply to test frames containing code groups between valid ESD and the start of Idle.

Workaround:

None.

Status: A1 NoFix

11. SDPx_1 Remains an Input in D3

Problem:

The software defined pins SDP0_1, SDP1_1, SDP2_1 and SDP3_1 normally function as an input and output signal. However when the device is in D3, the listed signals behave only as inputs until the device is reset.

Implication:

If using this signal as a Tx-disable in an SFP design the SFP module will not disable its transmitting function.

Workaround:

Use an external pull-up or pull-down resistor to keep the SDP line at the intended voltage level during D3.

Status: A1 NoFix

12. NC-SI Enable Global Multicast Filter Command - No Differentiation of Neighbor Discovery Modes

Problem:

The NC-SI Enable Global Multicast Filter Command defines two separate enable bits for IPv6 Neighbor Advertisement and IPv6 Router Advertisement packets filtering. The two types of packet are differentiated by their ICMPv6 header message type.

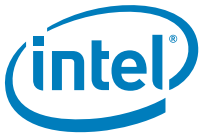
There is a single control that enables forwarding of both types of packets; the setting of one bit allows forwarding of both types of packets.

Implication:

When enabling either IPv6 Neighbor Advertisement or IPv6 Router Advertisement packet filtering, both type of packets are sent through the NC-SI interface to the BMC.

Workaround:

None.



Status: A1 NoFix

13. NC-SI: Select Package Command Increments NC-SI Commands Received Counter for Ports 0 and 1 and Not for Ports 2 and 3

Problem:

For package-based commands, the NC-SI Commands Received Counter should increase in all channels. When a select package command is received, the counter is increased only for ports 0 and 1 and not for ports 2 and 3.

Implication:

Inaccurate statistics.

Workaround:

None.

Status: A1 NoFix

14. VLAN Tags Byte-Swapped in Loopback Packets

Problem:

When operating in a Virtualization environment (SR-IOV or VMDq) and sending loopback packets, the VLAN tags in the Rx descriptors may be byte-swapped.

Implication:

Sending loopback packets, VF to VF or PF to PF on the same VLAN may result in a failure to correctly identify the packets.

Workaround:

The software driver can check the Rx descriptors for the loopback bit in the extended status field and byte swap the VLAN tags before processing the packet.

Status: A1 NoFix

15. DEV_OFF_N May Not Properly Disable the Device

Problem:

When asserting DEV_OFF_N with Power Down Enabled (0x01E:15 = 1) and a PCIe reset is taken, the device may not load the EEPROM image.

Implication:

The device may not be correctly configured after asserting the DEV_OFF_N signal and all PHYs are disabled.

Workaround:

Using the LAN_DIS_N pins is equally effective. However, if these are not available, enabling Pass-Through/Manageability and/or keeping the PHY enabled in D3 can avoid this condition. In addition, software can clear the CTRL_EXT (Extended Device Control) register, Power Down enable bit (20 = 0) or set 0x0F:6 = 0 in the EEPROM image.

Status: A1 NoFix

16. DMA Coalescing Reacts to MSI-X Vectors Instead of Interrupt Causes

Problem:

DMA coalescing should coalesce interrupt events only if interrupts reflect recurring events like packet receive or transmit events. Interrupts reflecting 'onetime events' should not be coalesced.

The DMA coalescing mechanism assumes that the first 16 MSI-X vectors reflect recurring events and the others reflect onetime events. A different mapping of events to MSI-X vector will not achieve the expected behavior.

Implication:

When DMA coalescing is activated, only specific mappings of events to MSI-X vectors are allowed.

Workaround:

Map recurring events to the first 16 vectors and other events to higher vectors.

Status: A1 NoFix

17. IEEE Std 802.3™-2008 Tx Distortion Marginality

Problem:

The I350 might not meet the IEEE Std 802.3™-2008 specification (40.6.1.2.4) that states that the Tx Distortion must meet the following criteria. "A PHY is considered to pass this test if the peak distortion is below 10 mV for at least 60% of the UI within the eye opening." The I350 might marginally fail this requirement when operating the device at a high temperature and high voltage corner.

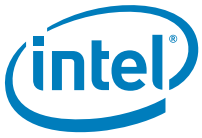
Implication

IEEE conformance is marginal.

The Tx distortion is less than 10 mV during the critical time when the signal is actually sampled therefore no impact on system performance is observed with the I350 due to this marginality.

Workaround:

None.



Status: A1 NoFix

18. Alternate MAC Address Port Order

Problem:

The datasheet specifies that the Alternate MAC Address Block in the EEPROM is ordered by PCIe function number. The firmware treats this block as being ordered by physical LAN port. These values are not identical if the LAN Function Sel bit (EEPROM word 0x21, bit 12) is 1b. In this case, the Alternate MAC Address values might be used in the wrong LAN port.

Implication:

Incorrect MAC addresses when the Alternate MAC Address Block is programmed in the EEPROM by software.

Workaround:

Program the Alternate MAC Address Block in LAN port order, rather than PCIe function order.

Status: A1 EEPROM Fix

Fixed in EEPROM version 1.52.

Contact your Intel representative to obtain updated EEPROM images.

19. Neighbor Discovery Offload: Override Bit not Set in Neighbor Advertisement Packet

Problem:

The Neighbor Advertisement packets transmitted by Powerville when providing the Neighbor Discovery protocol offload have the Override bit cleared. Since the protocol offload is functioning as a host system and not a router, the Override bit should be 1b.

Implication:

The requestor's Neighbor Cache table might not be updated correctly.

Microsoft certification for Windows 8 fails as a result of this issue.

Workaround:

None.

Status: A1 EEPROM Fix

Fixed in EEPROM v1.52.

Contact your Intel representative to obtain updated EEPROM images.

20. VFTA Register Access Might Not Be Performed

Problem:

There is a small probability that a read or write access to a VFTA register might not be performed. This can occur while traffic is being processed if any of the following settings are configured:

- OS-to-BMC and/or BMC-to-OS traffic is enabled by the BMC (NC-SI mode).
- VMDq loopback mode is enabled (MRQC.Multiple Receive Queues Enable is 011b and TXSWC.Loopback_en is 1b).
- Anti-spoofing filters are enabled (MRQC.Multiple Receive Queues Enable is 011b and any of the TXSWC.MACAS or TXSWC.VLANAS bits is set).

Implication:

VFTA reads are not reliable and the wrong data value might be returned.

VFTA writes are not reliable and the new value might not be written to the register.

Workaround:

For VFTA reads: Do not read the VFTA registers. Maintain a shadow copy of the VFTA registers in software.

For VFTA writes: Perform the write access multiple times to ensure that the value has been written with a sufficiently high probability. The probability of failure is less than 1-in-35, so 10 consecutive write accesses should bring the cumulative probability of failure to a negligible level ($<10^{-15}$).

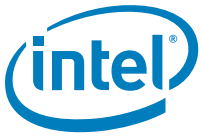
Intel LAN drivers implemented the above workaround in release 16.6 (igb v3.2.9)

An alternate workaround is to use the following procedure to write to the VFTA while packets are being transmitted.

1. Clear TCTL.EN.
2. Set DTXMXPKTSZ to 0xFFF.
3. Poll register 0x35B4 until bit (8 + LAN_ID) is 0b.
4. Write to VFTA.
5. Restore DTXMXPKTSZ value.
6. Set TCTL.EN.

Note: This workaround is not intended for use when NC-SI pass-through traffic is enabled.

Status: A1 NoFix



1.5 Software Clarification

1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB

Clarification:

The I350 supports 256 KB TCP packets; however, each buffer is limited to 64 KB since the data length field in the transmit descriptor is only 16 bits. This restriction increases driver implementation complexity if the operating system passes down a scatter/gather element greater than 64 KB in length. This can be avoided by limiting the offload size to 64 KB.

Investigation has concluded that the increase in data transfer size does not provide any noticeable improvements in LAN performance. As a result, Intel network software drivers limit the data transfer size in all drivers to 64 KB.

Please note that Linux operating systems only support 64 KB data transfers.

For further details about how Intel network software drivers address this issue, refer to Technical Advisory TA-191.

2. Serial Interfaces Programmed By Bit Banging

Clarification:

When bit-banging on a serial interface (such as SPI, I2C, or MDIO), it is often necessary to perform consecutive register writes with a minimum delay between them. However, simply inserting a software delay between the writes can be unreliable due to hardware delays on the CPU and PCIe interfaces. The delay at the final hardware interface might be less than intended if the first write is delayed by hardware more than the second write. To prevent such problems, a register read should be inserted between the first register write and the software delay, i.e. “write”, “read”, “software delay”, “write”.

§ §

NOTE: This page intentionally left blank.