Intel® Ethernet Controller X710/XXV710/XL710

Specification Update

February 2024

Revision 4.9
331430-029

Revision History

Revision	Date	Comments
4.9	February 7, 2024	**Errata added or updated:** • 106. Device is Not Enumerated on the PCIe Bus Following Certain Warm Reset Flows (Added)
4.8	August 10, 2023	**Specification Clarifications added or updated:** • 19. The 710-series device Does Not Support Timestamping of IEEE 1588 Packets with a Non-Zero Value of minorVersionPTP (Added) **Specification Changes added or updated:** • 25. PCIe REPLAY_TIMER Timeout Value Increase (Added) **Errata added or updated:** • 105. Incorrect Receive Length Errors (Added)
4.7	March 10, 2023	**Documentation Changes added or updated:** • 1. X710/XXV710/XL710 MDC Clock Frequency Min/Max Parameters (Added) **Errata added or updated:** • 104. Pre-boot Failure with Multiple Traffic Classes (Added)
4.6	September 15, 2022	**Specification Changes added or updated:** • 24. Reduction in the Size of the VSI Replication List Pool (Added) **Errata added or updated:** • 98. Memory Leak in Receive Packet Buffer (Updated)
4.5	April 1, 2022	**Errata added or updated:** • 100. Failure to Manipulate the Default MAN/VLAN Filter (Updated) • 102. Remove MAC/VLAN Pair AQ Command Failure (Updated) • 103. LAN-to-BMC Packets Delayed or Dropped in Pre-Boot Environment (Added)
4.4	November 18, 2021	**Errata added or updated:** • 97. Thermal Alarm Points for Pluggable Modules Are Using Fixed Values Instead of Reading from Module EEPROM (Updated)
4.3	November 4, 2021	**Errata added or updated:** • 100. Failure to Manipulate the Default MAN/VLAN Filter (Added) • 101. XXV710 PHY Lock-Up (Added) • 102. Remove MAC/VLAN Pair AQ Command Failure (Added)
4.2	August 4, 2021	**Errata added or updated:** • 98. Memory Leak in Receive Packet Buffer (Added) • 99. Single-Byte I^2C Writes Clear the Following Byte (Added)
4.1	February 12, 2021	**Specification Changes added or updated:** • 23. Transmit Scheduler Does Not Support Weighted Strict Priority Arbitration (Added) **Errata added or updated:** • 97. Thermal Alarm Points for Pluggable Modules Are Using Fixed Values Instead of Reading from Module EEPROM (Updated)
4.0	December 7, 2020	**Errata added or updated:** • 93. EMP Reset After Using Intel QCU Tool (Updated) • 97. Thermal Alarm Points for Pluggable Modules Are Using Fixed Values Instead of Reading from Module EEPROM (Added)
3.9	September 1, 2020	**Errata added or updated:** • 13. Common Clock Configuration Bit Specification Compliance (Updated) • 44. DCBx Resume of a Port Affects Other Ports (Updated) **Miscellaneous updates:** • Updated "Intel® Ethernet Controller XL710 for 40 GbE backplane" row in Table 1-2, "Device IDs"
3.8	May 21, 2020	**Errata added or updated:** • 96. No Length Error on VLAN Packets with BAD Type/Length Field (Added)
3.7	October 18, 2019	**Errata added or updated:** • 95. PCIe Replay Timer Can Occasionally be Lower Than PCIe Spec Requirements (Added)

Revision	Date	Comments
3.6	June 17, 2019	**Errata added or updated:** • 93. EMP Reset After Using Intel QCU Tool (Added) • 94. 25G-AUI Takes a Long Time to Link with Optical Modules and AoC (Added) **Specification Clarifications added or updated:** • 17. The 710-series device Packet Drop Rate is Limited to 27 MPPS (Added) • 18. Expansion ROM is Exposed in Blank Flash Programming Mode (Added) **Miscellaneous updates:** • Updated example component and marking diagram in Section 1.2, "Marking Diagrams".
3.5	November 28, 2018	**Specification Clarifications added or updated:** • 16. Small Packets Performance Degrade when Using Private VLAN (Added) **Specification Changes added or updated:** • 22. Fixed VCCD Voltage Supply (Added) **Errata added or updated:** • 91. Activity LED Might Blink Regardless if Link is Up or Down for a Port (Added) • 92. PCIe Phase 2 Fails to Timeout Under Certain Channel Conditions (Added)
3.4	July 16, 2018	**Specification Clarifications added or updated:** • 13. Intel® Ethernet Controller X710/XXV710/XL710 Throughput Limit (Updated)
3.3	December 12, 2017	**Specification Clarifications added or updated:** • 15. 1G Link on SFP Connections Using XXV710 (Added)
3.2	October 31, 2017	**Errata added or updated:** • 89. Receive IP Packets in a Low-Latency Traffic Class Are Not Fully Processed (Added) • 90. QSFP+ Modules with Tx Squelch (Added)

Revision	Date	Comments
3.1	August 31, 2017	**Specification Clarifications added or updated:** • 11. S-TAG and VLAN Ethernet Types (Updated) • 13. Intel® Ethernet Controller X710/XXV710/XL710 Throughput Limit (Updated) • 14. The BMC Needs to Follow DMTF v1.1.0 to Properly Link at 25G and 40G Using NCSI Set Link Command (Added) **Specification Changes added or updated:** • 5. Support of the Admin Queue Command "Set Loopback modes command (opcode:0x0618)" (Updated) • 17. GLQF_PCNT Counters (Updated) • 20. IEEE 1588 (TimeSync) over UDP is Not Supported in NVM 5.05 (Updated) • 21. Internal Thermal Diode Measurements are Not Reliable (Updated) **Errata added or updated:** • 2. PCIe Subsystem ID Incorrectly Reported for All PCI Functions Except Function 0 (Updated) • 20. Incorrect Optic Module Presence Reported to BMC (Updated) • 23. "Multi-speed module timeout" NVM Parameter Has No Effect (Updated) • 45. PHY Type Field is Not Reported when Link is Down (Updated) • 70. LLDP Disable Can Result in Incorrect Configuration of the Receive Packet Buffer (Updated) • 77. Filtering Restriction when Double VLAN is Enabled (Updated) • 78. Transmit Hang in 2x40 GbE Configuration When Flow Control is Enabled (Updated) • 80. SGMII Receiver Sensitivity (Added) • 81. IEEE 802.3 Clause 73 AN Does Not Support Parallel Detection (Added) • 82. IEEE 802.3 Clause 73 AN Echoed Nonce Field is Zero (Added) • 83. KR Transmitter Output Waveform Violations (Added) • 84. 10GBASE-KR wait_timer Value Smaller Than Specification (Added) • 85. Receive Queue Disable Can Get Stuck (Added) • 86. Add Cloud Filter Command Can Fail with Return Code ENOSPC (Added) • 87. First Tx Descriptor Not Processed (Added) • 88. Set DCB Parameters AQC (Opcode 0x303) Might Return EINVAL Even when It Succeeds (Added) Miscellaneous updates: • Updated "Compatible Silicon Device" column for Device ID 0x1584 in Table 1-2, "Device IDs" • Updated Table 1-4, "Nomenclature". • Added device-specific indicators to clarification, change, and errata summaries in Table 2-1, Table 2-2, and Table 2-4, respectively.

Revision	Date	Comments
3.0	January 10, 2017	**Specification Clarifications added or updated:** • 9. All NVM Versions Should be Updated to Version 5.05 or Later (Updated) • 10. L2 Padding and L4 Checksum Offloads (Added) • 11. S-TAG and VLAN Ethernet Types (Added) • 12. Malicious Driver Detection MAX_BUFF Event (Added) • 13. Intel® Ethernet Controller X710/XXV710/XL710 Throughput Limit (Added) **Specification Changes added or updated:** • 17. GLQF_PCNT Counters (Added) • 18. Flash CS Negation Time (Added) • 19. Parsing of MPLS Headers (Added) • 20. IEEE 1588 (TimeSync) over UDP is Not Supported in NVM 5.05 (Added) • 21. Internal Thermal Diode Measurements are Not Reliable (Added) **Errata added or updated:** • 73. Glitch on SDP Outputs During GLOBR (Added) • 74. Function-Level Reset Fails to Complete (Added) • 75. 10GBASE-T Link Issues (Added) • 76. Incorrect Flexible Payload Extraction from Flow Director Filter to Receive Descriptor (Added) • 77. Filtering Restriction when Double VLAN is Enabled (Added) • 78. Transmit Hang in 2x40 GbE Configuration When Flow Control is Enabled (Added) • 79. Aux Power Detected Bit Not Implemented (Added) **Miscellaneous updates:** • Added Figure 1-1, "Example Component with Identifying Marks". • Updated Table 1-2, "Device IDs" • Minor formatting changes to Table 1-1, "Markings" and Table 1-3, "MM Numbers".
2.9	June 1, 2016	**Specification Clarifications added or updated:** • 9. All NVM Versions Should be Updated to Version 5.05 or Later (Added) **Specification Changes added or updated:** • 16. Teredo UDP Tunneling Offload Support (Added)

Revision	Date	Comments
2.8	April 7, 2016	**Specification Clarifications added or updated:** • 3. Source Pruning of Packets with Unknown Source MAC Address (SA) (Updated) • 5. NVM Revision 4.53 Improvements (Added) • 6. I^2C Minimum Time Between Transactions (Added) • 7. LLDPDU without End of LLDPDU TLV (Added) • 8. Qualified Module Bit (Added) **Specification Changes added or updated:** • 8. VEB Statistics Disable NVM Bit (Updated) • 12. Force 40G Enabled Status Bit (Added) • 13. Receive Control Path Registers Admin Queue Commands (Added) • 14. Lane Ordering for Optical Breakout Cable (Added) • 15. Set Local LLDP MIB when DCBX Agent is Disabled or Stopped (Added) **Errata added or updated:** • 27. Jabber Packets are Not Counted in Jabber Packets Received Field of NC-SI Get Controller Packet Statistics Command (Updated) • 31. PRTDCB_RUP2TC and PRTDCB_TC2PFC are Not Writable (Updated) • 51. LPLU is Not Functional in KR Mode (Added) • 52. A Function-level Reset Might Affect Other Functions (Added) • 53. Rx Packet Drops Even with Priority Flow Control (Added) • 54. Configuration of Stopped DCBx Agent is Not Returned to Default (Added) • 55. B1: Incorrect Revision ID Reported to MC (Added) • 56. SMBUS: Wrong Driver Status Reported to MC After TCO Isolate Command (Added) • 57. Shutdown LLDP Agent Command (Added) • 58. NC-SI: A Link Change Event Might Cause an Endless Loop of AENs (Added) • 59. DCBx Configuration Might Change After LLDP Stops (Added) • 60. LLDP Receive-only Mode Malfunctions (Added) • 61. Non-optimal Receive Packet Buffer Partitioning (Added) • 62. QSFP Configuration Transition Requires Power Cycle (Added) • 63. MAC Link Flow Control Might Not be Updated After Auto-negotiation Restart (Added) • 64. PFR Flow Limitations (Added) • 65. FC Refresh Threshold Field Ambiguity (Added) • 66. Link Auto-negotiation Reporting Might be Incorrect (Added) • 67. Inefficient Shared Resources Allocation (Added) • 68. Prepare for Endpoint Discovery MCTP Command (Added) • 69. Delete Mirror Rule Admin Command Failure (Added) • 70. LLDP Disable Can Result in Incorrect Configuration of the Receive Packet Buffer (Added) • 71. PCIe Interrupt Status Bit (Added) • 72. Changes to Tx Scheduler Structure Can Cause the Device to Become Unusable (Added) **Software Clarifications added or updated:** • 2. VXLAN Guidance for VMware vSphere (Updated)
2.7	October 22, 2015	**Documentation Changes added or updated:** • Pre-boot Combo Image Version (Removed - Content added to Datasheet) **Specification Changes added or updated:** • 11. PCI Revision ID Update for B1 Stepping (Added) **Errata added or updated:** • 48. XL710 40 GbE MAC Transmit Packet Issue (Updated - Fixed in B1 Stepping) **Miscellaneous updates:** • Added B1 stepping information to the following tables: — Table 1-1, "Markings" — Table 1-3, "MM Numbers" • Other miscellaneous updates related to B1 stepping.

Revision	Date	Comments
2.6	August 18, 2015	**Specification Clarifications added or updated:** • 4. PCIe Re-timers Might Cause Replay Timer Timeout Correctable Errors (Added) **Documentation Changes added or updated:** • Pre-boot Combo Image Version (Added) **Errata added or updated:** • 28. INTENA_MSK Setting Might Clear Interrupt (Updated) • 49. Unrecognized Optical Modules Will Not Link with NVM Release 4.53 (Added) • 50. Get Link Status AQ Command Might Return Incorrect Status (Added)
2.5	July 8, 2015	**Specification Changes added or updated:** • 10. Input Reference Clock Rise/Fall Times (Added) **Errata added or updated:** • 47. Legacy SMBus: Failure to De-assert Alert Signal when Not Using ARA Cycle (Added) • 48. XL710 40 GbE MAC Transmit Packet Issue (Added)
2.4	June 22, 2015	**Specification Clarifications added or updated:** • 3. Source Pruning of Packets with Unknown Source MAC Address (SA) (Added) **Specification Changes added or updated:** • 6. Logging of PCIe Correctable Receiver Error (Updated) • 9. Unicast Hash Filtering Removal (Added) **Errata added or updated:** • 26. VLAN Insertion Limitation with Specific Tunneling Packet (Updated) • 28. INTENA_MSK Setting Might Clear Interrupt (Updated) • 45. PHY Type Field is Not Reported when Link is Down (Added) • 46. A Global SDP Might be Affected by a Specific Port Power State (Added)
2.3	April 15, 2015	**Specification Changes added or updated:** • 8. VEB Statistics Disable NVM Bit (Added) **Errata added or updated:** • 19. A Switching Table Might Reduce Small Packets Performance (Updated) • 34. No LAN-to-BMC Pass-through Traffic in Dr State (Updated) • 36. Priority Flow Control (PFC) is Not Fully Functional (Added) • 37. NC-SI Set Link Compliance (Added) • 38. VF's Admin Send Queue Might Hang for VFs Greater Than 31 (Added) • 39. NVM Update Admin Queue Command Responds with the Incorrect Error Code (Added) • 40. Outer User Priority Incorrectly Configured (Added) • 41. SFP+/QSFP+ Modules or an External 10GBASE-T PHY Might Not be Recognized (Added) • 42. SFP+ Optical Module Might Fail to Link in Specific Adapters (Added) • 43. VLAN Mirroring Admin Queue Command Writes Only the First VLAN in a List (Added) • 44. DCBx Resume of a Port Affects Other Ports (Added) **Software Clarifications added or updated:** • 2. VXLAN Guidance for VMware vSphere (Added)
2.2	December 9, 2014	**Specification Changes added or updated:** • 6. Logging of PCIe Correctable Receiver Error (Added) • 7. PRTPM_SAL and PRTPM_SAH are Re-loaded from NVM on PCIe Reset (Added) **Errata added or updated:** • 16. L2 Tag Stored in the Wrong RX Descriptor Field (Updated) • 32. AER Header Log Might be Invalid (Added) • 33. A CfgWr to a VF TLP with Error Might Generate an Error Message with Wrong VF Number (Added) • 34. No LAN-to-BMC Pass-through Traffic in Dr State (Added) • 35. MNG Packets are Dropped while a Function-Level Reset to PF 0 is in Progress (Added)

Revision	Date	Comments
2.1	October 16, 2014	**Specification Clarifications added or updated:** • 1. SFP+ Cable EEPROM Overwrite on Power Down (Added) • 2. QSFP+ to SFP+ Breakout Cables Support (Added) **Specification Changes added or updated:** • 3. NC-SI Get Controller Packet Statistics Command Limitations (Added) • 4. SMBus Minimum Packet Size (Added) • 5. Support of the Admin Queue Command "Set Loopback modes command (opcode:0x0618)" (Added) **Errata added or updated:** • 27. Jabber Packets are Not Counted in Jabber Packets Received Field of NC-SI Get Controller Packet Statistics Command (Added) • 28. INTENA_MSK Setting Might Clear Interrupt (Added) • 29. Manageability Checksum Filtering of IPv6 Packets (Added) • 30. Link Remains Up During Power Saving State (Added) • 31. PRTDCB_RUP2TC and PRTDCB_TC2PFC are Not Writable (Added) **Software Clarifications added or updated:** • 1. XL710 Option ROM Should Not be Integrated in the BIOS (Added)
2.0[1]	August 1, 2014	Initial release (Intel public).

1. There were no previous versions of this document released.

1. Introduction

This document applies to the Intel® Ethernet Controller X710/XXV710/XL710 (710-series device).

This document is an update to a published specification, the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*. It is intended for use by system manufacturers and software developers. All product documents are subject to frequent revision and new order numbers might apply. New documents might be added. Be sure you have the latest information before finalizing your design.

References to PCIe Express (PCIe) in this document refer to PCIe v3.0 (2.5GT/s, 5GT/s, and 8GT/s).

For more information on supported features, see the *Intel® Ethernet Controller X710/XXV710/XL710 Feature Support Matrix*. This document is updated periodically. Please ensure that you have the latest version.

1.1 Product Code and Device Identification

Product Codes: FTX710, FTXXV710, and FTXL710

The following tables and drawings describe the various identifying markings on each device package:

Table 1-1. Markings

Device	Stepping	Top Marking	S-Specification[1]	Description
XL710	B0	FTXL710-AM2	S R1ZL[2] S R1ZK[3]	Ethernet controller (2x40/1x40/4x10/2x10/1x10)
XL710	B0	FTXL710-AM1	S R1ZN[2] S R1ZM[3]	Ethernet controller (1x40/4x10/2x10/1x10)
X710	B0	FTX710-AM2	S R1ZQ[2] S R1ZP[3]	Ethernet controller (2x10/1x10)
XL710	B1	FTXL710-BM2	S LLK7[2] S LLK8[3]	Ethernet controller (2x40/1x40/4x10/2x10/1x10)
XL710	B1	FTXL710-BM1	S LLK9[2] S LLKA[3]	Ethernet controller (1x40/4x10/2x10/1x10)
X710	B1	FTX710-BM2	S LLKB[2] S LLKC[3]	Ethernet controller (2x10/1x10)
XXV710	B1	FTXXV710-AM1	S LLZ2[2] S LLZ4[3]	Ethernet controller (1x25/2x10/1x10)
XXV710	B1	FTXXV710-AM2	S LLZ3[2] S LLZ5[3]	Ethernet controller (2x25/1x25/2x10/1x10)

1. For Tray and Tape & Reel data, see Table 1-3.
2. Tray.
3. Tape and Reel.

Table 1-2. Device IDs

Branding String	Interface Type	Compatible Silicon Device[1]	Device ID	Vendor ID	Revision ID	
					B0	B1
N/A- Power on with Blank Flash	N/A	All	0x154B	0x8086	0x01	0x02
Intel® Ethernet Controller X710 for 10 GbE SFP+	SFI	FTXL710-BM2 (4x10/2x10/1x10) FTXL710-BM1(4x10/2x10/1x10) FTX710-BM2 (2x10/1x10)	0x1572	0x8086	0x01	0x02
Intel® Ethernet Controller XL710 for 40 GbE backplane	KR4	FTXL710-BM2 (2x40/1x40) FTXL710-BM1 (2x40)	0x1580	0x8086	0x01	0x02
Intel® Ethernet Controller X710 for 10 GbE backplane	KR/SFI	FTXL710-BM2 (4x10/2x10/1x10) FTXL710-BM1(4x10/2x10/1x10) FTX710-BM2 (2x10/1x10)	0x1581	0x8086	0x01	0x02
Intel® Ethernet Controller XL710 for 40 GbE QSFP+	XLPPI/CR4/SFI	FTXL710-BM2 (2x40/4x10)	0x1583	0x8086	0x01	0x02
Intel® Ethernet Controller XL710 for 40 GbE QSFP+	XLPPI/CR4/SFI	FTXL710-BM2 (1x40/4x10) FTXL710-BM1 (1x40/4x10)	0x1584	0x8086	0x01	0x02
Intel® Ethernet Controller X710/X557-AT 10GBASE-T	10GBASE-T with X557	FTXL710-BM2 (4x10/2x10/1x10) FTXL710-BM1(4x10/2x10/1x10) FTX710-BM2 (2x10/1x10)	0x1589	0x8086	0x01	0x02
Intel® Ethernet Controller XXV710 for 25 GbE backplane	25GBASE-KR/AUI C2C	XXV710	0x158A	0x8086	N/A	0x02
Intel® Ethernet Controller XXV710 for 25 GbE SFP28	25GBASE-CR/SR/ AUI C2M	XXV710	0x158B	0x8086	N/A	0x02

1. For X710 and XL710, the matching AM1/AM2 parts work if the BM1/BM2 parts work.

Table 1-3. MM Numbers

Product	S-Specification	Tray MM#	Tape and Reel MM#
FTXL710-AM2	S R1ZL[1]	936550	---
	S R1ZK[2]	---	936549
FTXL710-AM1	S R1ZN[1]	936552	---
	S R1ZM[2]	---	936551
FTX710-AM2	S R1ZQ[1]	936554	---
	S R1ZP[2]	---	936553
FTXL710-BM2	S LLK7[1]	947348	---
	S LLK8[2]	---	947349
FTXL710-BM1	S LLK9[1]	947351	---
	S LLKA[2]	---	947353
FTX710-BM2	S LLKB[1]	947359	---
	S LLKC[2]	---	947361
FTXXV710-AM1	S TKF[1]	954098	---
FTXXV710-AM2	S TKG[1]	954099	---
FTXXV710-AM1	S LLZ2[1]	954100	---
	S LLZ4[2]	---	954102
FTXXV710-AM2	S LLZ3[1]	954101	---
	S LLZ5[2]	---	954103

1. Tray.
2. Tape and Reel.

1.2 Marking Diagrams

Figure 1-1. Example Component with Identifying Marks

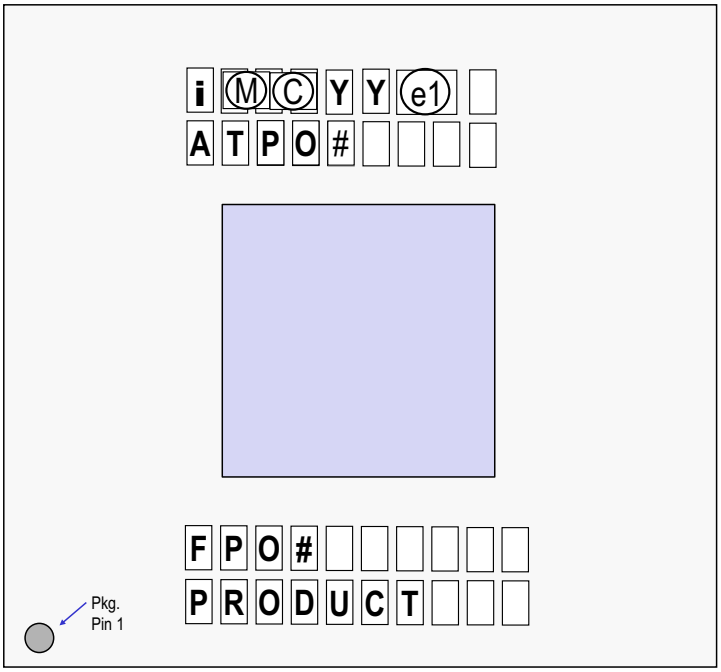

Figure 1-2. Marking Diagram

- LINE1: I, Maskwork, Copyright, YY, Pb-free
- LINE2: ATPO#
- LINE3: FPO#
- LINE4: Product

1.3 Nomenclature Used in This Document

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

Table 1-4. Nomenclature

Name	Description
A0, B0, etc.	Stepping to which the status applies.
Device	Where applicable, indicates the specific device to which an errata applies. Possible values are: • XL710 — 40 GbE device. • XXV710 — 25 GbE device. • X710 — 10 GbE device • All — All 710-series devices.
Doc	Document change or update that will be implemented.
Documentation Changes	Typos, errors, or omissions from the current published specifications. These changes will be incorporated in the next release of the specifications.
Errata	Design defects or errors. Errata might 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.
Eval	Plans to fix this erratum are under evaluation.
Fix Planned	This erratum is intended to be fixed in a future stepping of the component.
Fix Planned in NVM	This erratum is intended to be fixed in a future NVM version.
Fixed	This erratum has been fixed.
Fixed in NVM	This erratum has been fixed in NVM X.XX.
NoFix	There are no plans to fix this erratum.
Software Clarifications	Applies to Intel drivers, EEPROM loads.
Specification Changes	Modifications to the current published specifications. These changes will be incorporated in the next release of the specifications.
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.

2. Hardware Clarifications, Changes, Updates and Errata

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

Table 2-1. Summary of Specification Clarifications

Specification Clarification	Device	Status
1. SFP+ Cable EEPROM Overwrite on Power Down	X710/XL710	N/A
2. QSFP+ to SFP+ Breakout Cables Support	X710/XL710	N/A
3. Source Pruning of Packets with Unknown Source MAC Address (SA)	X710/XL710	N/A
4. PCIe Re-timers Might Cause Replay Timer Timeout Correctable Errors	All	N/A
5. NVM Revision 4.53 Improvements	X710/XL710	N/A
6. I^2C Minimum Time Between Transactions	All	N/A
7. LLDPDU without End of LLDPDU TLV	X710/XL710	N/A
8. Qualified Module Bit	All	N/A
9. All NVM Versions Should be Updated to Version 5.05 or Later	X710/XL710	N/A
10. L2 Padding and L4 Checksum Offloads	All	N/A
11. S-TAG and VLAN Ethernet Types	All	N/A
12. Malicious Driver Detection MAX_BUFF Event	X710/XL710	N/A
13. Intel® Ethernet Controller X710/XXV710/XL710 Throughput Limit	All	N/A
14. The BMC Needs to Follow DMTF v1.1.0 to Properly Link at 25G and 40G Using NCSI Set Link Command	XXV710/XL710	N/A
15. 1G Link on SFP Connections Using XXV710	XXV710	N/A
16. Small Packets Performance Degrade when Using Private VLAN	All	N/A
17. The 710-series device Packet Drop Rate is Limited to 27 MPPS	All	N/A
18. Expansion ROM is Exposed in Blank Flash Programming Mode	All	N/A
19. The 710-series device Does Not Support Timestamping of IEEE 1588 Packets with a Non-Zero Value of minorVersionPTP	All	N/A

Table 2-2. Summary of Specification Changes

Specification Change	Device	Status
1. Ingress Mirroring Cannot be Changed on the Fly	All	N/A
2. RSS Field Selection is Globally Defined	All	N/A
3. NC-SI Get Controller Packet Statistics Command Limitations	All	N/A
4. SMBus Minimum Packet Size	All	N/A
5. Support of the Admin Queue Command “Set Loopback modes command (opcode:0x0618)”	All	N/A
6. Logging of PCIe Correctable Receiver Error	X710/XL710	N/A

Table 2-2. Summary of Specification Changes [continued]

Specification Change	Device	Status
7. PRTPM_SAL and PRTPM_SAH are Re-loaded from NVM on PCIe Reset	All	N/A
8. VEB Statistics Disable NVM Bit	X710/XL710	N/A
9. Unicast Hash Filtering Removal	All	N/A
10. Input Reference Clock Rise/Fall Times	All	N/A
11. PCI Revision ID Update for B1 Stepping	All	N/A
12. Force 40G Enabled Status Bit	X710/XL710	N/A
13. Receive Control Path Registers Admin Queue Commands	All	N/A
14. Lane Ordering for Optical Breakout Cable	XL710	N/A
15. Set Local LLDP MIB when DCBX Agent is Disabled or Stopped	X710/XL710	N/A
16. Teredo UDP Tunneling Offload Support	All	N/A
17. GLQF_PCNT Counters	All	N/A
18. Flash CS Negation Time	All	N/A
19. Parsing of MPLS Headers	All	N/A
20. IEEE 1588 (TimeSync) over UDP is Not Supported in NVM 5.05	All	N/A
21. Internal Thermal Diode Measurements are Not Reliable	All	N/A
22. Fixed VCCD Voltage Supply	X710/XL710	N/A
23. Transmit Scheduler Does Not Support Weighted Strict Priority Arbitration	All	N/A
24. Reduction in the Size of the VSI Replication List Pool	All	N/A
25. PCIe REPLAY_TIMER Timeout Value Increase	All	N/A

Table 2-3. Summary of Documentation Changes

Documentation Update	Status
1. X710/XXV710/XL710 MDC Clock Frequency Min/Max Parameters	N/A

Table 2-4. Summary of Errata; Errata Include Steppings

Erratum	Device	Status
1. TX Performance Degradation for Small Cloud Packets	All	B0=Yes, B1=Yes; NoFix
2. PCIe Subsystem ID Incorrectly Reported for All PCI Functions Except Function 0	All	B0=Yes, B1=Yes; NoFix
3. Illegal Byte Error Statistical Counter Inaccuracy	All	B0=Yes, B1=Yes; NoFix
4. Receive Performance Degradation with Specific Cloud Header	All	B0=Yes, B1=Yes; NoFix
5. MCTP Discovery Error when Replacing Active PF	All	B0=Yes, B1=Yes; NoFix
6. RX Queue Disable is Reported Done Before It is Disabled	All	B0=Yes, B1=Yes; NoFix
7. TX Descriptor Might be Read Twice	All	B0=Yes, B1=Yes; NoFix

Table 2-4. Summary of Errata; Errata Include Steppings [continued]

Erratum	Device	Status
8. Immediate Interrupts are Delayed in Very Loaded System	All	B0=Yes, B1=Yes; NoFix
9. ECRC Bits are Not RO when ECRC is Disabled	All	B0=Yes, B1=Yes; NoFix
10. NC-SI I/Os Output Rise Slew Rate is Higher Than Specification	All	B0=Yes, B1=Yes; NoFix
11. TC Strict Priority Does Not Work as Expected	All	B0=Yes, B1=Yes; NoFix
12. Management-only Packets Cannot be Ignored for Wake-Up	All	B0=Yes, B1=Yes; NoFix
13. Common Clock Configuration Bit Specification Compliance	All	B0=Yes, B1=Yes; NoFix
14. Low Latency TC Might be Momentarily Starved	All	B0=Yes, B1=Yes; NoFix
15. Round Robin (RR) Bandwidth Distribution is Traffic Dependent	All	B0=Yes, B1=Yes; NoFix
16. L2 Tag Stored in the Wrong RX Descriptor Field	All	B0=Yes, B1=Yes; NoFix
17. Internal VLAN is Not Reflected in RX Descriptor	All	B0=Yes, B1=Yes; NoFix
18. Transmit Queue Group with Single Queue Enabled Performance	All	B0=Yes, B1=Yes; NoFix
19. A Switching Table Might Reduce Small Packets Performance	All	B0=Yes, B1=Yes; NoFix
20. Incorrect Optic Module Presence Reported to BMC	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
21. Set Binding Command is Not Functional for IPv4	All	B0=Yes, B1=Yes; NoFix
22. Get PHY Abilities Does Not Return 20GBASE-KR2	X710/XL710	B0=Yes, B1=Yes; NoFix
23. “Multi-speed module timeout” NVM Parameter Has No Effect	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
24. Cloud Traffic Over VEB is Transmitted to LAN	All	B0=Yes, B1=Yes; NoFix
25. VLAN Prune is Not Functional	All	B0=Yes, B1=Yes; NoFix
26. VLAN Insertion Limitation with Specific Tunneling Packet	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
27. Jabber Packets are Not Counted in Jabber Packets Received Field of NC-SI Get Controller Packet Statistics Command	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
28. INTENA_MSK Setting Might Clear Interrupt	All	B0=Yes, B1=Yes; NoFix
29. Manageability Checksum Filtering of IPv6 Packets	All	B0=Yes, B1=Yes; NoFix
30. Link Remains Up During Power Saving State	All	B0=Yes, B1=Yes; NoFix
31. PRTDCB_RUP2TC and PRTDCB_TC2PFC are Not Writable	All	B0=Yes, B1=Yes; NoFix
32. AER Header Log Might be Invalid	All	B0=Yes, B1=Yes; NoFix
33. A CfgWr to a VF TLP with Error Might Generate an Error Message with Wrong VF Number	All	B0=Yes, B1=Yes; NoFix
34. No LAN-to-BMC Pass-through Traffic in Dr State	All	B0=Yes, B1=Yes; NoFix
35. MNG Packets are Dropped while a Function-Level Reset to PF 0 is in Progress	All	B0=Yes, B1=Yes; NoFix
36. Priority Flow Control (PFC) is Not Fully Functional	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
37. NC-SI Set Link Compliance	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42

Table 2-4. Summary of Errata; Errata Include Steppings [continued]

Erratum	Device	Status
38. VF's Admin Send Queue Might Hang for VFs Greater Than 31	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
39. NVM Update Admin Queue Command Responds with the Incorrect Error Code	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
40. Outer User Priority Incorrectly Configured	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
41. SFP+/QSFP+ Modules or an External 10GBASE-T PHY Might Not be Recognized	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
42. SFP+ Optical Module Might Fail to Link in Specific Adapters	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
43. VLAN Mirroring Admin Queue Command Writes Only the First VLAN in a List	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.42
44. DCBx Resume of a Port Affects Other Ports	All	B0=Yes, B1=Yes; Fixed in NVM 4.53
45. PHY Type Field is Not Reported when Link is Down	All	B0=Yes, B1=Yes; Fixed in NVM 6.01
46. A Global SDP Might be Affected by a Specific Port Power State	All	B0=Yes, B1=Yes; NoFix
47. Legacy SMBus: Failure to De-assert Alert Signal when Not Using ARA Cycle	All	B0=Yes, B1=Yes; NoFix
48. XL710 40 GbE MAC Transmit Packet Issue	XL710	B0=Yes; NoFix. B1=No; Fixed
49. Unrecognized Optical Modules Will Not Link with NVM Release 4.53	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
50. Get Link Status AQ Command Might Return Incorrect Status	All	B0=Yes, B1=Yes; NoFix
51. LPLU is Not Functional in KR Mode	All	B0=Yes, B1=Yes; NoFix
52. A Function-level Reset Might Affect Other Functions	All	B0=Yes, B1=Yes; NoFix
53. Rx Packet Drops Even with Priority Flow Control	All	B0=Yes, B1=Yes; NoFix
54. Configuration of Stopped DCBx Agent is Not Returned to Default	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
55. B1: Incorrect Revision ID Reported to MC	X710/XL710	B0=No, B1=Yes; Fixed in NVM 5.02
56. SMBUS: Wrong Driver Status Reported to MC After TCO Isolate Command	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
57. Shutdown LLDP Agent Command	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
58. NC-SI: A Link Change Event Might Cause an Endless Loop of AENs	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
59. DCBx Configuration Might Change After LLDP Stops	All	B0=Yes, B1=Yes; NoFix
60. LLDP Receive-only Mode Malfunctions	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
61. Non-optimal Receive Packet Buffer Partitioning	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
62. QSFP Configuration Transition Requires Power Cycle	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
63. MAC Link Flow Control Might Not be Updated After Auto-negotiation Restart	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02

Table 2-4. Summary of Errata; Errata Include Steppings [continued]

Erratum	Device	Status
64. PFR Flow Limitations	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
65. FC Refresh Threshold Field Ambiguity	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
66. Link Auto-negotiation Reporting Might be Incorrect	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 4.53
67. Inefficient Shared Resources Allocation	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
68. Prepare for Endpoint Discovery MCTP Command	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
69. Delete Mirror Rule Admin Command Failure	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
70. LLDP Disable Can Result in Incorrect Configuration of the Receive Packet Buffer	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 6.01
71. PCIe Interrupt Status Bit	All	B0=Yes, B1=Yes; NoFix
72. Changes to Tx Scheduler Structure Can Cause the Device to Become Unusable	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.02
73. Glitch on SDP Outputs During GLOBR	All	B0=Yes, B1=Yes; NoFix
74. Function-Level Reset Fails to Complete	All	B0=Yes, B1=Yes; NoFix
75. 10GBASE-T Link Issues	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 5.03
76. Incorrect Flexible Payload Extraction from Flow Director Filter to Receive Descriptor	All	B0=Yes, B1=Yes; NoFix
77. Filtering Restriction when Double VLAN is Enabled	All	B0=Yes, B1=Yes; Fixed in NVM 6.01
78. Transmit Hang in 2x40 GbE Configuration When Flow Control is Enabled	XL710	B0=Yes, B1=Yes; Fixed in NVM 6.01
79. Aux Power Detected Bit Not Implemented	All	B0=Yes, B1=Yes; NoFix
80. SGMII Receiver Sensitivity	XXV710	B1=Yes; NoFix
81. IEEE 802.3 Clause 73 AN Does Not Support Parallel Detection	XXV710	B1=Yes; NoFix
82. IEEE 802.3 Clause 73 AN Echoed Nonce Field is Zero	XXV710	B1=Yes; NoFix
83. KR Transmitter Output Waveform Violations	XXV710	B1=Yes; NoFix
84. 10GBASE-KR wait_timer Value Smaller Than Specification	XXV710	B1=Yes; NoFix
85. Receive Queue Disable Can Get Stuck	All	B0=Yes, B1=Yes; NoFix
86. Add Cloud Filter Command Can Fail with Return Code ENOSPC	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 6.01
87. First Tx Descriptor Not Processed	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 6.01
88. Set DCB Parameters AQC (Opcode 0x303) Might Return EINVAL Even when It Succeeds	X710/XL710	B0=Yes, B1=Yes; NoFix
89. Receive IP Packets in a Low-Latency Traffic Class Are Not Fully Processed	All	B0=Yes, B1=Yes; NoFix
90. QSFP+ Modules with Tx Squelch	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 6.01

Table 2-4. Summary of Errata; Errata Include Steppings [continued]

Erratum	Device	Status
91. Activity LED Might Blink Regardless if Link is Up or Down for a Port	All	B0=Yes, B1=Yes; NoFix
92. PCIe Phase 2 Fails to Timeout Under Certain Channel Conditions	All	B0=Yes, B1=Yes; Fixed in NVM 6.80
93. EMP Reset After Using Intel QCU Tool	X710/XL710	B0=Yes, B1=Yes; Fixed in NVM 7.20
94. 25G-AUI Takes a Long Time to Link with Optical Modules and AoC	XXV710	B0=Yes, B1=Yes; Fixed in NVM 7.00
95. PCIe Replay Timer Can Occasionally be Lower Than PCIe Spec Requirements	All	B0=Yes, B1=Yes; NoFix
96. No Length Error on VLAN Packets with BAD Type/Length Field	X710/XL710	B0=Yes, B1=Yes; NoFix
97. Thermal Alarm Points for Pluggable Modules Are Using Fixed Values Instead of Reading from Module EEPROM	All	B0=Yes, B1=Yes; NoFix
98. Memory Leak in Receive Packet Buffer	All	B0=Yes, B1=Yes; Fixed in NVM 9.00
99. Single-Byte I^2C Writes Clear the Following Byte	All	B0=Yes, B1=Yes; Fixed in NVM 8.40
100. Failure to Manipulate the Default MAN/VLAN Filter	All	B0=Yes, B1=Yes; Fixed in NVM 8.60
101. XXV710 PHY Lock-Up	XXV710	B0=Yes, B1=Yes; Fixed in NVM 8.50
102. Remove MAC/VLAN Pair AQ Command Failure	All	B0=Yes, B1=Yes; Fixed in NVM 8.60
103. LAN-to-BMC Packets Delayed or Dropped in Pre-Boot Environment	All	B0=Yes, B1=Yes; NoFix
104. Pre-boot Failure with Multiple Traffic Classes	All	B0=Yes, B1=Yes; NoFix
105. Incorrect Receive Length Errors	All	B0=Yes, B1=Yes; NoFix
106. Device is Not Enumerated on the PCIe Bus Following Certain Warm Reset Flows	All	B0=Yes, B1=Yes; Fix Planned in NVM

2.1 Specification Clarifications

1. SFP+ Cable EEPROM Overwrite on Power Down

After PCIe Reset, the normal 710-series device operation might include I^2C transactions to the SFP+ cable EEPROM. Under certain timing conditions, these transactions might coincide with power ramping down. This could lead to an unintentional I^2C write command, causing the cable EEPROM contents to be overwritten on cables that do not have write protection, and making them inoperable.

NVM v4.25 provides updated timing of the I^2C transactions to avoid unintentional modification of the cable EEPROM contents.

2. QSFP+ to SFP+ Breakout Cables Support

When using QSFP+ to SFP+ breakout cables with NVM v4.24/4.25/4.26 configured for Cfg ID 6.0, the link is not established unless SDP3 pins of the four ports are connected to the ModPresL signal from the QSFP+ cage.

Note: Starting in NVM 4.42, it is no longer necessary to connect SDP3 of all 4 ports this way.

3. Source Pruning of Packets with Unknown Source MAC Address (SA)

When using NVMs prior to v5.02 and working in virtualization mode, the 710-series device can be configured to VEB mode, which enables traffic loopback from one VSI to another. The source pruning feature filters loopback packets by comparing the SA to the MAC Addresses stored in the switching forwarding tables. This prevents packets from being looped back to the sender.

Source pruning only filters packets which have an SA that is in the list of exact MAC Addresses supplied by the software device driver. If a packet is sent to the network with an SA that is not in the list of exact MAC Addresses, the packet might be looped back to the sender, which could cause unexpected reflections and/or confuse the software stack.

Situations in which this could occur include:

1. A VSI is configured to use a unicast hash filter for its MAC Address.
2. A VSI is configured to use the unicast promiscuous mode.

Note that loopback packets have the *LPBK* bit set in the receive descriptor. This could be used to assist software in filtering the packets.

LINUX / i40e driver:

Users can avoid this issue by changing the hardware L2 switch on each port from VEB mode to VEPA mode, which disables the loopback function. This can be done with the following command on each port:

```
bridge link set dev DEVICE hwmode vepa
```

i40e version 1.2.47 and later does this by default.

Windows:

VEB is disabled by default in currently-shipped Intel drivers. No further action should be required.

Notes:

1. The configuration of the hardware L2 switch in VEPA mode does NOT require an external VEPA enabled switch unless using SR-IOV with VFs.
2. The primary use of VEB mode is when using SR-IOV with VFs directly assigned to the VMs so a VM can communicate with other VMs on the same host. If this mode is not required, there should be no downside to using VEPA.

4. PCIe Re-timers Might Cause Replay Timer Timeout Correctable Errors

The addition of PCIe re-timers add to the total channel latency. According to PCI-SIG ECN extension devices, latency is defined as "the time from when the last bit of a Symbol is received at the input pins of one Pseudo Port to when the equivalent bit is transmitted on the output pins of the other Pseudo Port". The ECN allows for a maximum of 64 symbol x latency per PCIe re-timer for 8 GT/s speed.

The PCIe ACK/NACK round trip delay is incremented according to the number of re-timers used in Tx/Rx lanes. The extra delay added by a re-timer might cause the 710-series device Replay_Timer to expire, causing replay timer timeout correctable errors. 710-series device design does not take into consideration the extension devices ECN.

If a design must include re-timers, and if Replay_Timer timeout correctable errors are seen, please contact your Intel representative for support.

5. NVM Revision 4.53 Improvements

NVM revision 4.53 or greater should be used for an improved implementation of the following features:

- DCBx configuration.
- LLDP - MIB change events, ports isolation, LLDP admin status, CDCP TLV support.
- Firmware flow used by the Intel QSFP configuration tool.
- Firmware flow used by the Intel NVM update tool.
- Link establishment flow after a period of an intermittent link state, LPLU support, and PHY status collection.
- Switch mirroring and/or promiscuous modes configuration.
- GLOBR assertion by PCIR.

6. I^2C Minimum Time Between Transactions

The SFF-8431 Specification requires that the minimum time between STOP and START on an I^2C bus (Tbuf) should be at least 20 μs. The time measured in the 710-series device is less than required by the specification. No functional implication should be expected.

7. LLDPDU without End of LLDPDU TLV

According to the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, the end of LLDPDU TLV is a mandatory LLDPDU TLV, but there are some implementations that exist that do not send it. Therefore, the device does not require that an LLDPDU ends with this TLV.

Certain LLDPDUs without the end of LLDPU TLV are ignored. This limitation was removed in NVM 5.02.

8. Qualified Module Bit

According to the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, the Qualified Module bit in Admin Queue Get link status response (Byte 3, bit 7) is cleared when the module is not found in pre-configured list of qualified modules. In addition, this bit can be cleared in case that there is a contradiction between the module and device configuration. For example, NVM of BASE-T with external optical module.

9. All NVM Versions Should be Updated to Version 5.05 or Later

Intel has found a security vulnerability in all the NVM versions released prior to v5.05.

Update NVM image to NVM 5.05 or later.

For more detail, see the Security Advisories at:

https://security-center.intel.com/advisories.aspx

10. L2 Padding and L4 Checksum Offloads

When using UDP/TCP checksum offloading on Tx packets (L4T in the Tx descriptor is 01b or 11b), any L2 padding at the end of the packet must be all zeros.

When using SCTP CRC offloading on Tx packets (L4T in the Tx descriptor is 10b), L2 padding should not be used.

11. S-TAG and VLAN Ethernet Types

The default settings for the 710-series device are to use Ethernet Type 0x88A8 to identify an S-TAG. An outer VLAN is identified by Ethernet Type 0x8100 when enabled by the Set Port Parameters admin command. Using Ethernet Type 0x88A8 for an outer VLAN is not supported.

Starting with NVM 6.01, the software can set the tags in the Set Switch Configuration command (Opcode: 0x0205). See the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, Revision 3.2 or later for more details.

12. Malicious Driver Detection MAX_BUFF Event

When the first Tx descriptor of a TSO packet contains both header and payload, it is counted twice in the malicious detection of MAX_BUFFS. Therefore, an MDD event is reported if the first segment is spread over eight descriptors, while it should only cause an MDD event if there are more than eight descriptors.

This can result in spurious Malicious Driver Detection events.

Software drivers must limit the first segment of a TSO packet to seven descriptors instead of eight. This restriction has been implemented in Intel drivers Release 21.3.

13. Intel® Ethernet Controller X710/XXV710/XL710 Throughput Limit

Small Packet throughput limit:

For packets below 160 bytes there is a hardware packet processing limit for the entire device of ~37 Mpps. This results in limited throughput for:

- The Intel® Ethernet Controller X710 & XL710 (4x10 GbE mode) when using 3 or 4 port 10 GbE operation.
- The Intel® Ethernet Controller XXV710.
- The Intel® Ethernet Controller XL710 (1x40 GbE).

Large packet throughput limit:

For packets at or over 160 bytes:

- The Intel® Ethernet Controller XXV710 (2x25 GbE) has a total hardware throughput limit for the entire device of ~96-97% of dual-port 25 GbE line rate in each direction. As an example, the total payload throughput is limited to ~45.5 Gb/s in each direction for IPv4 TCP large packets (>1518 bytes) with standard 1500-byte MTU. Thus, while single-port 25 GbE throughput is not impacted, total simultaneous dual-port 25 GbE throughput is expected to be slightly lower than line rate.

Note: The Intel® Ethernet Controller XL710 (4x10 GbE, 1x40 GbE, 2x40 GbE) has an expected total throughput for the entire device of 40 Gb/s in each direction.

14. The BMC Needs to Follow DMTF v1.1.0 to Properly Link at 25G and 40G Using NCSI Set Link Command

When a BMC sends an NCSI Set Link command to the device, but is following the DMTF NCSI v1.0.0 specification, the highest speed mode is limited to 10G (Link Settings field Bit 04). To support 25G speed mode on XXV710 and 40G speed mode XL710, the BMC should follow the DTMF NCSI v1.1.0 specification to advertise 25G or 40G speed as the highest speed capability (Link Settings field, Bit 06 and Bit 07, respectively). If auto-negotiation is not used, the channel attempts to force the link to the specified setting.

15. 1G Link on SFP Connections Using XXV710

When using XXV710 devices in a 1G SFP connection, link might not be achieved. In NVM 6.01 for 1G SFP links, XXV710 operates in CL-37 AN mode and is unable to link with link partners in forced mode. In NVM 6.02 the 1G SFP operates in 1G forced mode, which is the same operation as NVM 5.51.

16. Small Packets Performance Degrade when Using Private VLAN

When using private VLAN, the device uses a VLAN pruning filter that slows down performance for small packets. For example, the total max MPPS (Million Packets Per Second) achievable drops from ~73 MPPS to ~39 MPPS.

Note: Depending on software driver capabilities, if only a single VLAN is applied, the driver might choose to use a port VLAN instead of a private VLAN and thus avoid the use of the VLAN pruning filter and associated performance penalty.

17. The 710-series device Packet Drop Rate is Limited to 27 MPPS

In NVM version 6.01 and later, the packet drop rate was limited to 27 MPPS instead of 37 MPPS to improve device functionality robustness.

Maximum drop packet rate is now 27 MPPS. In case of enabling a queue without handling its descriptors, there might be a case of massive packets drop (i.e. broadcast) which will effect the overall traffic bandwidth. This case is forbidden to keep the device operation at full bandwidth.

18. Expansion ROM is Exposed in Blank Flash Programming Mode

In blank flash programming mode, the expansion ROM will be exposed, but might point to invalid pre boot driver code.

19. The 710-series device Does Not Support Timestamping of IEEE 1588 Packets with a Non-Zero Value of minorVersionPTP

IEEE 1588 PTP packets that have a non-zero value in the *minorVersionPTP* field are not recognized by the 710-series device's timestamping logic. This can occur, for example, when using **ptp4l** with the PTP version set to 2.1.

When using **ptp4l**, set the *PTP_VERSION* value in *msg.h* to 2 before compiling:

```
#define PTP_VERSION 2
```

2.2 Specification Changes

1. Ingress Mirroring Cannot be Changed on the Fly

Changing of Ingress Mirroring configuration during traffic might cause a bad configuration.

2. RSS Field Selection is Globally Defined

RSS field selection is done globally and cannot be configured differently per PF or VF.

- Functions that require the Hash (RSS) filters on IPv4 packets should set all IPv4 PCTYPEs in the PFQF_HENA / VFQF_HENA (PCTYPEs 31, 33…36).
- Functions that require the Hash filters on IPv6 packets should set all IPv6 PCTYPEs in the PFQF_HENA / VFQF_HENA (PCTYPEs 41, 43…46).
- Functions that require the Hash filters on FCoE packets should set all FCoE PCTYPEs in the PFQF_HENA / VFQF_HENA (PCTYPEs 48…50).

3. NC-SI Get Controller Packet Statistics Command Limitations

Counter 0 of NC-SI "Get Controller Packet Statistics" command returns the value of "Valid Bytes Received" instead of "Total Bytes Received".

4. SMBus Minimum Packet Size

The minimum Ethernet packet size transmitted by BMC over SMBus supported by the 710-series device is 17 bytes.

5. Support of the Admin Queue Command "Set Loopback modes command (opcode:0x0618)"

Starting with NVM 6.01, "Set Loopback modes command (opcode:0x0618)" is supported. For full details see the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, Revision 3.2 or later. For prior NVM versions, this command should not be used.

6. Logging of PCIe Correctable Receiver Error

The optional error logging of correctable receiver error is disabled in the 710-series device, which is allowed as described in:

> *PCI Express® Base Specification, Revision 3.0, November 10, 2010, Section 7.10.5, Correctable Error Status Register Footnote 101*

PHY layer receiver error detection and recovery mechanisms are operational such that there is no functional implication to the device or system operation. Please note that both Correctable Error Status Register[0] and Correctable Error Mask Register[0] are implemented such that the 710-series device is a PCI-SIG compliant device.

This change is implemented in NVM 4.42.

7. PRTPM_SAL and PRTPM_SAH are Re-loaded from NVM on PCIe Reset

PRTPM_SAL and PRTPM_SAH registers are re-loaded from the NVM on PCIe reset. Therefore, only the station address values stored in the NVM can be used for Wake-On-LAN.

8. VEB Statistics Disable NVM Bit

A new NVM bit was added starting from 710-series device NVM 4.42:

EMP Settings Module Header Section - Word Offset #3 - Bit 0 - VEB Statistics Disable.

Description:

When set to 0b - VEB statistics are enabled.
When set to 1b - VEB statistics are disabled (default).

The statistics counters disabled by this bit are:

GLPRT_RUPP[0]
GLSW_GOTCH/L
GLVEBVL_GOTC_[n]
GLVEB_TCBCH/L[n]
GLVEB_TCPCH/L[n]
GLSW_UPTCH/L[n]
GLSW_MPTCH/L[n]
GLSW_BPTCH/L[n]
GLSW_GORCH/L[n]
GLVEB_VLBCH/L[n]
GLVEB_RCBCH/L[n]
GLVEB_RCPCH[n]
GLSW_UPRCH/L[n]
GLSW_MPRCH/L[n]
GLSW_BPRCH/L[n]
GLSW_RUPP_[n]
GLVEB_VLUPCH/L[n]
GLVEB_VLMPCH/L[n]
GLVEB_VLBPCH/L[n]

This bit can be used to disable VEB statistics and improve 64-byte packet performance in SR-IOV or any other configuration that has more than one VSI connected to a port.

Note: The EVB Protocols Enable bit originally mapped to bit 0 of the same NVM word is now mapped to bit 1.

Starting from NVM 5.02, the NVM bit is no longer used. VEB statistics are enabled/disabled on a per-VEB basis using a new flag in the Add VEB command. See the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* for details. Intel drivers that accompany this NVM release disable the VEB statistics on all VEBs by default.

Software drivers before Release 20.7 (i40e 1.4.25, ixl 1.4.26, VMware ESX i40e v1.4.26) should not be used with NVM 5.02 and above.

9. Unicast Hash Filtering Removal

Unicast Hash filtering is removed from the 710-series device switching elements options and should not be used. Unicast MAC Addresses can be filtered by Perfect filtering (up to 2K entries) or Promiscuous filtering.

10. Input Reference Clock Rise/Fall Times

The minimum input reference clock rise and fall times (Tr/Tf) in Table 14-23 in the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* changes from 300 ps to 50 ps. This is a relaxation of the electrical specification for external oscillators.

Note: The routing of the input clock between the external oscillator and the 710-series device balls must be routed as a differential pair. Target 100 Ω differential impedance.

11. PCI Revision ID Update for B1 Stepping

The PCIe Revision ID is determined by an XOR of the value in the following registers:

- GPLCI_REVID
- GLPCI_DREVID

The B1 stepping changed the value in GPLCI_REVID to 0x3. The value in GLPCI_DREVID did not change and is still 0x1.

The resulting Revision ID in PCIe config space is 0x2 for B1 silicon. It was 0x1 for B0.

12. Force 40G Enabled Status Bit

In the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, Table 3-55, "Get link status response data structure", Bit 4.4 is now defined as "Force 40G Enabled". This bit is set if link is configured to XLLPI, and should be cleared otherwise.

This is valid from NVM 5.02 and on.

13. Receive Control Path Registers Admin Queue Commands

Two new AQ commands - Receive Control Path Read/Write Register - have been added to NVM version 5.02. These commands should be used by the software device driver to access any of the following registers, since direct software access to these registers is not always reliable under Rx small-packet stress. Refer to the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* for the command syntax.

- PFQF_CTL_0
- PFQF_HENA
- PFQF_FDALLOC
- PFQF_HREGION
- PFLAN_QALLOC
- VPQF_CTL

- VFQF_HENA
- VFQF_HREGION
- VSIQF_CTL
- VSILAN_QBASE
- VSILAN_QTABLE
- VSIQF_TCREGION
- PFQF_HKEY
- VFQF_HKEY
- PRTQF_CTL_0
- GLQF_CTL
- GLQF_SWAP
- GLQF_HASH_MSK
- GLQF_HASH_INSET
- GLQF_HSYM
- GLQF_FD_MSK
- PRTQF_FD_INSET
- PRTQF_FD_FLXINSET
- PRTQF_FD_MSK

14. Lane Ordering for Optical Breakout Cable

Starting from NVM version 5.02, the 710-series device supports the use of two QSFP+ breakout cables with the same two channels on each cable (CfgID 6.3), in addition to the previously-supported mode that uses the top two channels on one breakout cable and the bottom two channels on the other (CfgID 6.4).

15. Set Local LLDP MIB when DCBX Agent is Disabled or Stopped

According to the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, if the Set Local LLDP MIB command (0x0A08) is received while the FW DCBX agent is disabled or stopped, the MIB is parsed by firmware and used to configure the local DCB settings of the port, with no DCB TLV exchange with the peer performed by firmware.

Starting from NVM 5.02, if the Set Local LLDP MIB command is received while the DCBx specific agent is stopped, the command returns an EPERM error. If the command is received while the LLDP agent is stopped, it sets the local MIB without exchanging LLDP with peer, and returns SUCCESS.

16. Teredo UDP Tunneling Offload Support

Starting in NVM 5.04, Intel removed support for UDP Teredo tunneling offload.

Note: This feature was not supported in Intel Drivers.

17. GLQF_PCNT Counters

GLQF_PCNT counters do not wrap around.

Software should periodically clear these counters by writing any value. Note that the counter might miss a few events during the clearing process.

Starting in i40e v 2.2.x the driver keeps a count in software. However, there is a possibility of missing a few counts. Other Intel drivers do not use the counter.

18. Flash CS Negation Time

The 710-series device's minimum value of the FLSH_CE_N High Time (tCS) specification was incorrectly reported in the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* as 25 ns. The actual minimum value of t_{CS} is 80 ns, except for the case of consecutive Read Status Register commands when the minimum value is 70 ns.

This specification is consistent with the requirements of the supported flash devices listed in Section 14 of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, except for the Micron/Numonyx M25PX32 and M25PX64, which require a minimum CS negation time (t_{CSH}) of 80 ns. No failures have been reported due to this specification mismatch.

19. Parsing of MPLS Headers

Starting from NVM image revision 5.02, the 710-series device identifies and skips up to 2 MPLS labels as described in Section 7.3.1.4 of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*.

20. IEEE 1588 (TimeSync) over UDP is Not Supported in NVM 5.05

Starting from NVM image revision 5.05, the 710-series device does not process IEEE 1588 (TimeSync) PTP packets in the UDP payload. It only processes PTP packets using the Layer 2 packet format.

Starting in NVM 6.01, this is supported.

21. Internal Thermal Diode Measurements are Not Reliable

Sections 13.6.7, 15.8 and 15.9 of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*, revisions 2.7 and earlier describe the use of the 710-series integrated thermal diodes. Intel discovered that the integrated thermal diodes are not reliable across process, voltage, and temperature corners as an on-die indicator of the silicon temperature. As such, the thermal diode functionality should not be used and is no longer supported by Intel. These sections have been removed from the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* and associated design collateral.

Customers designing their own boards who wish to measure the temperature of the 710-series parts might choose to use an external I^2C thermal sensor with a remote diode implemented as an external transistor. I^2C temperature sensors with 1 °C accuracy are available from major vendors. Please follow the vendor recommendations for the circuit implementation. The I^2C temperature sensor should be connected to the BMC or thermal management I^2C bus of the customer host system. Among other considerations, physical placement of the external transistor should be as close to the 710-series package as permissible. The temperature of the 710-series devices might be correlated through empirical lab measurements in customer implementations.

22. Fixed VCCD Voltage Supply

A new preferred Fixed VCCD Voltage Supply design option is available. The Fixed VCCD voltage supply simplifies designs using the 710-series device. The VCCD Voltage Scaling supply design option is still valid, and is known as "VCCD Legacy" voltage supply option. See Section 13.2.2 in the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*.

23. Transmit Scheduler Does Not Support Weighted Strict Priority Arbitration

The 710-series device transmit scheduler supports strict priority and weighted round-robin arbitration schemes, but it does not support weighted strict priority. Any TC that uses strict priority is assigned unlimited bandwidth.

24. Reduction in the Size of the VSI Replication List Pool

Starting from NVM version 9.00, the VSI Replication List Pool size is reduced from 6K VSI list entries to 2K VSI list entries.

VSI replication lists are used when a single switch filter routes packets to more than one VSI. For example, a broadcast promiscuous filter might route broadcast packets to all the VSIs of a port. Additionally, this change reduces the small-packet receive throughput if a large fraction of the arriving packets are replicated.

Most applications do not require more than a few hundred VSI list entries and only replicate a small fraction of the packets. Therefore, the effects of this change are not expected to be visible in most applications. Please contact your Intel representative if your application is affected by this change.

25. PCIe REPLAY_TIMER Timeout Value Increase

Starting with NVM version 9.30, the timeout value for the PCIe REPLAY_TIMER has been increased on Intel 710-series device-based NICs to prevent spurious timeouts on systems with retimers, as recommended by the *PCI Express Base Specification, Rev. 4.0*. The timeout values are 49-50 μs for PCIe 2.5GT/s and 5.0 GT/s links and 24-25 μs for PCIe 8.0 GT/s links.

2.3 Documentation Changes

1. X710/XXV710/XL710 MDC Clock Frequency Min/Max Parameters

In the MDIO AC Specification section of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* (Section 13.6.2.6, "MDIO AC Specification"), the values of MDC Clock Frequency should be corrected to Min: 2.3 MHz and Max: 14 MHz.

The corrections are shown as follows in **red bold** text:

Table 13-19. MDIO I/F Timing Parameters

Symbol	Parameter	Min	Typ	Max	Units	Note
t_{MCLK}	MDC Clock Frequency	**2.3**		**14**	MHz	
t_{MH}	MDIO Hold Time	10			ns	
t_{MSU}	MDIO Setup Time	10			ns	
t_{MPR}	MDIO Propagation Delay	10		30	ns	

2.4 Errata

1. TX Performance Degradation for Small Cloud Packets

Problem:

Happening for GRE+IPv6+TCP without payload. Degradation is expected to give 33 Gb/s instead of 34 Gb/s.

Implication:

This is seen if GRE+IPV6+TCP Packet is transmitted with no payload. This not typical packet format, and is not expected is most use cases.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

2. PCIe Subsystem ID Incorrectly Reported for All PCI Functions Except Function 0

Problem:

All PCIe functions except Function 0 report a Subsystem ID of 0x0000 in the configuration space (including related Virtual Functions) regardless of the value programmed in the NVM.

Implication:

No functional impact to the device or drivers. However, this might impact the branding of the device if the Subsystem ID is used to select the device branding string.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

3. Illegal Byte Error Statistical Counter Inaccuracy

Problem:

Short packets with bad symbols that arrive back-to-back might not be counted by GLPRT_ILLERRC.

Implication:

GLPRT_ILLERRC is inaccurate.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

4. Receive Performance Degradation with Specific Cloud Header

Problem:

A small performance degradation is expected when receiving back-to-back GRE+IPv6+TCP cloud frames with 128-byte Header and almost no payload.

Implication:

Expected 33 Gb/s instead of 34 Gb/s.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

5. MCTP Discovery Error when Replacing Active PF

Problem:

MCTP Discovery might respond with a wrong PF ID when BMC is replacing the active PF. Expected to be a rare scenario on specific machines.

Implication:

PF replacement might not work for MCTP.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

6. RX Queue Disable is Reported Done Before It is Disabled

Problem:

RX Queue disable is reported done before it is disabled.

Implication:

An RX Hang could result if the software re-enables the queue too early.

Workaround:

An RX Queue should be reused only after a minimum delay of 50 ms. This workaround is implemented in Intel Software Release 19.3.

Status: B0=Yes, B1=Yes; NoFix

7. TX Descriptor Might be Read Twice

Problem:

A TX Descriptor might be read more than once in corner case conditions.

Implication:

Negligible.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

8. Immediate Interrupts are Delayed in Very Loaded System

Problem:

In a case where there are ten or more active queues in the system, and some of the queues are assigned with immediate interrupts, the interrupt delay might exceed the value specified in the Datasheet (”ITR and immediate interrupts jitter” table).

Implication:

Low performance impact

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

9. ECRC Bits are Not RO when ECRC is Disabled

Problem:

ECRC bits in PCIe AER registers are writable even when ECRC is disabled.

Implication:

Specification compliance issue.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

10. NC-SI I/Os Output Rise Slew Rate is Higher Than Specification

Problem:

NC-SI I/Os output rise time might be as low as 400 ps, while the NC-SI Specification requires a minimum of 500 ps.

Implication:

Specification compliance issue.

Workaround:

Place a 30 Ω resistor in serial to the pad.

Status: B0=Yes, B1=Yes; NoFix

11. TC Strict Priority Does Not Work as Expected

Problem:

An UP might not get exclusive priority if PCIe bandwidth is insufficient (although gets higher priority).

Implication:

RX TC strict priority limitation.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

12. Management-only Packets Cannot be Ignored for Wake-Up

Problem:

Due to a “NoTCO” wake-up capability malfunction, a wake event might be issued for packets that are expected to be routed to the BMC exclusively.

Implication:

Management-only packets cannot be ignored for wake-up.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

13. Common Clock Configuration Bit Specification Compliance

Problem:

Common clock configuration bit should be writable for all PFs, but it is not always writable for a PF > 0.

Implication:

Specification compliance issue. No functional impact.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

14. Low Latency TC Might be Momentarily Starved

Problem:

Low Latency TC might be momentarily starved under TPB Non-Strict Priority (RR) policy when both Bulk and Low Latency traffic are pending.

Implication:

Low Latency TC impact.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

15. Round Robin (RR) Bandwidth Distribution is Traffic Dependent

Problem:

Under RR RX Policy, RX bandwidth might be distributed unevenly among ports and TCs if PCIe bandwidth is smaller than incoming traffic, or traffic is a stream of small packets (smaller than 128 bytes).

Implication:

Uneven traffic distribution under RR.

Workaround:

Use Strict Priority policy instead of Round Robin.

Status: B0=Yes, B1=Yes; NoFix

16. L2 Tag Stored in the Wrong RX Descriptor Field

Problem:

If two L2 tags (for example: VLAN and S-TAG) are programmed to be extracted to the receive descriptor, and the receive packet includes only a single L2 tag, the extracted L2 tag is always posted in the L2TAG1 field if L2TSEL is set to 1b, or to L2TAG2 if L2TSEL is set to 0b.

Implication:

In the following cases there are no implications:

1. If the receive data flow always includes two L2 tags.
2. If the receive data might include packets with a single L2 tag, but are always the same tag type (first or second).

If the receive data flow that might include packets with only one L2 tag (which can be either the first or the second tag), software cannot identify which of the two enabled L2 tags were extracted to the receive descriptor.

Workaround:

If the receive data flow includes packets with only one L2 tag, and software is not able to identify if it is the first or the second tag, it should not enable more than a single L2 tag to be extracted to the receive descriptor.

Status: B0=Yes, B1=Yes; NoFix

17. Internal VLAN is Not Reflected in RX Descriptor

Problem:

When SHOWIV field is set in the receive queue context, the internal VLAN is stripped, but it is not inserted in the RX descriptor as expected.

Implication:

Internal VLAN is not reflected in RX descriptor.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

18. Transmit Queue Group with Single Queue Enabled Performance

Problem:

A transmit queue Group with single Queue enabled might have performance limitations when scheduling consecutive packets.

Implication:

TX Performance issue.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

19. A Switching Table Might Reduce Small Packets Performance

Problem:

If the switching table is relatively full, it might reduce performance with a continuous stream of packets smaller than 160 bytes. A data stream that includes a mix of small and big packets should not experience any degradation.

Implication:

Small packets performance impact.

Workaround:

Avoid switching table fullness.

Status: B0=Yes, B1=Yes; NoFix

20. Incorrect Optic Module Presence Reported to BMC

Problem:

When optic module (SFP+/QSFP+) is removed, NC-SI Get Link Status reports Media Available.

Implication:

Removing the optical module is not reported in Get Link Status.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

21. Set Binding Command is Not Functional for IPv4

Problem:

Set Binding command is not functional for IPv4.

Implication:

No manageability traffic after command.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

22. Get PHY Abilities Does Not Return 20GBASE-KR2

Problem:

Get PHY Abilities command response does not set 20GBASE-KR2 PHY type bit (bit 30), even if set in NVM.

Implication:

Backplane setups with 20 GbE KR2 settings lack the report in Get PHY Abilities.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

23. "Multi-speed module timeout" NVM Parameter Has No Effect

Problem:

During multi-speed module link establishment flow, the "Multi-speed module timeout" defined in NVM is not used.

Implication:

When using multi-speed module, TTL when moving from one speed to another speed, is at least 5 seconds.

Workaround:

Force a desired speed by disabling the other speed using the "set phy config" admin command.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

24. Cloud Traffic Over VEB is Transmitted to LAN

Problem:

Cloud traffic over VEB is transmitted to LAN.

Implication:

Cloud traffic over VEB is transmitted to LAN.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

25. VLAN Prune is Not Functional

Problem:

Default action for VLAN Prune table is not set after "Add VLAN AQ" command.

Implication:

VLAN Prune is not functional.

Workaround:

Additional VLAN Prune configuration should be done by software.

Status: B0=Yes, B1=Yes; NoFix

26. VLAN Insertion Limitation with Specific Tunneling Packet

Problem:

VLAN cannot be inserted in the Inner L2 Header of the following tunneled packet:

L2 Header	IPV4 Header - No Checksum	UDP Header	Inner L2 Header	Inner IPv4 Header - No Checksum	Inner Payload or Encapsulation

Implication:

VLAN Tag cannot be inserted.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

27. Jabber Packets are Not Counted in Jabber Packets Received Field of NC-SI Get Controller Packet Statistics Command

Problem:

Upon issuing NC-SI Get Controller Packet Statistics command, return value counter 12 - Jabber Packets Received should reflect the number of packets received which are larger than maximum frame size. This counter does not work, and the return value is always 0.

Implication:

Cannot get the number of jabber packets received using the NC-SI Get Controller Packet Statistics command.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

28. INTENA_MSK Setting Might Clear Interrupt

Problem:

A write access to a xxINT_DYN_CTLx CSR with INTENA_MSK bit (bit 31) set to 0 clears the corresponding interrupt bit in the PBA array.

Implication:

There is a possibility of an Interrupt missing. However, current Intel software implementation has this bit set to 1 except when enabling or disabling interrupts.

Workaround:

INTENA_MSK should be set in all CSR write accesses other than INTENA bit change.

Status: B0=Yes, B1=Yes; NoFix

29. Manageability Checksum Filtering of IPv6 Packets

Problem:

The IPv6 checksum calculation could be incorrect for received packets that contain either a Routing (Type 2) Extension Header or a Destination Options Extension Header that includes a Home Address option.

Implication:

If the manageability filtering is configured to drop packets with checksum errors, IPv6 manageability packets with the extension headers described above could be incorrectly dropped.

Workaround:

Do not enable checksum filtering for manageability if the IPv6 Extension Headers described above are used on manageability traffic.

For SMBus: The Enable Xsum Filtering to MNG bit should be 0b in the Update Management Receive Filter Parameters command and in the Set Common Filters Receive Control Bytes command if these commands are used.

For NCSI: Do not use the Enable Checksum Offloading Command (Intel OEM Command 0x23).

Status: B0=Yes, B1=Yes; NoFix

30. Link Remains Up During Power Saving State

Problem:

Intel X710/XL710-based devices might maintain link, regardless of system power state, as long as power is provided to the device.

Implication:

Link remains up during power saving state.

Workaround

None

Status: B0=Yes, B1=Yes; NoFix

31. PRTDCB_RUP2TC and PRTDCB_TC2PFC are Not Writable

Problem:

PRTDCB_RUP2TC (0x1C09A0) and PRTDCB_TC2PFC (0x001C0980) CSRs cannot be written directly by software when CSR protection is enabled.

Implication:

Programming this CSR is required if the software is configuring DCB on the device.

Workaround:

For PRTDCB_RUP2TC: Write to PRTDCB_RUP2TC as usual, then use a Direct Admin command with the following values to complete the write transaction.

For PRTDCB_TC2PFC: Write to PRTDCB_TC2PFC as usual, then use a Direct Admin command with the following values to complete the write transaction.

Field	Byte	Value PRTDCB_RUPTC	Value PRTDCB_TC2PFC
Flags	0-1	0x0	0x0
Opcode	2-	0xFF04	0xFF04
Data Length	4-5	0x0	0x0
Return Value/VFID	6-7	0x0	0x0
Cookie	8-15	Arbitrary value defined by software	Arbitrary value defined by software
Param 0	16,--19	0	0
Param 1	20-23	(0x000AC440 + 0x4 * PRT)	(0x000AC200 +0x4 * PRT)
Data Address High	24-27	0	0
Data Address Low	28-31	<CSR Write Data>	<CSR Write Data>

Status: B0=Yes, B1=Yes; NoFix

32. AER Header Log Might be Invalid

Problem:

If more than two uncorrectable function-specific errors are reported to VFs connected to the same PF, the Advanced Error Reporting (AER) Header Log (PCIe Configuration Registers offset 0x11C... 0x128) might be invalid.

This occurs only in case that one or more of the two errors have been cleared by the host, and a 3rd one arrives later for a VF connected to the same PF. In this case, the header log of this last error might be corrupted.

Implication:

Error source debug limitation. Uncommon systems suffering from multiple uncorrectable errors might have invalid AER Header Log.

Workaround:

PCIe trace data collected by a protocol analyzer can alternatively be used to recognize the TLP that is causing the error.

Status: B0=Yes, B1=Yes; NoFix

33. A CfgWr to a VF TLP with Error Might Generate an Error Message with Wrong VF Number

Problem:

When a CfgWr TLP that is poisoned or has a parity error is received by the 710-series device, an error message with the wrong VF number might be generated. Note that the status is correctly reported in the respective VF Status registers.

Implication:

PCIe error message with wrong Requester ID.

Workaround:

When the OS gets an error message and the status registers bits are cleared, it should poll the other VFs' status registers.

Status: B0=Yes, B1=Yes; NoFix

34. No LAN-to-BMC Pass-through Traffic in Dr State

Problem:

While in Dr state and pass-through is enabled, the 710-series device should keep pass-through functionality active. However, LAN-to-BMC traffic is not functional in Dr state.

Implication:

Cannot maintain manageability pass-through traffic while the system is in Soft Off G2/S5 state.

Workaround:

An NVM workaround is available in NVM 4.42.

Status: B0=Yes, B1=Yes; NoFix

35. MNG Packets are Dropped while a Function-Level Reset to PF 0 is in Progress

Problem:

When Function-Level Reset (FLR) is applied to PF 0, it also resets the LAN-to-BMC pass-through flow.

Implication:

LAN-to-BMC pass-through traffic stops while FLR is applied to PF 0.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

36. Priority Flow Control (PFC) is Not Fully Functional

Problem:

PFC is not fully functional. Tx does not pause when receiving PFC.

Implication:

Possible buffer overflows in the link partner's receive path.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

37. NC-SI Set Link Compliance

Problem:

- Set Link does not fail when many speeds are selected without auto-negotiation.
- Set Link Pause Capability bit implementation is inverted (should be 0b to enable).

Implication:

BMC flow modification.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

38. VF's Admin Send Queue Might Hang for VFs Greater Than 31

Problem:

For VFs greater than 31, the 710-series device might ignore a pending command directed to a specific VF and not execute the command.

Implication:

The Admin Send Queue (ASQ) of VFs 32-62, 64-94, 96-126 might be ignored and fail to execute.

Workaround:

Do not use affected VFs.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

39. NVM Update Admin Queue Command Responds with the Incorrect Error Code

Problem:

NVM Update AQC returns the wrong value when module authentication fails (EAGAIN instead of EACCESS).

Implication:

NVM Update AQC responds with the incorrect error code.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

40. Outer User Priority Incorrectly Configured

Problem:

When applying the Set Port Parameters Admin Queue command with Enable_Double_VLAN=0b, the Outer VLAN handling might be incorrect.

Implication:

Incorrect L2 Tag handling.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

41. SFP+/QSFP+ Modules or an External 10GBASE-T PHY Might Not be Recognized

Problem:

SFP+/QSFP+ Modules or an external 10GBASE-T PHY might not be recognized after GLOBR.

Implication:

No link.

Workaround:

Reset the 710-series device.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

42. SFP+ Optical Module Might Fail to Link in Specific Adapters

Problem:

During a specific sequence of SFP+ laser enable and module recognition, I^2C might be busy and the SFP+ optical module qualification might fail in some adapters.

Implication:

Link is down.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

43. VLAN Mirroring Admin Queue Command Writes Only the First VLAN in a List

Problem:

When a VLAN mirroring AQ command with multiple VLAN entries is sent, only the first VLAN is configured.

Implication:

VLAN misconfiguration.

Workaround:

Use only one VLAN per AQ command.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.42

44. DCBx Resume of a Port Affects Other Ports

Problem:

When DCBx resumes a port's traffic, done after port draining is performed, traffic might also be resumed for other ports.

Implication:

A port might be unintentionally resumed.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.53

45. PHY Type Field is Not Reported when Link is Down

Problem:

PHY Type field in the Get PHY Abilities AQ response is not reported when link is down.

Implication:

Software reports PHY type as unknown when link is down.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

46. A Global SDP Might be Affected by a Specific Port Power State

Problem:

When a GPIO is defined as a global SDP and its behavior is unrelated to any specific port, the GLGEN_GPIO_CTL.*PRT_NUM_NA* bit should be set, and the SDP value should be controlled by the GLGEN_GPIO_SET register regardless of port state.

However, the *PRT_NUM_NA* bit does not take effect, and SDP output is tri-stated or driven high (depending on GLGEN_GPIO_CTL.*OUT_CTL*) when the port specified at the *PRT_NUM* field is in power-down state.

Implication:

A global SDP unrelated to any specific port is disabled according to a port power state.

Workaround:

This should be taken in account in the board design. In some cases, it might just be a matter of inverting the polarity.

Status: B0=Yes, B1=Yes; NoFix

47. Legacy SMBus: Failure to De-assert Alert Signal when Not Using ARA Cycle

Problem:

In legacy SMBus mode, the MC might get an indication of outstanding events through the SMBALRT_N line. The MC should then do an ARA cycle to get the indicating function. It can instead read the status of all functions. If the MC fails to do this, and reads only a single function status, the SMBALRT_N line will never de-assert, even if the timeout expires.

Implication:

SMBALRT_N is not de-asserted.

Workaround:

Poll status of all functions.

Status: B0=Yes, B1=Yes; NoFix

48. XL710 40 GbE MAC Transmit Packet Issue

Problem:

In the XL710 40 GbE MAC, if the packet to be transmitted contains a specific 8-byte sequence within the header or payload, an invalid packet is transmitted and results in a CRC error at the receiver.

This issue impacts all 40 GbE modes that the XL710 supports. No 10 GbE modes are impacted. See the following table for impacted SKUs and modes.

Device	Stepping	Master Material Number	Impacted Modes	Modes Not Impacted
FTXL710-AM1	B0	MM# 936551 (T&R) MM# 936552 (Tray)	1x40 QSFP+: XLPPI, CR4 Backplane: KR4, XLAUI	4x10/2x10/1x10 SFP+: SFI Backplane: KR, KX4, KX
FTXL710-AM2	B0	MM# 936549 (T&R) MM# 936550 (Tray)	2x40/1x40 QSFP+: XLPPI, CR4 Backplane: KR4, XLAUI	4x10/2x10/1x10 SFP+: SFI Backplane: KR, KX4, KX

The following 10 GbE SKU is not impacted.

Device	Stepping	Master Material Number	Impacted Modes	Modes Not Impacted
FTX710-AM2	B0	MM# 936553 (T&R) MM# 936554 (Tray)	None	2x10/1x10 SFP+: SFI Backplane: KR, KX4, KX

Implication:

The device might transmit a packet incorrectly, causing the packet to be dropped by the receiver.

Workaround:

None for 40 GbE operation.

Devices used in 10 GbE mode do not exhibit this behavior.

Status: B0=Yes; NoFix. B1=No; Fixed

49. Unrecognized Optical Modules Will Not Link with NVM Release 4.53

Problem:

Supported optical modules are identified based on compliance code fields as defined in the SFP+ or QSFP+ specifications. When the 710-series device does not recognize an optical module (QSFP+ or SFP+), it continually tries to establish link by setting the PHY to XLPPI or SFI (Opportunistic Link Setup).

In NVM release 4.53, Opportunistic Link Setup flow is blocked.

Implication:

Unrecognized optical modules cannot establish link with NVM release 4.53, even though they linked in previous releases.

Workaround:

Use QSFP+/SFP+ modules recognized by the 710-series device, or do not upgrade to NVM release 4.53.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

50. Get Link Status AQ Command Might Return Incorrect Status

Problem:

If there is an I^2C access error when executing the Get Link Status AQ command, the 710-series device might falsely provide a link down response.

Implication:

A transient error in accessing the external module via I^2C causes the software device driver to report a link flap to the system.

Workaround:

If a Get Link Status response shows a link de-assertion, the Get Link Status command should be repeated.

Status: B0=Yes, B1=Yes; NoFix

51. LPLU is Not Functional in KR Mode

Problem:

If LPLU is enabled in KR mode, a link is not established at D3/Dr state.

Implication:

If LPLU is used in KR mode, the link is down on D3/Dr. If LPLU is disabled, the power consumption impact is negligible.

Workaround:

Disable LPLU.

LPLU is enabled by default in NVM v4.42 images. In NVM v5.02 LPLU is disabled in backplane NVM images.

Status: B0=Yes, B1=Yes; NoFix

52. A Function-level Reset Might Affect Other Functions

Problem:

When a function-level reset is applied (PFR, VFR or VMR), under rare conditions it might affect the Tx of a different function.

Implication:

Tx hang.

Workaround:

To prevent the failure, ensure that all queues belonging to the entity to be reset are disabled before initiating the reset. If this cannot be ensured, the failure could occur and the software device driver should use a CORER to recover from a Tx hang that cannot be cleared by a function-level reset.

Status: B0=Yes, B1=Yes; NoFix

53. Rx Packet Drops Even with Priority Flow Control

Problem:

When using flow control, the expectation is for no Rx packet drops caused by a Receive Packet Buffer overflow. In the situation where Priority Flow Control (PFC) is enabled on some traffic classes, but not on all enabled traffic classes, there is a possibility for the Receive Packet Buffer to fill up and drop packets belonging to any traffic class.

Implication:

PFC is not completely effective in preventing Receive Packet Buffer overflows under small-packet stress conditions.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

54. Configuration of Stopped DCBx Agent is Not Returned to Default

Problem:

The 710-series device should return to its default configuration when a DCBX agent has stopped, but it does not.

Implication:

Configuration of stopped DCBx agent is not returned to default.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

55. B1: Incorrect Revision ID Reported to MC

Problem:

In B1 stepping, the silicon revision ID that is reported back to the MC is the same as B0.

Implication:

Incorrect revision ID reported to MC.

Workaround:

None.

Status: B0=No, B1=Yes; Fixed in NVM 5.02

56. SMBUS: Wrong Driver Status Reported to MC After TCO Isolate Command

Problem:

The TCO Isolate command disables PCIe write operations to the LAN port. Once TCO isolate is set, the software device driver is supposed to be disabled, but the Read Status command is retrieving a status of “LAN Driver Is Up”.

Implication:

Wrong driver status reported to MC after TCO Isolate command.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

57. Shutdown LLDP Agent Command

Problem:

Executing the Stop LLDP Agent AQ command with the command field set to Shutdown LLDP Agent results in the internal LLDP agent transmitting a shutdown frame. Starting the agent afterwards results in transmitting an additional shutdown frame.

Implication:

Spurious LLDP shutdown frame.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

58. NC-SI: A Link Change Event Might Cause an Endless Loop of AENs

Problem:

A link change event should cause an NC-SI Asynchronous Event Notification (AEN). If the software device driver is down, an endless loop of AENs might be generated.

Implication:

Endless AEN messages.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

59. DCBx Configuration Might Change After LLDP Stops

Problem:

After a Stop LLDP Agent AQ command, the LLDP agent should be stopped but DCBx configuration should stay unchanged. However, if a CORER or GLOBR is asserted when the LLDP agent is stopped, the configuration might be changed.

Implication:

DCBx configuration unstable when LLDP agent stops.

Workaround:

Prior to a Stop LLDP Agent AQ command, software should read the MIB (Get LLDP MIB AQ command). After LLDP stops, software should write a previous MIB (Set Local LLDP MIB AQ command).

Status: B0=Yes, B1=Yes; NoFix

60. LLDP Receive-only Mode Malfunctions

Problem:

When LLDP mode configured in the NVM LLDP Admin Status word is set to 1, the LLDP agent should be enabled to only receive LLDPDUs. However, when this setting is used:

- LLDPDUs are transmitted.
- The Stop LLDP Agent AQ command fails.

Implication:

The LLDP Rx-only mode is not functional.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

61. Non-optimal Receive Packet Buffer Partitioning

Problem:

Receive Packet Buffer (RPB) partitioning is a function of port link and traffic classes definition. In some corner case the partitioning is not optimal.

Implication:

Slightly increased probability of packet drops in stress cases.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

62. QSFP Configuration Transition Requires Power Cycle

Problem:

When performing a QSFP configuration change, the device should function with the new configuration after a reboot. However, when the change reduces the number of active ports, such as 4 x 10 GbE -> 2 x 40 GbE, a full power cycle might be required for the device to be functional in the new mode.

Implication:

Some QSFP configurations are not functional after a reboot.

Workaround:

Perform a full power cycle after using the QCU tool to change the QSFP configuration to a configuration with a smaller number of ports.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

63. MAC Link Flow Control Might Not be Updated After Auto-negotiation Restart

Problem:

If link is up and the link partner restarts auto-negotiation with new advertised flow control settings, the new advertised settings might be ignored.

Implication:

Incorrect flow control settings.

Workaround:

Software device driver should restart auto-negotiation if there is a chance that the link partner flow control settings have changed.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

64. PFR Flow Limitations

Problem:

1. Once the NVM resource ownership is held by a PF, a PFR on another PF will fail because the firmware attempts to request the NVM semaphore and fails.
2. PFR might not fully clear the internal switch table entries related to the PF.

Implication:

1. Device might be unavailable.
2. Partial reset.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

65. FC Refresh Threshold Field Ambiguity

Problem:

The FC Refresh Threshold field is the amount of time before the expiration of the current XOFF where another pause frame must be transmitted. This is how it is implemented for 10 GbE and lower link speeds.

In case of a 40 GbE link, the FC Refresh Threshold field in the Set MAC Config AQ command is incorrectly interpreted as the amount of time after which another flow control pause frame must be transmitted.

Implication:

Field ambiguity might cause misconfiguration.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

66. Link Auto-negotiation Reporting Might be Incorrect

Problem:

Information about the use of Auto-negotiation (AN) when establishing an Ethernet link might be incorrect in the following cases:

- The Get PHY Abilities AQ command and the NC-SI Get Link command report AN disabled when using CR4.
- The AN information reported by Get Link Status might be incorrect if the link is down or if the current link was established without AN.

Implication:

The auto-negotiation information is not always reliable.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 4.53

67. Inefficient Shared Resources Allocation

Problem:

Due to an inefficient resource sharing algorithm, an Add MAC, VLAN Pair AQ Command might return ENOSPC, indicating a lack of resources, when many resources are assigned as “Dedicated”.

Implication:

Shared resources starvation.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

68. Prepare for Endpoint Discovery MCTP Command

Problem:

Prepare for Endpoint Discovery MCTP command does not work.

Implication:

The MCTP bus owner might not be able to set up communication with the 710-series device.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

69. Delete Mirror Rule Admin Command Failure

Problem:

The mirror rule might not be removed by a Delete Mirror Rule Admin command using Rule Type 010b, virtual port egress mirroring.

Implication:

Delete Mirror Rule Admin command does not take effect. Mirroring could continue.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

70. LLDP Disable Can Result in Incorrect Configuration of the Receive Packet Buffer

Problem:

The default operation of the device is LLDP enabled. LLDP can be disabled by an AQ command or by a custom NVM obtained from Intel.

Disabling LLDP when using NVM 4.53 or 5.x can result in the device firmware not configuring the Receive Packet Buffer according to the link mode and flow control settings.

Implication:

If flow control is disabled, a hang of the receive data path could occur wherein all received packets are dropped.

If flow control is enabled, the incorrect settings could cause packet drops despite the flow control, making the flow control ineffective.

Also, throughput might be sub-optimal in some cases.

Workaround:

When using NVM 4.53 or 5.x, do one of the following:

- Do not disable LLDP or DCBx. Leave the device in default operation mode.
- Use the Set Local LLDP MIB admin queue command to apply the (default) DCB configuration during driver initialization and following any link-up event.

When using NVM 6.01 or later, the receive data path does not hang even if the FW LLDP is disabled.

Note: When the firmware DCBx agent is disabled, it is still necessary to apply the DCB configuration to the hardware. The software driver can use the Set Local LLDP MIB admin queue command following each link up event, or the software driver can use the Set DCB Parameters AQC (opcode 0x303) during initialization to instruct the firmware to apply the DCB configuration as required. If a non-standard configuration is required, the software driver can instead apply the DCB configuration by writing directly to the hardware CSRs.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

To ensure proper operation with Intel drivers, Intel recommends using NVM 6.01 or later and drivers from Release 22.6 or later.

71. PCIe Interrupt Status Bit

Problem:

The *Interrupt Status* bit in the Status register of the PCIe configuration space is not implemented and is not set as described in the PCIe specification.

Implication:

When using shared legacy PCI interrupts, software might use this bit to determine if the 710-series device has a pending interrupt. Since the bit is not implemented, the software might not handle the interrupt, resulting in a continuous interrupt assertion.

There is no implication when using MSI or MSI-X.

Workaround:

The *Interrupt Status* bit should not be used. Avoid using shared legacy PCI interrupts.

Status: B0=Yes, B1=Yes; NoFix

72. Changes to Tx Scheduler Structure Can Cause the Device to Become Unusable

Problem:

Certain changes to the Tx scheduler structure can cause the scheduler to hang, possibly resulting in an EMP reset. This has been observed, for example, when unloading a PF driver on a device that uses SR-IOV. The possibility of failure depends on the internal structure of the Tx scheduler tables and cannot easily be determined.

Implication:

Device hang, eventual EMP reset that removes the manageability configuration.

Workaround:

If there is a Tx hang detected that is not cleared by a PFR, software should attempt a CORER. Also, if a PFR times out, attempt a CORER.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.02

73. Glitch on SDP Outputs During GLOBR

Problem:

GPIO pins that are defined as SDP outputs (*PIN_FUNC* is 000b and *PIN_DIR* is 1b in GLGEN_GPIO_CTL) can have a high-to-low glitch during GLOBR if *OUT_CTL* is 0b.

The same applies when the port specified in GLGEN_GPIO_CTL.*PRT_NUM* is enabled/disabled.

Implication:

The implication depends on the use of the SDP. For example, an SDP used as a QSFP+ reset signal might cause the module to malfunction due to a short reset assertion.

Workaround:

One of the following:

- If the SDP is supposed to be high during GLOBR, set *OUT_CTL* to 1b.
- For a general-purpose 2-state SDP output (*PHY_PIN_NAME* is 0x3F), set *PIN_FUNC* to 001b (LED) and use the *LED_MODE* field (0000b or 1111b) to control the output value.

Status: B0=Yes, B1=Yes; NoFix

74. Function-Level Reset Fails to Complete

Problem:

In rare cases, the hardware activity of a function-level reset (PFR, VFR, or VMR) might fail to complete.

Implication:

PFR: Software times out while waiting for the PFR to complete. The firmware gets stuck and the firmware watchdog timer expires, triggering an EMPR.

VFR/VMR: Software times out while waiting for the reset to complete.

Workaround:

PFR: Software should re-initialize the device after the EMPR.

VFR/VMR: After a timeout waiting for the reset to complete, software should retry the reset by clearing and then setting the reset trigger bit (GLGEN_VFRTRIG.*VFSWR* for VFR or VSIGEN_RTRIG.*VMSWR* for VMR) and then restarting the polling for reset completion. After three retry attempts, abort with an error.

Status: B0=Yes, B1=Yes; NoFix

75. 10GBASE-T Link Issues

Problem:

The following issues can occur when using a 10GBASE-T link.

- If the link partner supports auto-negotiation but only advertises 10 GbE, a 10 GbE link is established even when Low-Power Link Up (LPLU) is enabled.
- No link with some switches when Low-Power Link Up (LPLU) is enabled.
- If the link partner supports auto-negotiation but does not advertise 100 Mb/s, no link is established when Low-Power Link Up (LPLU) is enabled.
- The link speed is sometimes 1 GbE instead of 100 Mb/s when in LPLU mode.
- If the Keep PHY Link Up (MNG_VETO) mode is set from the MC on one port, there might be no link on the other ports after a PCIe reset.
- If the Keep PHY Link Up (MNG_VETO) mode is set from the MC while in LPLU mode, a 10 GbE link might be established while still in the low-power state.

Implication:

Unreliable 10GBASE-T link establishment.

Workaround:

Avoid the scenarios described above.

Status: B0=Yes, B1=Yes; Fixed in NVM 5.03

76. Incorrect Flexible Payload Extraction from Flow Director Filter to Receive Descriptor

Problem:

When programming a Flow Director filter, if *FD_STATUS* is 10b, the FLEXOFF value provided in the programming descriptor is used incorrectly, and the wrong bytes are extracted to the receive descriptor.

Implication:

Incorrect descriptor content.

Workaround:

To get four bytes starting from offset N of the flexible payload to the receive descriptor, the value N-2 should be used for FLEXOFF. Byte offsets 0 and 1 cannot be extracted with *FD_STATUS* of 01b.

Status: B0=Yes, B1=Yes; NoFix

77. Filtering Restriction when Double VLAN is Enabled

Problem:

When Double VLAN is enabled using the Set Port Parameters AQ command, the L3 and L4 headers from any single VLAN or double VLAN packet are not available for filtering.

Implication:

When Double VLAN is enabled, RX packet filtering of packets containing a VLAN based on L3 and L4 headers is not functional.

Workaround:

Execute the following Direct AQ command before enabling Double VLAN:

Field	Byte	Value
Flags	0-1	0
Opcode	2-3	0xFF04
Data Length	4-5	0
Return Value	6-7	0
Cookie	8-15	Arbitrary value defined by software.
Param 0	16-19	0
Param 1	20-23	0x0026C7A0
Param 2	24-27	0
Param 3	28-31	0x000000A8

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

78. Transmit Hang in 2x40 GbE Configuration When Flow Control is Enabled

Problem:

When using a 2x40 GbE configuration, the transmit pipeline can hang if the flow control setting is changed on Port 0 while Port 1 is transmitting

Implication:

If the Port 0 link goes up or down while data is being transmitted from Port 1 and the Port 0 flow control setting changes as a result, the transmission could hang.

Workaround:

Disable flow control on Port 0.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

79. Aux Power Detected Bit Not Implemented

Problem:

The *Aux Power Detected* bit in the Device Status register of the PCIe Configuration Space is not implemented. The bit is always 0b.

Implication:

PCIe specification compliance, but this issue is not detected by the existing compliance testing.

It is not expected that any software uses this bit. If it is being used, the workaround should be implemented by the software.

Workaround:

Use the most-significant bit of PMCR.*PME_Support* instead.

Status: B0=Yes, B1=Yes; NoFix

80. SGMII Receiver Sensitivity

Problem:

The SGMII specification requires a maximum receiver sensitivity of 100 mV peak-to-peak. The XXV710 receiver sensitivity can be as high as 190 mV peak-to-peak.

Implication:

No expected implication, since the input signal voltage would normally be high enough.

Workaround:

Ensure that the input signal is strong enough when using an SGMII connection.

Status: B1=Yes; NoFix

81. IEEE 802.3 Clause 73 AN Does Not Support Parallel Detection

Problem:

When using Clause 73 auto-negotiation, parallel detection is not supported.

Implication:

Inability to link with legacy devices that do not have Clause 73 AN enabled.

Workaround:

Use the firmware-based 25G Link Enablement State Machine (LESM).

Status: B1=Yes; NoFix

82. IEEE 802.3 Clause 73 AN Echoed Nonce Field is Zero

Problem:

During the IEEE 802.3 Clause 73 auto-negotiation process, the last message page has the Echoed Nonce field set to 00000b even if the ACK bit is 1b.

Implication:

With certain link partners this might cause auto-negotiation failures.

Workaround:

None.

Status: B1=Yes; NoFix

83. KR Transmitter Output Waveform Violations

Problem:

The KR transmitter does not meet the IEEE 802.3 Clause 72.7.1.11 transmitter output waveform requirements for R_{pre} when both C(1) and C(-1) are disabled and c(0) is maximum.

Implication:

Conformance issue. Not expected to impact functionality.

Workaround:

None.

Status: B1=Yes; NoFix

84. 10GBASE-KR wait_timer Value Smaller Than Specification

Problem:

The 10GBASE-KR wait_timer is defined by IEEE 802.3 to have a value between 100 and 300 training frames. The actual value is 75 training frames.

Implication:

Potential training failure with some link partners.

Workaround:

None.

Status: B1=Yes; NoFix

85. Receive Queue Disable Can Get Stuck

Problem:

If there are no descriptors available for a receive queue that belongs to a no-drop TC and the queue is disabled at the same time that a packet arrives for the queue, the queue disable can get stuck.

Implication:

Head-of-line blocking continues despite an attempt to disable the queue.

Workaround:

To avoid this situation, the driver should try to ensure that there are always Rx descriptors available, especially when disabling an Rx queue.

If a head-of-line blocking situation does occur, it is handled as usual when the PFCTIMER expires.

Status: B0=Yes, B1=Yes; NoFix

86. Add Cloud Filter Command Can Fail with Return Code ENOSPC

Problem:

In some cases, an Add Cloud Filter admin queue command can fail with return code ENOSPC although the maximum number of filters has not been reached. Generally, this failure can occur after deleting some filters and then adding more filters.

Implication:

Reduction in the effective number of filters supported.

Workaround:

Perform a Core Reset and reconfigure the device.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

87. First Tx Descriptor Not Processed

Problem:

If there is a Tx queue belonging to a VSI that belongs to a VF or VM and supports multiple TCs, and the Tx queue is disabled and then re-enabled, the first Tx descriptor could be skipped.

Implication:

Detected as a Tx hang.

Workaround:

Avoid the problematic scenario.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

88. Set DCB Parameters AQC (Opcode 0x303) Might Return EINVAL Even when It Succeeds

Problem:

The Set DCB Parameters AQ command (opcode 0x303) might return EINVAL even when it succeeds.

Implication:

Software driver does not know whether the command succeeded.

Workaround:

Ignore the return code.

Status: B0=Yes, B1=Yes; NoFix

89. Receive IP Packets in a Low-Latency Traffic Class Are Not Fully Processed

Problem:

Receive packets that contain an IP header and belong to a low-latency traffic class (as defined by PRTDCB_RETSC.*LLTC*) are not fully processed by the 710-series device. The following processing is not performed on these packets:

- Validating the IP checksum.
- Validating the L4 checksum.
- Stripping/extracting the VLAN from a tunneled packet.

Implication:

Performing this data processing in software results in lower overall performance of the product.

Workaround:

When using DCBx, ETS should be enabled for all active TCs.

Status: B0=Yes, B1=Yes; NoFix

90. QSFP+ Modules with Tx Squelch

Problem:

The 710-series device could fail to link when using QSFP+ modules that implement Tx squelch. Intel-branded QSFP+ modules do not implement Tx squelch and are not affected.

Implication:

Failure to obtain a link.

Workaround:

Use an Intel-branded QSFP+ module.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.01

91. Activity LED Might Blink Regardless if Link is Up or Down for a Port

Problem:

710-series device Activity LEDs toggle as a result of BMC/HOST transmit packets regardless of the port link state. Activity LEDs are MAC_ACT or FILTER_ACT (set by the field *LED_MODE* - 1101 or 1110 respectively).

Implication:

The Activity LED might be blinking even if link is down.

Workaround:

BMC/HOST should transmit packets only when link is up.

Status: B0=Yes, B1=Yes; NoFix

92. PCIe Phase 2 Fails to Timeout Under Certain Channel Conditions

Problem:

With certain channel conditions it was found that the PCIe Phase 2 coefficient evaluation timer is too short, causing the device to enter a persistent failure state.

Implication:

If Phase 2 EQ fails to timeout, it will stay in this state for an unlimited amount of time, and the device will either fail to show up on the bus or cause a fatal error.

If the issue occurs, the 710-series device either fails to establish a PCIe link, or causes a PCIe fatal error.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 6.80

93. EMP Reset After Using Intel QCU Tool

Problem:

EMP reset occurs after changing the device configuration using Intel QCU tools. For example, changing from 2x40 to 4x10.

Implication:

Device hang and later EMP reset that removes the manageability configuration.

Workaround:

None.

Status: B0=Yes, B1=Yes; Fixed in NVM 7.20

94. 25G-AUI Takes a Long Time to Link with Optical Modules and AoC

Problem:

When using 25G-AUI mode with optical modules or AoC, the PHY might not link achieve link or might take a long time to link. It can take from a few seconds up to few minutes to link.

Implication:

No link or long time to link, ranging from a few seconds to minutes.

Workaround:

Update NVM to NVM 7.00.

Status: B0=Yes, B1=Yes; Fixed in NVM 7.00

95. PCIe Replay Timer Can Occasionally be Lower Than PCIe Spec Requirements

Problem:

PCIe Replay Timer can occasionally be lower than PCIe spec requirements

Implication:

PCIe Replay Timer Timeouts and Replay_Num rollover correctable errors. Due to this error, reduced PCIe performance is possible.

These errors can only be observed on platforms with downstream port ACK latencies that are beyond maximum PCIe spec limits.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

96. No Length Error on VLAN Packets with BAD Type/Length Field

Problem:

The 710-series device, when the link speed is 10G or lower, will not assert length error for VLAN packets that have a bad type/length field in the MAC header.

Implication:

No impact on system level performance. The packets are posted to the host as any other packets.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

97. Thermal Alarm Points for Pluggable Modules Are Using Fixed Values Instead of Reading from Module EEPROM

Problem:

Thermal alarm points for pluggable modules are using fixed values defined by the NVM instead of reading from module EEPROM.

Implication:

Incorrect thermal alarm points that could result false alarm events or failures to raise an alarms when necessary.

Workaround:

None.

Status: B0=Yes, B1=Yes; NoFix

98. Memory Leak in Receive Packet Buffer

Problem:

In rare circumstances, the memory used to store a packet in the Receive Packet Buffer is not released after the packet has been processed. Each time this occurs, the effective size of the Receive Packet Buffer is reduced.

Factors involved in the failure:

- Packets that are replicated to multiple VSIs, for example broadcast packets.
- Packets that are dropped due to disable queues or a lack of receive descriptors.

Implication:

If this failure occurs repeatedly, the effective buffering of the receive traffic is reduced over time.

If the effective fill level of the Receive Packet Buffer gets high enough, all input packets will be dropped on one or more ports. See Section 7.7.1.2.3 of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet* for the conditions under which this can occur.

Workaround:

To reduce the probability of failure:

- Remove unnecessary replications of packets to unused VFs/queues.
- Ensure that there are always receive descriptors available for all active queues.
- Increasing the Receive Packet Buffer thresholds can increase the time before packets start to be dropped.

Recovery:

- A reboot restores the operation of the Receive Packet Buffer.

Status: B0=Yes, B1=Yes; Fixed in NVM 9.00

99. Single-Byte I^{2}C Writes Clear the Following Byte

Problem:

When using the Set PHY Register admin queue command to write via the I^{2}C interface, the 710-series device performs a 2-byte I^{2}C write. The data value provided in the command is written to the byte address provided in the command, and the value 0 is written to the following byte address.

Implication:

If the address following the target address is also writable in the module EEPROM, the value of that location could be corrupted.

Workaround:

If the I^{2}C write can be done as a one-time configuration, use the bit-bang interface instead of the Set PHY Config command. Disable firmware link management before doing so.

Status: B0=Yes, B1=Yes; Fixed in NVM 8.40

100. Failure to Manipulate the Default MAN/VLAN Filter

Problem:

If the default MAC/VLAN filter (PF MAC Address with VLAN ID 0) is removed and then added back, a subsequent attempt to remove this filter is only partially executed. Further attempts to add or remove this filter malfunction.

Implication:

Inconsistent filter state if the default MAC/VLAN filter has been removed more than once.

Workaround:

Do not remove the default MAC/VLAN filter more than once between PF resets.

Status: B0=Yes, B1=Yes; Fixed in NVM 8.60

101. XXV710 PHY Lock-Up

Problem:

While attempting to establish a network link, the XXV710 Ethernet PHY locks up. No link can be established until a firmware reset or power cycle is performed. This is typically observed when restarting the link partner to which the XXV710 is connected, resulting in repeated attempts to establish a link.

NVM releases starting from version 8.00 are susceptible to the issue.

Implication:

Loss of link.

Workaround:

Disabling the link from software before restarting the link partner can significantly reduce the probability of this failure occurring.

When using the *i40e* Linux driver, the link can be disabled by setting the *link-down-on-close* flag and unloading the driver.

Status: B0=Yes, B1=Yes; Fixed in NVM 8.50

102. Remove MAC/VLAN Pair AQ Command Failure

Problem:

If more than 2048 MAC/VLAN filters have been added to the device, a Remove MAC/VLAN Pair AQ command for the filters beyond the first 2048 fails with the ENOENT error.

Implication:

Failure to remove MAC/VLAN filters. This erratum only applies to NVM version 8.40 and later.

Workaround:

Do not use more than 2048 MAC/VLAN filters in the device.

or

Use PF reset to clear MAC/VLAN filters instead of removing them individually.

Status: B0=Yes, B1=Yes; Fixed in NVM 8.60

103. LAN-to-BMC Packets Delayed or Dropped in Pre-Boot Environment

Problem:

In the pre-boot environment, the 710-series device is configured in “PXE Mode” (see Section 8.3.1.1.1 of the *Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*) due to memory limitations of some pre-boot drivers. To avoid dropping packets when a burst arrives during a pre-boot session, the 710-series device is usually operated in no-drop mode, meaning that incoming packets are buffered, rather than being dropped, if no Rx descriptors are available.

However, there are situations in the pre-boot environment, such as when a user menu is displayed, in which the application does not process Rx traffic from the 710-series device and does not refresh the Rx descriptors. When this occurs, received packets are backed up into the Rx packet buffer. If there is an active manageability pass-through session, the LAN-to-BMC packets might be blocked behind host packets in the queue. A dynamic drop mode has been implemented to occasionally drop host packets that are stuck for a long time. However, the drop algorithm does not cover all cases, and there can still be situations where the pass-through traffic is delayed.

Implication:

When the LAN-to-BMC traffic is delayed/dropped in the Rx packet buffer, the manageability session could be lost.

Workaround:

Avoid pre-boot user menus that wait indefinitely for user input. Keep the time during which Rx packets are not processed to a minimum.

For specific environments, there are NVM settings (PXE PFC Timer Value, PXE GPC High Threshold Value) that can be used to tune the dynamic drop algorithm. Contact your Intel representative for additional information.

Status: B0=Yes, B1=Yes; NoFix

104. Pre-boot Failure with Multiple Traffic Classes

Problem:

The Intel pre-boot drivers only use Traffic Class 0. If the Ethernet switch sends a DCBx configuration that includes multiple traffic classes, the 710-series device firmware might choose a traffic class other than 0 for the default VSI, in which case the pre-boot driver is unable to transmit.

Implication:

Preboot failure with certain switch configurations.

Workaround:

Disable DCBx on the switch or on the 710-series device.

OR

Configure the switch to map both User Priority 0 and User Priority 3 to Traffic Class 0.

Status: B0=Yes, B1=Yes; NoFix

105. Incorrect Receive Length Errors

Problem:

If an Ethernet packet is received with a packet length larger than 64 bytes and the *Length*/*Type* field contains a value that is smaller than the number of bytes remaining in the packet, an Rx length error is detected. This type of packet could be generated if a short packet is padded to 64 bytes and then a VLAN tag is inserted without removing any padding bytes.

When an Rx length error is detected, the packet is considered a bad packet and is discarded unless storing bad packets is enabled. Also, the GLPRT_RLEC counter is incremented.

Implication:

Packets with the structure defined above are dropped. These packets are often local diagnostic packets that might be discarded by the software anyway.

The driver reports the error count to the OS, possibly resulting in a high error rate that is reported by the OS.

Workaround:

The driver should not report the Rx length error count to the OS. This is implemented in Intel drivers starting from Release 28.2.

Status: B0=Yes, B1=Yes; NoFix

106. Device is Not Enumerated on the PCIe Bus Following Certain Warm Reset Flows

Problem:

PERST# assertion following a Hot Reset (SBR) cycle may cause the 710-series device to stay in reset or fail to be enumerated.

Some platforms perform a Hot Reset followed by a PERST# assertion as part of the warm reset sequence. On such platforms there is a small probability of this failure occurring on an OS reboot.

Implication:

PCIe link stays down or device is not enumerated on the PCIe bus. A full reset of the 710-series device is required to recover from this issue.

Workaround:

The following condition should be met to avoid any risk of observing this errata:

Ensure that PERST# is not asserted within 200 us of the 710-series device exiting Hot Reset.

Status: B0=Yes, B1=Yes; Fix Planned in NVM

3. Software Clarifications

Table 3-1. Summary of Software Clarifications

Software Clarification	Status
1. XL710 Option ROM Should Not be Integrated in the BIOS	N/A
2. VXLAN Guidance for VMware vSphere	N/A

1. XL710 Option ROM Should Not be Integrated in the BIOS

Previous generations of Intel networking controllers allowed the Option ROM to be stored in the flash attached to the device, or in the BIOS flash. The 710-series device requires the Option ROM, if one is used, to be stored in the flash attached to the 710-series device. This is done to maintain alignment of the pre-boot code with the internal 710-series device firmware when upgrades are necessary.

2. VXLAN Guidance for VMware vSphere

For VXLAN traffic in production VMware vSphere environments with the 710-series device, use the 1.3.38 ESXi driver or later. For the latest driver version currently available, please reference the VMware Compatibility Guide.

NOTE: ***This page intentionally left blank.***

LEGAL