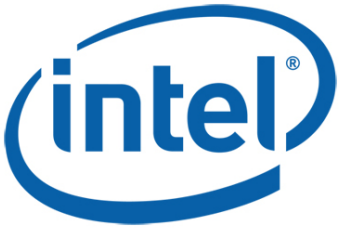

Intel® 82575EB Gigabit Ethernet Controller Specification Update

LAN Access Division (LAD)

Revision: 2.69
September 2012

Legal

Revisions

Revision	Date	Description
0.25	8/2006	Initial release.
0.30	10/2006	Added errata 4 through 14
0.40	1/2007	Added errata 15 and 16
1.00	8/2007	Removed errata fixed in current stepping, previously numbered 4-12 and renumbered the remainder. Added new errata 8 through 16.
1.1	10/2007	Added "SerDes" to application types. Corrected Fiber/SerDes Device ID to 0x10A9.
2.0	10/2008	Updated Erratum 13; added errata 17 and 18.
2.1	3/2009	Added erratum 19.
2.2	5/15/2009	Added specification clarification [2.].
2.3	2/26/2010	Added or updated the following Errata: • 21. MNG: Critical Session (Keep PHY Link Up) Mode Doesn't block All PHY Resets Caused by PCIe Resets • 20. JTAG: Instruction Register Functionality Doesn't Meet IEEE Std 1149-1-2001 • 22. PCIe: Missing Replay Due to Recovery During TLP Transmission • 23. PCIe: LTSSM Moves from L0 to Recovery Only When Receiving TS1/TS2 on All Lanes • 24. SERDES/SGMII: No Link Reported When Auto-negotiation Is Enabled from EEPROM • 25. MSI-X violation of PCIe Posted-Posted Rule • 26. NoFix. PCIe Elastic Buffer Noise Immunity Is Not Optimized • 27. PCIe: Completion Timeout Settings Not Loaded from EEPROM to GCR • 28. PCIe: Wrong Byte Enable Bit Used for Completion Timeout Disable Bit in Device Control 2 Register • 29. PCIe: Completion with UR/CA Status Causes Unexpected Completion and Completion Timeout Errors to be Reported • 30. Tx Packet Lost After PHY Speed Change Using Auto-Negotiation • 31. PCIe: Hot Reset Can Lead to a Firmware Hang
2.4	3/9/2010	25. MSI-X violation of PCIe Posted-Posted Rule - Workaround text updated.
2.5	12/3/2010	Spec Clarifications added: • 3. SerDes: AN_TIMEOUT Only Works When the Link Partner is Idle • 4. Use of Wake on LAN Together with Manageability • 5. Receiver Detection Circuit Design and Established Link Width Errata added: • 32. PCIe: Spurious SDP/STP Causes Packets to be Dropped • 33. I2C Data Out Hold Time Violation Spec Change added: • 1. Update to PBA Number EEPROM Word Format

2.6	12/14/2010	Software Clarifications: • 1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB - Moved from errata to software clarification section. Errata: • 19. LED Remains On In D3 Power State in SerDes Mode - Text updated: "The LED may remain on in D3 power state when SerDes power down is enabled (EEPROM word 0xf, **bit 2**; register CTRL_EXT 0x0018, bit18)..."
2.61	1/28/2011	Doc. title updated to match brand string. Spec. Change added: • 2. Updates to PXE/iSCSI EEPROM Words Errata added: • 34. SerDes: RXCW.RxConfigInvalid Set Incorrectly
2.62	2/1/2011	Existing Spec. Change updated: • 2. Updates to PXE/iSCSI EEPROM Words
2.63	7/20/2011	Spec Clarifications added: • 6. LED Modes Based on Link Speed Only Work in Copper (Internal PHY) Mode Spec. Changes added: • 3. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4) • 4. SerDes Forced Mode Override EEPROM Setting • 5. PCIe: Device Control 2 Register Should not be Written While DMA is Enabled • 6. Update to Section 13.7, Copper/Fiber Switch; in the Intel® 82575 Gigabit Ethernet Controller Software Developer's Manual Errata added: • 35. NC-SI: Get Link Status Command Might Cause Corruption of PHY Registers • 36. Packets Received with an L2+L3 Header Length Greater than 512 Bytes Can Incorrectly Report a Checksum Error
2.64	8/22/2011	Software Clarification added: • 2. Serial Interfaces Programmed By Bit Banging
2.65	9/23/2011	Errata added: • 36. Packets Received with an L2+L3 Header Length Greater than 512 Bytes Can Incorrectly Report a Checksum Error
2.66	12/2/2011	• Table 2, Product Code and Device Identification. Contains updated product ID information. The table replaces two older tables and introduces the A3 stepping. Note that A3 parts are functionally the same as A2 parts. The update was driven by a change in the manufacturing process.
2.67	12/13/2011	• Table 2, Product Code and Device Identification. One of the table notes was updated. The new text is: "There is no die change for parts listed as A3. There are no Form, Fit, or Function changes to this silicon. Intel anticipates no impact to customers. This is an internal package change to provide a material solution that is RoHS compliant; Intel qualified and certified this change in the same way as it does for all products supplied to customers."
2.68	4/26/2012	Spec. Changes added: • 7. CRC8 Fields of Analog Initialization Structures in the EEPROM Image are not Checked by the Device • 8. CTRL.SLU Should be Set by Software Following Device Reset
2.69	9/1/2012	Specification Clarification added: • 7. Dynamic LED Modes Can Only Be Used in an Active Low Configuration

1. Introduction

This document applies to the Intel® 82575EB Gigabit Ethernet Controller. It is an update to published documentation. All product documents are subject to frequent revision. Be sure you have the latest information before finalizing or trying to support your design.

1.1 Nomenclature

Table 1. Definitions

Name	Description
Specification Changes	Modifications to the current published specifications. These changes will be incorporated in the next release of the specifications.
Errata	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 that have not been completely confirmed or root caused.
Specification Clarifications	Greater detail or further highlights concerning a specification's impact to a complex design situation.
Documentation Changes	Typos, errors, or omissions from the current published specifications. These changes will be incorporated in the next release of the specifications.
Doc	Covered by document change or update.
Fix	This erratum is intended to be fixed in a future stepping of the component.
Fixed	This erratum has been fixed.
NoFix	There are no plans to fix this erratum.
Eval	Plans to fix this erratum are under evaluation.
(No mark) or (Blank box)	This erratum is fixed in listed stepping or specification change does not apply to listed stepping.
DS	Datasheet
DG	Design Guide
SDM	Software Developer's Manual
EDS	External Data Specification
AP	Application Note

1.2 Product Code & Device Identification

Product Code: JL82575EB (unleaded), HL82575EB (leaded).

The following tables and drawings describe markings on each package:

Table 2. Product Code and Device Identification

Device*	Step**	Top Marking		Spec	Description	Media***		MM#	
82575EB	A2	JL82575EB			Production (lead free)	T&R		890990	
82575EB	A2	JL82575EB			Production (lead free)	Tray		890991	
82575EB	A2	HL82575EB			Production (lead)	T&R		891621	
82575EB	A2	HL82575EB			Production (lead)	Tray		891622	
82575EB	A3**	JL82575EB		S LJBS	Production (lead free)	T&R		916837	
82575EB	A3	JL82575EB		S LJBT	Production (lead free)	Tray		916838	
82575GB	A3	JL82575GB****		S LJBU	Production (lead free)	T&R		916839	
82575GB	A3	JL82575GB		S LJBV	Production (lead free)	Tray		916840	

* These devices can have a "GB" marking; these devices are used only on Intel network interfaces. The "GB" is functionally equivalent to the EB version.

** **There is no die change for parts listed as A3. There are no Form, Fit, or Function changes to this silicon. Intel anticipates no impact to customers. This is an internal package change to provide a material solution that is RoHS compliant; Intel qualified and certified this change in the same way as it does for all products supplied to customers.**

*** Tray or T&R (Tape and reel).

**** Devices can have a "GB" marking; these devices are used only on Intel network interface adapters. The "GB" is functionally equivalent to the "EB" version.

Table 3. Vendor ID, Device ID, Revision ID

Device	Vendor ID	Device ID	Revision ID*
82575EB (Copper Applications)	8086	0x10A7	0x1
82575EB (Fiber/SerDes) Applications)	8086	0x10A9	0x1

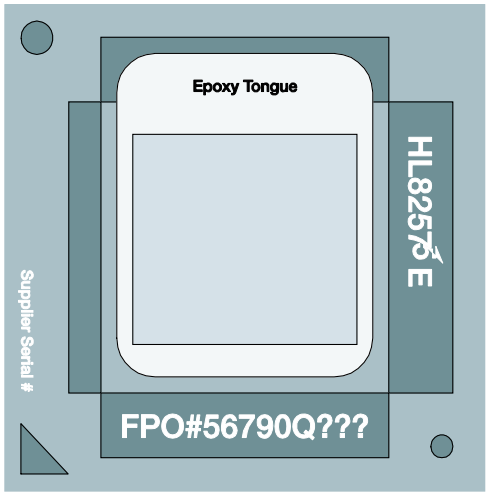

Figure 1. Device Identifying Marks (Similar)

Lead-free parts have "JL" as the prefix for the product code (vs. "HL") and that the "Q" designator refers to the Q Specification number in the table.

Devices can also have a "GB" marking; these devices are used only on Intel network interface adapters. The "GB" is functionally equivalent to the "EB" device.

There is an internal designator visible as a dash on the die side of some packages. The location may differ. The mark (looks like a dash, approximately 300X1000 µm in size) does not impact function.

2. Sightings, Clarifications, Changes, and Errata

Table 1. Summary of Sightings, Clarifications, Changes and Errata

Sightings	Status
None	
Specification Clarifications	**Status**
1. PCIe Partial Memory Write Requests Actually Writing Full DW	NA
2. PCIe* Completion Timeout Mechanism Compliance	NA
3. SerDes: AN_TIMEOUT Only Works When the Link Partner is Idle	NA
4. Use of Wake on LAN Together with Manageability	NA
5. Receiver Detection Circuit Design and Established Link Width	NA
6. LED Modes Based on Link Speed Only Work in Copper (Internal PHY) Mode	NA
7. Dynamic LED Modes Can Only Be Used in an Active Low Configuration	NA
Specification Changes	**Status**
1. Update to PBA Number EEPROM Word Format	NA
2. Updates to PXE/iSCSI EEPROM Words	NA
3. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)	NA
4. SerDes Forced Mode Override EEPROM Setting	NA
5. PCIe: Device Control 2 Register Should not be Written While DMA is Enabled	NA
6. Update to Section 13.7, Copper/Fiber Switch; in the Intel® 82575 Gigabit Ethernet Controller Software Developer's Manual	NA
7. CRC8 Fields of Analog Initialization Structures in the EEPROM Image are not Checked by the Device	NA
8. CTRL.SLU Should be Set by Software Following Device Reset	NA
Errata	**Status**
1. PCIe: Poisoned TLP Reported In All Functions Instead Of Only Target Function	A2,A3 NoFix
2. 10BASE-T: IDL Template Failure.	A2,A3 NoFix

Table 1. Summary of Sightings, Clarifications, Changes and Errata

3. 10BASE-T: Link Pulse Hits Template Mask Due to Voltage Ripple/Glitch.	A2,A3 NoFix
4. Power Management Is Version 2; Should Be Version 3	A2,A3 NoFix
5. L0s Exit Latency in Link Cap Register Is Not Updated for Common Clock Configuration.	A2,A3 NoFix
6. Wakeup Event Occurs on Magic Packet that Doesn't Pass Address Filter	A2,A3 NoFix
7. PCIe Differential Return Loss More than Specified Value	A2,A3 NoFix
8. PCIe: SKP Ordered Set Resets Training Sequence Counter	A2,A3 NoFix
9. Preamble Error Reception and Recovery	A2,A3 NoFix
10. Device Doesn’t Correctly Implement Master/Slave Resolution	A2,A3 NoFix
11. Device Improperly Implements Auto-Negotiation Advertisement Register	A2,A3 NoFix
12. SGMII MODE: Counters Incorrectly Increment on Collision	A2,A3 NoFix
13. Link LED Remains On after System Power Down and Cable Removed	A2,A3 NoFix
14. Improperly Implements Sample Timer	A2,A3 NoFix
15. Uses INTB Line When in Single-Function Mode	A2,A3 NoFix
16. In Small Number of Cases, Sends “Abilities” Instead of “Break Link”	A2,A3 NoFix
17. Corrupted Packets in Rx PB When RXEN Is 0	A2,A3 NoFix
18. PCIE: Reception of Completion That Should Be Dropped Can Occasionally Cause Device Hang or Corruption	A2,A3 NoFix
19. LED Remains On In D3 Power State in SerDes Mode	A2,A3 NoFix
20. JTAG: Instruction Register Functionality Doesn't Meet IEEE Std 1149-1-2001	A2,A3 NoFix
21. MNG: Critical Session (Keep PHY Link Up) Mode Doesn’t block All PHY Resets Caused by PCIe Resets	A2,A3 NoFix
22. PCIe: Missing Replay Due to Recovery During TLP Transmission	A2,A3 NoFix
23. PCIe: LTSSM Moves from L0 to Recovery Only When Receiving TS1/TS2 on All Lanes	A2,A3 NoFix
24. SERDES/SGMII: No Link Reported When Auto-negotiation Is Enabled from EEPROM	A2,A3 NoFix
25. MSI-X violation of PCIe Posted-Posted Rule	A2,A3 NoFix
26. NoFix. PCIe Elastic Buffer Noise Immunity Is Not Optimized	A2,A3 NoFix
27. PCIe: Completion Timeout Settings Not Loaded from EEPROM to GCR	A2,A3 NoFix
28. PCIe: Wrong Byte Enable Bit Used for Completion Timeout Disable Bit in Device Control 2 Register	A2,A3 NoFix
29. PCIe: Completion with UR/CA Status Causes Unexpected Completion and Completion Timeout Errors to be Reported	A2,A3 NoFix
30. Tx Packet Lost After PHY Speed Change Using Auto-Negotiation	A2,A3 NoFix
31. PCIe: Hot Reset Can Lead to a Firmware Hang	A2,A3 NoFix
32. PCIe: Spurious SDP/STP Causes Packets to be Dropped	A2,A3 NoFix
33. I2C Data Out Hold Time Violation	A2,A3 NoFix

Table 1. Summary of Sightings, Clarifications, Changes and Errata

34. SerDes: RXCW.RxConfigInvalid Set Incorrectly	A2,A3 NoFix
35. NC-SI: Get Link Status Command Might Cause Corruption of PHY Registers	A2,A3 NoFix
36. Packets Received with an L2+L3 Header Length Greater than 512 Bytes Can Incorrectly Report a Checksum Error	A2,A3 NoFix
37. NC-SI: Get NC-SI Pass-through Statistics Response Format	A2,A3 NoFix
Specification Clarifications	**Status**
1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB	NA
2. Serial Interfaces Programmed By Bit Banging	NA

2.1 Sightings

None.

2.2 Specification Clarifications

1. PCIe Partial Memory Write Requests Actually Writing Full DW

Clarification: The PCIe specification allows a device to not accept certain requests. This is under the "programming model" case. A device needs to issue a Completer Abort error if the specific request violates the programming model. As part of its programming model, the 82575EB does not support writes and reads with Byte enables to specific memory addresses. Such writes will be full executed and will not be treated as completion abort.

CSR writes and reads with partial (or zero) Bytes Enables will be executed (In specific address ranges). This scenario will not happen when using the device driver. This functionality is also not needed for the normal operation of the design.

Note: This can be avoided by having no partial or zero bytes enable writing to the 82575EB.

2. PCIe* Completion Timeout Mechanism Compliance

Clarification: PCIe Completion Timeout Value Must Be Properly Set

The Completion Timeout Value(3:0) must be properly set by the system BIOS in PCIe Configuration Space Device Control 2 Register (0xC8; RW). Failure to do so can cause unpredictable system behavior.

The 82575EB 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 82575EB defaults to a range of 500us – 1ms for PCIe capabilities version 1 and 2. The PCIe 2.0 Specification also strongly recommends that the default timeout value be such that the completion timeout mechanism not expire in less than 10ms.

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 device is installed. This will ensure that the 82575EB 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. In the event of a completion timeout, the device assumes the original completion is lost, and resends the

original request, by default. In this condition, if the completion for the original request arrives at the 82575EB, this will result in 2 completions arriving for the same request, which may cause unpredictable system behavior.

As long as the Completion Timeout value is properly programmed by the system the completion timeout mechanism works without issue. It is expected that the system BIOS will set this value appropriately for the system.

Workaround: Alternatively a device driver could ensure the completion timeout value is set above 10ms (in order to follow the recommendation of the PCIe 2.0 specification). The driver would modify the timeout value, if and only if the default timeout value remains in configuration space. This will not impact BIOSs already changing the timeout value since the driver will not override any non-default setting of the timeout value. For extra protection against unpredictable system behavior in case the timeout setting is incorrect, it is recommended to disable the resend of the request. This can be done by clearing the Completion_Timeout_Resend bit in the GCR Register.

The latest Intel drivers implement this workaround by modifying the completion timeout value in config space if the timeout value is still set to a value of 0x0 when the driver loads. They also clear the Completion_Timeout_Resend bit in the GCR Register.

Release14.4 includes this fix.

3. SerDes: AN_TIMEOUT Only Works When the Link Partner is Idle

Clarification: The auto-negotiation timeout mechanism (PCS_LCTL.AN_TIMEOUT_EN) only works if the SerDes partner is sending idle code groups continuously for the duration of the timeout period, which is the usual case. However, if the partner is transmitting packets, an auto-negotiation timeout will not occur since auto-negotiation is restarted at the beginning of each packet. If the partner has an application that indefinitely transmits data despite the lack of any response, it is possible that a link will not be established. If this is a concern, the auto-negotiation timeout mechanism may be considered unreliable and an additional software mechanism could be used to disable auto-negotiation if sync is maintained without a link being established (PCS_LSTS.SYNC_OK=1b and PCS_LSTS.LINK_OK=0b) for an extended period of time.

Note: To Disable Auto-Negotiation when using an Intel driver, use the SerDes Forced Mode Override bit as described in Specification Change #4.

4. Use of Wake on LAN Together with Manageability

Clarification: The Wakeup Filter Control Register (WUFC) contains the NoTCO bit, which affects the behavior of the wakeup functionality when manageability is in use. Note that if manageability is not enabled, the value of NoTCO has no effect.

When NoTCO contains the hardware default value of 0b, any received packet that matches the wakeup filters will wake the system. This could cause unintended wakeups in certain situations. For example, if Directed Exact Wakeup is used and the manageability shares the host's MAC address, IPMI packets that are intended for the BMC wake the system, which might not be the intended behavior.

When NoTCO is set to 1b, any packet that passes the manageability filter, even if it also is copied to the host, is excluded from the wakeup logic. This solves the previous problem since IPMI packets do not wake the system. However, with NoTCO=1b, broadcast

packets, including broadcast magic packets, do not wake the system since they pass the manageability filters and are therefore excluded.

Table 2. Effects of NoTCO

WoL Type	NoTCO	Shared MAC Address	Unicast Packet	Broadcast Packet
Magic Packet	0b	-	OK	OK
Magic Packet	1b	Y	No wake	No wake
Magic Packet	1b	N	OK	No wake
Directed Exact	0b	Y	Wake even if MNG packet. No way to talk to BMC without waking host.	N/A
Directed Exact	0b	N	OK	N/A
Directed Exact	1b	-	OK	N/A

Intel Windows* drivers set NoTCO by default.

If this is not the desired behavior, the EnableWakeOnManagmentOnTCO registry entry can be used to change it starting with Intel LAN driver software release 15.5. Setting this registry entry to 1b causes the driver to program NoTCO to 0b. A tool to modify the registry entry can be provided.

Contact Intel for access.

5. Receiver Detection Circuit Design and Established Link Width

Clarification: The receiver detection circuit was designed according to the PCIe Specification Rev. 1.1, which requires that an un-terminated receiver have an input impedance of at least 200 Kohm. PCIe Specification Rev. 2.0 allows the input impedance to be as low as 1 Kohm at input voltages in the range -150 - 0 mV and does not specify a minimum input impedance below -150 mV. As a result, a powered-down receiver lane with low input impedance at negative voltages could be compliant to Rev 2.0 and yet be falsely detected by the device as a terminated lane.

This is normally not an issue since any connnected lanes should be properly terminated within 5 ms after fundamental reset according to the PCIe Specification. However, there are some chipset devices that require significantly more time to prepare the termination and expect the link partner to remain in the LTSSM Detect state as long as none of the lanes are terminated. When used with such devices, the 82575EB might falsely detect a receiver on one or more lanes and leave the Detect state. This can lead to establishing a link that is less than full width.

In this case, it is recommended that a Hot Reset be performed after a link has been established in order to force the 82575EB to detect the receivers again when they are properly terminated. As a result, a full-width link can be established.

6. LED Modes Based on Link Speed Only Work in Copper (Internal PHY) Mode

Clarification: LED modes based on LINK speed work only in copper mode, not SerDes/SGMII modes. This includes the modes LINK_10/1000. LINK_100/1000, LINK_10, LINK_100, LINK_1000 and COLLISION. Designs using SerDes and SGMII modes requiring a Link-up indication should use LINK_UP or LINK/ACTIVITY LED modes. Using these modes results in no issues in using the LEDs to properly indicate the link is up.

7. Dynamic LED Modes Can Only Be Used in an Active Low Configuration

Clarification: In any of the dynamic LED modes (FILTER_ACTIVITY, LINK/ACTIVITY, COLLISION, ACTIVITY, PAUSED), LED blinking should only be enabled if the LED signal is configured as an active low output.

2.3 Specification Changes

1. Update to PBA Number EEPROM Word Format

Change: PBA Number Module — Word 0x8-0x9

The nine-digit Printed Board Assembly (PBA) number used for Intel manufactured Network Interface Cards (NICs) is stored in EEPROM.

Through the course of hardware ECOs, the suffix field is incremented. The purpose of this information is to enable customer support (or any user) to identify the revision level of a product.

Network driver software should not rely on this field to identify the product or its capabilities.

PBA numbers have exceeded the length that can be stored as HEX values in two words. For newer NICs, the high word in the PBA Number Module is a flag (0xFAFA) indicating that the actual PBA is stored in a separate PBA block. The low word is a pointer to the starting word of the PBA block.

The following shows the format of the PBA Number Module field for new products.

PBA Number	Word 0x8	Word 0x9
G23456-003	FAFA	Pointer to PBA Block

The following provides the format of the PBA block; pointed to by word 0x9 above:

Word Offset	Description
0x0	Length in words of the PBA Block (default is 0x6)
0x1 ... 0x5	PBA Number stored in hexadecimal ASCII values.

The new PBA block contains the complete PBA number and includes the dash and the first digit of the 3-digit suffix which were not included previously. Each digit is represented by its hexadecimal-ASCII values.

The following shows an example PBA number (in the new style):

PBA Number	Word Offset 0	Word Offset 1	Word Offset 2	Word Offset 3	Word Offset 4	Word Offset 5
G23456-003	0006	4732	3334	3536	2D30	3033
	Specifies 6 words	G2	34	56	-0	03

Older NICs have PBA numbers starting with [A,B,C,D,E] and are stored directly in words 0x8-0x9. The dash in the PBA number is not stored; nor is the first digit of the 3-digit suffix (the first digit is always 0b for older products).

The following example shows a PBA number stored in the PBA Number Module field (in the old style):

PBA Number	Byte 1	Byte 2	Byte 3	Byte 4
E23456-003	E2	34	56	03

2. Updates to PXE/iSCSI EEPROM Words

Change: Words 0x30, 0x34; bits 2:0 and bit 5 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.				

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

Change: Port Identification LED blinking (Word 0x04)

Driver software provides a method to identify an external port on a system through a command that causes the LED's to blink. Based on the setting in word 0x4, the LEDs 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. SerDes Forced Mode Override EEPROM Setting

Change: Bit 14 in EEPROM Compatibilty Word 0x3 is no longer reserved; it is defined as follows:

Bit	Name	Default	Description
14	SerDes Forced Mode Override	0	SerDes Forced Mode Enable: 0 = Normal operation Intel Driver will enable PCS_LCTL.AN_ENABLE 1 = Forced Mode enable. Intel Driver will not set PCS_LCTL.AN_ENABLE

This bit is used by Intel drivers starting with Intel LAN Driver Release 16.4. This bit is read by the Intel driver and used to determine if the device should be operated in SerDes forced mode. When the 82575EB is not in SerDes mode, the bit has no effect. The recommended setting for this bit is 0.

Note: Use the SerDes Forced Mode Override bit to allow the driver to disable Auto-Negotiation, in order to work around the issue described in Specification Clarification #3.

5. PCIe: Device Control 2 Register Should not be Written While DMA is Enabled

Change: The Device Control 2 Register should only be written during initialization. When a port is enabled to transmit or receive data this register should not be written, even if the value is not changed.

6. Update to Section 13.7, Copper/Fiber Switch; in the Intel® 82575 Gigabit Ethernet Controller Software Developer's Manual

Search for "Internal PHY to SerDes Transition" procedure. Use the updated steps below:

1. Disable the receiver by clearing RCTL.RXEN.
2. Disable the transmitter by clearing TCTL.EN.
3. Ensure Smart Power Down is not enabled in the PHY. EEPROM word 0xF bit 1 or PHY register 25d bit 0.
4. Verify the 82575EB has stopped processing outstanding cycles and is idle.
5. Set CTRL.SPEED=10, CTRL.FRCSPD=1, CTRL_EXT.SPD_BYPS=1.
6. Modify LINK mode to SerDes or SGMII by setting CTRL_EXT.LINK_MODE to 11b or 10b, respectively.
7. Delay a minimum of 10-20µs.
8. Clear CTRL.FRCSPD, CTRL_EXT.SPD_BYPS.
9. Set up the link as described in Section 3.7.3, MAC/SerDes Link Setup (CTRL_EXT.LINK_MODE = 11b)* or Section 3.7.4, MAC/SGMII Link Setup (CTRL_EXT.LINK_MODE = 10b)*.
10. Set up Tx and Rx queues and enable Tx and Rx processes.

* Links (section numbers) can change in a specific document revision. Use the links provided by the revision you are using.

7. CRC8 Fields of Analog Initialization Structures in the EEPROM Image are not Checked by the Device

Changes: Section 6.4 of the datasheet describes the analog initialization structures. The CRC8 fields of these structures are not checked by the device. The CRC_DIS EEPROM bit (word 0x23, bit 6) must be set to 1b.

8. CTRL.SLU Should be Set by Software Following Device Reset

Change: As documented, the CTRL.SLU bit is cleared during reset and then set to 1b during EEPROM auto-load if the APM_Enable EEPROM bit is 1b.

Also as documented, the APM_Enable bit is in word 0x14/0x24 and is not read following a device reset initiated by software writing to CTRL.RST.

Therefore, CTRL.SLU is not automatically set following a software device reset. This bit should be explicitly set by software in order to establish a link.

Exceptions:

- If manageability is enabled, CTRL.SLU is set automatically.
- If a PHY reset is performed after a device reset, CTRL.SLU is set automatically.

2.4 Errata

1. PCIe: Poisoned TLP Reported In All Functions Instead Of Only Target Function

Problem: A fatal message is incorrectly sent in response to a poisoned Transaction Layer Packet (TLP).

Implication: The 82575EB treats all poisoned memory requests as non-function specific. Instead of reporting in the target function, a fatal error is reported in all functions. The correct action is to report poisoned requests per function.

Workaround: None.

Status: A2,A3 NoFix

2. 10BASE-T: IDL Template Failure.

Problem: The 10base-T TP_IDL waveform fails the template test on twisted-pair model combined with test load 2.

Implication: There is no impact on system level performance.

Workaround: None.

▼ *Return to Summary*

Status: A2,A3 NoFix

3. 10BASE-T: Link Pulse Hits Template Mask Due to Voltage Ripple/Glitch.

Problem: The 10base-T link pulse touches the specification template due to voltage ripple/glitch.

Implication: There is no effect at system level.

Workaround: None.

Status: A2,A3 NoFix

4. Power Management Is Version 2; Should Be Version 3

Problem: Power management version in PMC register is “2” instead of “3.”

Implication: No impact at functional system level.

Workaround: None.

Status: A2,A3 NoFix

5. L0s Exit Latency in Link Cap Register Is Not Updated for Common Clock Configuration.

Problem: The L0s Exit Latency from theLink Capabilities register (offset 0xAC) will be the same whether the Link Control reg. Common Clock Configuration bit is set or not.

Implication: This function will remain with the default, which is the non-common mode.

Workaround: None.

Status: A2,A3 NoFix

6. Wakeup Event Occurs on Magic Packet that Doesn't Pass Address Filter

Problem: The 82575EB receives a magic packet that didn't pass address filtering. The 82575EB will generate a wakeup event at the next packet, if the next received packet (non-magic packet) is accepted according to address filtering scheme.

Implication: The 82575EB may wake the system on a non-wakeup packet.

Workaround: None.

Status: A2,A3 NoFix

7. PCIe Differential Return Loss More than Specified Value

Problem: The PCIe transmitter's differential return loss is -9.7 dB instead of the -10 dB requirement.

Implication: The out-of-specification return loss adds noise to the Tx line. Performance is not affected.

Workaround: None.

Status: A2,A3 NoFix

8. PCIe: SKP Ordered Set Resets Training Sequence Counter

Problem: If a SKP ordered set is received during a TS1 or TS2 sequence, the TS counter is cleared. This will generally not be a problem since the upstream device should transmit at least 16 TS2 ordered sets, and the 82575EB only needs to detect 8 consecutive TS2 ordered sets to complete the Recovery process, so a single reset of the counter will not cause a failure. A failure can occur if the upstream device is non-compliant and transmits fewer than 16 TS2 ordered sets. In this case, the 82575EB could fail to compete the Recovery process and then the PCIe link would go down.

Implication: There should be no failure when the upstream device functions according to the PCIe spec. If the upstream device is non-compliant, this issue could result in a Surprise Down error.

Workaround: None.

Devices should recover.

Status: A2,A3 NoFix

9. Preamble Error Reception and Recovery

Problem: Requirement is for the 82575EB to be tolerant of the content of the preamble, that is, any content should be accepted as long is it is not a SFD (Start of Frame Delimiter). The 82575EB improperly discards frames with variations in the preamble pattern.

Implication: This problem will only be seen if the link partner is sending a frame with incorrect preambles.

Workaround: None.

Status: A2,A3 NoFix

10. Device Doesn't Correctly Implement Master/Slave Resolution

Problem: For some configurations, the 82575EB resolves to Master instead of Slave (even though it is in forced-Slave mode). This condition can be detected during the initiation of training by the 82575EB. This indicates that the 82575EB did not remain silent (defined in Slave mode) and that the 82575EB resolved to Master mode.

Implication: The equipment that is used in these tests will identify activity as 1 Gb activity if it senses any activity on all four channels. Also, it does not identify 1 Gb idles as 1 Gb training signals. Monitoring the line showed that no training was initiated by the 82575EB when the it was resolved to Slave mode.

The activity detected is illegal data that the 82575EB transmits during the transition from 10Mbps mode (Auto-Negotioation) to 1000Mbps mode (after Master/Slave resolution is completed). Internal indications (the PHY registers) show that the 82575EB complies with IEEE 802.3, Table 40-5; for any configuration, the 82575EB resolved to the correct defined mode.

Workaround: None.

Status: A2,A3 NoFix

11. Device Improperly Implements Auto-Negotiation Advertisement Register

Problem: The 82575 improperly transmits the Link Code Word due to a write to register 4. The Link Code Word improperly switch immediately, which corresponds to a write to register 4. Link Code Word bits behaved as required with the following notes.

Implication: Bits 4.7 and 4.8: Always set in the base page transmission.

Bit 4.9: This bit represents 100BASE-T4 support by the local device. The 82575 does not support T4. It is unlikely that the Auto-Negotiation feature of the 82575EB would be used in an implementation to advertise the presence of a separate T4 physical device within the system implementation. Therefore, the fact that this device does not allow T4 to be advertised is insignificant.

Bit 4.15: the 82575EB always supports Next Page (regardless the value of bit 4.15). When bit 4.15 is set to "one," the 82575EB requires Register 7 (AN Next Page Transmit Register) to be written to complete the Next Page Exchange. In this case however, the 82575EB's Next Pages do not correspond to Register 7, but contain valid 1000BASE-T Next Pages.

Workaround: Any write to register 4 should be followed with a restart of Auto-Negotiation by setting bit 0.9.

Status: A2,A3 NoFix

12. SGMII MODE: Counters Incorrectly Increment on Collision

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

Name	Definition	Location
RLEC	Lenght error counter	0X4040

CRCERRS	CRC error counter	0x4000
RFC	receive frame counter	0x40A8

Implication: Error counters may not be accurate.

Workaround: None.

Status: A2,A3 NoFix

13. Link LED Remains On after System Power Down and Cable Removed

Problem: The link-up LED will be set during link-down if a software reset (CSR 0X0000 bit number 26) occurs. The LED will remain on even though power has been removed. This occurs only when the Software Power Down (SPD) is enabled in the EEPROM. The *Smart Power Down Enable* bit is word 0xF, bit 1, which is set by default to 1, which enables SPD.

Implication: The LED will show link-up even when link is down.

Workaround: Disable SPD.

Status: A2,A3 NoFix

14. Improperly Implements Sample Timer

Problem: The Auto-Crossover State Machine (Auto-MDIX) has two states: MDI_MODE and MDI-X_MODE. The time that should be spent in each mode is defined as a multiple of a pseudo-random number and A sample timer, which is defined to be 62 ms ± 2 ms. This violation occurs in ~10% of the runs.

Implication: The time that the PHY is in MDI mode will have a slight deviation from the specified definition.

Workaround: None.

Status: A2,A3 NoFix

15. Uses INTB Line When in Single-Function Mode

Problem: Problem occurs when the 82575EB is configured with lan 0 to use INTA and lan 1 to use INTB. When lan0 is disabled, there is a switch between the functions and func0 works with lan 1 (single function mode). When interrupt pin register address (0x3D) is read, it shows that the function use INTB, which violates the PCI-e standard of one interrupt line for a single-function device or connector. For a single-function device only INTA# can be used (the other three interrupt lines, used in multi-function mode, have no meaning).

Implication: Interrupt is tagged with INTB instead of INTA.

Workaround: When working with LAN0 disable, change EEPROM Word 0x14, bit 12 and bit 11 to “0”.

Status: A2,A3 NoFix

16. In Small Number of Cases, Sends "Abilities" Instead of "Break Link"

Problem: When the 82575EB is receiving continuous ill-formed /C/ order sets, it should transmit "break link" and not "abilities." During testing, about 14% of the time, the 82575EB sent "abilities" following the link timer expiration.

Implication: Minor specification violation—no system-level performance impact.

Workaround: None.

Status: A2,A3 NoFix

17. Corrupted Packets in Rx PB When RXEN Is 0

Problem: Due to a bug in the receive packet buffer control logic, a corrupted packet or partial packet may be written to the packet buffer before RCTL.RXEN is set. Once RXEN has been set, the first packet provided to the host may have an invalid descriptor and/or corrupted data. In some cases, the packet buffer control is also corrupted, and invalid packets will continue to be transferred to the host indefinitely. This problem occurs when either of the following conditions is true: Manageability is enabled or Wake on LAN is enabled

Implication: When a single corrupted packet is received, higher layers of the protocol stack will usually discard it, and the issue will not affect the application. However, when running Linux, an invalid descriptor that indicates a packet length larger than the MTU can cause a kernel panic. If the packet buffer control is corrupted, the receive path will be unusable until it is reset.

Workaround: Implement the following modifications to the initialization sequence: 1. Before setting RCTL.RXEN: - Clear RXDCTL.ENABLE for all queues that are in use. - Write 0x0000 to RLPML.LPML. - Set RCTL.LPE and clear RCTL.SBP. - Clear RFCTL.LEF. 2. Set RCTL.RXEN. 3. Wait 2 ms. 4. Restore RXDCTL, RLPML, RCTL, and RFCTL to the intended initial values.

Status: A2,A3 NoFix

Intel drivers implement the SW workaround.

18. PCIE: Reception of Completion That Should Be Dropped Can Occasionally Cause Device Hang or Corruption

Problem: This erratum can occur when the 82575EB's PCIe logic receives a completion that should be dropped while the 82575EB is starting a new request with the same TAG as the completion.

On an error-free PCIe link, this situation should never occur (since the 82575EB does not assert a second request with the same tag as an outstanding request).

Errors that could cause this failure include:

- The TAG of a completion is corrupted due to noise on the line. This completion packet will be dropped due to LCRC error, but could cause a failure if by chance a new request is asserted with the corrupted TAG value at the same time.
- On some platforms, it has been observed that when the upstream switch port transitions the link to L0s, the 82575EB occasionally responds with a NAK as a result of noise on the line. This NAK can cause a completion to be replayed. The 82575 will drop the duplicate packet based on the sequence number. However, the failure can occur if a new request is being asserted with the same TAG as the duplicate completion.

- An edge case of ACK timers results in a replay of a completion. This can cause the same case as above.

Implication: When the failure occurs, the actual completion data from the new request will be corrupted. The implications of this corruption of the read data depend on the type of request the 82575EB was starting to send. The type and result are listed below:

- TX descriptor with TSO – the 82575EB offload machine may hang.
- TX data or Tx descriptor without offload – the 82575EB will transmit a packet on the network with invalid data but a valid CRC
- RX descriptor – the 82575EB will DMA a receive packet to the wrong Memory address.

Workaround: Disabling L0s in the switch port to which the 82575EB is connected will prevent the duplicate completions caused by L0s.

Keeping bit 13 "ACK/NACK Scheme", word 0x1A " PCIe Initialization Configuration 3" set to 0 in the EEPROM image will minimize the chances of an ACK timeout.

Status: A2,A3 NoFix

19. LED Remains On In D3 Power State in SerDes Mode

Problem: The LED may remain on in D3 power state when SerDes power down is enabled (EEPROM word 0xf, **bit 2**; register CTRL_EXT 0x0018, bit18). If a link is established when the device enters the D3 power state and the LED mode is programmed to reflect LINK indication, the LED remains on, even though the SerDes interface powers down.

Implication: The LED incorrectly reflects link is up when there is no link(as SerDes is powered off).

Workaround: Clear PCS_CFG.PCS Enable(0x4200, bit 3) before putting a function in D3. This will bring the link down and turn off the LED; this bit will be reset when the device transitions back to D0. When using fiber adapters from Intel that are based on the 82575EB, the driver should disable optics (when going to D3) by setting SDP3 to 1.

Status: A2,A3 NoFix

20. JTAG: Instruction Register Functionality Doesn't Meet IEEE Std 1149-1-2001

Problem: If UPDATE_IR directly follows CAPTURE_IR, the active instruction is DEVID. This is not as specified by the INSTRUCTION_CAPTURE attribute.

Implication: The value from CAPTURE_IR cannot be relied upon.

Workaround: Shift in an instruction; do not rely on the value from CAPTURE_IR.

Status: A2,A3 NoFix

▼ *Return to Summary*

21. MNG: Critical Session (Keep PHY Link Up) Mode Doesn't block All PHY Resets Caused by PCIe Resets

Problem: D3 to D0 transition will cause a PHY reset even in Keep PHY Link Up mode. When Critical Session Mode (Keep PHY Link Up) is enabled (via the NC-SI Set Intel Management

Control command or the SMBUS Management Control command), PCIe resets should not cause a PHY reset. However, the following event will still cause a PHY reset:

- Transition from D3 to D0 without general PCIe reset (i.e. PMCSR[1:0] changed from 11 to 00 by configuration write).

Implication: Loss of link can cause a loss of the MNG session. These events do not normally occur during a reboot cycle, so no effect will be seen in most circumstances.

Workaround: None; not necessary since problem does not occur under normal conditions.

Status: A2,A3 NoFix

22. PCIe: Missing Replay Due to Recovery During TLP Transmission

Problem: If the replay timer expires during the transmission of a TLP and the LTSSM moves from L0 to Recovery during the transmission of the same TLP, the expected replay does not occur. Additionally, the replay timer is disabled, so no further replays will occur unless a NAK is received.

Implication: This situation should not occur during normal operation. If it does occur while the upstream switch is waiting for a replay, the result would be a Surprise Down error which might halt the system.

Workaround: Not needed.

Status: A2,A3 NoFix

23. PCIe: LTSSM Moves from L0 to Recovery Only When Receiving TS1/TS2 on All Lanes

Problem: According to the PCIe specification, the LTSSM should move from L0 to Recovery if a TS1 or TS2 ordered set is received on any configured Lane. The 82575 LTSSM only moves from L0 to Recovery if a TS1 or TS2 ordered set is received on all configured lanes.

Implication: This situation should not occur during normal operation since the upstream switch will transmit the TS1 or TS2 ordered sets on all lanes at the same time. If it does occur due to a broken lane, the result would be a Surprise Down error which might halt the system.

Workaround: None; not necessary since problem does not occur under normal conditions.

Status: A2,A3 NoFix

24. SERDES/SGMII: No Link Reported When Auto-negotiation Is Enabled from EEPROM

Problem: When using SerDes or SGMII mode and enabling auto-negotiation by setting the ANE bit in EEPROM word 0x0F, no Ethernet traffic can be passed on the link. Although a link is established, both the LEDs and the Status register indicate that the link is down.

Implication: ANE must be 0 in the EEPROM. The 82575EB cannot communicate over SerDes or SGMII before a driver is loaded to enable auto-negotiation and establish a link. Specifically, the BMC does not have access to the LAN before the driver is loaded.

Workaround: EEPROM dev starter version 2.1 and above have workaround.

Status: A2,A3 NoFix

25. MSI-X violation of PCIe Posted-Posted Rule

Problem: According to the PCIe Specification: "the acceptance of a Posted Request must not depend upon the transmission of any TLP from that same Upstream Port within the same traffic class. " The 82575EB has a dependency between downstream posted requests to its MSI-X table and upstream MSI-X packets (MSI-X interrupt messages) that differs from this rule.

Implication: Under specific stress scenarios, the upstream device might stop providing posted credits to the 82575EB. If the 82575EB has a MSI-X message to send out and it runs out of posted credits, any upstream device access to the MSI-X table (read/write) does not complete until credits are renewed. Under this condition, the 82575EB stops releasing posted credits to the upstream device, and posted data transfer stops in both directions resulting in a link deadlock. If the upstream device is able to renew its credit release flow, the 82575EB is not susceptible to this erratum.

Workaround: Use MSI instead of MSI-X interrupts. This can be accomplished via registry edits in Windows*.

Intel has a tool that automatically makes the required registry edits in Windows.

For Linux* this can be accomplished with added parameters at driver load by modifying InterruptType in /etc/modprobe.conf. InterruptType=0,0 means set both port 0 and port 1 to legacy interrupts (1,1 is MSI for both ports, 2,2 is MSI-X for both ports). Full details can be found in the Linux driver README.

Status: A2,A3 NoFix

26. NoFix. PCIe Elastic Buffer Noise Immunity Is Not Optimized

Problem: The PCIe elastic buffer is used to synchronize between the clock generated by the clock recovery circuit and the internal clock. During electrical idle, in the absence of an input signal, the clock recovery circuit can be disturbed by noise and move the elastic buffer fill level away from the optimum value. In the 82575EB, EEPROM control bits were implemented to maintain stability during electrical idle. In the default EEPROM image provided for the 82575EB, these bits were not set correctly.

Implication: In cases of increased noise levels during Electrical Idle, elastic buffer instability may, in rare instances, cause a link training failure when exiting from L1 state. Failure to exit L1 results in a Surprise Down error which may be a fatal error in operating systems that fully support Advanced Error Reporting. This issue is not relevant to L0s exit since L0s should be disabled in the downstream direction due to another erratum.

Workaround: Surprise Down Error reporting can be masked in the system. The system will recover after the link is re-established.

Workaround implemented in revised EEPROM version number V2.2 for the 82575EB; setting PCIe Init Configuration 3 word - Bits 4 & 5.

Status: A2,A3 NoFix

27. PCIe: Completion Timeout Settings Not Loaded from EEPROM to GCR

Problem: The following GCR fields should be loaded from EEPROM word 0x15 but are not:

- Completion_Timeout_Resend
- Completion_Timeout_Value (capability version 1 only)

- Completion_Timeout_Disable (capability version 1 only)

Implication: EEPROM settings are ignored. Hardware default values are used.

Workaround: Software should set the values in the GCR if non-default values are required. This can be done by the BIOS or the driver.

Status: A2,A3 NoFix

28. PCIe: Wrong Byte Enable Bit Used for Completion Timeout Disable Bit in Device Control 2 Register

Problem: BE[1] is used to enable the write to the Completion Timeout Disable bit in Device Control 2 register in the configuration space. It should be BE[0] since it is bit 4 in the register.

Implication: If a byte write is used, this bit is not updated since BE[1] is 0b.

The bit could be incorrectly written if a byte write to the high byte is performed. However, this is unlikely since bits 15:8 are all reserved.

Workaround: Use only word or Dword accesses to the Device Control 2 register.

Status: A2,A3 NoFix

29. PCIe: Completion with UR/CA Status Causes Unexpected Completion and Completion Timeout Errors to be Reported

Problem: When the 82575EB receives a PCIe completion with Unsupported Request (UR) or Completer Abort (CA) status in response to a request it generated, it reports an Unexpected Completion error. Because the completion timer is not disabled, a completion timeout error is reported when the timer expires.

Implication: This situation should not occur in systems that are operating correctly; since all requests generated by the 82575EB are supported.

If an UR/CA completion is received, the completion timeout error can bring down the operating system when it is reported.

Workaround: Not required for systems that are operating correctly. Note that reporting completion timeout errors can be masked in the Uncorrectable Error Mask register.

Status: A2,A3 NoFix

30. Tx Packet Lost After PHY Speed Change Using Auto-Negotiation

Problem: If the PHY establishes a link at 10/100 Mb/s and then auto-negotiation is re-started and a link is established at 1 Gb/s without resetting the PHY in between, the first 1-to-3 Tx packets provided by the MAC might not be transmitted.

Implication: This situation is generally seen during testing where the speed of the link partner is intentionally changed.

During normal operation, the packet loss could occur if the cable was moved to a different port. In most cases, the higher layers would handle the packet loss and it would not be visible to the end user.

Workaround: If it is critical that no packets be lost, the software driver could be modified to perform a PHY reset each time it is notified of a speed change.

Status: A2,A3 NoFix

31. PCIe: Hot Reset Can Lead to a Firmware Hang

Problem: A PCIe hot reset prevents the firmware from accessing internal registers. When the hot reset occurs while the firmware is performing initializations, the firmware might fail to clear a semaphore bit. As a result, the firmware goes into an infinite polling loop when the next initialization sequence occurs.

Implication: PHY initialization from the EEPROM is performed by firmware, so if the firmware hangs, the initialization is not performed and the Ethernet link might not operate correctly.

Additionally, the software driver might fail to load as a result of a hung semaphore bit.

When using manageability, the firmware hang causes a watchdog timer to expire and then resets the firmware, requiring re-initialization of the manageability.

Workaround: Use new EEPROM images that include a firmware workaround for this issue; 82575 2.5 or later.

Also, do not keep the device in hot reset for more than 20 ms. Hot reset should be performed by setting and then immediately clearing the Secondary Bus Reset bit in the upstream PCIe device.

Status: A2,A3 NoFix

32. PCIe: Spurious SDP/STP Causes Packets to be Dropped

Problem: When a spurious SDP or STP symbol is received without a corresponding END symbol, the alignment of the received data presented to the link layer might be incorrect. In this case, any following DLLPs or TLPs are dropped. This situation continues until there is an END symbol received that is not immediately followed by an SDP or STP symbol.

During normal operation, the SKP Ordered Sets that are inserted between packets guarantee that the proper alignment is restored within a short time.

This issue only occurs when the PCIe link width is x1 or x2.

Implication: Usually, this issue causes nothing more than a replay of a few TLPs; the 82575EB recovers from this situation autonomously.

If the 82575EB is connected to an ICH7, a spurious SDP or STP symbol that occurs just before entering L1 could cause a hang of the PCIe link since the ICH7 does not insert SOS when transmitting PM_Request_ACK DLLPs, so the 82575EB does not receive them and never enters L1.

Workaround: If the 82575EB is connected to an ICH7, ASPM L1 should be disabled. Otherwise, no workaround is required.

Status: A2,A3 NoFix

33. I2C Data Out Hold Time Violation

Problem: The 82575EB 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 I2C CLK and DATA signals are reasonably matched on the board.

Workaround: None.

Status: A2,A3 NoFix

34. SerDes: RXCW.RxConfigInvalid Set Incorrectly

Problem: When the device has been receiving a continuous stream of /C/ ordered sets for an extended period of time, the RXCW.RxConfigInvalid bit may be set as the result of an internal FIFO overflow even if all input symbols are valid.

Implication: False indication of invalid symbols may cause the driver to disable the link when there is no problem.

This only applies when the Link Mode is Legacy SerDes (not the recommended mode of operation for SerDes for the 82575EB).

Workaround: Use the New SerDes Link mode (CTRL_EXT.LINK_MODE = 11b). This is the recommend mode for SerDes operation on the 82575EB.

Status: A2,A3 NoFix

35. NC-SI: Get Link Status Command Might Cause Corruption of PHY Registers

Problem: The firmware code that executes the NC-SI Get Link Status command does not acquire the PHY semaphore bit when accessing PHY registers. Therefore, there could be corruption of the PHY registers if the driver accesses PHY registers in parallel with the firmware.

Implication: Corruption of the PHY registers could result in a link failure and/or a driver failure if NC-SI Get Link Status command is used.

Workaround: None.

Status: A2,A3 NoFix

36. Packets Received with an L2+L3 Header Length Greater than 512 Bytes Can Incorrectly Report a Checksum Error

Problem: L2/L3 packets with long/multiple next header extensions incorrectly report a Receive checksum error when the length from Destination Address (DA) to the beginning of the TCP/UDP header is greater than 512 bytes.

Implication: A receive checksum error can incorrectly be reported by the device, even if there is no checksum error.

Workaround: When the driver receives a packet with a checksum error reported by the hardware, software should check the L2/L3 header length. If the L2/L3 header length is 512 bytes or greater, software should verify the checksum.

The Intel Windows & Linux drivers address this issue by passing packets with bad checksums to the network stack for further examination.

Status: A2,A3 NoFix

37. NC-SI: Get NC-SI Pass-through Statistics Response Format

Problem: The NC-SI Specification, version 1.0.0a defines the Pass-through Tx Packets counter contained in the Get NC-SI Pass-through Statistics Response Packet to be an 8-byte field. The 82575EB provides this counter as a 4-byte field.

Implication: A BMC that uses the Get NC-SI Pass-through Statistics command and expects the response format as described in the NC-SI Specification will not parse the response as

Workaround: The BMC can account for the different format provided by the 82575EB and parse the response accordingly.

Status: A2,A3 NoFix

2.5 Software Clarifications

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

Clarification: The 82575EB 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 64KB 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”.

§ §