

10/40 Gigabit Ethernet Rack Mountable Switch Standard

Author: Michael Miller

1 Scope

This standard provides board-specific information detailing the features and functionality of a general-purpose 10/40 gigabit Ethernet rack mountable switch supporting 48 x SFP+ and 4 x QSFP+ ports for adoption by the Open Compute Project community.

THE FOLLOWING IS A LIST OF MERELY REFERENCED TECHNOLOGY: Microprocessor technology, semiconductor manufacturing technology, operating system technology (including without limitation networking operating system technology), emulation technology, graphics technology, video technology, integrated circuit packaging technology and the like, compiler technologies, object oriented technology, optical/RF communications technology including chip I/O and driver technology, bus technology, memory chip technology (including, without limitation, NAND memory, NOR memory, resistive RAM (RRAM), seek scan probe (SSP) memory, nonvolatile memory (including without limitation, memory based on chalcogenide materials, phase change memory (PCM), one or more stacked layers of memory cells, embedded PCM memories, non-volatile cache memory, solid state drives, SRAM, embedded DRAM, ferro-electric memory, and polymer memory)) and/or health-related and medical technology. IMPLEMENTATION OF THESE TECHNOLOGIES MAY BE SUBJECT TO THEIR OWN LEGAL TERMS.

2 Table of Contents

Contents

1	Scope	2
2	Table of Contents	2
3	Overview	5
4	License	5
5	Compliance to the Standard	5
6	Feature Requirements	6
7	Mechanical	8
8	Block Diagram	10
9	IO and Connectors	11
ç	9.1 SFP+	11
g	9.2 QSFP+	11
g	9.3 Ethernet Management and Serial Console	11
	9.3.1 RJ-45 Ethernet Connector	11
	9.3.2 RS232 Serial Console Connector	11
g	9.4 USB	12
ç	9.5 Switch Board Power Input Connectors	12

9	0.6 AC Inlet and Filter	13
9	0.7 Fan Connector	13
9	0.8 JTAG Connector	13
9	9.9 I2C / SPI Debug Connector	14
9	0.10 Glue Logic FPGA Flash Programming Header	14
7	7.11 Glue Logic CPLD Programming Header	14
10	IO and Connectors	15
1	0.1 Management	15
1	0.2 SFP+	15
1	0.3 QSFP+	15
1	0.4 LEDs	15
11	Power System, Fans, and Monitors	17
1	1.1 PSUs	17
1	1.2 Power Connections to the OCP Open Rack	17
1	1.3 Fan Control and Monitors	17
12	Management	19
1	2.1 Reset	19
1	2.2 Switch VR	20
1	2.3 I2C / SMBus	21
1	2.4 Console	22
1		22
-	2.5 Ethernet Management Port	22
1	2.5 Ethernet Management Port2.6 USB	22 22 22
1	2.5 Ethernet Management Port2.6 USB2.7 Glue Logic CPLD and FPGA	22 22 22 22
1 1 1 1	 2.5 Ethernet Management Port 2.6 USB 2.7 Glue Logic CPLD and FPGA 2.8 Interrupts and Alerts 	22 22 22 22
1 1 1 13	 2.5 Ethernet Management Port	22 22 22 22 22
1 1 1 13 1	 2.5 Ethernet Management Port	22 22 22 22 22 23 23
1 1 13 1 1	 2.5 Ethernet Management Port	22 22 22 22 22 23 23 23
1 1 13 1 1 1 1	 2.5 Ethernet Management Port	22 22 22 22 23 23 23 23
1 1 13 1 1 1 1 1	 2.5 Ethernet Management Port	22 22 22 22 23 23 23 23 23
1 1 13 1 1 1 1 1 1 1 1	 2.5 Ethernet Management Port 2.6 USB 2.7 Glue Logic CPLD and FPGA 2.8 Interrupts and Alerts Switch Subsystem 3.1 Switch Port Distribution 3.2 PCIe 3.3 I2C / SMBus 3.4 SPI 3.5 GPIO 	22 22 22 22 23 23 23 23 23 23
1 1 1 13 1 1 1 1 1 1 1 1 1	 2.5 Ethernet Management Port 2.6 USB 2.7 Glue Logic CPLD and FPGA 2.8 Interrupts and Alerts Switch Subsystem 3.1 Switch Port Distribution 3.2 PCIe 3.3 12C / SMBus 3.4 SPI 3.5 GPIO 3.6 Switch Device LED Output Stream 	22 22 22 22 23 23 23 23 23 23 23

OCP Specification

14.3	1 Connector Labeling	. 25
14.2	2 Documentation	. 25
14.3	3 Disallowed Components	. 25
14.4	4 Capacitors and Inductors	. 25
15	Testability	.26
15.3	1 JTAG / Boundary Scan	.26
15.2	2 SMBUS / I2C / PMBus / MDIO Debug	.26
16	Appendix – Engineering Workshop Features	. 27

3 Overview

The purpose of this document is to define 10/40 gigabit Ethernet switch that is capable of deployment in scale-out data centers as well as in traditional data centers with 19" rack enclosures. Considerations are made in the specification for suitable 10/40 gigabit Ethernet switch products that were in production when the specification was released. This document is not intended to be used solely as a basis for a procurement of OCP compatible products. The OCP community may have additional requirements. These incremental requirements can be captured in additional procurement documentation. Products that claim to comply with this specification SHALL provide, at a minimum, all features defined as mandatory by the use of the keyword SHALL. Such products may also provide recommended features associated with the keyword SHOULD, and permitted features associated with the keyword MAY.

4 License

As of November 8, 2013, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0:

Intel Corporation

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at http://opencompute.org/licensing/, which may also include additional parties to those listed above.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, noninfringement, fitness for a particular purpose, or title, related to the Specification. The entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

5 Compliance to the Standard

Products making the claim of compliance with this specification **SHALL** provide, at a minimum, all features defined as mandatory by the use of the keyword "**SHALL**". Such products may also provide recommended features associated with the keyword "**SHOULD**" and permitted features associated with the keyword "**MAY**".

6 Feature Requirements

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard **SHALL** have the following required features present for the product to comply with this standard.

Feature	Description		
Mechanical	Fits into 19" EIA rack (1 RU height)		
Switch Device	 Integrated wire-speed 10/40 Gbps Ethernet switch silicon device Minimum 64 Ethernet lanes supporting 10GbE serial Maximum port to port latency of 1us Maximum variation between port pairs of 50% Compliance with the following specifications: SFI 5m DA copper (10GBase-CR) 40GBase-CR4 40GBase-SR4 		
Processor / Chipset	Minimum of 2 processor cores		
	Minimum of 2 memory channels		
	Minimum of one 1-lane PCI Express Gen 2 root complex		
	External ports:		
	 Minimum of one UART port 		
	 Minimum of two I2C/SMBus ports 		
	 Minimum of one 1GbE port 		
	 Minimum of one USB full-speed port 		
Memory • Minimum 4GB DDR3 DRAM			
	Support ECC		
Storage	Minimum 8GB NAND Flash		
	Minimum 4MB NOR Flash		
External	Minimum of one 1GbE RJ45 port		
Management	Minimum of one RS232 management port		
	Minimum of one USB (mini, micro, or standard) full-speed		
	port		
External High Speed	48 x SFP+ and 4 x QSFP+		
Ports	 For SFP+, compliance with SFF-8431 ver 4.1, SFF-8472 		
	ver 10.4		
	 For QSFP+, compliance with SFF-8436 ver 4.1 		
LEDs	Minimum one LED per SFP+ port and one per QSFP+ port		
	Minimum one LED for the 1GbE management port		
Power Supplies	Two pluggable PSUs		
	Load-sharing and redundant (under sustained loads)		
Fans	Front to back		
	Managed fans		

TABLE 1: Required Features

The following table includes additional desirable features which MAY be included.

Feature	Description		
Mechanical	 May fit into an Open Rack (1 OU height) 		
Switch Device	Additional Ethernet connections between switch device and OPL subsystem: 100h5, XAUL 250h5, SOMU		
	CPU subsystem: TUGDE, XAUI, 2.5GDE, SGIVIII		
	Compliance with the following specifications		
	o SGMII		
	○ 1000Base-X		
	o XAUI		
	 Support for 2.5GbE (not a standard) 		
Processor / Chipset	 May reside on a separate PCB 		
	Additional Ethernet connections between switch device and		
	CPU subsystem: 10GbE, XAUI, 2.5GbE, SGMII		
	 Additional PCIe root complex ports 		
	 Additional lanes per PCIe port (x4, x8) 		
	SATA port		
Storage	Include a uSATA drive		
	 Include redundant NOR flash 		
Environment	 Support Open Network Install Environment (ONIE) 		
	Support network boot		
LEDs	 Second LED per SFP+ port 		
	Second LED for 1GbE management port		
Power	Support for 12V DC bus bar input (with power distribution		
	board)		

7 Mechanical

The three figures below show examples of a mechanical form factor.

Figure 1: Example Front Pane	
	 <u>74 76 76 76 76 76 76 76 76 76 76 76 76 76 </u>
USB ETH	

Figure 2: Example Rear Panel with Dual Power Supplies

|--|--|

Figure 3: Example 3-D Top Level View



A product that meets the 10/40 Gigabit Ethernet Rack Mountable Switch Standard that is designed to fit into a 19" EIA standard rack **SHOULD** comply with the following layout options.



Figure 4: Width / Depth Measurements and Example Placement

8 Block Diagram





9 IO and Connectors

9.1 SFP+

There **SHALL** be 48 ports of SFP+ Connectors on the switch board. They **MAY** be grouped as three 2x8 gangs on the front panel (eg. Molex 76352-7001). These are SFP and SFP+ compliant interfaces that each require one of the following: SFP or SFP+ optical module or active optical cable, or SFP or SFP+ passive or active direct attach copper cable.

9.2 QSFP+

There **SHALL** be 4 QSFP+ Connectors on the switch board. They **MAY** be grouped as two 1x2 gangs on the front panel (eg. Molex 76871-0006). These are QSFP and QSFP+ compliant interfaces that each require one of the following: QSFP or QSFP+ optical module or active optical cable, or QSFP or QSFP+ passive or active direct attach copper cable.

9.3 Ethernet Management and Serial Console

There **SHALL** be at least one Ethernet management port and one console port, where the Ethernet port is exposed as an RJ45. A dual RJ45 stacked connector **MAY** be used for the RJ45 management interface and the console port (eg. XMultiple XRJD-2-21-B-BX).

9.3.1 RJ-45 Ethernet Connector

The 10/100/1G Ethernet ports on the switch board **SHALL** be compliant with IEEE 802.3. The interface **SHALL** be auto MDI-X, auto-negotiation, and auto-polarity inversion.

TABLE 3: RJ45 Ethemet Port Pinout		
PIN	DIR	NAME
1	AUTO	TX D1+
2	AUTO	TX D1-
3	AUTO	RX D2+
4	AUTO	BI D3+
5	AUTO	BI D3-
6	AUTO	RX D2-
7	AUTO	BI D4+
8	AUTO	BI D4-

TABLE 3: RJ45 Ethernet Port Pinout

9.3.2 RS232 Serial Console Connector

The RS232 SHOULD be compliant with EIA/TIA 561.

TADLE A. Example DIAE Canada Dart Direct

TABLE 4. Example RJ45 Console Fort Fillout		
PIN	DIR	NAME
1	IN	DSR – data set ready
2	IN	DCD – data carrier detect
3	OUT	DTR – data terminal ready

OCP Specification

4	GND	Ground
5	IN	RD – rx data
6	OUT	TD – tx data
7	IN	CTS – clear to send
8	OUT	RTS – request to send

9.4 USB

There **SHALL** be a standard, mini, or micro USB Connector on the switch board (eg. Molex 47346-0001). This USB **SHALL** be compatible with the USB specification, and at a minimum support full-speed USB.

9.5 Switch Board Power Input Connectors

The product **SHALL** support two redundant, load-sharing power supplies. There are two power input connectors (eg. Molex 45984-0007, which can be used with power supplies such as the Emerson DS460S-3 load-sharing power supply).

All onboard switch board power **SHALL** be derived from +12V.

	TABLE 5. Example FSU Com
PIN	NAME
PB1	+12V
PB2	+12V
PB3	+12V
PB4	+12V Return
PB5	+12V Return
PB6	+12V Return
PB7	+12V Return
PB8	+12V Return
PB9	+12V
PB10	+12V
S1	Standby (+12V)
S2	Standby (+12V)
S3	Reserved
S4	Interrupt#
S5	Present#
S6	PSOK
S7	I-Mon
S8	PSON#
S9	SCL
S10	SDA
S11	GND
S12	I2C A0
S13	I2C A1
S14	I2C A2
S15	GND

TABLE 5: Example PSU Connector Pinout

9.6 AC Inlet and Filter

When the product is designed for implementation into a 19" EIA rack, the product **SHALL** support a standard AC inlet and filter built into PSU. This **SHALL** accept a minimum range of 100-250VAC up to 30A inrush.

9.7 Fan Connector

The product **SHALL** support managed fans. The fans **MAY** be powered through a +12V source and support PWM and tachometer, which is connected to the on-board fan controller.

9.8 JTAG Connector

A JTAG header **MAY** be provided (eg. Molex 87832-2020) which links all devices available for boundary scan. An example pinout is shown in the table below.

PIN	DIR	NAME
1	IN	TRST#
2	GND	Ground
3	OUT	TDO
4	GND	Ground
5	IN	TDI
6	GND	Ground
7	IN	TMS
8	GND	Ground
9	IN	ТСК
10	GND	Ground
11	Power	VPP
12	GND	Ground
13	Power	AWR
14	GND	Ground
15	IN/OUT	USER0
16	GND	Ground
17	IN/OUT	RDY/BUSY#
18	GND	Ground
19	IN/OUT	USER1
20	NC	NC

TABLE 6: Example JTAG Header Pinout

9.9 I2C / SPI Debug Connector

A debug and programming header **MAY** be provided which is compatible with a suitable tool (eg. Totalphase Aardvark debug tool - Samtec TSM-105-01-T-DV). The header can be used to monitor the I2C Bus and to program local Flash and EEPROM devices.

PIN	NAME
1	SCL
2	GND
3	SDA
4	NC
5	MISO
6	NC
7	SCLK
8	MOSI
9	SS
10	GND

TABLE 7: Example I2C/SPI Debug Header Pinout

9.10 Glue Logic FPGA Flash Programming Header

If a glue logic FPGA is used, a programming header **SHOULD** be provided which is compatible with FPGA programming tools (eg. Altera USB Blaster programming tool - Samtec TSM-105-01-T-DV). The header can be used to program the Flash or EEPROM connected to the FPGA.

PIN	NAME
1	CLK
2	GND
3	DONE
4	+3.3V
5	CONFIG_N
6	CE_N
7	DATA
8	CSO
9	ASDO
10	GND

TABLE 8: Example FPGA Programming Header Pinout

7.11 Glue Logic CPLD Programming Header

If a glue logic CPLD is used, a programming header **SHOULD** be provided which is compatible with CPLD programming tools (eg. Altera USB Blaster programming tool - Samtec TSM-105-01-T-DV). The header can be used to program the CPLD directly.

PIN	NAME
1	TCK
2	GND
3	TDO
4	+3.3V
5	TMS
6	NC
7	NC
8	NC
9	TDI
10	GND

TABLE 9: Example CPLD Programming Header Pinout

10 IO and Connectors

10.1 Management

The SFP+ and QSFP+ interfaces **SHALL** be managed via an I2C bus interface from the CPU subsystem. The I2C interface **MAY** be done indirectly via glue logic. The I2C interface **SHOULD** use I2C fanout muxes and expanders (refer to Management chapter below).

10.2 SFP+

The SFP+ ports **SHALL NOT** use a PHY; they **SHALL** be directly connected to the switch device. The product **SHALL** have 48 ports of SFP+, these **MAY** be ganged using connector/cage assemblies (eg. Molex 76352-5001 2x8 ganged SFP+ connector/cage assemblies).

10.3 QSFP+

The QSFP+ ports **SHALL NOT** use a PHY; they **SHALL** be directly connected to the switch device. The product **SHALL** have 4 ports of QSFP+, these **MAY** be ganged using connector/cage assemblies (eg. Molex 76871-0008 2x1 ganged QSFP+ connector/cage assemblies).

10.4 LEDs

Each SFP+ port **SHALL** have at least one LED. Each port **SHOULD** have two LEDs. When there is one LED then it **SHOULD** light solid for Link Up and blink for Traffic (RX or TX). When there are two LEDs the left LED **SHOULD** be used for Link Up (solid) and **MAY** also be used for Port Training (blinking slowly). The right LED **SHOULD** be used for Traffic (RX or TX) (blinking).





Each QSFP+ port **SHALL** have at least one LED. Each QSFP+ is intended to be interpreted as a 40G lane. The LED **SHOULD** light solid for link up, and blink for traffic.





The LEDs **SHALL** be controlled from one of two alternate schemes.

- 1. LEDs driven based on port status which is output from the switch device with software masking.
- 2. LEDs driven fully under software control.

For scheme #1, the switch device outputs port status information, for example an encoded serial bit stream, which is then interpreted by a CPLD or FPGA. The software **SHOULD** provide basic masking control for each port, independent for the SFP+ Link and Traffic LEDs, common for the QSFP+ LED.

For scheme #2, software directly controls the behavior of each LED. The software **SHOULD** be capable of turning each LED solid on/off or blinking. This may be accomplished using glue logic.

11 Power System, Fans, and Monitors

11.1 PSUs

When installed in a standard 19" EIA rack, the 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** provide for two power supplies with the output(s) connected directly to the switch board. The power supplies **SHALL** meet the following minimum criteria:

TABLE 10: PSU Requirements				
INPUT	NAME			
Input Range	100-250 VAC			
Frequency	50-60 Hz			
Inrush Current	30 A max			
Efficiency	80% at 50% load			
Leakage Current	1 mA @ 240 VAC			
OUTPUT	DESCRIPTION			
DC Output Voltage	+12 V @ 36 A			
Overcurrent	Trip 120-200% rated current for 1 second			
Turn-on Delay	0 – 2 seconds			
Load-sharing				
Redundant	(at sustained load)			
CONTROL	DESCRIPTION			
PMBus	PMBus control interface support			
Fault Detection	Fault detection sideband interrupts			

11.2 Power Connections to the OCP Open Rack

When used in the OCP Open Rack, the product enclosure **SHOULD** contain a power distribution board (PDB) that provides an electrical interconnector between the 12VDC bus bars in the OCP rack and the 10/40 Gigabit Ethernet Rack Mountable Switch Standard product switch board PSU inputs. A cable harness may also be used to provide an interconnect

The Open Rack website provides a comprehensive list of open rack specifications and design guidelines.

11.3 Fan Control and Monitors

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** use managed fans, and **SHALL** provide temperature monitoring for the CPU, switch device, and a minimum of two local-ambient temperature points on the PCB.

As an example, a fan control / monitor circuit using the Maxim MAX31785ETL controller is shown below. This is designed for management of up to six PWM fans and monitoring of up to three temperature points. An example fan which may be used is the SanyoDenki SanAce 9CRA0412P4K03, capable of 32CFM.



Figure 8: Example Fan Controller Configuration

12 Management

12.1 Reset

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** support at least three independent reset domains:

- 1. Master reset
- 2. Switch reset
- 3. I2C reset

Master reset is the primary reset for the product; it resets all devices and causes assertion of all other reset domains. It **SHOULD** be driven by control logic (eg. Maxim 16062E supervisor) at power on or under CPU control. For example, the supervisor can monitor the primary board voltage rails, if they are not within their expected limit then master reset can be asserted, this happens automatically at power on. Additionally there is an input which can be driven under software control which can cause master reset.

Switch reset **SHALL** be driven under software control. It is used to reset the switch subsystem.

I2C reset **SHALL** be driven under software control. At a minimum it **SHOULD** be used to reset the I2C fanout mux and expander tree to the SFP+ and QSFP+ ports. It **SHOULD** also be used to reset any other reset-capable I2C devices.

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **MAY** contain additional reset domains, for example:

- 1. USB reset
- 2. Fan controller reset
- 3. Glue logic reset

If included, each of these **SHALL** be driven under software control.

The below figure shows an example of a supervisor circuit used for controlling master reset.





- Monitor primary power rails
- Manage MASTER_RESET_N as primary reset for the switch
- Allow software to perform a master reset

12.2 Switch VR

The voltage regulators used for the switch device **SHOULD** have the following requirements:

- 1. Power-down input
- 2. SMBus management interface
- 3. Adjustable voltage output level

20

- 4. Voltage and Current status information
- 5. Fault status information

12.3 I2C / SMBus

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** provide CPU management capability over I2C or SMBus; glue logic is acceptable to facilitate management. PMBus **SHOULD** also be available.

I2C or SMBus **SHALL** be used for front panel SFP+ and QSFP+ management as well as management of the fan controller, PSUs (PMBus), and the VRs for the switch device. The bus **SHOULD** be fanned out into multiple branches; see the figure below for an example.

The first stage of fanout is a 1x4 mux (eg. NXP PCA9545B) which branches to a set of seven 40-bit parallel bus expanders (eg. NXP PCA9505DGG) (six for SFP+ and one for QSFP+) which allow access to the parallel signals of each module. A second stage is a set of 6 1x8 muxes (eg. NXP PCA9548A) and one 1x4 mux (eg. NXP PCA9545B) which fan out the I2C bus to the I2C access points on each SFP+ and QSFP+ module.





12.4 Console

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** provide a RS232 connection from the CPU subsystem to a front panel header for console management.

12.5 Ethernet Management Port

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** provide a 1GbE Ethernet host management port via an RJ45 on the front panel. This interface **MAY** utilize a PHY (eg. Marvell 88E1111 Ethernet BASET PHY) configured to operate in SGMII mode

12.6 USB

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** provide a micro, mini, or standard USB port compliant with USB full-speed operation.

12.7 Glue Logic CPLD and FPGA

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **MAY** include CPLD and/or FPGA glue logic. If included there **SHOULD** be headers provided to program the CPLD and FPGA flash/eeprom. There **SHOULD** be a software path to reprogram the device dynamically.

This glue logic may include:

- 1. SMBus / I2C glue logic
- 2. LED management
- 3. Reset assist control
- 4. Host Ethernet MDIO management
- 5. Interrupt aggregation

12.8 Interrupts and Alerts

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** manage the following interrupts and alerts:

- 1. Overtemp alert (CPU, switch, or local ambient)
- 2. VR altert and interrupt
- 3. SFP+ / QSFP+ port interrupts
- 4. Switch interrupt
- 5. PSU alert and interrupt
- 6. Host Ethernet interrupt

Overtemp **SHALL NOT** be maskable and will cause action within 200ms of reporting. This action **SHOULD** include reset and shutdown of select power supplies. Immediate actions **MAY** be taken by glue logic and reported to the CPU after the fact. Other events **MAY** be maskable.

These events **MAY** be aggregated via glue logic and delivered to the CPU via sideband signals or PCIe MSI messages.

13 Switch Subsystem

13.1 Switch Port Distribution

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product switch **SHALL** support a minimum of 64 lanes connected between the switch device and the front panel and organized as 48 x SFP+ and 4 x QSFP+.

Additional lanes **MAY** be connected between the switch device and CPU subsystem. These **MAY** support 10GbE, XAUI, 2.5GbE, and/or SGMII.

The lanes going to the front panel **MAY** be shared with the lanes going to the CPU subsystem. For example either a QSFP+ OR the XAUI may be enabled at any given time, the other would be disabled.

The design **SHOULD** be constructed to reduce the number of PCB layers required to route the interfaces by avoiding cross-overs among ports. The design **SHALL** meet applicable signaling specs.

13.2 PCIe

PCIe **SHALL** be the primary interface from the CPU Subsystem to the Switch device used for management. It **SHALL** be at minimum x1 link directly routed between the CPU and the Switch device. This PCIe bus **SHALL** be fully compliant to the PCI Express Base Specification Rev 2.0.

The PCIe connection MAY be x4 or x8 and may be Gen 3 compliant.

13.3 I2C / SMBus

I2C or SMBus **SHOULD** be available as a secondary interface from the CPU subsystem to the Switch device. The CPU **SHOULD** have full management access to the switch via I2C.

13.4 SPI

SPI **MAY** be available as a boot option for the switch device, for example a SPI flash (eg. Atmel AT45DB321) can be used to boot and configure the switch.

If this is provided, it **SHOULD** be programmable via a board header.

13.5 GPIO

Switch GPIO MAY be connected to glue logic or the CPU for sideband interfacing.

13.6 Switch Device LED Output Stream

The switch device **MAY** include LED outputs to indicate port status. These outputs **MAY** be used to drive the front panel LEDs via glue logic.

14 Design Considerations

The 10/40 Gigabit Ethernet Rack Mountable Switch Standard product is expected to conform to design standards, specifications, schematics, PCB stack-up, layout constraints, and thermal and mechanical constraints provided by any of the device manufacturers used in the design of the product.

14.1 Connector Labeling

All ports, connectors, and memory slots **SHALL** be clearly labeled.

14.2 Documentation

The supplier of the 10/40 Gigabit Ethernet Rack Mountable Switch Standard product **SHALL** make available documentation that contains the following information:

- Block diagram
- List of tools and commands that allow the flashing of firmware

14.3 Disallowed Components

The following components **SHALL NOT** be used in the design of the motherboard:

- Components disallowed by the European Union's Restriction of Hazardous Substances Directive (RoHS 6)
- Trimmers and/or potentiometers
- Dip switches

14.4 Capacitors and Inductors

The following limitations apply to the use of capacitors and inductors:

- All capacitors **SHALL** have a predicted life of at least 50,000 hours at 45°C inlet air temperature, under the worst conditions
- Tantalum capacitors using manganese dioxide cathodes SHALL NOT be used
- SMT ceramic capacitors with a case size > 1206 SHOULD NOT be used (size 1206 are still allowed when installed far from the PCB edge and with a correct orientation that minimizes the risk of cracking)
- Ceramic material for SMT capacitors **SHOULD** be X5R or better material (COG or NP0 type are used in critical portions of the design)
- Only SMT inductors SHALL be used.
- Through-hole inductors **SHALL NOT** be used.

15 Testability

15.1 JTAG / Boundary Scan

The product **MAY** be provided with a full JTAG chain connecting all devices that have boundary scan functionality.

15.2 SMBUS / I2C / PMBus / MDIO Debug

Headers **SHOULD** be provided for access to these various buses for debug purposes.

16 Appendix – Engineering Workshop Features

Workshop Request	10/40 Gigabit Ethernet Rack Mountable
	Switch Standard Product Support
FORM FACTOR, PORT COUNT, CABLING	
2 Form Factors: (a) 19" - 1RU or 2RU(b) OCP	Supports 1RU for both 19" and 21" form factor
Rack depth max appx 24-26"	Supports max depth of 15.6"
Aggregate Switching BW between 540G & 650G; 40-48 x 10GE Connectivity to Server, 4-6 x 40GE Connectivity for uplink	Supports 640G aggregate switching bandwidth to front panel Supports 48 x 10Gbe SFP+ ports and 4 x 40GbE
DAC cabling for Server Connectivity; SR optics for ToR Uplinks	Support for both DAC and optical cabling
CPU MANAGEMENT	
Recommend accommodate Modular CPU	Optionally supports a modular CPU option
Need to be able to boot from network (PXE), console access, flash the bios.	Supported
POWER, THERMAL	
Fixed, Single 12V DC or 200-277V AC	Supports 90-264 VAC Optionally supports 12V DC using adapter
Prefer Dual, HotSwappable, Redundant	Supports dual, load-sharing, redundant PSUs Optionally PSUs are hot-swappable
Power Allotment: 300-400W per RU	Supports 300-400W under sustained operating conditions
Prefer 80-90% loading on primary PS, secondary PS on "Hot" Standby	Supported
Anticipate power draw to be at 30-40% during idle, 91% peak loading	Supported
Airflow: - Front to back only	Supported

TABLE 11: Workgroup Feature Mapping