

Intel® 82576EB Gigabit Ethernet Controller Specification Update

LAN Access Division (LAD)

**Revision 2.85
September 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.

Copyright © 2008, 2009, 2010, 2011, 2012; Intel Corporation. All Rights Reserved.



Revisions

Date	Revision	Description
January 2008	1.0	Initial release.
January 2008	1.1	Internal work copy.
March 2008	1.2	Updated to to show additional errata and errata addressed by A1.
March 2008	1.3	Internal work copy.
May 2008	1.9	Updated for Production stepping. Removed information that does not apply to the production stepping removed errata fixed by documentation, mature silicon.
July 2008	2.0	Resolved documentation issues were removed. Intel Confidential stamp removed; prepared for release on Developer.
February 2009	2.2	Non-security product information provided: 1.20 Product Code and Device Identification .
May 15, 2009	2.3	Spec clarification added: 3. PCIe: Completion Timeout Mechanism Compliance
August 6, 2009	2.4	Spec clarification added: 3. PCIe: Completion Timeout Mechanism Compliance Errata updated or added: 6. Critical Session (Keep PHY Link Up) Mode Does Not Block All PHY Resets Caused by PCIe Resets 18. JTAG: Instruction Register Functionality Doesn't Meet IEEE Std 1149-1-2001 20. PCIe Elastic Buffer Noise Immunity Is Not Optimized 21. PCIe: Missing Replay Due to Recovery During TLP Transmission 22. PCIe: LTSSM Moves from L0 to Recovery Only When Receiving TS1/TS2 on All Lanes 24. PCIe: Missing Completion on D3 to D0 Transition 25. PCIe: Completion Timeout Settings Not Loaded from EEPROM to GCR 26. MSI-X: Descriptor Write-back Not Triggered by EITR Expiration in MSI-X Mode
March 5, 2010	2.5	Errata updated or added. 19. LED Stays On When SerDes Is Powered Down 27. Tx Packet Lost After PHY Speed Change Using Auto-Negotiation 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. PCIe Hot Reset Can Lead to a Firmware Hang 31. MACSec: Replay Protect Not Working In Check Mode 32. MACSec: Packets With E=0, C=1 Should Not Be Handled as a Authenticated MACSec Packet 33. MACSec: Packets With PN = 0 In The SECTAG Are Not Dropped 34. MACSec: SA Creation Doesn't Clear Frame Verification Statistics 35. MACSec: LSECRXNUSA and LSECRXUNSA Statistic Counters Not Provided 36. MACSec: When MC Is MACSec owner, MAC Reset Still Clears Keys
March 8, 2010	2.6	36. MACSec: When MC Is MACSec owner, MAC Reset Still Clears Keys - Additional information added; editorial changes made for clarity.
March 26, 2010	2.7	Specification clarification added: 4. PCIe: Receiver Dtection Circuit Design and Established Link Width
June 11, 2010	2.71	LinkSec references changed to MACSec.



Date	Revision	Description
August 11, 2010	2.72	Spec Clarifications added: 5. Use of Wake on LAN Together with Manageability Spec Changes added: 1. Update to PBA Number EEPROM Word Format Errata added: 37. TimeSync: Missing Tx timestamps in SerDes mode 38. Virtualization: Dropped Packets When Using VM-to-VM Switching
September 9, 2010	2.73	Spec Clarification added: 6. Critical Session (Keep PHY Link Up) Mode Does Not Block All PHY Resets Caused by PCIe Resets - Moved from Errata. Spec Change, language updated: 1. Update to PBA Number EEPROM Word Format Explanation expanded to address confusion about number format.
October 14, 2010	2.74	Software Clarification: 1. While In TCP Segmentation Offload, Each Buffer is Limited to 64 KB
December 15, 2010	2.75	Spec Clarification added: 7. SerDes: AN_TIMEOUT Only Works When Link Partner Idle Errata added: 39. I2C Data Out Hold Time Violation
April 20, 2011	2.76	Spec Changes added: 2. PCIe: Device Control 2 Register Should Not Be Written While DMA Is Enabled 3. Updates to PXE/iSCSI EEPROM Words Errata added: 40. TSYNC: Auxiliary Timestamp from SDP is Unreliable
May 5, 2011	2.77	Spec Change added: 4. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)
July 20, 2011	2.78	Spec Changes added: 5. SerDes Forced Mode Override EEPROM Setting 6. Update to the Sequence in Switching Between Media Errata added: 41. NC-SI: Get Link Status Command Might Cause Corruption of PHY Registers
August 2, 2011	2.79	Software Clarification added: 2. Serial Interfaces Programmed By Bit Banging
September 14, 2011	2.80	Errata added: 42. NC-SI: Get NC-SI Pass-through Statistics Response Format
November 28, 2011	2.81	- Table 1, Product Identification and Packaging is an update. It replaces two older tables. Introduces the A2 stepping. A2 is functionally the same as A1. Errata added: 43. PCIe: Receiver Overflow Error .
December 7, 2011	2.82	- Added A2 information to status data for each errata. - Added the following text in Section 1.20: "There is an internal designator visible as a dash on the die side of some packages. It's location may differ. The mark (looks like a dash, approximately 300X1000 µm in size) does not impact function."
December 13, 2011	2.83	Table 1, Product Identification and Packaging. A table note has been updated. The new text is: "There is no die change for parts listed as A2. 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."



Date	Revision	Description
January 30, 2012	2.84	Spec Clarifications added: 8. Padding on Transmitted SCTP Packets
September 4, 2012	2.85	Spec Clarification: 9. Dynamic LED Modes Can Only Be Used in an Active Low Configuration Spec Change added: 7. CRC8 Fields of Analog Initialization Structures in the EEPROM Image are not Checked by the Device Errata Added 44. PF's MSI TLP Might Contain the Wrong Requester ID When a VF Uses MSI-X Software Clarification 3. PF/VF Drivers Should Configure Registers That Are Not Reset By VFLR



NOTE: *This page intentionally left blank.*



1.10 Introduction

This document applies to the 82576.

This document is an update to a published specification, the *Intel® 82576EB Gigabit Ethernet Controller Datasheet*. It is intended for use by system manufacturers and software developers. All product documents are subject to frequent revision and new order numbers may apply. New documents may be added. Be sure you have the latest information before finalizing your design.

References to PCIe* in this document refer to PCIe v2.0 (2.5GT/s).

1.20 Product Code and Device Identification

The following table describes identification and packaging information.

Table 1. Product Identification and Packaging

Lead (Pb)									
Device	Step*	Top Marking [Product Code]		Spec	Description	Media**		MM#	
82576EB	A2*	HL82576EB		S LJBC	PRODUCTION	T&R**		916949	
82576EB	A2	HL82576EB		S LJBD	PRODUCTION	TRAY		916950	
82576EB	A2	HL82576EB		Q PE4	ENGINEERING	TRAY		916530	
82576EB	A1	HL82576EB			PRODUCTION	TRAY		898304	
82576EB	A1	HL82576EB			PRODUCTION	T&R		898303	
* There is no die change for parts listed as A2. 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).									
Lead (Pb) Free									
Device	Step*	Top Marking [Product Code]		Spec*	Description	Media**		MM#	
82576EB	A2*	JL82576EB		S LJBG	PRODUCTION	T&R**		916955	
82576EB	A2	JL82576EB		S LJBH	PRODUCTION	TRAY		916956	
82576EB	A2	JL82576GB		S LJBL	PRODUCTION	T&R		916959	
82576EB	A2	JL82576GB		S LJBM	PRODUCTION	TRAY		916960	
82576EB	A2	JL82576NS		S LJBN	PRODUCTION NO SECURITY	T&R		916961	
82576EB	A2	JL82576NS		S LJBP	PRODUCTION NO SECURITY	TRAY		916962	



Table 1. Product Identification and Packaging (Continued)

82576EB	A2	JL82576EB		Q PE6	ENGINEERING	TRAY		916652	
82576EB	A2	JL82576GB		Q PE8	ENGINEERING	TRAY		916654	
82576EB	A2	JL82576NS		Q PE9	ENGINEERING	TRAY		916655	
82576EB	A1	JL82576EB			PRODUCTION	TRAY		897983	
82576EB	A1	JL82576EB			PRODUCTION	T&R		897979	
82576EB	A1	JL82576NS			PRODUCTION NO SECURITY	TRAY		899826	
82576EB	A1	JL82576NS			PRODUCTION NO SECURITY	T&R		899825	
* There is no die change for parts listed as A2. 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).									

Table 2. Device ID

82576 Device ID Code	Vendor ID	Device ID	Revision ID*
82576 (Copper Applications)	0x8086	0x10C9	1
82576 (Fiber Applications)	0x8086	0x10E6	1
82576 (SerDes Backplane Applications)	0x8086	0x10E7	1
82576 (Non-Security Copper Applications)	0x8086	0x150A	1

1.30 Marking Diagram

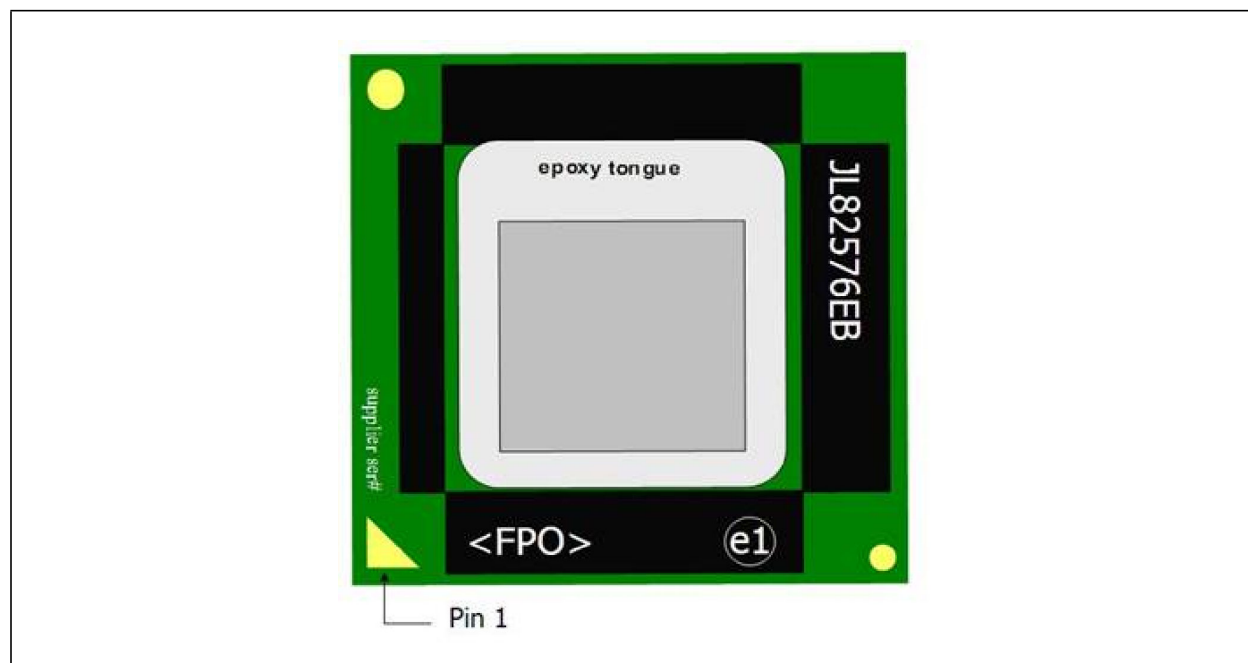


Figure 1. Example With Identifying Marks

Lead-free parts will have "JL" as the prefix for the product code (vs. "HL").

The devices can also have a "GB" marking, instead of "EB". These are functionally equivalent and only used on Intel network interface adapters.

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



1.40 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 3](#) for a description.

Table 3. 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.
Documentation Changes	Typos, errors, or omissions from the current published specifications. These changes will be incorporated in the next release of the specifications.
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.
EEPROM/NVM Fix	This indicates the Errata was in the EEPROM/NVM and is fixed in an updated version.
NoFix	There are no plans to fix this erratum.
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.

1.50 Sightings, Clarifications, Changes, Errata and Software Clarifications

See [Section 1.40](#) for an explanation of terms, codes, and abbreviations.

Table 4. Summary of Sightings, Clarifications, Changes, Errata, Software Clarifications; Errata Include Steppings

Sightings	Status
None.	NA
Specification Clarifications	Status
1. PCIe: End Point Request of I/O Space After Initialization	NA
2. PCIe: Partial Memory-Write Requests Actually Writing Full DW	NA
3. PCIe: Completion Timeout Mechanism Compliance	NA
4. PCIe: Receiver Detection Circuit Design and Established Link Width	NA


Table 4. Summary of Sightings, Clarifications, Changes, Errata, Software Clarifications; Errata Include Steppings (Continued)

5. Use of Wake on LAN Together with Manageability	NA
6. Critical Session (Keep PHY Link Up) Mode Does Not Block All PHY Resets Caused by PCIe Resets	NA
7. SerDes: AN_TIMEOUT Only Works When Link Partner Idle	NA
8. Padding on Transmitted SCTP Packets	NA
9. 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. PCIe: Device Control 2 Register Should Not Be Written While DMA Is Enabled	NA
3. Updates to PXE/iSCSI EEPROM Words	NA
4. Updated Definition of SW EEPROM Port Identification LED Blinking (Word 0x4)	NA
5. SerDes Forced Mode Override EEPROM Setting	NA
6. Update to the Sequence in Switching Between Media	NA
7. CRC8 Fields of Analog Initialization Structures in the EEPROM Image are not Checked by the Device	NA
Errata	Status
1. Internal Copper PHY: Improperly Implements Auto-Negotiation Advertisement Register	A1, A2 NoFix
2. PCIe: Differential Return Loss More than Specified Value	A1, A2 NoFix
3. SGMII Counters Incorrectly Increment on Collision	A1, A2 NoFix
4. Internal Copper PHY: Test Equipment May Report Master/Slave Device Doesn't Correctly Implement Master/Slave Resolution	A1, A2 NoFix
5. Internal Copper PHY: Auto-MDX Improperly Implements Sample Timer	A1, A2 NoFix
6. SCTP CRC Check Incorrect	A1, A2 NoFix
7. TLP: Poisoned TLP Reported In All Functions Instead Of Only Target	A1, A2 NoFix
8. Internal Copper PHY: 10BASE-T IDL Template Failure	A1, A2 NoFix
9. Internal Copper PHY: 10BASE-T Link Pulse Hits Template Mask Due To Voltage Ripple/Glitch	A1, A2 NoFix
10. MAC: Wakeup Event Occurs On Magic Packet That Doesn't Pass Address Filter	A1, A2 NoFix
11. PCIe: L0s Exit Latency In Link Capabilities Register Not Updated For Common Clock Configuration	A1, A2 NoFix
12. Time SYNC: Reserved Bits Must Be Zero In PTP Header	A1, A2 NoFix
13. Internal Copper PHY: No Link In Forced Mode	A1, A2 NoFix
14. Duplicate. See 18.	



Table 4. Summary of Sightings, Clarifications, Changes, Errata, Software Clarifications; Errata Include Steppings (Continued)

15. Internal Copper PHY: IEEE 1411.10.03 – 10BASE-T Harmonic Content Fails at Low Voltage	A1, A2 NoFix
16. NC-SI: Additional Multicast Packets May Be Forwarded To MC	A1, A2 NoFix
17. SMBus: Unread Packets Received On One Port May Cause Loss of Ability To Receive on Other Port	A1, A2 NoFix
18. JTAG: Instruction Register Functionality Doesn't Meet IEEE Std 1149-1-2001	A1, A2 NoFix
19. LED Stays On When SerDes Is Powered Down	A1, A2 NoFix
20. PCIe Elastic Buffer Noise Immunity Is Not Optimized	Fixed.
21. PCIe: Missing Replay Due to Recovery During TLP Transmission	A1, A2 NoFix
22. PCIe: LTSSM Moves from L0 to Recovery Only When Receiving TS1/TS2 on All Lanes	A1, A2 NoFix
23. This entry moved from the Errata to Specification Clarification section (above).	N/A
24. PCIe: Missing Completion on D3 to D0 Transition	A1, A2 NoFix
25. PCIe: Completion Timeout Settings Not Loaded from EEPROM to GCR	A1, A2 NoFix
26. MSI-X: Descriptor Write-back Not Triggered by EITR Expiration in MSI-X Mode	A1, A2 NoFix
27. Tx Packet Lost After PHY Speed Change Using Auto-Negotiation	A1, A2 NoFix
28. PCIe: Wrong Byte Enable Bit Used for Completion Timeout Disable Bit in Device Control 2 Register	A1, A2 NoFix
29. PCIe: Completion with UR/CA Status Causes Unexpected Completion and Completion Timeout Errors to be Reported	A1, A2 NoFix
30. PCIe Hot Reset Can Lead to a Firmware Hang	A1, A2 NoFix
31. MACSec: Replay Protect Not Working In Check Mode	A1, A2 NoFix
32. MACSec: Packets With E=0, C=1 Should Not Be Handled as a Authenticated MACSec Packet	A1, A2 NoFix
33. MACSec: Packets With PN = 0 In The SECTAG Are Not Dropped	A1, A2 NoFix
34. MACSec: SA Creation Doesn't Clear Frame Verification Statistics	A1, A2 NoFix
35. MACSec: LSECRXNUSA and LSECRXUNSA Statistic Counters Not Provided	A1, A2 NoFix
36. MACSec: When MC Is MACSec owner, MAC Reset Still Clears Keys	A1, A2 NoFix
37. TimeSync: Missing Tx timestamps in SerDes mode	A1, A2 NoFix
38. Virtualization: Dropped Packets When Using VM-to-VM Switching	A1, A2 NoFix
39. I2C Data Out Hold Time Violation	A1, A2 NoFix
40. TSYNC: Auxiliary Timestamp from SDP is Unreliable	A1, A2 NoFix
41. NC-SI: Get Link Status Command Might Cause Corruption of PHY Registers	A1, EEPROM/NVM Fix
42. NC-SI: Get NC-SI Pass-through Statistics Response Format	A1, A2 NoFix
43. PCIe: Receiver Overflow Error	A1, A2 NoFix



Table 4. Summary of Sightings, Clarifications, Changes, Errata, Software Clarifications; Errata Include Steppings (Continued)

44. PF's MSI TLP Might Contain the Wrong Requester ID When a VF Uses MSI-X	A1, A2 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
3. PF/VF Drivers Should Configure Registers That Are Not Reset By VFLR	NA

1.5.1 Sightings

None.

1.5.2 Specification Clarifications

1. PCIe: End Point Request of I/O Space After Initialization

Clarification: The 82576 requests I/O space if EEPROM bit 14, word 0x19 is set. When this EEPROM bit is set, I/O Space is always requested.

The specification does not define a way to signal that IO BAR usage is done. When PCIe compliance tests are run, this may cause a test failure.

Failure when running PCI SIG compliance tests with EEPROM bit 14, word 0x19 set.

Workaround: Disable I/O BAR requests via EEPROM bit 14, word 0x19. Since various pre-boot SW tools require the I/O Space be requested, the bit is enabled by default in EEPROM images.

▼ [Return to Summary](#)

2. PCIe: Partial Memory-Write Requests Actually Writing Full DW

Clarification: The PCIe specification allows a device not to accept certain requests. This is under "programming model" cases. The device needs to issue a Completer Abort error if specific request violates the programming model. As part of its programming model, the 82576 does not support writes and reads with Byte Enables to specific memory addresses. Such writes will be fully executed and will not be treated as completion abort.

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

Workaround: No partial (zero) Byte Enables writing and reading to the device.

▼ [Return to Summary](#)

3. PCIe: Completion Timeout Mechanism Compliance

Clarification: PCIe Completion Timeout value must be properly set.

The 82576 Completion Timeout Value(3:0) must be properly set by the system BIOS in Intel PCIe Configuration Space Device Control 2 Register (0xC8; RW). Failure to do so can cause unpredictable system behavior.



The 82576 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 82576 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 82576 is installed. This will ensure that the 82576 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 82576 device, 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.

Release 14.4 includes this fix.

▼ [Return to Summary](#)

4. PCIe: Receiver Detection Circuit Design and Established Link Width

Clarification: The 82576 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 82576 as a terminated lane.

This is normally not an issue since any connected 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 82576 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 82576 to detect the receivers again when they are properly terminated. As a result, a full-width link can be established.

Workaround: Not Applicable.

▼ [Return to Summary](#)



5. 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 5. 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

The Intel Windows* drivers set NoTCO by default.

If this is not the desired behavior, the EnableWakeOnManagementOnTCO 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 your Intel representative for access.

Workaround: N/A.

▼ [Return to Summary](#)

6. Critical Session (Keep PHY Link Up) Mode Does Not Block All PHY Resets Caused by PCIe Resets

Clarification: 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 events 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)
- Function-level reset



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

Workaround: N/A.

▼ [Return to Summary](#)

7. SerDes: AN_TIMEOUT Only Works When Link Partner 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 #5.

Workaround: N/A

▼ [Return to Summary](#)

8. Padding on Transmitted SCTP Packets

Clarification: When using the 82576 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 82576 pads the packets after performing the CRC calculation.

Workaround: N/A

▼ [Return to Summary](#)

9. 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.

Workaround: N/A

▼ [Return to Summary](#)

1.5.3 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

[▼ Return to Summary](#)



2. 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.

▼ [Return to Summary](#)

3. Updates to PXE/iSCSI EEPROM Words

Change: For Gigabit Main Setup Options Word 0x30, 0x34; bits 2:0 and bit 5 are now defined as indicated in the table below.

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

▼ [Return to Summary](#)



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

Change: 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.

See the following table.

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.

▼ [Return to Summary](#)

5. SerDes Forced Mode Override EEPROM Setting

Change: Bit 14 in EEPROM Compatibility Word 0x3 is no longer reserved and 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 only used by Intel drivers starting with Intel LAN Driver Release 16.4. This bit is read by Intel driver to determine if the device should be operated in SerDes forced mode. When the 82576 is not in SerDes mode this 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 #7.

▼ [Return to Summary](#)

6. Update to the Sequence in Switching Between Media

This is an update to the "Transition to the SerDes/SGMII Mode" procedure in the Datasheet Search for the procedure in Chapter 3 (Section 3.5.8.2.1*). Use the following updated steps.



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 8276 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 4.6.7.3, MAC/SERDES Link Setup (CTRL_EXT.LINK_MODE = 11b)* or Section 4.6.7.4, MAC/SGMII Link Setup (CTRL_EXT.LINK_MODE = 10b)*.
10. Set up Tx and Rx queues and enable Tx and Rx processes.

*Links can change in a specific Datasheet revision. Use the links provided by the revision you are using.

▼ [Return to Summary](#)

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

Change: See the "SerDes/PHY/PCIe/PLL/CCM Initialization Structures" section in Chapter 6 of the Datasheet. This section describes 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.

▼ [Return to Summary](#)



1.5.4 Errata

1. Internal Copper PHY: Improperly Implements Auto-Negotiation Advertisement Register

Problem: The 82576 incorrectly transmits the link code word due to a write to register 4.

Implication: See the following notes on link code word bits:

Bits 4.7 and 4.8: Are always set in the base page transmission.

Bit 4.9: Represents 100BASE-T4 support. The 82576 does not support T4. It is unlikely that the auto-negotiation feature of the 82576 will be used in an implementation to advertise the presence of a T4 physical device. Therefore, the fact that this device does not allow T4 to be advertised is not significant.

Bit 4.15: The 82576 always supports Next Page (regardless the value of bit 4.15). When bit 4.15 is set to "one," the 82576 requires Register 7 (AN Next Page Transmit Register) to be written to complete the Next Page Exchange. In this case, the 82576'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: A1, A2 NoFix

▼ [Return to Summary](#)

2. 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: A1, A2 NoFix

▼ [Return to Summary](#)

3. 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, A2 NoFix

▼ [Return to Summary](#)



4. Internal Copper PHY: Test Equipment May Report Master/Slave Device Doesn't Correctly Implement Master/Slave Resolution

Problem: When the internal Copper PHY is operating in 1000 Mbps forced slave mode, illegal data may be detected from the device during the transition from 10 Mbps mode (auto negotiation) to 1000 Mbps mode after master/slave resolution is complete.

Implication: Test equipment checking for compliance of Master/Slave resolution may report failures when the device is in Force Slave mode. In Forced Slave mode, the device should not transmit any 1000 Mbps signals, which it does not. However some test equipment looks for any activity sent from the device in forced slave mode and considers this a failure instead of looking for valid 1000 Mbps signals. Therefore, the illegal data may result in failures reported by test equipment.

Internal validation shows the device complies with IEEE 802.3 Table 40-5; for all configurations, the device resolves to the correct defined mode

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

5. Internal Copper PHY: Auto-MDX 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 $\pm$ 2 ms.

A violation occurs in $\sim$ 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: A1, A2 NoFix

▼ [Return to Summary](#)

6. SCTP CRC Check Incorrect

Problem: For SCTP CRC check; if the SCTP packet is smaller than 64 bytes and padding is added, the SCTP CRC calculation is wrong.

Implication: The 82576 could incorrectly indicate the SCTP CRC as incorrect.

Workaround: If a 64-byte SCTP packet is received and the 82576 indicates the SCTP CRC is wrong (RDESC.ERRORS.L4E is set), the driver should re-calculate the CRC.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

7. TLP: Poisoned TLP Reported In All Functions Instead Of Only Target

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

Implication: The 82576 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: A1, A2 NoFix

▼ [Return to Summary](#)

8. Internal Copper PHY: 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: Specification violation. There is no impact on system level performance.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

9. Internal Copper PHY: 10BASE-T Link Pulse Hits Template Mask Due To Voltage Ripple/Glitch

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

Implication: Specification violation. There is no effect at system level.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

10. MAC: Wakeup Event Occurs On Magic Packet That Doesn't Pass Address Filter

Problem: If the 82576 receives an magic packet that doesn't pass address filtering, it generates a wakeup event upon receiving the next packet if the next received packet (non-magic packet) is accepted according to the address filtering scheme.

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

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

11. PCIe: L0s Exit Latency In Link Capabilities Register Not Updated For Common Clock Configuration

Problem: L0s Exit Latency from the Link Capabilities register (0xAC) remains the same, whether the Link Control register's Common Clock Configuration bit is set or not.

Implication: This function remains in the default (non-common mode).

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)



12. Time SYNC: Reserved Bits Must Be Zero In PTP Header

Problem: Receiver expects the reserved bits located in octet 1's high nibble part of PTP message header to be zero; future compatibility could be an issue if the specification changes and uses those bits.

Implication: Specification violation. If these reserved bits are changed in future specifications, there could be an issue.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

13. Internal Copper PHY: No Link In Forced Mode

Problem: Link may fail if Auto MDI-X is enabled during forced 100BASE-TX mode operation.

Implication: Disable Auto MDI-X for Forced 100BASE-TX operation.

Workaround: Since the 82576 does not disable this function automatically, drivers must perform this step. Auto MDI-X can be disabled by clearing PHYREG18.12. Intel software drivers are implemented this way.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

14. Duplicate. See 18.

15. Internal Copper PHY: IEEE 1411.10.03 – 10BASE-T Harmonic Content Fails at Low Voltage

Problem: IEEE 1411.10.03 – 10BASE-T harmonic content fails at low voltage and the results are < 27dBC.

Implication: None. Specification violation.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

16. NC-SI: Additional Multicast Packets May Be Forwarded To MC

Problem: If the MC enables Multicast filtering for IPv6 Neighbor Advertisement and/or IPv6 Router Advertisement, additional multicast packets will be forwarded to the MC. The additional packets forwarded are:

1. Packets with the ICMPv6 header's Message Type: 135,137
2. IPv6 Neighbor Advertisement
3. IPv6 Router Advertisement

Implication: Additional packets may be forwarded to the MC.

Workaround: The MC should filter the different multicast packets.



Status: A1, A2 NoFix

▼ [Return to Summary](#)

17. SMBus: Unread Packets Received On One Port May Cause Loss of Ability To Receive on Other Port

Problem: The device's two ports share memory. When packets are received by one of the ports and not read, they are stored in the shared memory. When this memory fills up, no more packets may be received.

Implication: The MC should be aware of the above behavior.

Workaround: 1. Make use of a SMBus alert timeout mechanism.
2. Read data and/or momentarily disable receives by the other port.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

18. 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: A1, A2 NoFix

▼ [Return to Summary](#)

19. LED Stays On When SerDes Is Powered Down

Problem: The LED may remain on in D3 power state when SerDes power down is enabled (EEPROM word 0xf, bit 1; register CTRL_EXT 0x0018, bit 18). If a link is established when the device enters 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: LED incorrectly reflects link is up when there is no link (as SerDes is powered off).

Workaround: When using the Intel Fiber adapters based on the 82576, the driver should disable optics (when going to D3) by setting SDP3 to 1.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

20. 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 EEPROM, control bits were implemented to maintain stability during electrical idle. In the default EEPROM image provided, these bits were not set correctly.

Implication: In cases of increased noise levels during Electrical Idle, elastic buffer instability may cause a link training failure when exiting from L0s or (in rare instances) L1 state.



Failure to exit L0s, resulting in a timeout and recovery, can cause a reduction in performance. In rare cases, the large number of replays may cause a PCIe Hang. 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.

Workaround: Fix implemented in revised EEPROM version number v1.5 by setting PCIe Init Configuration 3 word - Bits 4 & 5.

Status: Fixed.

▼ [Return to Summary](#)

21. 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: A1, A2 NoFix

▼ [Return to Summary](#)

22. 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 Intel®82576 GbE Controller 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.

23. This entry moved from the Errata to Specification Clarification section (above).

The move makes the issue's presentation consistent with the positioning in other products.

24. PCIe: Missing Completion on D3 to D0 Transition

Problem: When both ports are in D3 and software transitions one port to D0 using a PMCSR write access which arrives at a specific time relative to exiting L1: completion for the PMCSR write is correctly provided, but the following transaction does not get a completion.

This bug only occurs when the IOV Enabled EEPROM bit is 1.

Implication: This situation has been observed when using certain operating systems that do not support IOV during a small fraction of S1 to S0 transitions. If the problem does occur, it will likely result in a system hang.

Workaround: Do any of the following:



- If IOV is not required, clear the IOV Enabled bit in the EEPROM (word 0x25, bit 0).
- Do not use S1. Use S3 instead.
- Make a BIOS modification to avoid the problematic event timing by providing 30 us of quiet before and after a PMCSR write access that transfers the port from D3 to D0. This could be done by adding an ACPI _PS0 method that includes STALLs surrounding a STORE to PMCSR. Contact Intel for more information.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

25. 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: A1, A2 NoFix

▼ [Return to Summary](#)

26. MSI-X: Descriptor Write-back Not Triggered by EITR Expiration in MSI-X Mode

Problem: When using MSI-X mode with multiple interrupt vectors, in many cases the descriptor write-back is not triggered by the expiration of the EITR to which the queue is mapped. This applies to both Tx and Rx queues.

Implication: Receive queues: If RXDCTL.WTHRESH is greater than 1, large latencies may be seen on received packets since no descriptor write-back occurs until WTHRESH is reached.

Transmit queues using descriptor write back: If TXDCTL.WTHRESH is greater than 1, buffers are not released by the driver until WTHRESH is reached.

Transmit queues using Head Write Back: The WB on EITR bit is ineffective and only descriptors with RS set cause a head write-back.

Workaround: One of the following:

- The default value of WTHRESH should be used in TXDCTL and RXDCTL. If using Head Write Back, set the RS bit in all descriptors that should trigger a head write back.
- Program all 32 entries of the IVAR array to point to valid interrupt vectors that have a non-zero EITR value programmed.
- Don't use MSI-X or use only a single vector, and clear GPIE.Multiple_MSIX.
- Use the SWFLUSH bit of TXDCTL and RXDCTL to periodically flush the descriptors based on an external timer.

Status: A1, A2 NoFix

▼ [Return to Summary](#)



27. 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: A1, A2 NoFix

▼ [Return to Summary](#)

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: A1, A2 NoFix

▼ [Return to Summary](#)

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

Problem: When the 82576 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 82576 are supported.

If an UR/CA status completion is received, the completion timeout error can bring down the operating system when 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: A1, A2 NoFix

▼ [Return to Summary](#)



30. 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: 1.9 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: A1, A2 NoFix

▼ [Return to Summary](#)

31. MACSec: Replay Protect Not Working In Check Mode

Problem: When working in check mode, replay packets are not dropped and they are counted in MACSec RX Octets Validated LSECRXOCTP (0x4320).

Implication: In check mode, replay protect is not working and LSECRXOCTP counts incorrectly.

Note: In strict mode, which is the normal operation mode, there is no problem.

Workaround: Drivers can provide a workaround by working in "no-header strip" mode and drop the replay packet in the driver while correcting the LSECRXOCTP and the LSECRXLATE statistics (maintained in the driver as 64-bit counters).

Status: A1, A2 NoFix

▼ [Return to Summary](#)

32. MACSec: Packets With E=0, C=1 Should Not Be Handled as a Authenticated MACSec Packet

Problem: MACSec Packets with E=0, C=1 is a reserved future mode, and should not be handled as a Authenticated MACSec packet.

Implication: If MACSec Packets with E=0, C=1 are sent it will be handled as a Authenticated MACSec packet. This a reserved mode and there is no implication as long as the specification does not change in the future.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

33. MACSec: Packets With PN = 0 In The SECTAG Are Not Dropped

Problem: According to the MACSec specification, frames with PN=0 in the sectag should be counted as bad tags/packets. The 82576 will accept such packets. Setting the first expected PN to



1 will cause the packet to be dropped but for the wrong reason (late packet instead of bad tagged).

Implication: These packets are not processed correctly and may cause statistic differences (when compared to specification requirements).

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

34. MACSec: SA Creation Doesn't Clear Frame Verification Statistics

Problem: SA (Security Association) creation doesn't zero frame verification statistics. The creation of the RX SA should reset the following statistics (per SA):

- 1.) InPktsOK (LSECRXOK 0x4360-0x4364)
- 2.) InPktsInvalid (LSECRXINV 0x4380-0x4384)
- 3.) InPktsNotValid(LSECRXNV 0x43A0-0x43A4)
- 4.) InPktsNotUsingSA(LSECRXNUSA 0x43C0)
- 5.) InPktsUnusedSA (LSECRXUNSA 0x43D0)

Implication: Statistics are not cleared out as expected by the specification.

Workaround: Read frame verification statistics before validating the SA (reading clears the statistics).

Status: A1, A2 NoFix

▼ [Return to Summary](#)

35. MACSec: LSECRXNUSA and LSECRXUNSA Statistic Counters Not Provided

Problem: InPktsNotUsingSA (LSECRXNUSA) and InPktsUnusedSA (LSECRXUNSA) should be defined per SA. In this implementation, these are counted in one counter.

Implication: Statistics defined in the MACSec standard cannot be provided.

Workaround: None.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

36. MACSec: When MC Is MACSec owner, MAC Reset Still Clears Keys

Problem: When MC is the owner of the MACSec and a MAC reset occurs, MACSec settings are lost.

A MAC reset (CTRL.RST) can clear out the MACSec - 82576 FW will give an indication to MC [using the MACSec AEN (NC-SI) or in the MACSec Parameters and MACSec Status Indication (SMBus)] when this occurs so the MC can reconfigure the keys.

Implication: MAC reset can clear the MACSec setting at anytime.

Workaround: MC must recognize the indication from the FW and reconfigure the keys when this occurs.



Status: A1, A2 NoFix

▼ [Return to Summary](#)

37. TimeSync: Missing Tx timestamps in SerDes mode

Problem: When transmitting a TimeSync packet in SerDes mode, there is a probability that the timestamp will not be sampled in the Tx Timestamp Value registers and thus TSYNCTXCTL.TXTT will not be set.

There is no issue when using 10/100/1000 BASE-T(Copper) mode.

Implication: Missing timestamps make it difficult for the software to effectively implement the TimeSync functionality.

Workaround: The software should implement a timeout while waiting for TSYNCTXCTL.TXTT to be set after transmitting a TimeSync packet. In case no timestamp has been captured, the software may resend the TimeSync packet until the timestamp capture is successful, as indicated by TSYNCTXCTL.TXTT being read as 1. The receiving software must be able to handle these repeated packets.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

38. Virtualization: Dropped Packets When Using VM-to-VM Switching

Problem: In a virtualized environment using IOV or VMDq, the 82576 supports up to 8 VMs with up to 2 queues per VM. As specified in the datasheet, when a packet is switched between VMs, RSS is not used for Rx queue selection within a VM.

However, there is also no default queue defined. Instead, the Tx queue number (0 or 1) within the source VM is used as the Rx queue number within the destination VM. If a packet is provided to Tx Queue 1 within a VM and it is switched to a VM in which only one queue is used and therefore Rx Queue 1 is not defined, the packet is dropped.

Implication: There might be no communication between VMs on the same 82576 device.

This only applies to a virtualized environment using IOV or VMDq.

Workaround: Use only one Tx queue per VM.

OR

Ensure that two Rx queues are defined on all VMs.

The first option above is implemented in Intel igb/igbvf 2.6.36 kernel drivers.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

39. I²C Data Out Hold Time Violation

Problem: The 82576 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: NA



Status: A1, A2 NoFix

▼ [Return to Summary](#)

40. TSYNC: Auxiliary Timestamp from SDP is Unreliable

Problem: SDP inputs to the timestamp logic are not properly synchronized. As a result, both the Auxiliary Timestamp Register values and the Auxiliary Timestamp Taken bits in TSAUXC are sometimes loaded incorrectly.

Implication: The auxiliary timestamp feature should be considered unreliable.

Workaround: For 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.

In addition, the following method can be used to filter out incorrect values:

- Connect the input signal to two SDP inputs for the same port.
- Using the TSSDP register, assign one of the SDP inputs to AUX0 and the other SDP input to AUX1.
- When reading the TSAUXC register to check for new samples, check that both AUTTO and AUTT1 are set. Otherwise, discard the sample.
- Read both the AUX0 and AUX1 timestamp values and compare the values. Discard the values if they differ by more than the sampling uncertainty -- 8 ns if the SDP inputs are balanced externally or slightly higher if the external trace lengths differ significantly.

Using this method, along with a SW filter for expected values, almost all errors can be filtered out with the remaining samples having a very high probability of being correct.

Status: A1, A2 NoFix

▼ [Return to Summary](#)

41. 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. See status.

Status: A1, EEPROM/NVM Fix

Fixed in EEPROM version 2.1 or later.

▼ [Return to Summary](#)

42. 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 82576 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 intended by the 82576, and will obtain inaccurate statistics.

Workaround: The BMC can account for the different format provided by the 82576 and parse the response accordingly.

Status: A1, A2 NoFix

[▼ Return to Summary](#)

43. PCIe: Receiver Overflow Error

Problem: When a MemWr transaction with 64-bit address, 64-bit data, and ECRC is received, the posted data (PD) CREDITS_RECEIVED flow control counter is not incremented.

Implication: No immediate effect. However, after 4064 of these packets, the CREDITS_ALLOCATED counter value wraps around and passes the CREDITS_RECEIVED counter value, resulting in a Receive Overflow error. This is an uncorrectable error that generally causes a system hang or crash.

This issue can only occur on a system that generates ECRC. Typically, Intel platforms do not support ECRC generation and are not susceptible to this issue.

Workaround: Do one of the following:

- Do not perform QWORD (64-bit) writes to the device.
- Disable ECRC generation on the root port to which the device is connected.
- Use 32-bit addressing for any BAR through which a QWORD MemWr transaction may be performed. This can be guaranteed by setting EEPROM word 0x21, bits 11:10 to 00b and (for SR-IOV systems) EEPROM word 0x25, bit 2 to 0.

Status: A1, A2 NoFix

[▼ Return to Summary](#)

44. PF's MSI TLP Might Contain the Wrong Requester ID When a VF Uses MSI-X

Problem: When using IOV, if a PF uses MSI interrupts and one or more VFs use MSI-X interrupts, some of the MSI TLPs for the PF might contain the wrong Requester ID.

Implication: There could be missing interrupts on the PF since the incorrect Requester ID could result in the virtualization mechanism misrouting or dropping TLPs.

Workaround: If any VFs use MSI-X, all PFs should also use MSI-X.

Status: A1, A2 NoFix

[▼ Return to Summary](#)



1.5.5 Software Clarifications

Applies to Intel drivers.

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

Clarification: The 82576 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.

▼ [Return to Summary](#)

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".

▼ [Return to Summary](#)

3. PF/VF Drivers Should Configure Registers That Are Not Reset By VFLR

Clarification: The following registers are not reset by VFLR and need to be configured by PF or VF in case of a change to a new configuration (such as VF OS transition):

RDH/T, TDH/T, PSRTYPE, SRRCTL, RXDCTL, TXDCTL, TDWBAL/H, RXCTRL, TXCTRL

▼ [Return to Summary](#)

§ §