Temperature Reporting Interfaces Support for Manageability

Application Note

Ethernet Products Group (EPG)

March 2024

Revision 1.2
630684-001

Revision History

Revision	Date	Comments
1.2[1]	March 22, 2024	Initial public release.
1.1	August 19, 2021	Updates include the following: • Updated NC-SI interface support on X710. • Updated Table 6, "Interfaces Supported to Report ASIC Temperature". • Updated Table 7, "Interfaces Supported to Report Plug Module Temperature". • Added Appendix A, "NC-SI OEM Commands".
1.0	July 22, 2020	Initial release (Intel Confidential).

1. There are no previous publicly-available versions of this document.

Contents

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

1.0 Introduction

This document provides an overview for the temperature reporting interface available in Intel Ethernet controllers for server applications. Thermal control plays an important role among the server operation. These interfaces provide a mechanism by which dedicated Management Controllers (MCs) can query the thermal data and status.

1.1 Terminology

Table 1. Acronyms and Definitions

Term	Definition
ASIC	Application Specific Integrated Circuit
BMC	Baseboard Management Controller
DMTF	Distributed Management Task Force
LOM	LAN on Motherboard
MC	Management Controller (also referred to as BMC)
MCTP	Management Component Transport Protocol
NC-SI	Network Controller Sideband Interface
NIC	Network Interface Controller
OCP	Open Compute Project
PCIe®	PCI Express®
PLDM	Platform Level Data Model
RBT	RMII Based Transport
RMII	Reduced Media Independent Interface
SFI	SerDes Framer Interface
SFP+	Small Form Factor Pluggable
SMBus	System Management Bus

1.2 Reference

Table 2. Reference Documents and Resources

Doc ID	Title	Location
613875	*Intel® Ethernet Controller E810 Datasheet*	https://www.intel.com/content/www/us/en/design/resource-design-center.html
596333	*Intel® Ethernet Controller X710-TM4/AT2 Datasheet*	
332464	*Intel® Ethernet Controller X710/XXV710/XL710 Datasheet*	
333369	*Intel® Ethernet Controller X550 Datasheet*	
336626	*Intel® Ethernet Controller I350 Datasheet*	
333016	*Intel® Ethernet Controller I210 Datasheet*	
DSP0222	*DMTF Standard. Network Controller Sideband Interface (NC-SI) Specification.*	https://www.dmtf.org/standards/pmci
DSP0236	*DMTF Standard. Management Component Transport Protocol (MCTP) Base Specification.*	
DSP0240	*DMTF Standard. Platform Level Data Model (PLDM) Base Specification.*	
	PCI Express Base Specification	https://pcisig.com/
	OCP Mezzanine card 2.0 Design Specification	https://www.opencompute.org/
	OCP NIC 3.0 Specification	

2.0 Overview of NIC Form Factor

Server designs have different considerations to fulfill their requirement of density, scalability and complexity. Intel® Ethernet controllers are adopted in different form factors to provide the communication interfaces. The hardware interfaces, like SMBus, PCIe, and RMII, are mostly leveraged for the manageability application. The section provides an overview for the hardware interfaces supported on the standard NIC form factors.

2.1 PCIe Add-In Card

PCI-SIG provides the PCIe pinout definition as shown in Table 3. Beside SMBus, the transport protocol over PCIe physical interface is used to communicate with the NIC for the management application, like MCTP, PLDM over MCTP, or NC-SI over MCTP.

Table 3. PCIe Pinout Definition

Pin #	Side B Connector		Side A Connector	
	Name	Description	Name	Description
1	+12V	+12 volt power	PRSNT#1	Hot plug detect
2	+12V	+12 volt power	+12V	+12 volt power
3	+12V	+12 volt power	+12V	+12 volt power
4	GND	Ground	GND	Ground
5	SMCLK	SMBus clock	JTAG2	TCK
6	SMDAT	SMBus data	JTAG3	TDI
7	GND	Ground	JTAG4	TDO
8	+3.3V	+3.3 volt power	JTAG5	TMS
9	JTAG1	+TRST#	+3.3V	+3.3 volt power
10	3.3Vaux	3.3 volt power	+3.3V	+3.3 volt power
11	WAEK#	Link Reactivation	PWRGD	Power Good
Mechanical Key Notch				
12	RSVD	Reserved	GND	Ground
13	GND	Ground	REFCLK+	Reference Clock Differential pair
14	HSOp(0)	Transmitter Lane 0 Differential pair	REFCLK-	
15	HSOn(0)		GND	Ground
16	GND	Ground	HSIp(0)	Receiver Lane 0 Differential pair
17	PRSNT#2	Hot plug detect	HSIn(0)	
18	GND	Ground	GND	Ground
PCI Express x1 end				

2.2 OCP 2.0

The OCP 2.0 specification extends the card mechanical and electrical interface to enable new uses cases for users in the OCP community. This supports the demands for different I/O types, increases bandwidth of data (x16 PCIe lane to baseboard) and management, and supports higher power controller IC. The specification makes changes on as-needed base to maximize backward compatibility to existing OCP platforms.

OCP 2.0 defines two management interfaces for the BMC's out-of-band (OOB) communication.

- I^2C/SMBus sideband interface for compatibility with baseboard.
- RMII-based NC-SI.

The thermal reporting interface is defined to be accessed from the SMBus interface. For Mezzanine cards with thermal design power (TDP) larger than 5 Watts, this interface implementation is required. Whether emulated method or remote on-die sensing method, each method is treated by the baseboard management controller as a TI/TMP421 thermal sensor.

2.3 OCP 3.0

The OCP NIC 3.0 specification extends the design of the OCP Mezzanine 2.0 specification to provide a broader solution space for the NIC and system vendors. There are three types of management requirements for OCP NIC 3.0 implementations, which are RBT+MCTP Type, RBT Type, and MCTP Type. Each type depends on the physical sideband management interfaces, transports, and traffic supported over different transports.

OCP 3.0 management functions provided on the primary bay are listed in Table 4:

Table 4. OCP 3.0 Management Functions on the Primary Bay

Function	Comment
Sideband Management Interface and Transport	• RBT+MCTP • MCTP • RBT
Power Management and Status Reporting	• Power brake for emergency power reduction. • State change control.
Control/Status Serial Bus	• NIC-to-Host status. • Host-to-NIC configuration Information.
Multi-Host PCIe Support Signals	• 2x PCIe resets. • 2x Reference clocks.
PCIe Wake Signal	• WAKE_N
Scan Chain Thermal Signals	• TEMP_WARN_N • TEMP_CRIT_N • FAN_ON_AUX

For the system management, the ASIC die temperature reporting is required with TDP larger than (or equal to) 5 Watts. The pluggable transceiver module temperature is independent of the ASIC die temperature reporting requirements. It is important for proper system operation to know the presence and temperature of the plug modules.

3.0 Temperature Reporting Interfaces

This section provides a summary of the manageability interface support matrix on Intel® Ethernet Controllers. In a system environment, the temperature of a component is a function of both the system and component thermal characteristics. System-level thermal constraints consist of the local ambient temperature at the component. With the supported interface, the MC can monitor the ASIC or plug module temperature to proceed the thermal control.

3.1 Sideband Support Matrix

Intel® Ethernet Controllers with a manageability sideband interface support the SMBus and the DMTF defined NC-SI, MCTP, and PLDM specifications. Table 5 summarizes the protocol and interface support. It is needed to decide the manageability interface (for example, SMBus and NC-SI), as they cannot be used at the same time with earlier products. Refer to the respective product collateral and NVM image release notes for the feature support.

Table 5. Protocol and Interface Support

Controller	Legacy SMBus	MCTP over SMBus	MCTP over PCIe	PLDM	NC-SI
Intel® Ethernet Controller E810		X	X	X	X
Intel® Ethernet Controller X710-TM4/AT2	X	X	X	X	X
Intel® Ethernet Controller X710/XXV710/XL710	X	X	X	X	X
Intel® Ethernet Controller X550	X	X	X		X
Intel® Ethernet Controller X540	X	X			X
Intel® Ethernet Controller I350	X	X			X
Intel® Ethernet Controller I210	X[1]	X[1]	X[1]		X[1]
Intel® 82599 10 GbE Controller	X				X

1. Not supported in Flash-less I210 operation.

3.2 ASIC Temperature Reporting Interface

The thermal management objective is to ensure that all system component temperatures are maintained within functional limits. The functional temperature limit is the range in which the electrical circuits are expected to meet specified performance requirements. Operation outside the functional limit can degrade system performance, cause logic errors, or cause device and/or system damage. The MC can monitor the ASIC temperature and maintain the case/junction temperature at or below the specified limits through the supported interface.

Table 6. Interfaces Supported to Report ASIC Temperature

Controller on which NIC is Based	PCIe[1]	OCP 2.0	OCP 3.0	LOM
Intel® Ethernet Controller E810	PLDM NC-SI	SMBus[2] PLDM NC-SI	PLDM NC-SI Scan Chain	PLDM NC-SI
Intel® Ethernet Controller X710-TM4/AT2	PLDM NC-SI AQ Command	N/A	PLDM NC-SI AQ Command Scan Chain	PLDM NC-SI AQ Command
Intel® Ethernet Controller X710/XXV710/XL710	N/A	SMBus[2]	PLDM[2] Scan Chain[2]	N/A
Intel® Ethernet Controller X550	NC-SI PCIe Register	N/A	N/A	NC-SI PCIe Register
Intel® Ethernet Controller I350	NC-SI PCIe Register	N/A	NC-SI PCIe Register	NC-SI SMBus PCIe Register

1. Refer to the respective PCIe adapter collateral for its feature support.
2. It monitors a thermal sensor on the adapter, instead of on-die temperature.

3.3 Plug Module Temperature Reporting Interface

The temperature of the transceiver module plays an important role and affects the function of communication system. Each plug module has its operating temperature range. The MC can monitor the plug module temperature and maintain the case/junction temperature at or below the specified limits through the supported interface

Table 7. Interfaces Supported to Report Plug Module Temperature

Controller on which NIC is Based	PCIe[1]	OCP 2.0	OCP 3.0	LOM[2]
Intel® Ethernet Controller E810	PLDM NC-SI	PLDM NC-SI	PLDM NC-SI Scan Chain	PLDM NC-SI
Intel® Ethernet Controller X710-TM4/AT2	PLDM NC-SI AQ Command	N/A	N/A[3]	PLDM NC-SI AQ Command
Intel® Ethernet Controller X710/XXV710/XL710	PLDM NC-SI	N/A	PLDM NC-SI Scan Chain	PLDM NC-SI
Intel® Ethernet Controller X550	N/A	N/A	N/A	N/A
Intel® Ethernet Controller I350	N/A	N/A	N/A[3]	N/A

1. Refer to the respective PCIe adapter collateral for its feature support.
2. Hardware schematic review is required to build up the thermal sensor profile for a LOM design.
3. The adapter is BASE-T solution and not using pluggable transceiver module.

Appendix A NC-SI OEM Commands

A.1 OEM Commands Summary

Table 8 lists Intel OEM NC-SI commands defined to provide an interface for reading the X710/E810 temperatures measured by the available thermal sensors.

Table 8. OEM Commands Summary

Intel Command	Parameter	Command Name	Supported in MCTP without PT	Reference
0x4B	N/A	Get ASIC Temperature	Yes	Section A.2
	0x2	Get SFF Module Temperature	Yes	Section A.3

A.2 Get ASIC Temperature Command (Intel Command 0x4B)

The Get ASIC Temperature command allows the Management controller to query for temperature values from the Controller's on-chip thermal sensor(s), or alternately from attached devices.

Currently this command is only supported as a channel command. Meaning that the command must be addressed to a specific channel in the package (channel 0x1F) and not to the package.

	Bits			
Bytes	**31:24**	**23:16**	**15:8**	**7:0**
0...3	NC-SI Header (0x50)			
4...7				
8...11				
12...15				
16...19	Manufacturer ID (Intel 0x157)			
20...23	0x4B	Padding		
24...27	Checksum (3..2)		Checksum (1..0)	
28...31	Padding			
32...35				
36...39				
40...43				
44...45	Padding			

A.2.1 Get ASIC Temperature Response

The package, in the absence of a checksum error or identifier mismatch, always accepts the Get ASIC Temperature Command and sends a response.

	Bits			
Bytes	**31:24**	**23:16**	**15:8**	**7:0**
0...3	NC-SI Header (0xD0)			
4...7				
8...11				
12...15				
16...19	Response Code		Reason Code	
20...23	Manufacturer ID (Intel 0x157)			
24...27	0x4B	Reserved	Maximum Temperature	Current Temperature
28...31	Checksum (3..2)		Checksum (1..0)	
32...35	Padding			
36...39				
40...43				
44...45	Padding			

Where:

- ***Maximum Temperature Value*** — This value is the maximum T-Diode temperature limit in °C at which controller can operate at full load for its rated service lifetime. The value should be de-rated to take measurement tolerance into account. The value is reported as a hexadecimal integer number.
- ***Current Temperature Value*** — This value is the current real-time temperature of the chip in °C. The value is reported as a hexadecimal integer number.

A.3 Get SFF Module Temperature Command (Intel command 0x4B, Parameter 0x02)

The Get SFF Module Temperature Sensor command allows the Management controller to query for the real-time temperature value and thresholds of the optical transceiver attached to the channel.

	Bits			
Bytes	**31:24**	**23:16**	**15:8**	**7:0**
0...3	NC-SI Header (0x50)			
4...7				
8...11				
12...15				
16...19	Manufacturer ID (Intel 0x157)			
20...23	0x4B	0x2	Padding	
24...27	Checksum (3..2)		Checksum (1..0)	
28...31	Padding			
32...35				
36...39				
40...43				
44...45	Padding			

A.3.1 Get SFF Module Temperature Response

The Get SFF Module Temperature response frame contains the current temperature of the attached module and the high side temperature thresholds.

	Bits			
Bytes	**31:24**	**23:16**	**15:8**	**7:0**
0...3	NC-SI Header (0xD0)			
4...7				
8...11				
12...15				
16...19	Response Code		Reason Code	
20...23	Manufacturer ID (Intel 0x157)			
24...27	0x4B	0x2	Reserved	
28...31	Temp High Alarm Threshold		Temp High Warning Threshold	
32...35	Temperature Value		Padding	

Definitions and interpretation of the data fields in the response are defined in the relevant SFF or MSA specification (for example, SFF-8472, SFF-8436, SFF-8636, and so on). 16-bit values are encoded as one contiguous entity with the most significant bit in bit 15 (or 31) and least significant bit in bit 0 (or 16) in the response packet.

Devices supporting SFF-8472 diagnostics are identified by a non-zero value at address 94 on I^2C address 0xA0. For the devices supporting SFF-8636/8436, it is assumed that on customer transceivers, the temperature data is always available.

Temperature/Threshold Fields

For SFP+ modules - SFF-8472 specification references:

- **Temp High Alarm Threshold** — The contents of address 0xA2 offset 0-1.
- **Temp High Warning Threshold** — The contents of address 0xA2 offset 4-5.
- **Temperature Value** — The contents of address 0xA2 offset 96-97.

For QSFP+ modules - SFF-8636/8436 specification references:

- **Temp High Alarm Threshold** — The contents of upper page 0x03 bytes 128-129.
- **Temp High Warning Threshold** — The contents of upper page 0x03 bytes 132-133.
- **Temperature Value** — The contents of lower page 0x00 bytes 22-23.

In case QSFP transceiver does not support optional page 0x03 (“Flat Memory”), Threshold values should be set to 0xFFFF, which indicates an “unknown” threshold. Flat Memory transceivers can be identified by Bit 2 (*Flat_mem*) under the *Status* field at Byte 2 of page 0x00.

According to SFF-8636, the Warning Threshold is optional, but since there is no indication about it, it is reported as is and it is up to the BMC whether to use this information and how to interpret it.

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

LEGAL