OPEN Compute Project

MSX1410-OCP

Rev. 1.3

SwitchX®-2 based 10GbE/40GbE, 1U Open Ethernet switch with 48 SFP+ ports and 12 QSFP+ ports

Author: Shy Zimmerman

Rev 1.3

Revision History

Revision	Date	Description of Change
1.3	April 2015	Added Table 1 Added Sections 3.1.3.1, 3.1.3.2
1.2	May 2014	Updated port speed to 10GbE and 40GbE in figures Added Abbreviations Minor editing of Figure 15
1.1	May 2014	Updated System Figures Added Long Switch Dimensions Updated Sx Layout Diagram Updated Main Block Diagram Added Split Configurations Updated I²C Tree Updated to 3 CPLD's (instead of 4) Updated CPLD Field Upgrade Updated Red LED to Amber LED Added Operating Temperature Added CCC Support Added Measured Units in Power Added CPU Basic Configuration
1.0	January 2014	Initial release

Scope

This document defines the technical specifications for the MSX1410-OCP switch used in Open Compute Project

Contents

Overview

The MSX1410-OCP switch is a top of rack switch with 48 ports of 10GbE and 12 uplink ports of 40GbE for non-blocking throughput between rack and aggregation layer. Based on advanced hardware design, this switch packs 48 SFP+ and 12 QSFP interfaces in an ultra-dense 1U form factor. The MSX1410-OCP features latency of 250ns and power efficiency, while providing optimal performance for open rack based enterprise data centers, financial services, Web 2.0, high performance computing and cloud computing applications.

License

Semiconductor Devices Licenses

All semiconductor devices that may be referred to in this specification, or required to manufacture products described in this specification, will be considered referenced only, and no intellectual property rights embodied in or covering such semiconductor devices shall be licensed as a result of this specification or such references.

Notwithstanding anything to the contrary in the CLA, the licenses set forth therein do not apply to the intellectual property rights included in or related to the semi without limitation the reference to devices listed below. For clarity, no patent claim that reads on such semiconductor devices will be considered a “Granted Claim” under the applicable Contributor License Agreement for this specification.

Manufacturer	Description
Mellanox	SwitchX-2, 36 Port 40GbE Switch IC MT51136A2-CLCR-B
Lattice	CPLD 4320LUT 3.3V LCMXO2-4000HC-4BG332CCA2
Intel	C QM77 EXPRESS CHIPSET BD82QM77-SLJ8A
Intel	IC CPU 1047UE 2M 1.40GHZ AV8063801116300-SR10E
Intel	IC GIGABIT ETHERNET CONTROLLER 82583 WG82583V-SLGVD
Windbond	MEMORY FLASH SPI SERIAL 32MBIT 4KBYTEX1024 W25Q32FVSSIGT
Windbond	MEMORY FLASH SPI SERIAL 64MBIT 4KBYTEX2048 W25Q64FVSSIGT
Windbond	SUPER IO LPC INTERFACE WITH UART NCT5577D
Intel	LHA5 "LEWISVILLE" GIGABIT ETHERNET LAN CONTROLLER HURON WG82579LM-SLHA6
INNODISK	MODULE MSATA SSD 16GB MLC 0..70 DEMSR-16GD07SC2DC-92
APACER	MODULE SO-DIMM DDR3 ECC 4GB 1600MBS 800MHZ SDRAM 204PIN 78.B2GCS.AT00C

1 Abbreviations

ToR – Top of Rack Switch

MNG – Management

SWB – Switch Board

FRU – Field Replaceable Unit

PWR – Power

HS – High Speed

SE – Single Ended

OCP – Open Compute Project

WD – Watch Dog

2 Mechanical Drawings

2.1 Mechanical System Overview

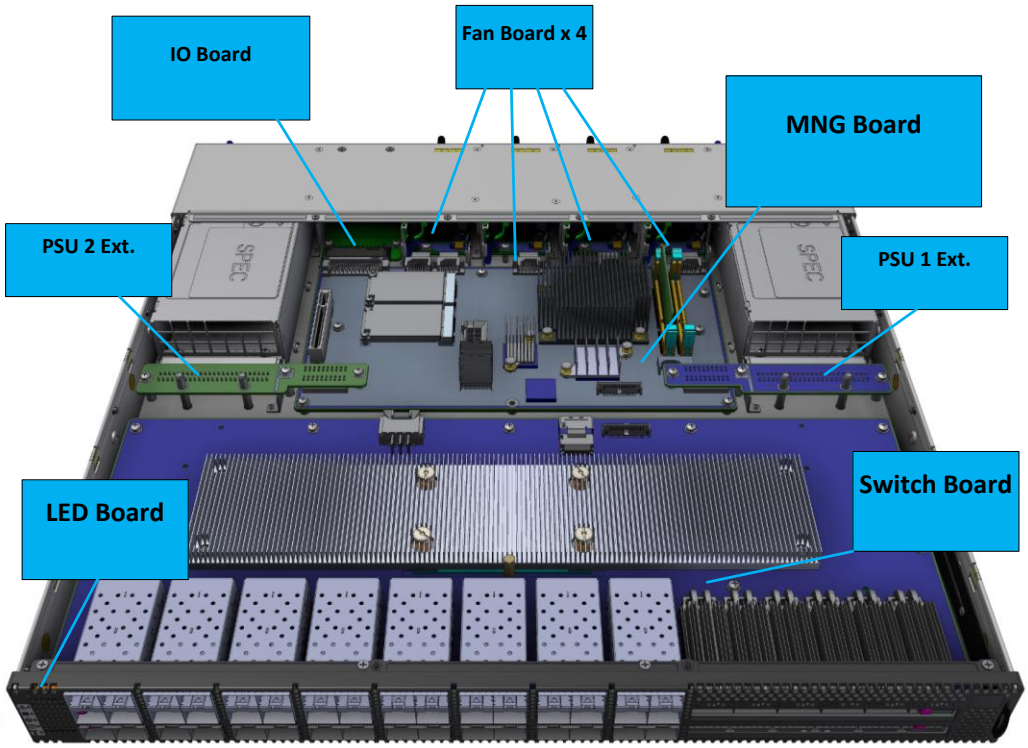

Figure 1 – Main Components

PCB Function	PCB Layer	Dimension (mm x mm)
Switch Board	18	413 x 184
MNG Board	16	250 x 134
IO Board	4	51 x 80.7
FAN Board	2	40.2 x 18.7
PSU Extender	4	127 x 25
LED Board	8	37.6 x 26.85

Table 1: PCB Layer Dimensions

2.2 Switch Dimensions (in mm)

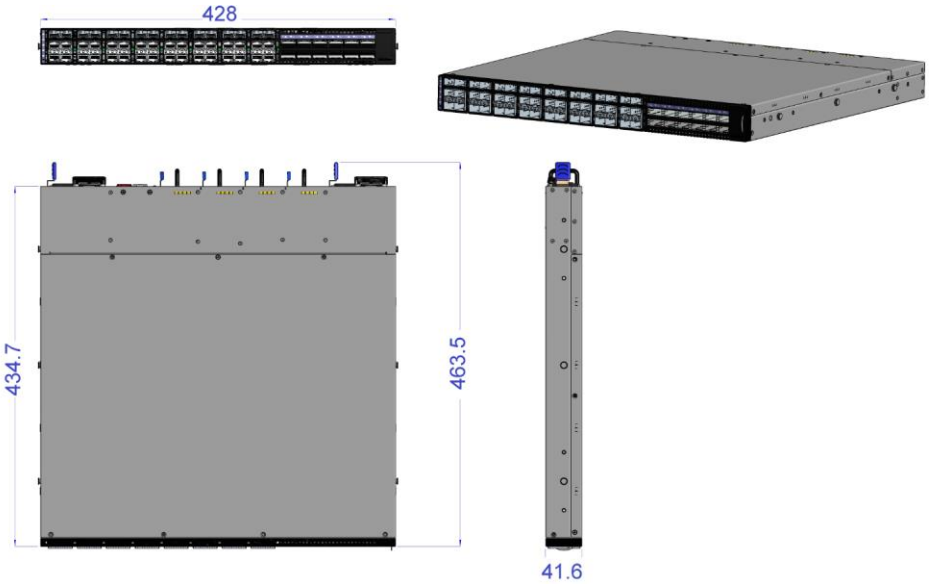

Figure 2 – Short Depth 17" Switch Dimensions

2.3 Fan Unit Dimensions (in mm)

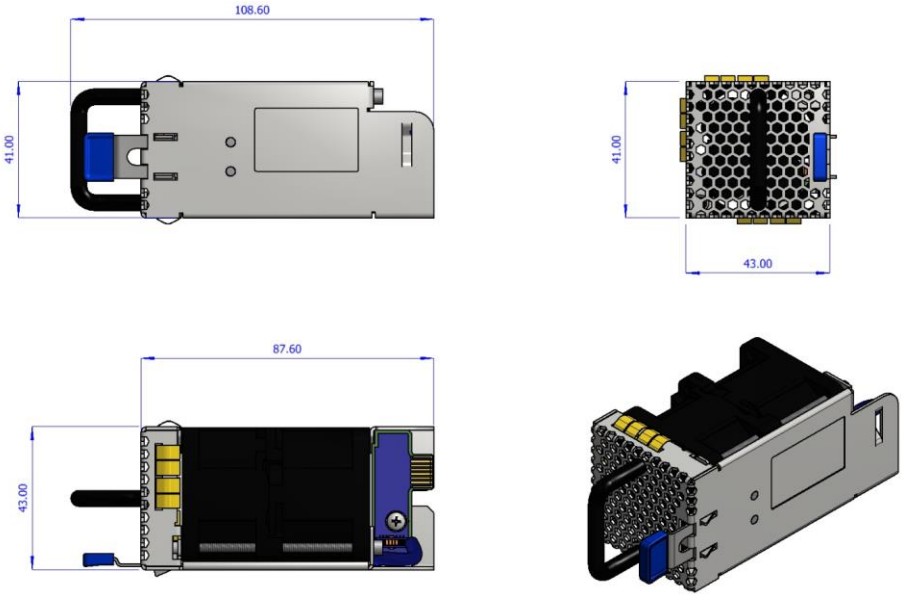

Figure 3: Fan Unit Dimensions

2.4 Power Supply Dimensions (in mm)

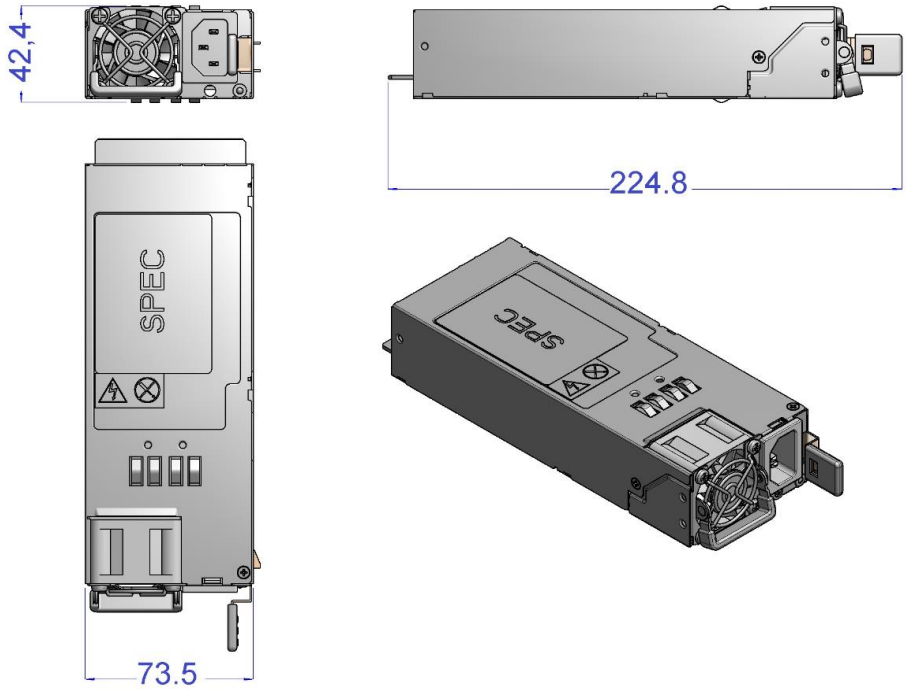

Figure 4: Power Supply Dimensions

2.5 SFP+ Port Assembly

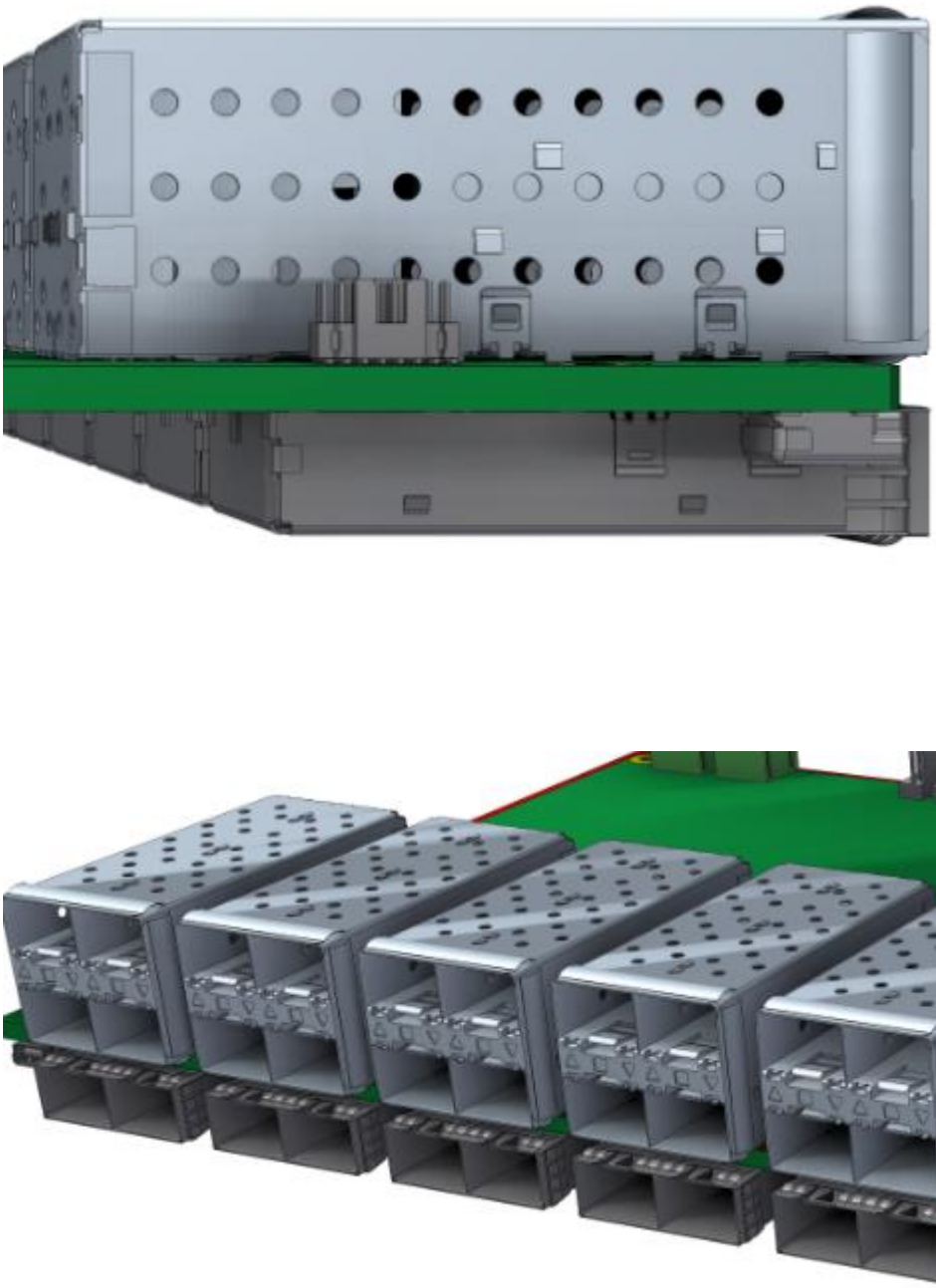

Figure 5: SFP+ Assembly on MSX1410-OCP

3 MSX1410-OCP Features

3.1 Electrical Features

3.1.1 Port Configurations – Based on SwitchX-2™

With each new generation of technology, server sizes shrink. Whereas in the past all servers occupied the entire rack width (19”), it is quite common today to find different racking solutions that allow fitting two servers on the same rack shelf, yielding a total of 56, 60 or more servers in the same rack.

Since most commercially available ToR switches have 48 10GbE ports, connecting more than 48 servers in this rack would require two ToR switches. These two switches increase the rack cost, latency and energy consumption, and occupy space that could have been allocated for servers.

This problem is solved by the port flexibility of the MSX1410-OCP switch. By simple configuration, its 40GbE ports can be configured to 10GbE and support more servers. For example, the MSX1410-OCP can be configured with 60 ports of 10GbE and 4 ports of 40GbE. This configuration connects 60 servers with an uplink of 160Gb/s, or oversubscription of 3.75:1. If configuring the four uplink ports to operate at 56GbE, the oversubscription ratio improves to 2.68:1. Alternatively, the MSX1410-OCP can be configured with 56 ports of 10GbE and 8 ports of 40GbE. This configuration connects 56 servers with an uplink of 320Gb/s, or oversubscription of 1.75:1.

A listing of all possible port configurations can be found in ***Table 2***.

Possible Port Configurations							
40GbE/56GbE ports	12	10	8	6	4	2	0
10GbE ports	48	52	56	58	60	62	64

Table 2: MSX1410-OCP Ports Configuration

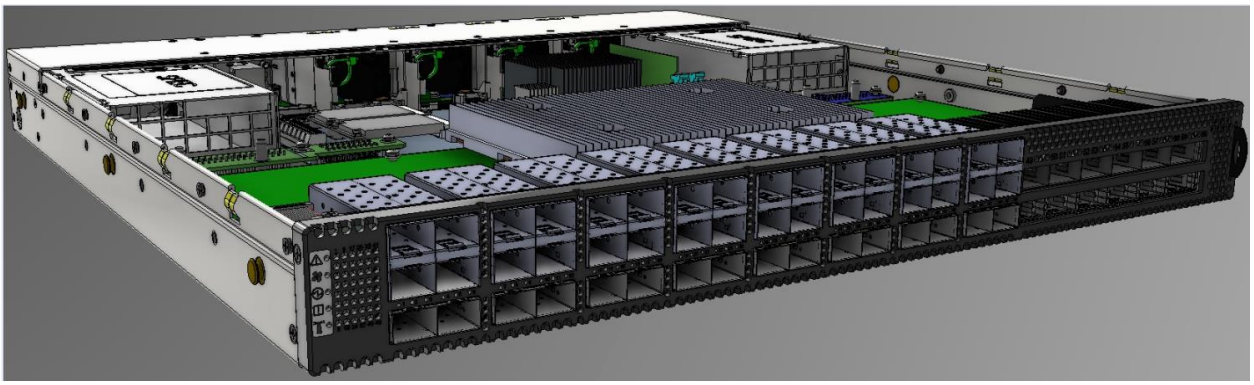

Figure 6: MSX1410-OCP Ports Panel

3.1.1.1 Cables Support

- SR
- SR4
- LR4
- DAC

3.1.1.2 Split Capabilities

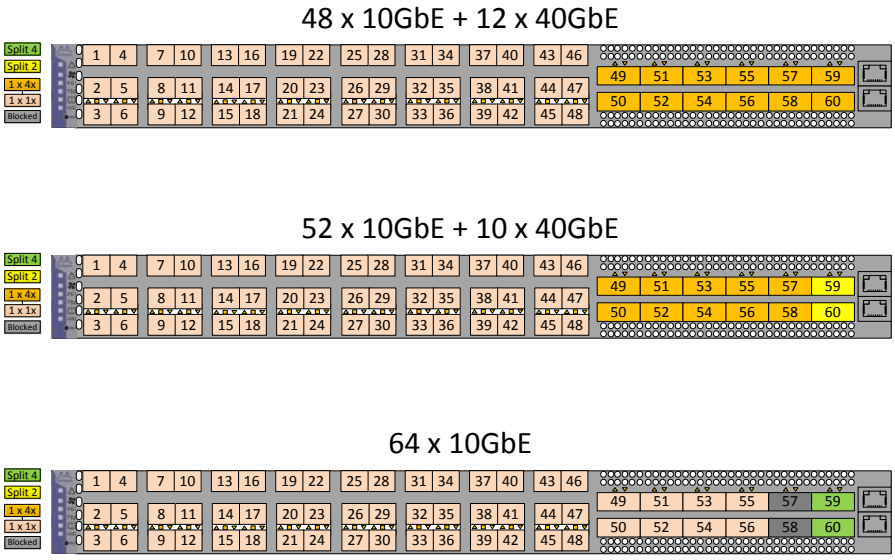

Figure 7: Port Split Configurations

3.1.2 Management I/O Interfaces

- 2x 10/100/1000BASE-T RJ-45 Ethernet port
- 1x RS232/I²C RJ-45 Console connector
- 1x USB connector

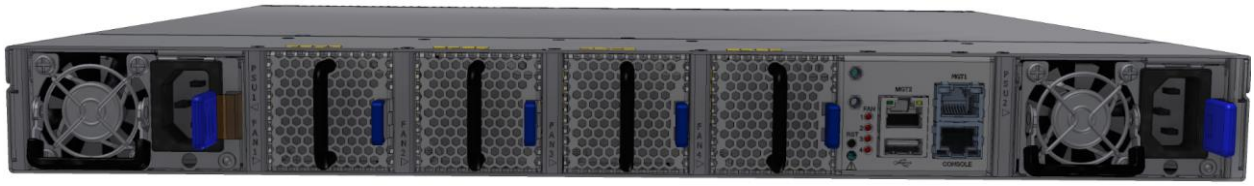

Figure 8: Management I/O interfaces

3.1.3 Field Replaceable Units - FRU

- 3+1 redundant fan field replaceable modules. Can support a dual fan pack in each module
- 1+1 redundant field replaceable power supplies. Two optional power supplies:
    - AC/DC 90-240Vac inlet - 460W
    - DC/DC 48vdc inlet - 800W

3.1.3.1 AC Power Supply Units

The AC power supply units (DPS-460KBJ_REV.03) offer the following features:

- Input voltage: 100 to 240 VAC
- Frequency: 50 to 60 Hz
- Efficiency: 89 to 91% at 220V

Table 3 describes the power supply connector pinout.

Pin #	Description	Pin #	Description
A1~9	GND	**B1~9**	GND
A10~18	+12V	**B10~18**	+12V
A19	PMBus SDA	**B19**	A0 (SMBus Address)
A20	PMBus SCL	**B20**	N/A
A21	PSON	**B21**	12VSB

A22	SMBAlert#	**B22**	Smart)on
A23	Return Sense	**B23**	12VLS
A24	+12V Remote Sense	**B24**	No Connect
A25	PWOK	**B25**	N/A

Table 3: Power Supply Connector Pinout

3.1.3.2 Fan Units

The fan FRU is Delta Module P/N FA121A04-D40. ***Table 4*** describes the fan unit connector pinout. Figure 18 is a mechanical drawing of the unit.

Pin #	Description	Pin #	Description
1	12V	**9**	12V
2	12V	**10**	12V
3	Fan PWM	**11**	FAN TACH
4	Fan PWM	**12**	FAN TACH
5	GND	**13**	GND
6	GND	**14**	GND
7	FAN Present	**15**	I2C SCL
8	3.3V	**16**	I2C SDA

Table 4: Fan Unit Connector Pinout

3.1.4 Management

- CPU: Intel Ivy Bridge
  - Supports all flavors:
    - Celeron
    - i3
    - i5
    - i7
- Chipset: Intel QM77 (Panther Point)
- Up to two 1600MTs, 8GB (Each), DDR3, ECC SO-DIMM modules
- Supports up to two SATA-slim 4/8/16/64GB SLC or MLC SSD modules
- BIOS Field upgrade with fail safe
- Real time clock 24H-SuperCap (optional battery)
- 3x CPLDs with field upgrade and fails safe capabilities

- Default configuration: Celeron / Single DDR3 4GB SO-DIMM / Single SSD 16GB MLC

4 SX1410-OCP Blocks

4.1 Main Block Diagram

The following block diagram specifies the main blocks in the MSX1410-OCP switch:

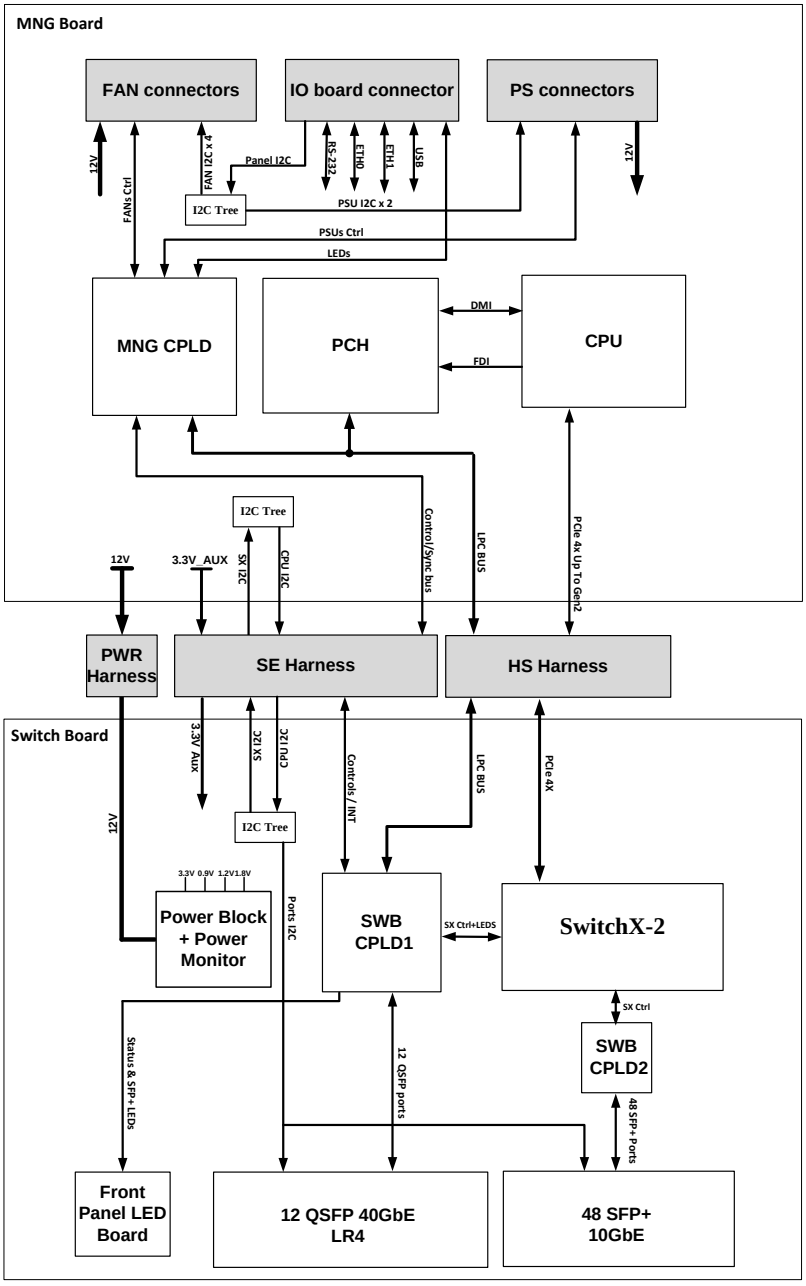

Figure 9: MSX1410-OCP Main Block Diagram

4.2 SwitchX-2™Block

There are 12x40GbE QSFP ports and 48x10GbE SFP+ ports towards the panel from the SwitchX-2. The following diagrams show SwitchX-2 breakout capabilities and the general fan-out of ETH signals on the MSX1410-OCP accordingly:

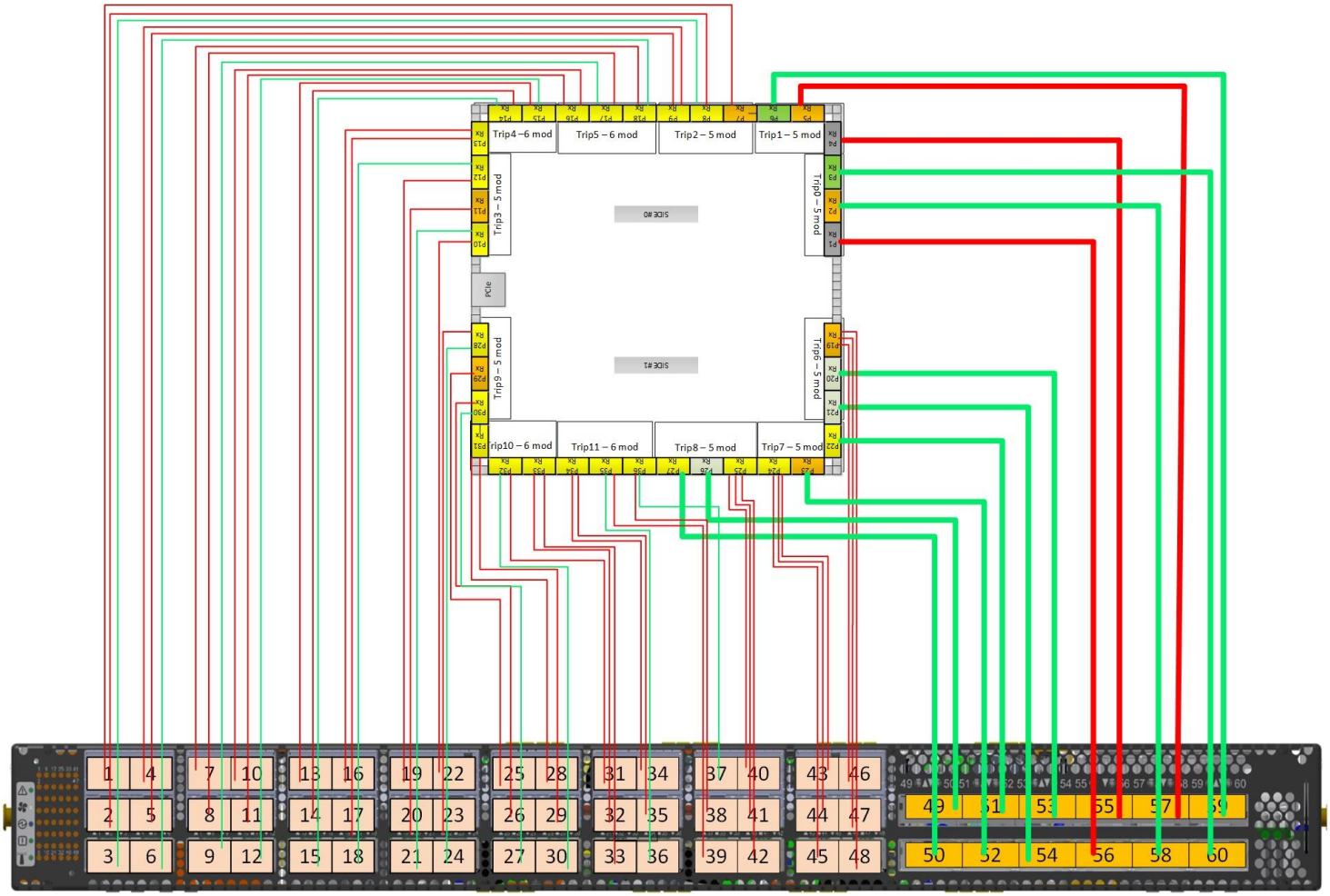

Figure 10: SwitchX-2 Breakout Capabilities and Layout

HW Port	Panel Port	QSFP/SFP+	HW Port	Panel Port	QSFP/SFP+
1	56	QSFP	20	53	QSFP
2	58	QSFP	21	54	QSFP
3	60	QSFP	22	51	QSFP
4	55	QSFP	23	52	QSFP
5	57	QSFP	24	43	SFP+
6	59	QSFP	24	44	SFP+
7	1	SFP+	24	45	SFP+
8	2	SFP+	25	40	SFP+
8	3	SFP+	25	41	SFP+
9	4	SFP+	25	42	SFP+
9	5	SFP+	26	49	QSFP
10	21	SFP+	27	50	QSFP
10	22	SFP+	28	23	SFP+
11	20	SFP+	28	24	SFP+
12	18	SFP+	29	25	SFP+
12	19	SFP+	30	26	SFP+
13	16	SFP+	30	27	SFP+
13	17	SFP+	31	28	SFP+
14	14	SFP+	31	29	SFP+
14	15	SFP+	32	30	SFP+
15	12	SFP+	32	31	SFP+
15	13	SFP+	33	32	SFP+
16	10	SFP+	33	33	SFP+
16	11	SFP+	34	34	SFP+
17	8	SFP+	34	35	SFP+
17	9	SFP+	35	36	SFP+
18	6	SFP+	35	39	SFP+
18	7	SFP+	36	38	SFP+
19	46	SFP+			
19	47	SFP+	36	37	SFP+

Table 5 : Port Mapping

4.2.1 SFP+ Interface

48 ports of the MSX1410-OCP have SFP+ connectors.
Each SFP+ connector supports the following:

Signal Name	Description	Type
TD+/TD-	High Speed Transmission pair	CML from SX to connector
RD+/RD-	High Speed Receive pair	CML from connector to SX
SCL/SDA	I²C slave interface address 0x50 or 0x51	Open Drain from SX to connector
TX Fault	Module transmitter fault	LVTTL from connector to CPLD
TX Disable	Turn off transmitter output (optical)	LVTTL from CPLD to connector
Mod ABS	Module absent (logic high)	Short to GND in module
RS0/RS1	RS0 – Receiver rate select RS1 – Transmitter rate select	LVTTL from CPLD to connector (short together)
RX LOS	Receiver loss of signal indication	LVTTL from connector to CPLD

Table 6 : SFP+ Interface

A single SFP+ interface on the MSX1410-OCP switch is illustrated in ***Figure 11***.

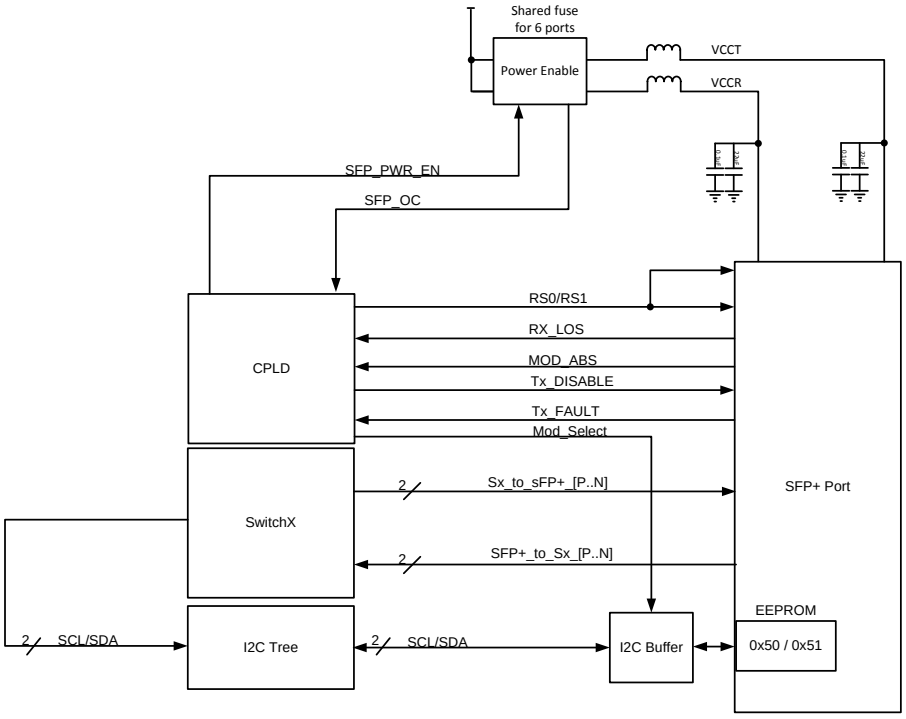

Figure 11 : SFP+ Interface

4.2.2 QSFP Interface

12 ports of the MSX1410-OCP are QSFP connectors.

Each QSFP connector supports the following:

Signal Name	Description	Type
TD+/TD-	High Speed Transmission pairs (x4)	CML from SX to connector
RD+/RD-	High Speed Receive pairs (x4)	CML from connector to SX
SCL/SDA	I²C slave interface address 0x50 or 0x51	Open Drain from SX to connector
ModSel	Module select for I²C communication to prevent bus conflicts	LVTTL from CPLD to connector
ResetL	Module reset use to reset all user module settings to their default state	LVTTL from CPLD to connector
IntL	Use to indicate the host for a fault or status critical.	OC from Connector to CPLD
ModPrsL	Module present	Short to GND in module
LPMod	When the module is in a low power mode it has a maximum power consumption of 1.5W	LVTTL from CPLD to connector

Table 7 : QSFP Interface

A single QSFP interface in the MSX1410-OCP board is presented in ***Figure 12***.

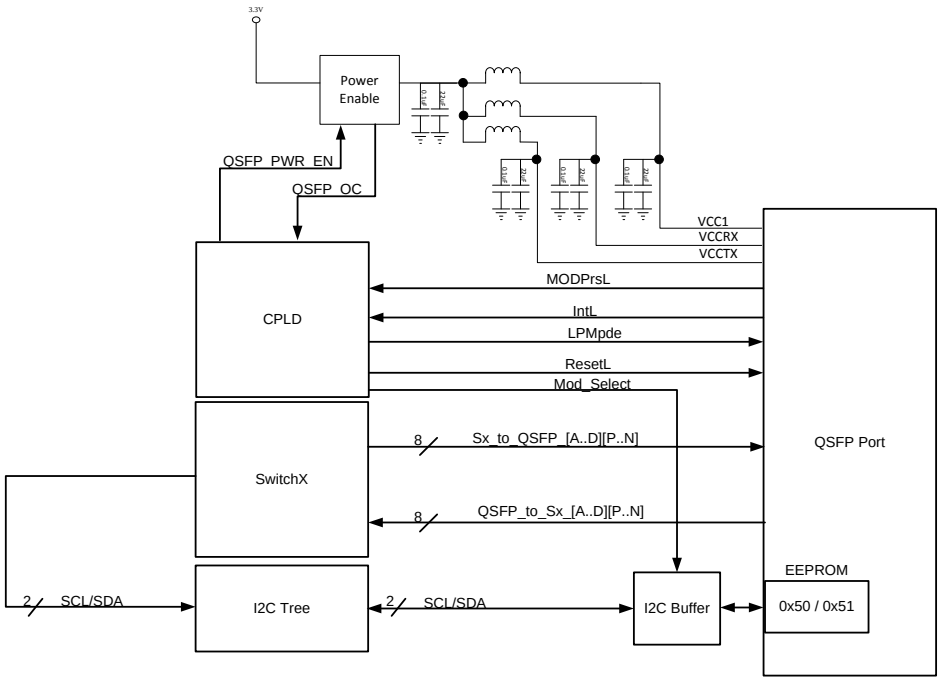

Figure 12 : QSFP Interface

4.2.3 Port over Current Protection

The power for every 4 QSFP/ 6 SFP modules is controlled by a current limited power distribution switch. All of these power limiting switches are controlled by QSFP_PWR_EN signals (default enabled), originated from the CPLD. When the output load exceeds the current limit threshold, the power switch limits the current to a safe level. In addition, the SFP/QSFP_OC signals are asserted low and create an interrupt event towards the CPU from the CPLD.

4.2.4 Port I²C Interface

CPLD controls the I²C switch and connected to SwitchX-2 secondary I²C bus. 60 different I²C buses are connected from I²C switch to the 60 MSX1410-OCP ports. MOD_SEL register selects which I²C bus will be open– each time only one bus can be open.

The main reason for using this CPLD buffering method is that all the QSFP modules have the same address (0x50). By separating the buses in the CPLD, Switch-X2 can select a specific module to communicate with. Another reason for selecting a different bus for each port is to isolate problematic modules from the others. This way, if a specific port is corrupted, it does not affect the other ports.

4.3 I²C

- MSX1410-OCP I²C tree has 4 optional masters:
  1. SW (LPC to I²C module)
  2. FW (SwitchX2)
  3. External connector
  4. BIOS (SMbus)
- The PCH does not support I²C, therefore, LPC to I²C module was implemented inside the MGMT CPLD.
- I²C switching is done by external switch devices. These devices are controlled by an I²C machine inside the CPLD. I²C channel access is controlled via relevant CPLD registers.
- The CPU and the I²C external connector are sharing the same switches.
- The I²C external connector is not connected directly to the head of the CPU I²C tree. Access to CPU I²C is granted by request to CPLD 0x80.
- When connection between the I²C external connector and the head of the tree is enabled, the CPU I²C connection to the head of the tree is disabled.
- In case of collision between masters, the following hierarchy will be implemented (decreasing order):
  1. Connector
  2. SW I²C register
  3. SW LPC register
  4. FW
- The I²C and the RS232 are sharing a single RJ45 connector.

The following diagram illustrates an I²C tree on MSX1410-OCP:

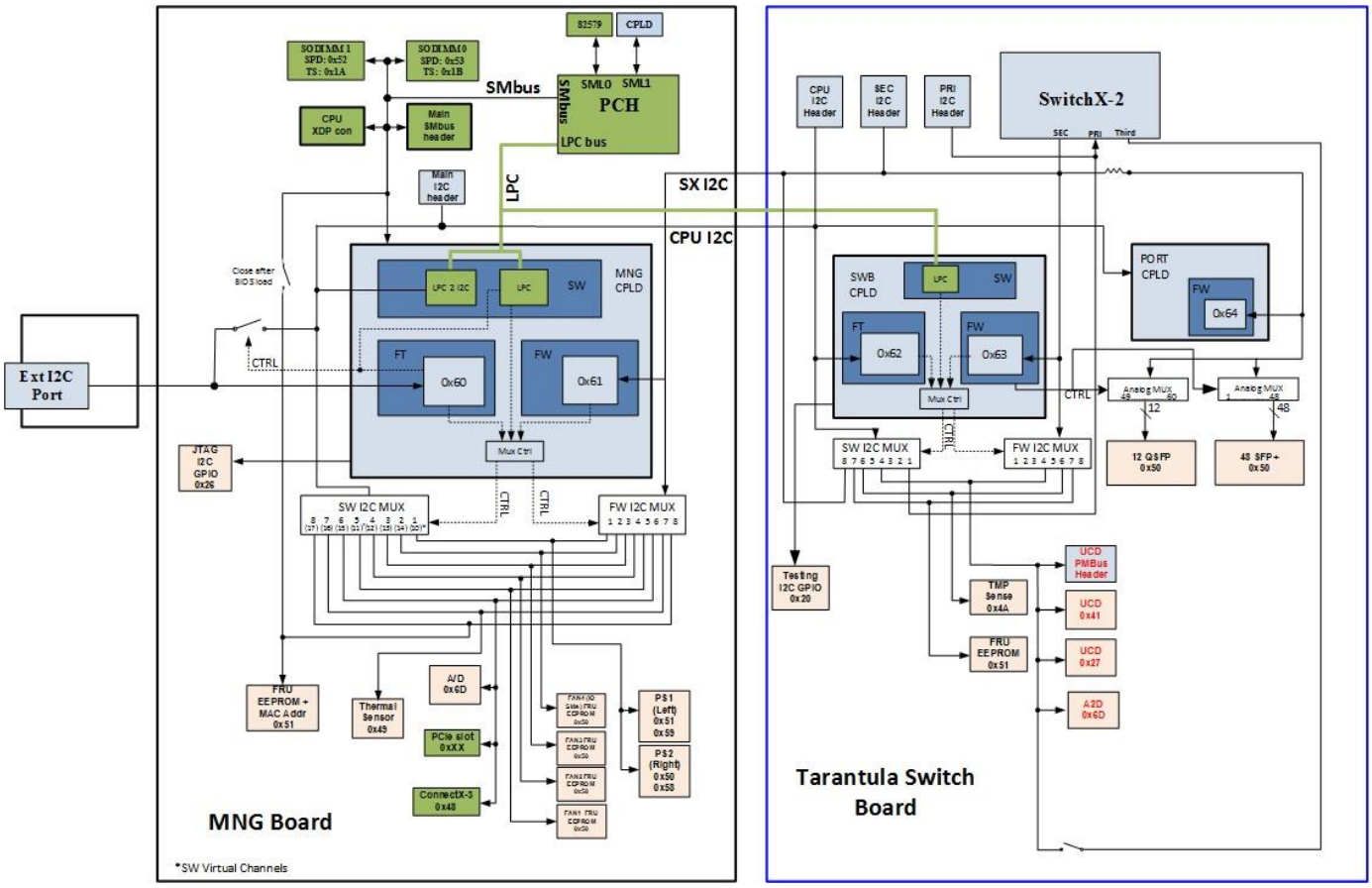

Figure 13: I²C Tree

4.4 CPLD

Three CPLD devices on the MSX1410-OCP provide monitoring and controlling capabilities:

1. MGMT CPLD: located on the MGMT board. Controls fans, PS, power sequence, software register map, software interrupt controller, firmware interrupt controller, reset controller, I²C switching, LEDs, CPU power and safe BIOS.
2. Ports CPLDs (1-2): controls LEDs, SwitchX-2 reset logic, power monitor, SwitchX-2 power sequence, I²C switching and ports control and monitor signals.

All CPLDs are powered from the auxiliary power; therefore, they are still functioning during main 12V rail shutdown.

4.4.1 CPLD Features

- Lattice XO2 CABGA332 device
- 4320 Logic elements
- 271 I/Os in four banks
- 96K bit UFM User flash memory
- 332-Pin FineLine BGA
- Fail-safe mechanism using SPI flash memory

4.4.2 CPLD Field Upgrade

All Three CPLD devices can be field upgraded via SwitchX-2 JTAG emulation. Fail safe feature is supported in case of field upgrade failure for all CPLDs.

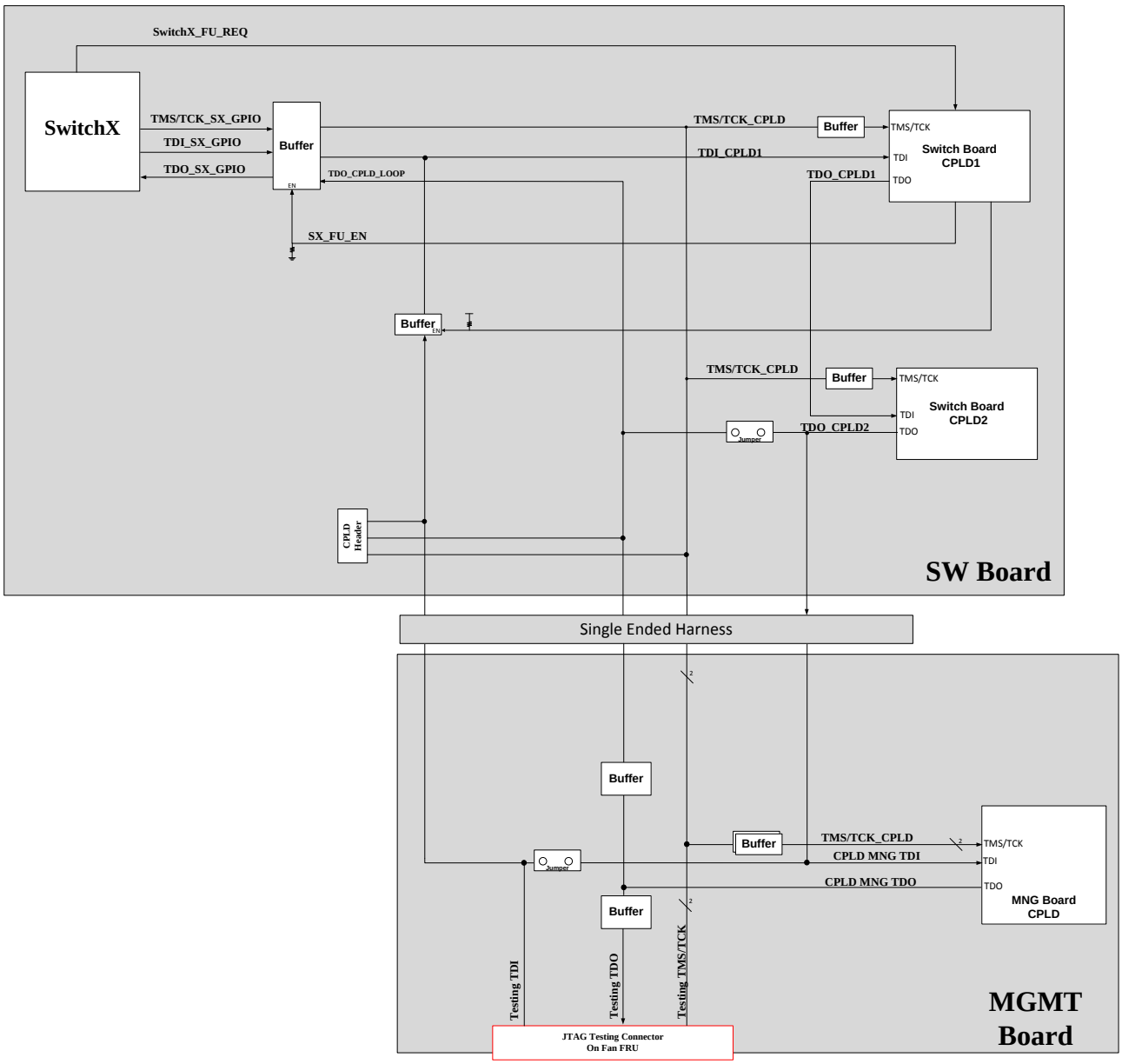

Figure 14: CPLD Field Upgrade

.

4.5 SPI and Safe BIOS Mechanism

The Serial Peripheral Interface on the chipset supports two 64MB (each) flash devices, storing a Unified BIOS Code.

The MGMT board has two pairs of SPI flash devices: one serves as the default pair and the other serves as the safe BIOS pair.

The BIOS SPI interface is described in ***Figure 15***.

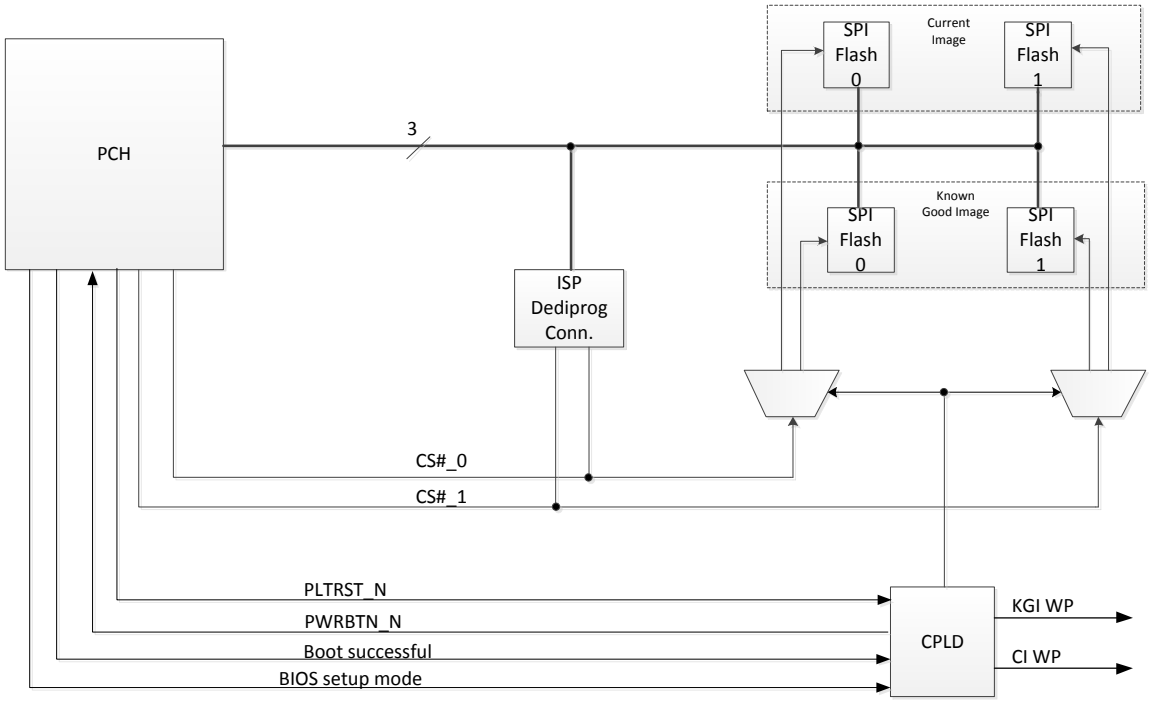

Figure 15: Safe BIOS Mechanism

The system has two SPI Flash pairs:

- Known Good Image (KGI) pair
- Current Image (CI) pair

Both of these pairs are loaded with identical images at the factory. The factory programmed images are the KGI images. The KGI image is never intended for update. The KGI provides an image to always recover the system.

The system always chooses the CI as the default boot device after power up. This selection is made by the CPLD.

Safe BIOS Events flow:

1. Initial boot: the default is CI.
2. When PLTRST_N is asserted (logic low) and then de-asserted, a timer starts inside the CPLD (this timer value can be a variable, based upon boot time of the system).
3. The CPLD then waits for “Boot Successful” bit driven by the BIOS.
4. If Boot Successful is high before the timer has been expired, then this is a normal boot.
5. If not, the mechanism starts working.
6. CPLD selects the KGI pair.
7. Then PWRBTN_N is asserted for 6 seconds, de-asserted for 5 seconds, asserted for 1 second and then released. This will power cycle the platform (PCH and CPU).
8. The system then boots from KGI.
9. After a successful boot, the CPLD selects the CI again.
10. At this stage, it is possible to recover the CI using BIOS FU mechanism.

In addition to the BOOT_SUCCESS signals, two other signals are connecting the PCH to the CPLD:

i. BIOS_SETUP_MODE - Asserted high by the PCH in case the user entered the setup menu. As long as this signal is high, the CPLD counter for boot success is paused.

ii. BIOS_STARTED - Asserted high by the PCH once BIOS load was started.

BOOT_SUCCESS and BIOS_STARTED are used for status LED control.

4.6 Ethernet

The MSX1410-OCP provides two Ethernet MDI 10/100/1000MB/s ports. One port is connected through internal PCH MAC and Intel 82579 PHY (port 1). The other port is connected through Intel 82583 MAC/PHY (port 2).

These two MDI ports are routed to two RJ45 ports on the FRU panel.

The Ethernet interface is displayed in ***Figure 16***.

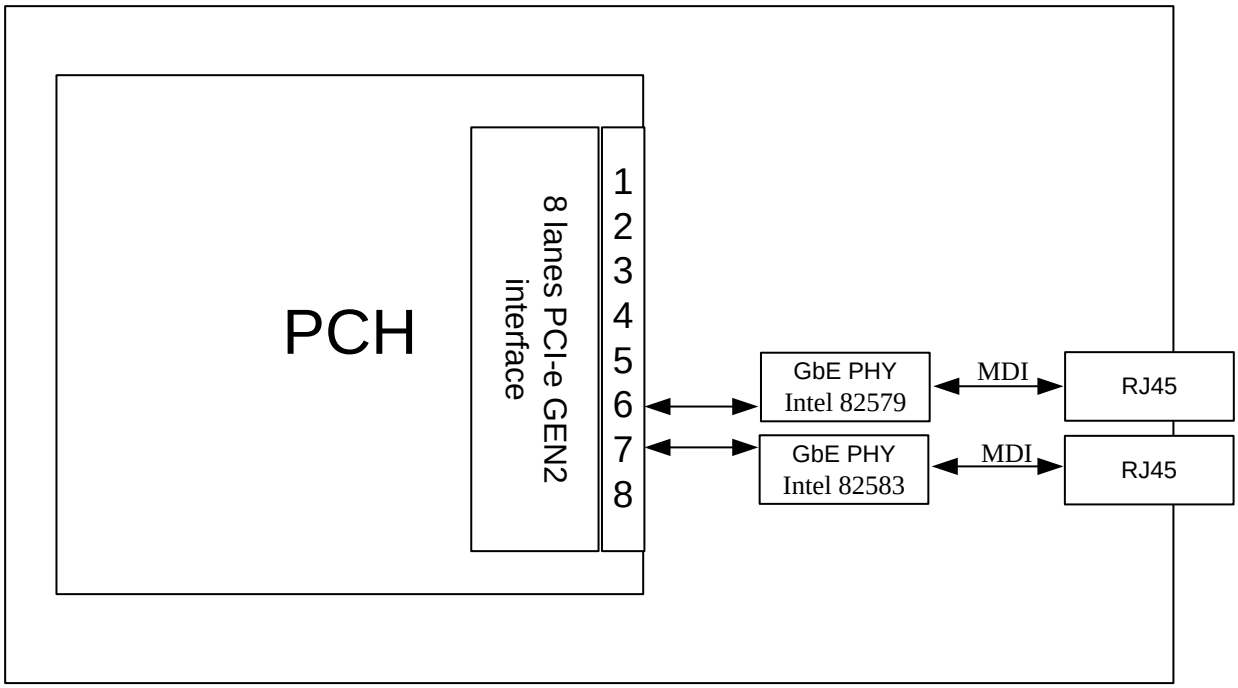

Figure 16: Ethernet Interface

4.7 RS232 Interface

The 8-pin RJ-45 connector provides CLI connectivity (RS-232) to the CPU. The I²C and the RS232 are sharing the same RJ45 connector. The RS-232 interface is shown in ***Figure 17***.

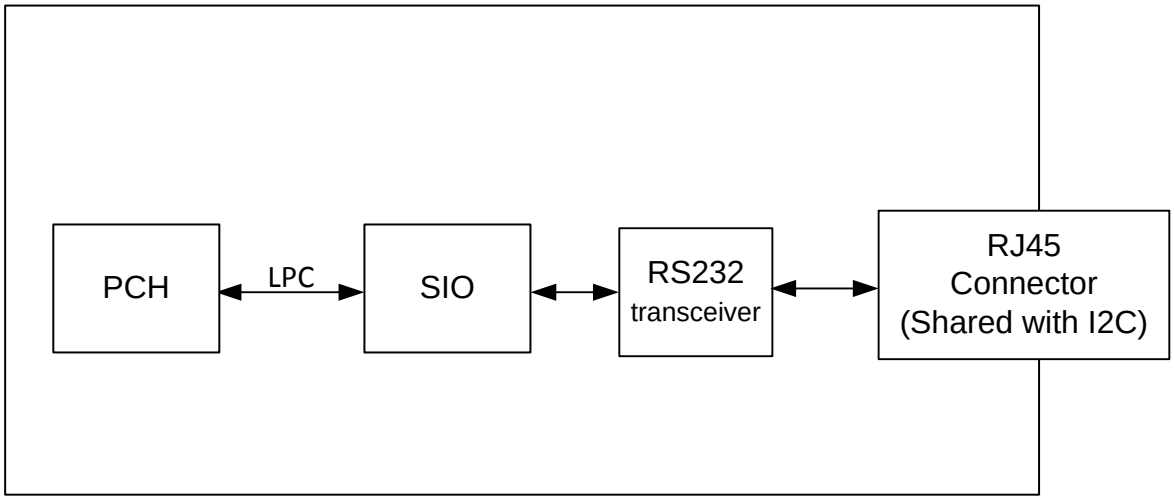

Figure 17: RS-232 Interface

4.8 USB

The system has one external USB 2.0 host interface for general use. The USB port can supply 500mA @ 5V and has internal current limiter with fault signal. ***Figure 18*** describes the USB interface.

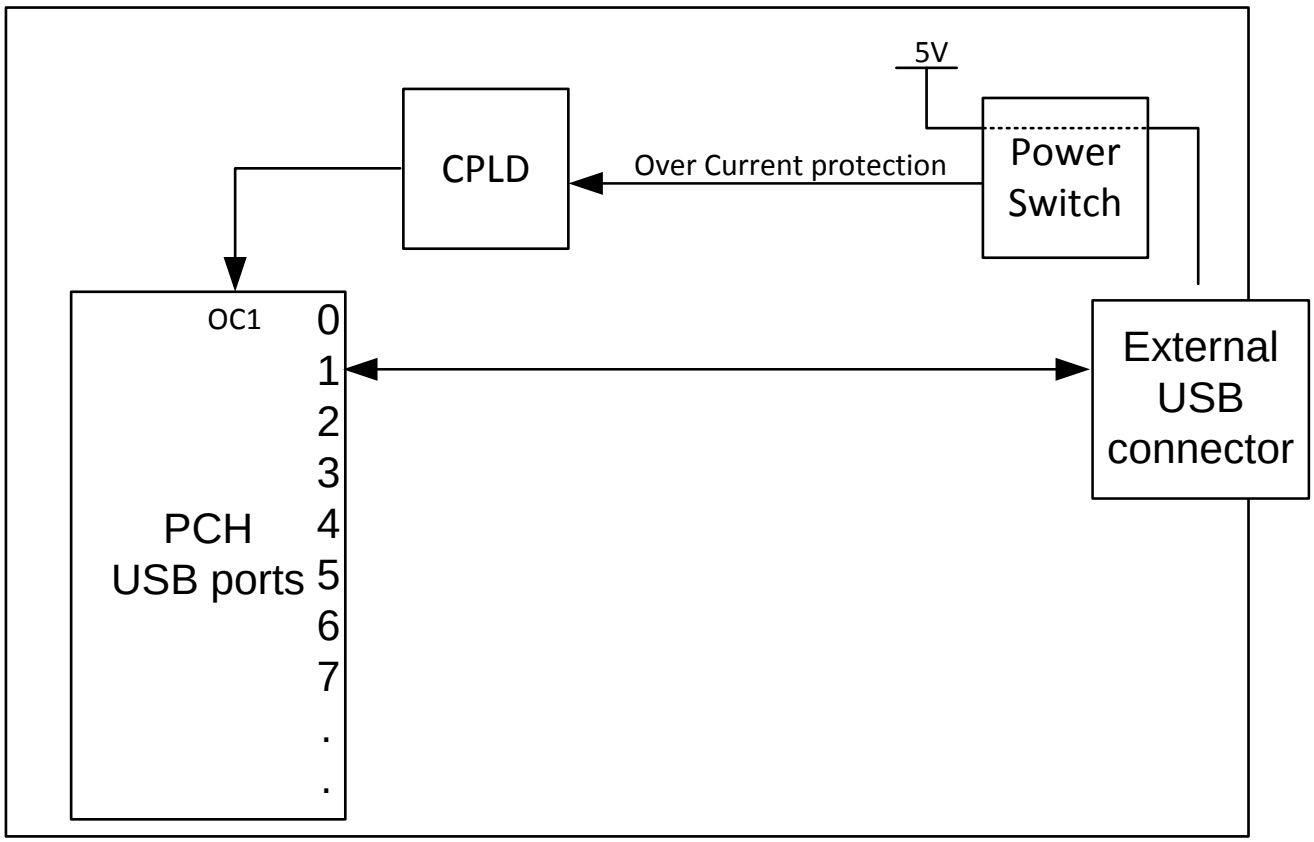

Figure 18: USB Interface

4.9 LPC

The PCH implements an LPC interface, as described in the Low Pin Count Interface Specification, Revision 1.1. The LPC interface from the PCH is shown in ***Figure 19***.

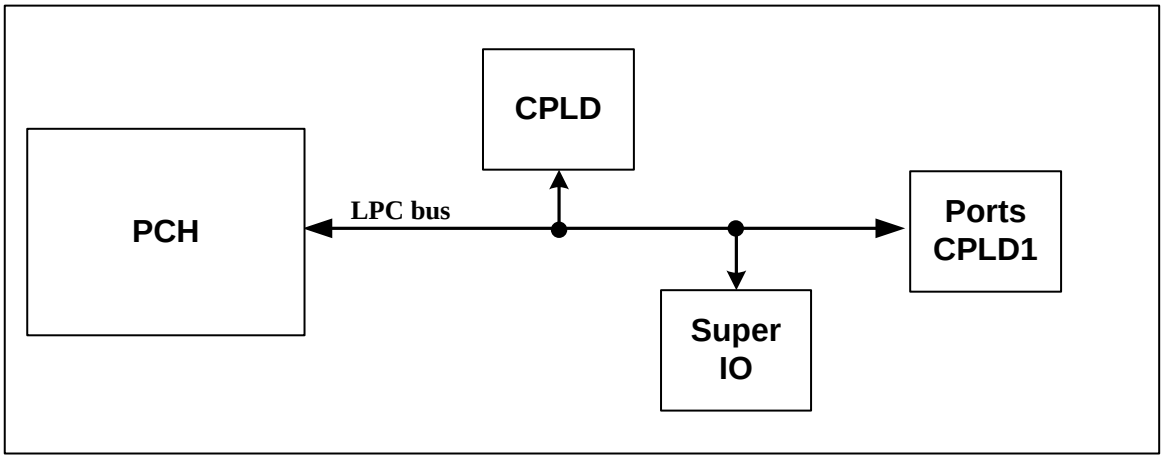

Figure 19: LPC Interface

4.10 SATA

Up to two mSATA SSD (Solid State Drive) modules are supported in the system. Each SSD supports 4-64GB SLC or MLC.

On power down sequence, both of the SSD modules are powered from Super capacitor, in order to perform a graceful shutdown. The super capacitors provide power to the SSD modules while the rest of the system is in shutdown mode.

The SATA interface is described in ***Figure 20***.

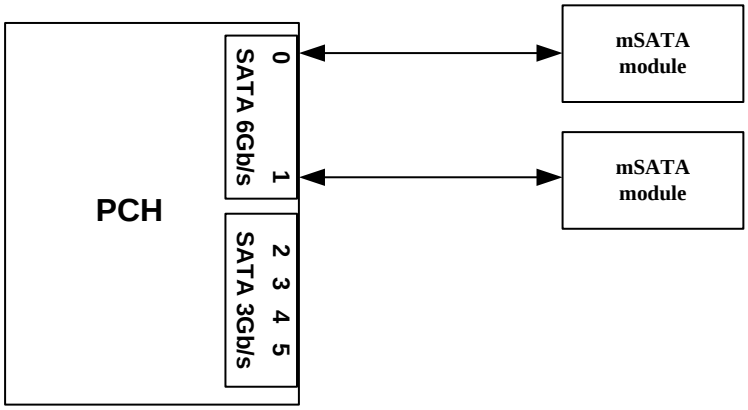

Figure 20: SATA Interface

4.11 PCIe

MSX1410-OCP implements 4 PCIe interfaces:

1. 1X from PCH to 82579
2. 1X from PCH to 82583
3. 8X (gen 2 for Celeron/i3, Gen3 for i5/i7) from CPU to 8X slot
4. 4X gen 2 connection between CPU and SwitchX -2.

The PCIe interface is shown in ***Figure 21***.

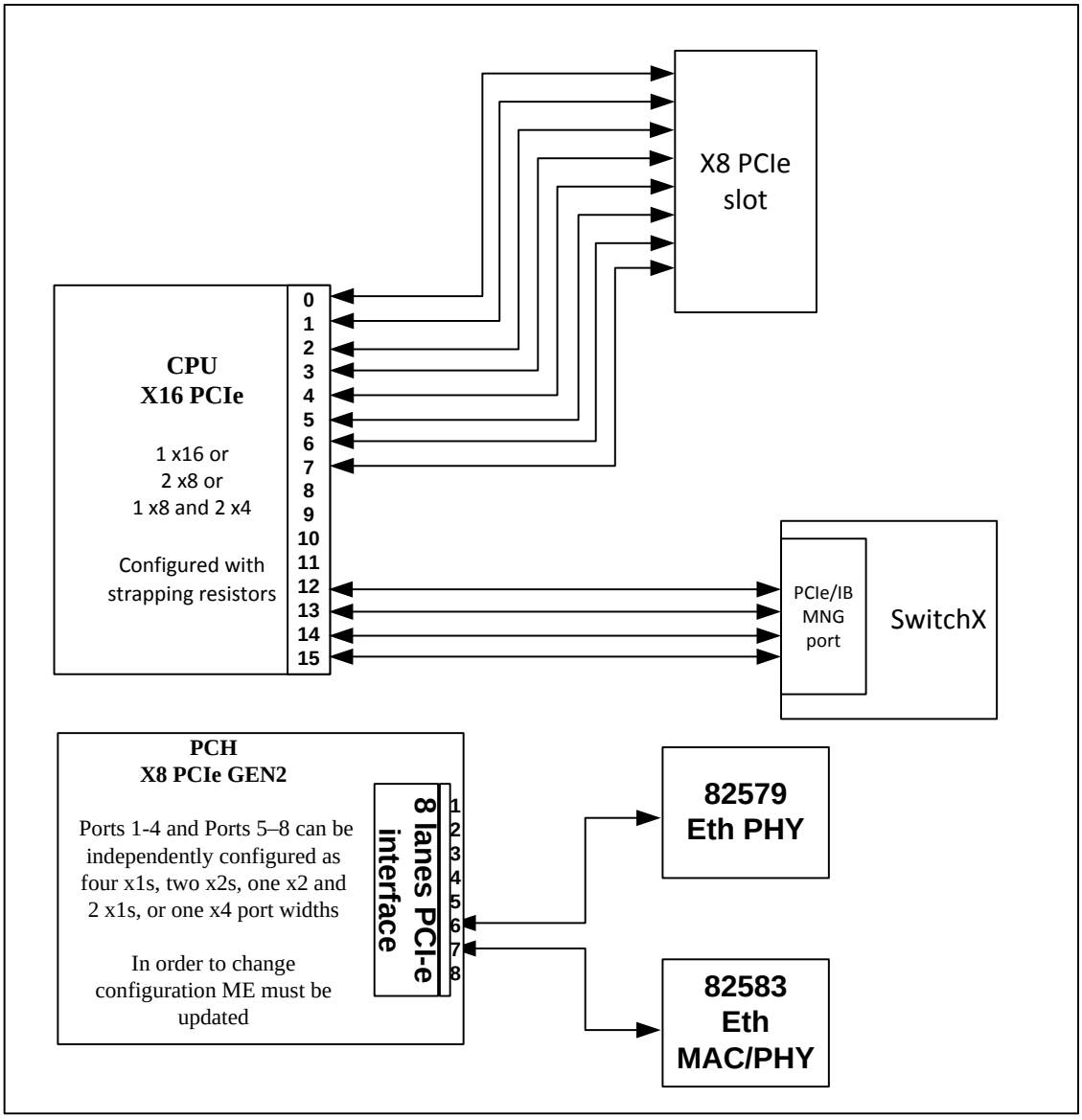

Figure 21: PCIe Interface

5 Power

5.1 Power Consumption

5.1.1 Switch Board Max Power Consumption

Configuration	3.3_SX (3.3V)[W]	1.8_SX (1.8V) [W]	1.2_SX (1.2V) [W]	VCORE_SX (0.9V) [W]	Total 12v Including DC/DC Efficiency
36 Copper (0W per port)	5.9	9.5	20.6	41.2	85.8
36 SR4 (3.5W per port)	77.9	9.5	20.6	41.2	165.8
36 LR4 (2W per port)	131.9	9.5	20.6	41.2	225.8
36 Active cables (1.5W per port)	59.9	9.5	20.6	41.2	145.8

Table 8: Switch Board Max Power Consumption

5.1.2MGMT Board Power Consumption

	Current (Amp)	Power[Watt]															Total power
voltage	voltage	0.90	1.23	1.05	1.05	1.35	0.675	0.85	5.00	3.30	1.90	1.05	1.80	0.90	1.20	1.80	consumption
Device	Qty.	CPU Core	CPU GFX	CPU VCCIO	PCH Vtt	Vddq DDR3	Vtt DDR3	VCCSA	5V	3.3V	1.9V	1.05	CPU PLL	ADIR Core	ADIR	ADIR	per part
CPU	1	47.700	0.000	8.925	0.000	6.750	0.000	5.400	0.000	0.000	0.000	0.000	2.160	0.000	0.000	0.000	35.000
PCH	1	0.000	0.000	0.000	6.152	0.225	0.000	0.000	0.010	0.878	0.000	0.000	0.004	0.000	0.000	0.000	4.000
82583	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.505	0.205	0.000	0.000	0.000	0.000	0.727
82579	1	0.000	0.000	0.000	0.320	0.000	0.000	0.000	0.000	0.310	0.000	0.000	0.000	0.000	0.000	0.000	0.630
CPLD	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.330	0.000	0.000	0.000	0.000	0.000	0.000	0.330
SODIMM	1	0.000	0.000	0.000	0.000	4.050	0.506	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	4.573
mSATA	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.990	0.000	0.000	0.000	0.000	0.000	0.000	0.990
ConnectX-3	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.957	0.000	0.000	0.000	2.124	1.032	0.504	4.617
PCIe slot	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
USB	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	2.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	2.500
CRT	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SPI flash	2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.132	0.000	0.000	0.000	0.000	0.000	0.000	0.132
SIO	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.132	0.000	0.000	0.000	0.000	0.000	0.000	0.132
OSC	4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.264	0.000	0.000	0.000	0.000	0.000	0.000	0.264
LEDs	15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.594	0.000	0.000	0.000	0.000	0.000	0.000	0.594
3.3->1.9V (Integrated in 82583)	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.878	0.000	0.000	0.000	0.000	0.000	0.000	0.878
3.3->1.05V (Integrated in 82583)	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.644	0.000	0.000	0.000	0.000	0.000	0.000	0.644
3.3-> 1.8V CPU	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	3.967	0.000	0.000	0.000	0.000	0.000	0.000	3.967
3.3V-> 1.8 ADIR	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.924	0.000	0.000	0.000	0.000	0.000	0.000	0.924
	Total	47.70	0.00	8.93	6.47	11.03	0.51	5.40	2.51	11.03	0.51	0.20	2.16	2.12	1.03	0.50	

	Power[W]
12V (including 90% efficiency)	51.66
12V_AUX (including 90% efficiency)	12.26
12V PCIe slot	0.00
Total	63.91

	CPU power [W]
t> 2sec	35
10ms<t<2s	45
t<10mS	71

Table 9: MGMT Board Power Consumption

5.1.3System Power Consumption

Part	Power	Current	Comments
Fans	92.2	7.7	---
MNG Board	51.7	4.3	---
SWB	225.8 (with 36 LR4)	18.8	---
12V Total	369.6	30.8	From 460W
12V AUX MNG	12.3	1.0	---
12V AUX SWB	3.5	0.3	---
12 AUX Total	15.7	1.3	From 24W

Table 10: System Max Power Consumption

5.2 Power Monitoring and Distribution

5.2.1 Switch Board Power Monitoring and Distribution

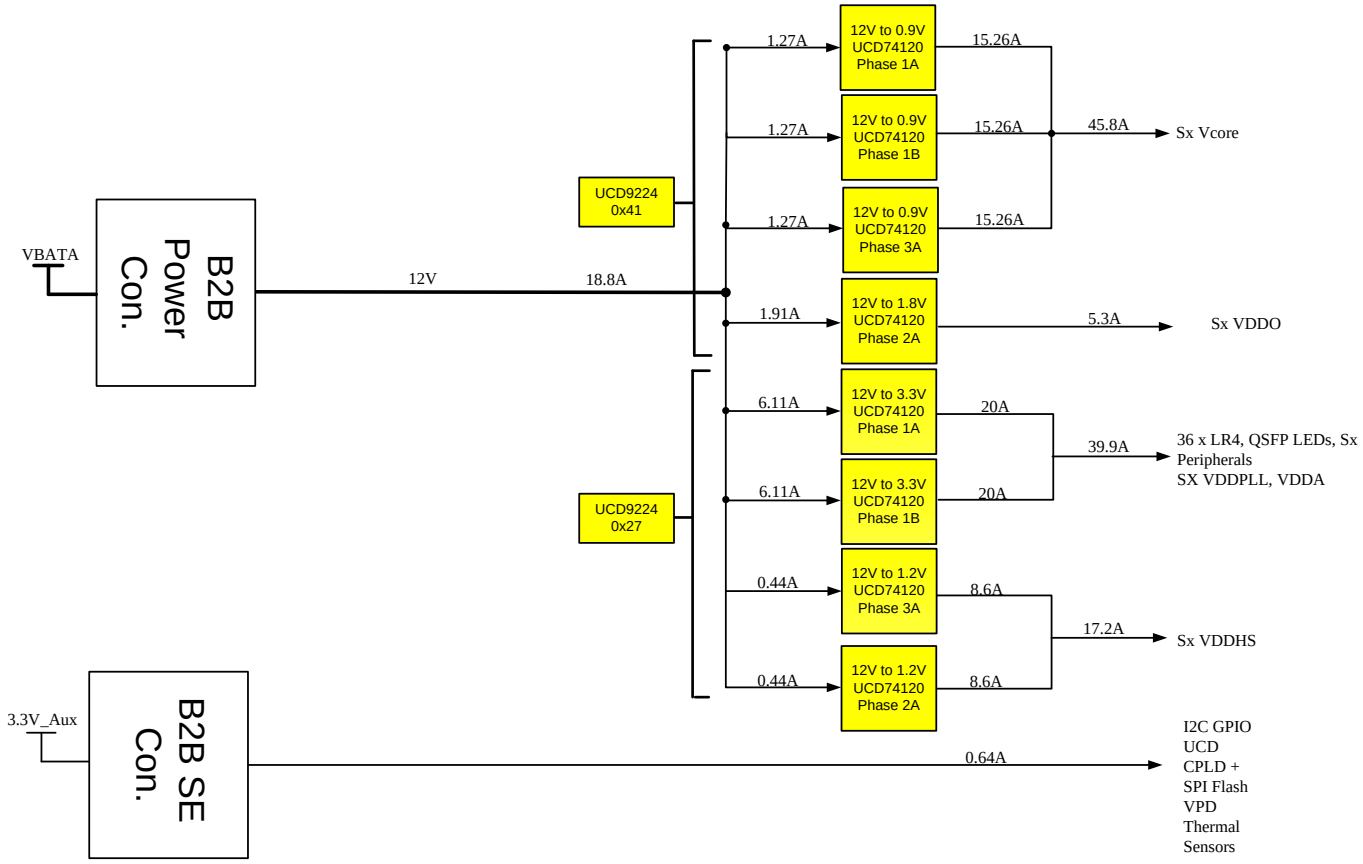

Figure 22: Switch Board Power Distribution

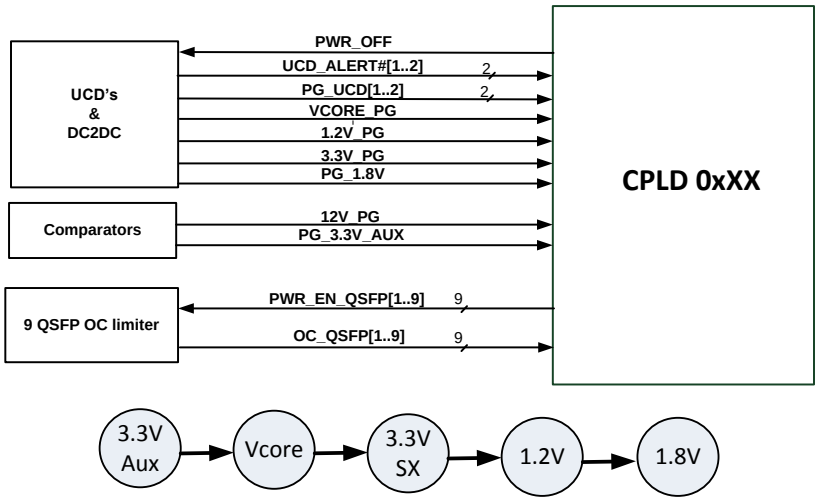

Figure 23: Switch Board Power Monitor and Power Sequence

5.2.2MGMT Board Power Monitoring and Distribution

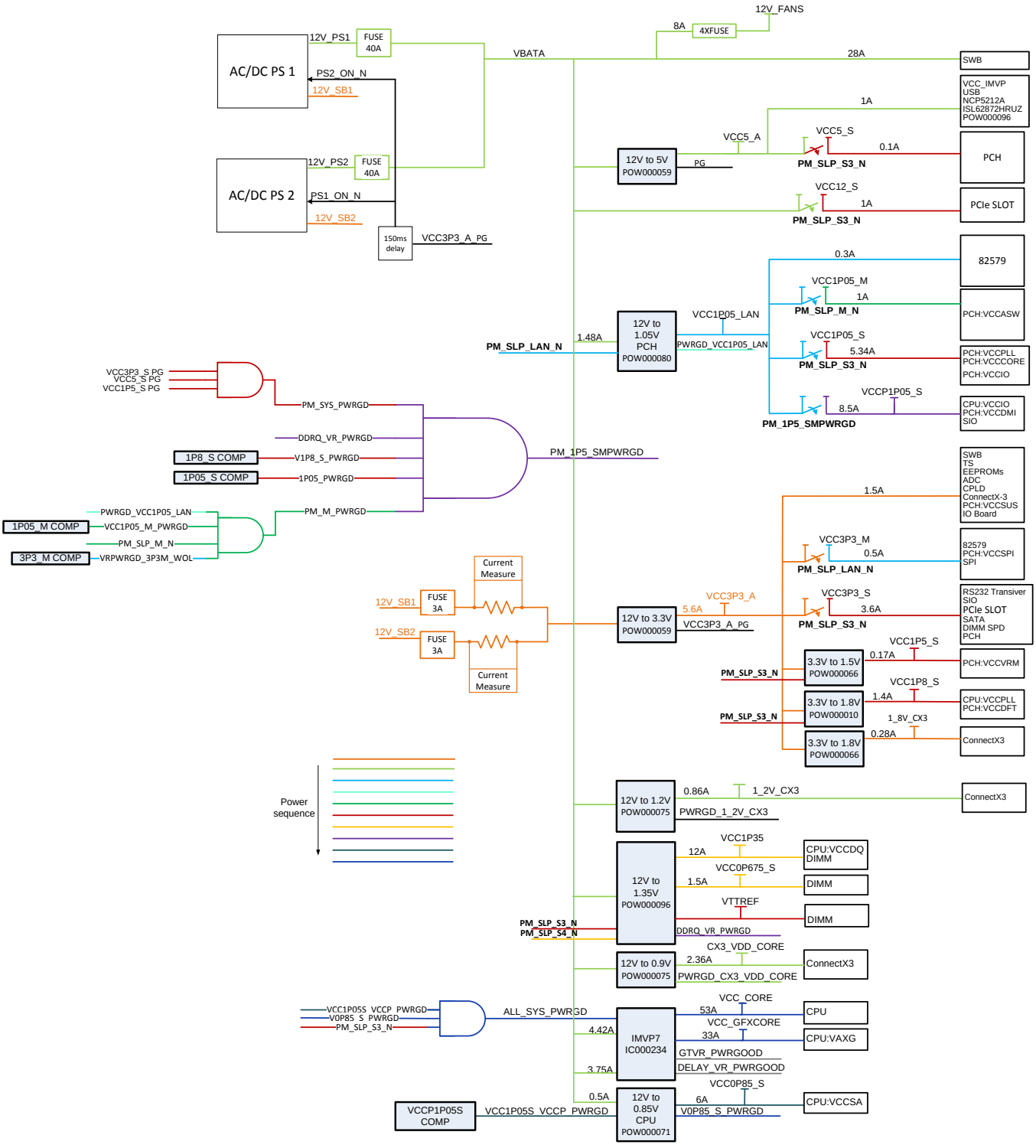

Figure 24: System and MGMT Board Power Distribution, Monitoring and Sequence

5.2.3 System and MGMT Board Voltages and Currents Sensing

Certain voltages and currents in the system can be read by SW/FW via an 8-channel I²C analog to digital converter. Voltage/current inputs to the A2D channels 2-7 are MUXed by 1:2 MUX. Therefore, two values can be read per channel. The MUX select signal is derived by the CPLD and can be controlled by SW and FW. The main 12V voltages and currents can be read directly from the AC/DC PS.

The following table describes the voltage rail in each channel and the expected reading.

The acceptable range for each rail is +-10% from nominal value.

A/D Channel	MUX Select Signal	Monitored Voltage/Current		"B" Parameter
0	NA	DDR3 VTT 0.675V (VCC0P675_S)		0.008 (See note 2)
1	NA VCC0P85_S (See Note 1)	**Rail Level**	**[VID0,VID1]**	0.008 (See note 2)
		0.675	[1,1]	
		0.725	[1,0]	
		0.8	[0,1]	
		0.9	[0,0]	
2	0	12V_AUX1 (12VSB1 current)		0.008 (See note 2)
	1	12V_AUX2 (12VSB2 current)		0.008 (See note 2)
3	0	3.3V (VCC3P3_A)		0.016 (See note 2,3)
	1	5V (VCC5_A)		0.024 (See note 2,3)
4	0	CPU 1.8V (VCC1P8_S)		0.008 (See note 2)
	1	CX3 1.2V (1_2V_CX3)		0.008 (See note 2)
5	0	CPU/PCH 1.05V (VCC1P05_LAN)		0.008 (See note 2)
	1	CPU 1.5V (VCC1P5_S)		0.008 (See note 2)
6	0	CPU 1.05V (VCCP1P05_S)		0.008 (See note 2)
	1	CX3 1.8V (1_8V_CX3)		0.008 (See note 2)
7	0	DDR3 1.35V (VCC1P35)		0.008 (See note 2)
	1	CX3 0.9V (CX3_VDD_CORE)		0.008 (See note 2)

Table 11: A/D Channels

Notes:

1. This rail has 4 valid values controlled by the CPU, using 2 lines which are connected to the voltage regulator - VCCSA_SEL[0:1].
   In order to allow the testing of all 4 levels in JTAG tests, these two controls are routed through the CPLD (CPLD_VCCSA_SEL[0:1]).
2. The B parameter is used to calculate the real voltage/current from I²C reading, using the following formula: V=B*R, where V is the real voltage and R is the I²C reading.
3. This voltage is divided by voltage divider before it is sampled; therefore its B parameter contains the divider value as well.

5.3 FRU Control

5.3.1 Fan Control

The MGMT CPLD includes a fan controller which is controlled by the SwitchX-2 FW. The fans speed is controlled by 8 PWM signals from the MGMT CPLD to the fans and they are monitored by 8 Tacho signals from the fans towards the CPLD.

PWM:

- A single PWM output signal from SwitchX-2 is connected to the MGMT CPLD. This signal is reflected to all 8 fans as default.
- 8 PWM registers (register per fan) are available via the FW I²C (part of the FW register map). Writing into one of these registers will override the PWM signal into the specific mated fan

Tacho:

- 8 Tacho RO registers will be available via the FW I²C (part of the FW register map).
- FW will be interrupted in case PWM ≠ Tacho.

Fan present signal can be monitored by SW and FW via the CPLD.

Inside the FRU, I²C is connected only to one EEPROM. FW and SW can accesses the fans I²C buses.

5.3.2 Power Supply Control

The PS is controlled via PMbus and a few signals. SW and FW can access the PMbus.

PS fans can be controlled via the PMbus. By default, PS fans are controlled automatically by the PS. PS fans speed can only be increased through FW/SW, but can never be decreased.

All other PS control signals can be controlled (for PS inputs) and monitored (for PS outputs) via FW and SW registers maps.

The MSX1410-OCP FRU control is illustrated in ***Figure 25***.

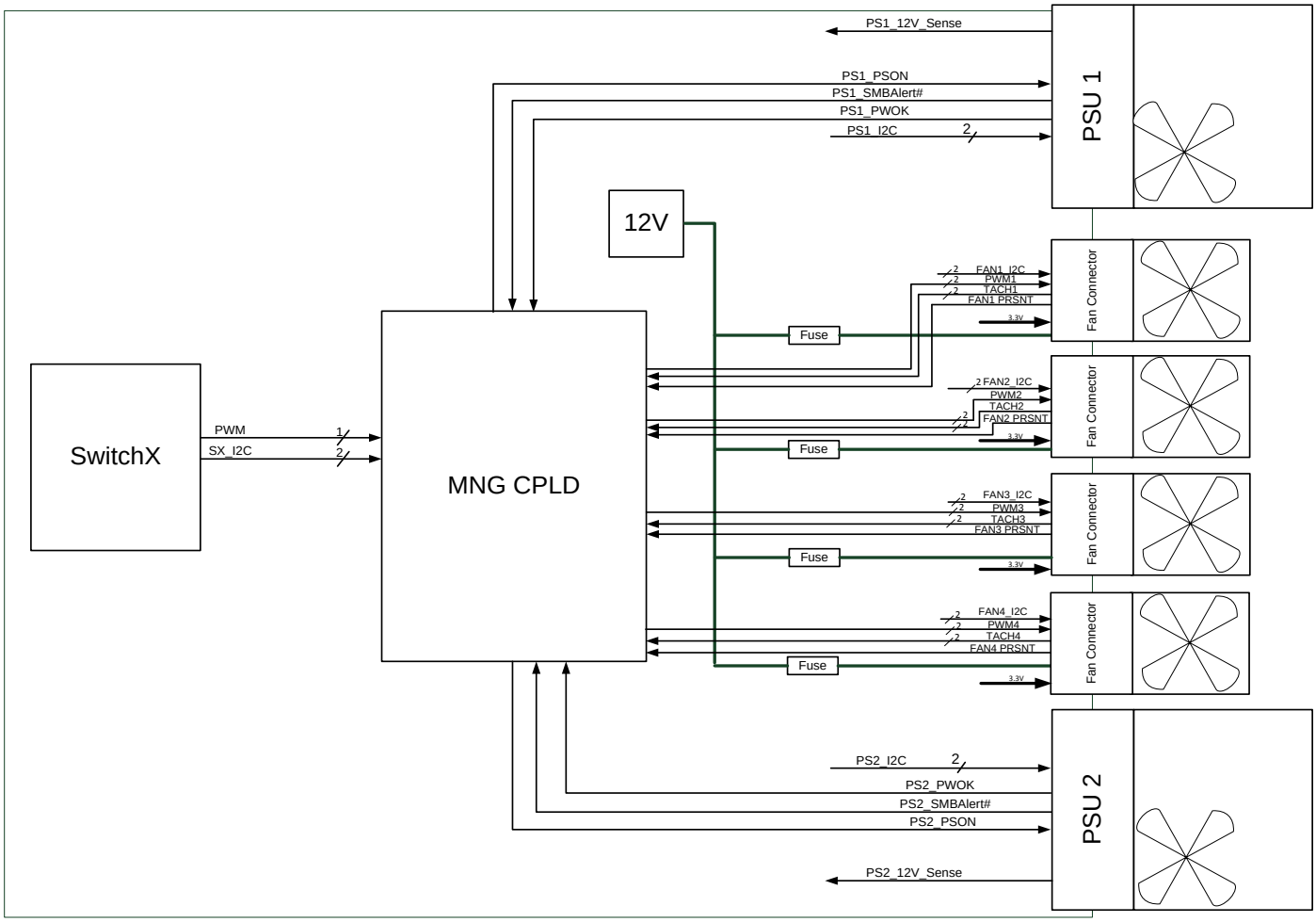

Figure 25: MSX1410-OCP FRU Control

5.4 Reset

Table 12 specifies the reset scenarios and power monitoring in the system:

#	Cause	Source	Target	Description
1.	Power monitor circuit indication	CPLD	Entire system	The power monitor circuit verifies that the local voltages are in the range of ±10% of the nominal value.
2.	External push button Short assertion	Push Button	Entire system	User reset from panel. Assertion of more than ½ sec and less than 12sec.
3.	External push button Long assertion	Push Button	Entire system	User reset from panel. Restores CPU factory defaults. Assertion of more than 12sec.
4.	SW RST	SW	Entire system SW has the ability to reset specific components via CPLD	SW request RST.
5.	FW RST	FW	Entire system SW has the ability to reset specific components CPLD	FW request RST
6.	WD expired	CPLD	Entire system WD is disabled as default.	WD mechanism is implemented in the CPLD device. The CPU periodically writes counter value via LPC to 4 WD registers in the SW board 0xXX CPLD – if the CPLD does not recognize a counter value change, board reset is initiated.
7.	I²C Reset request	CPU\SX	I²C tree	CPU\SX requests the CPLD to reset the internal I²C Switches.

Table 12: MSX1410-OCP Reset Matrix

The MSX1410-OCP reset mechanism is displayed in ***Figure 26.***

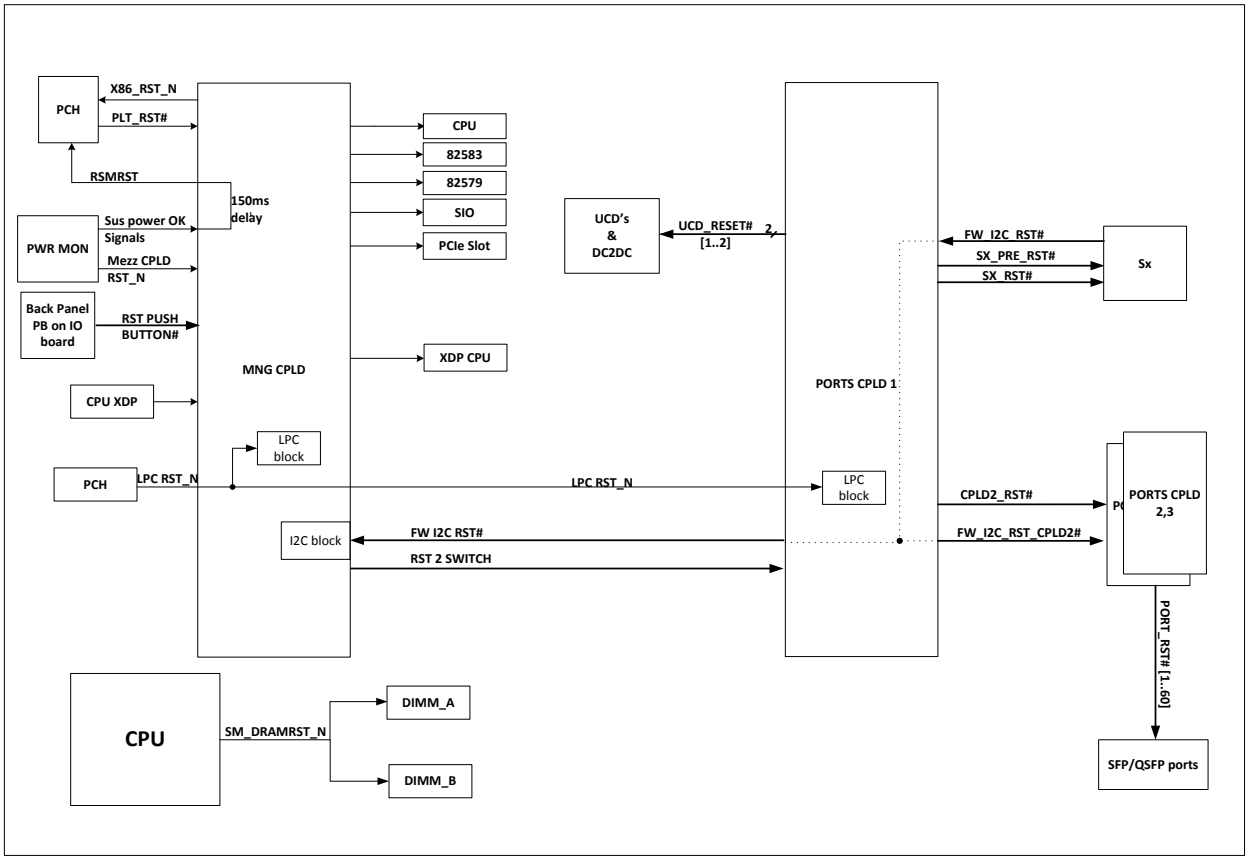

Figure 26: MSX1410-OCP Reset Tree

5.5 Temperature Monitor

The following temperature sensors can be read by SW, FW or both:

- Two external sensors provide temperature sensing:
  - In front of fan FRUs: sense the ambient temperature when fans are pushing air into the system (Air in system)
  - In front of the ports panel: sense the ambient temperature when fans are sucking air from the system (Air out system)
- The Switch-X2 has an internal temperature sensing diode. The temperature sensed by the diode is processed and according to it, two signals may be asserted: Over Thermal Warning and Over Thermal Shutdown. The temperature above / below which these signals are activated /deactivated is controlled by firmware via the INI file, and it must be within the operational temperature range of the device.
- The CPU has an internal sensor that can be monitored only by SW. The CPU can reduce its power (the performance might be reduced as well) in order to prevent from reaching its maximum value.
- The PCH has an internal sensor that can be monitored only by SW.
- Each SODIMM has one module thermal sensor that can be monitored only by SW.

5.6 Clock Distribution

5.6.1Switch Clock Distribution

Figure 27 specifies the clock distribution in the switch board.

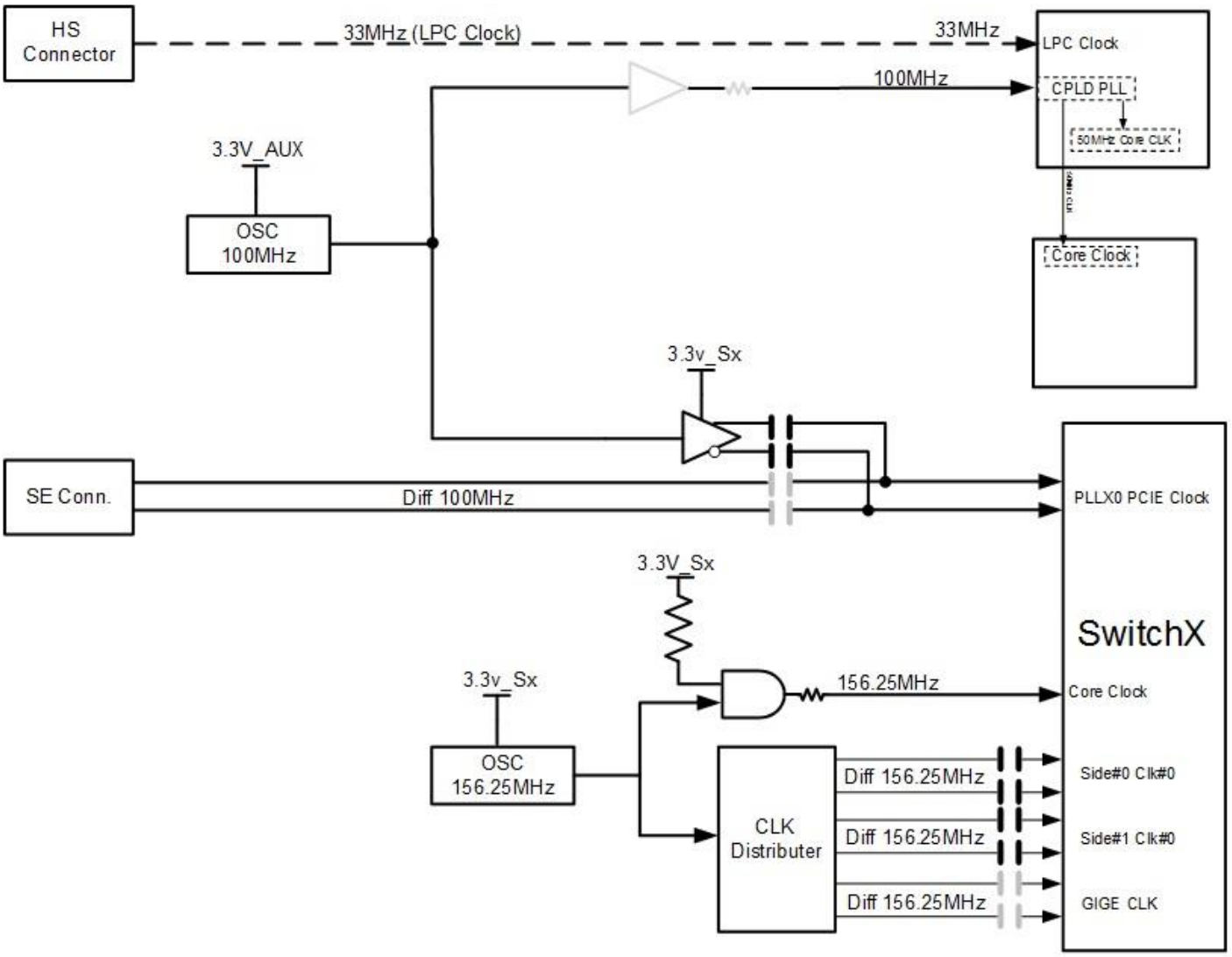

Figure 27: Switch Clock Distribution

5.6.2MGMT Clocks

Most of the MGMT clocks are generated from two crystals and distributed by the PCH. In addition, one 25MHz oscillator is buffered and distributed to the ETH PHYs and to the CPLD.

Figure 28 describes the clock distribution for the MGMT.

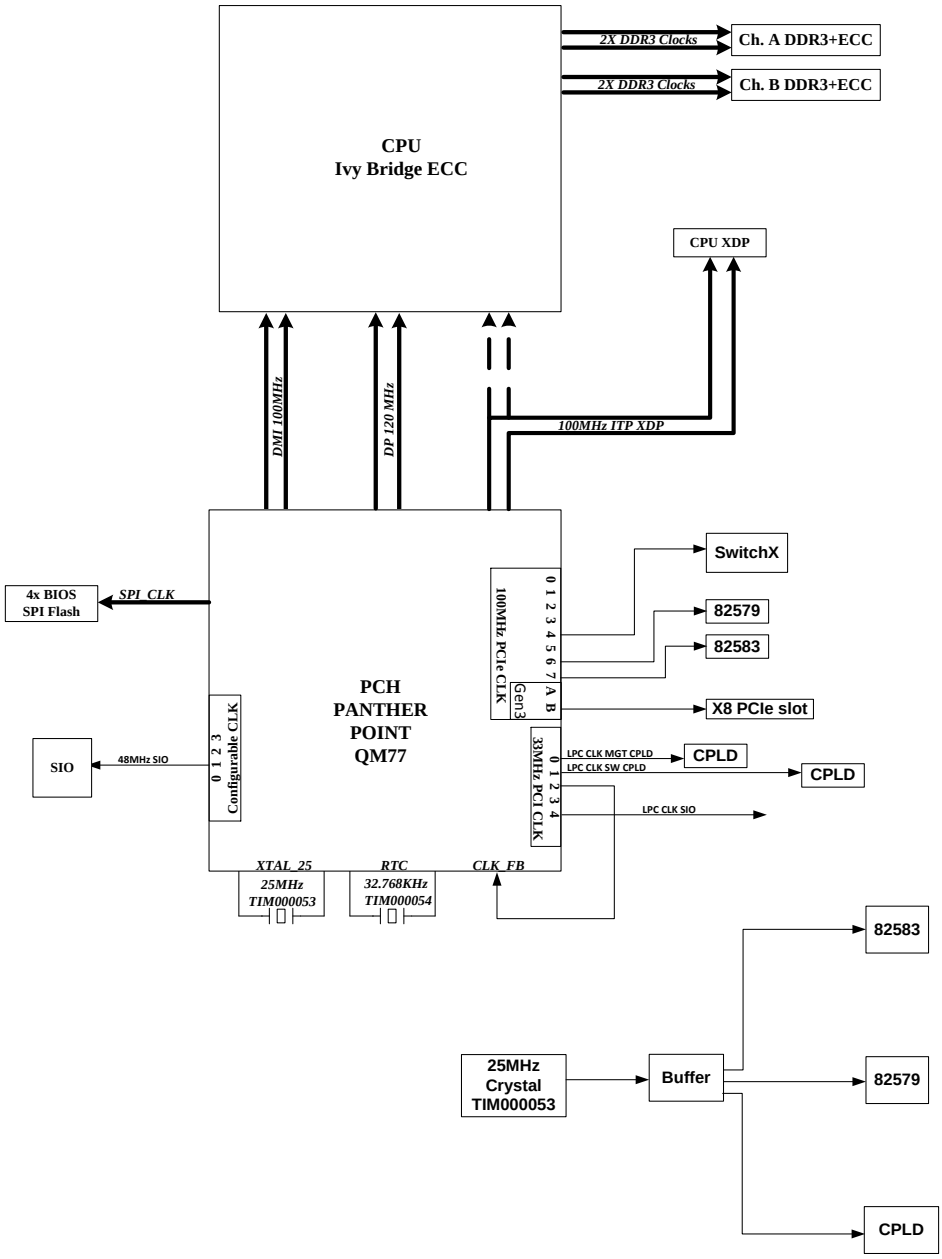

Figure 28: MGMT Clock Distribution

5.7 LEDs

T

Table 13 describes the indication LEDs on the MSX1410-OCP.

LED	Qty.	Color	Description		Originator	Driver	Location
QSFP LEDs	12	Green/ Amber	Ethernet: Off – Link down Solid Green – Link up ,no traffic Blinking Green – Traffic (Received packet) Blinking Amber – Physical errors		Switch-X2	Above QSFP Ports	QSFP LEDs
SFP	12	Green/ Amber	Ethernet: Off – Link down Solid Green – Link up ,no traffic Blinking Green – Traffic (Received packet) Blinking Amber – Physical errors		Switch-X2	Above SFP Ports	SFP
Status LED	1	Amber / Green	Off – No power Solid Amber – Fault Solid Green – Normal operation Blinking Green – Boot		CPU	On Ports Side and on FRU side	Status LED
General Fans LED	1	Amber / Green	Off – No power Green – All fans are operating Amber – Fan failure		CPLD	On Ports Side	General Fans LED
Fans 1-4 LEDs	4	Amber / Green	Off – No power Green – Fan# is operating Amber – Fan failure		CPLD	FRU side	Fans 1-4 LEDs
UID LED	1	Blue	Static – The operator has activated this LED to identify this module. Blinking – The Operator is instructing to replace this module		SwitchX-2	Ports Side	UID LED
Bad Port indication	1	Green / Amber	Blinking Amber – Bad port indication		Switch-X2	Ports Side	Bad Port indication
PSU LED	1	Green/ Amber	Green – Both PSU OK Amber –PSU is Faulty		CPLD	Ports Side	PSU LED
PSU LED	1 per PS	Green/ Amber	Output ON and OK	GREEN	PS	PS	PS FRU
			No AC power to all	OFF			

	FRU		power supplies				
			AC present / Only 12VSB on (PS off) or PS in Smart on state	1Hz Blink GREEN			
			AC cord unplugged or AC power lost; with a second power supply in parallel still with AC input power.	AMBER			
			Power supply warning events where the power supply continues to operate; high temp, high power, high current, slow fan.	1Hz Blink Amber			
			Power supply critical event causing a shutdown; failure, OCP, OVP, Fan Fail	AMBER			
			Power Supply FW updating	2Hz Blink GRENN			

T

Table 13: MSX1410-OCP LEDs

5.8 Power

5.8.1Power Consumption

5.8.1.1 Celeron Based CPU, 48xDAC+12xSR4

	Voltage [V]	Current [A]	Power [W]
Fans (normal speed)	12V	3.2	38.4
CPU + periphery	12V	4.3	51.7
Switch + periphery	12V	6.4	76.6
SFP – DAC	12V	0	0
QSFP – SR4	12V	2.2	26.6
Total 12V power	12V	16.1	193.3
Auxilary power	12VSB	1.3	15.6

Table 14: Celeron, 48xDAC, 12xSR4

5.8.1.2 Celeron based CPU, 48xSR+12xLR4

	Voltage [V]	Current [A]	Power [W]
Fans (high speed)	12V	4.8	57.6
CPU + periphery	12V	4.3	51.7
Switch + periphery	12V	6.4	76.6
SFP – SR	12V	4.4	53.3
QSFP – LR4	12V	3.9	46.6
Total 12V power	12V	22.2	285.8
Auxilary power	12VSB	1.3	15.6

Table 15: Celeron, 48xSR, 12xLR4

5.8.2Power Distribution – Switch

The power rails of SwitchX-2 and its periphery are driven from TI UCD9224 with UCD74120 power stages.

Figure 29 specifies the power distribution of the Switch power:

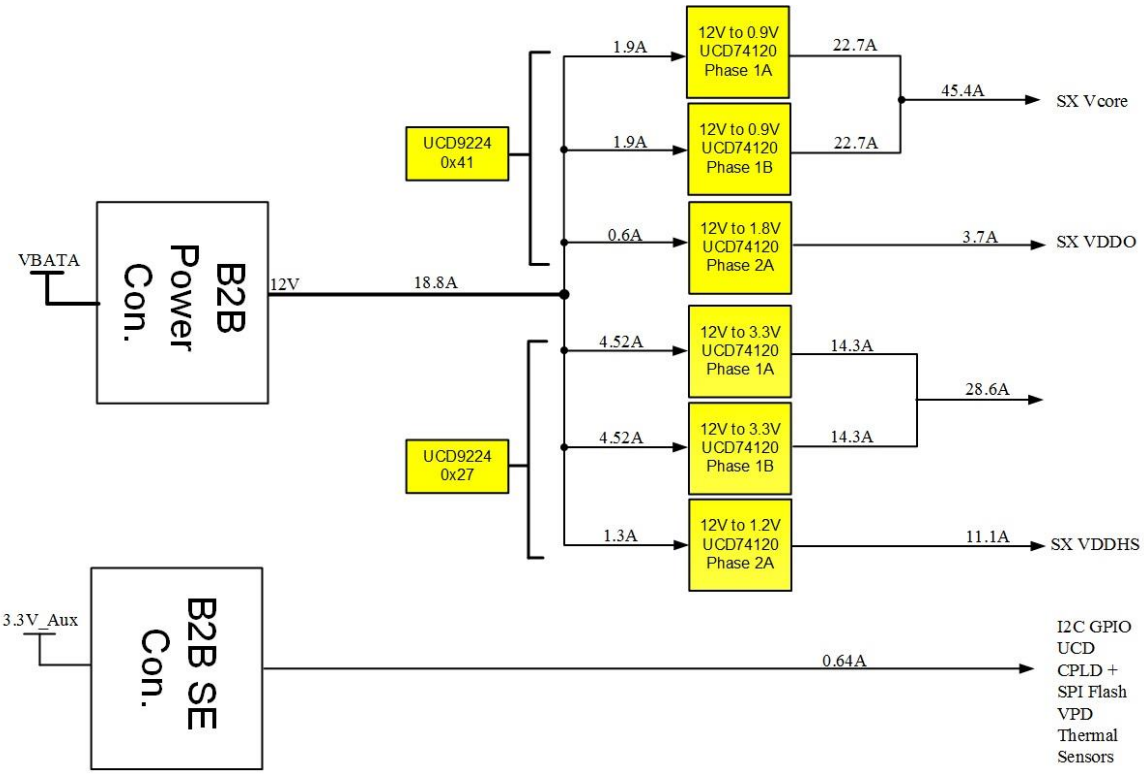

Figure 29: Power Distribution - Switch

5.8.3Power Monitoring – Switch

Power monitoring on SwitchX-2 power rails is done in Port CPLD 1.

Figure 30 specifies Switch power monitoring.

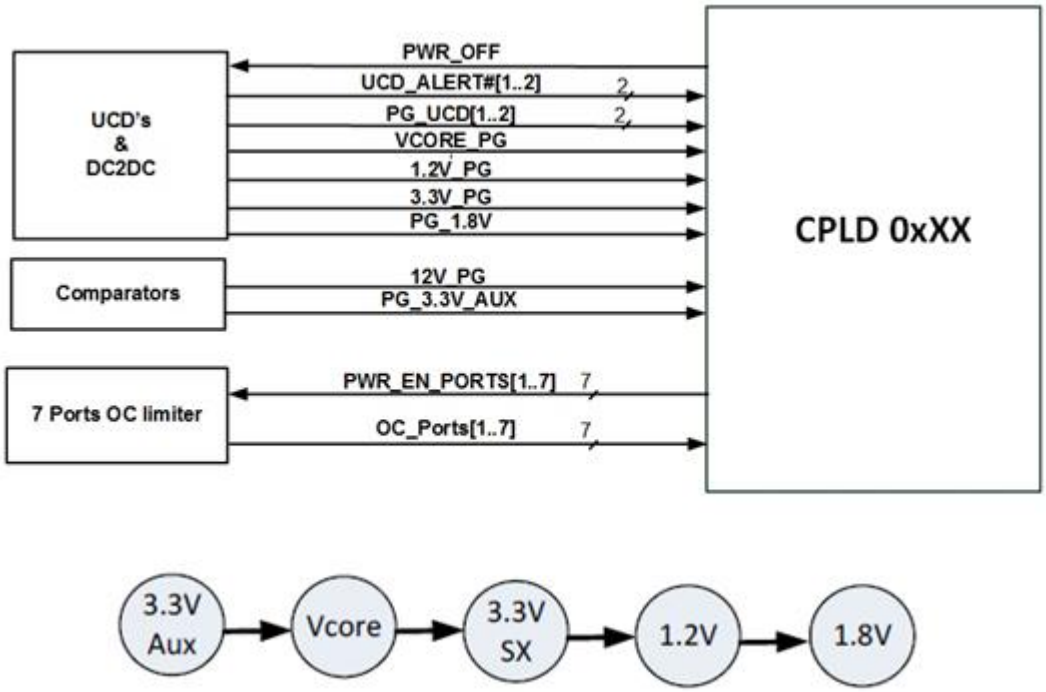

Figure 30: Power Monitoring and Power Sequence- Switch

5.8.4Power Distribution and Monitoring – Management

Figure 31 specifies power distribution on CPU, chipset, and fans.

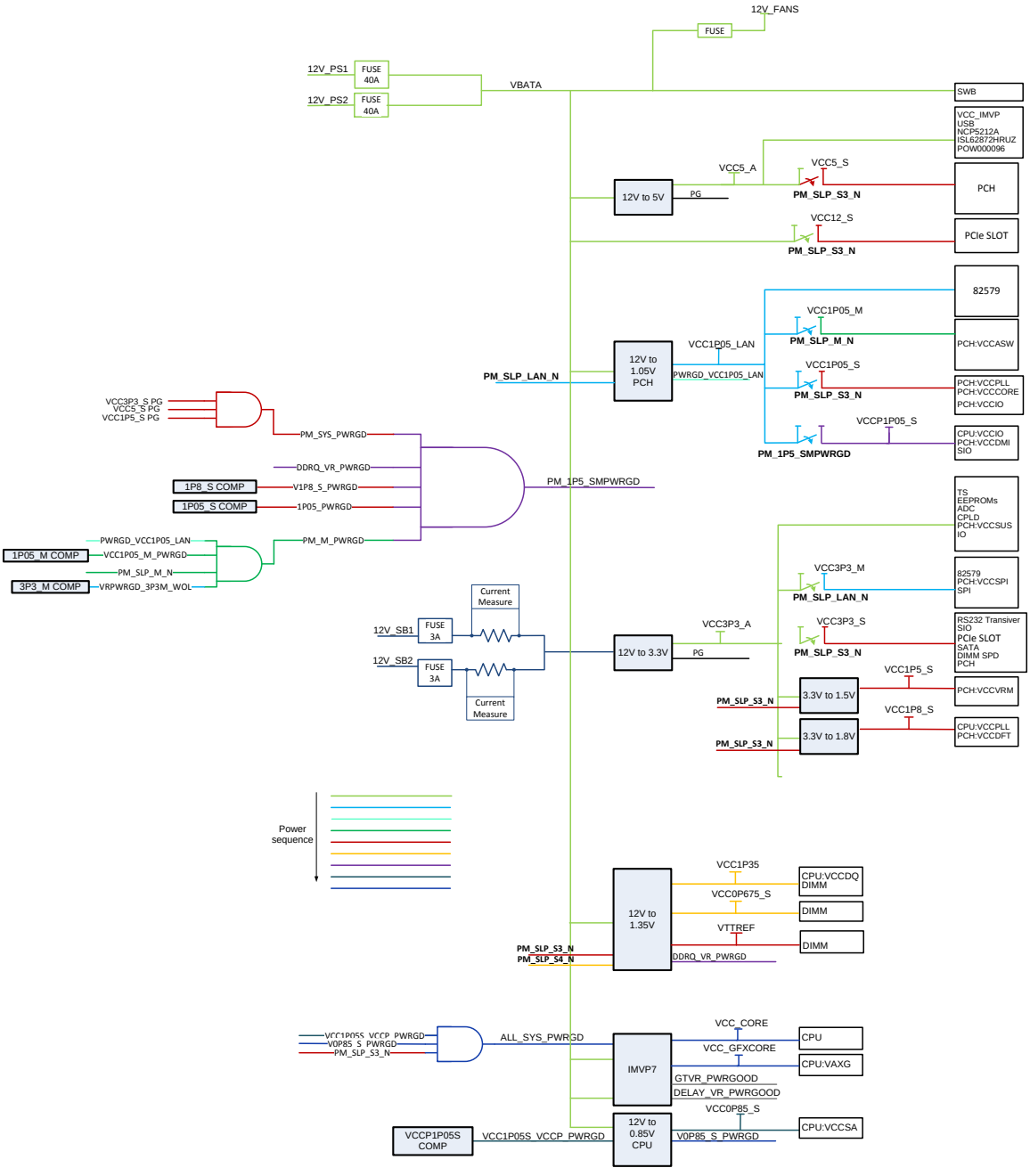

Figure 31: Management Power Distribution and Monitoring

5.8.5System Voltages and Currents Sensing

System voltages and currents are accessible by SW/FW via an 8 channel I²C analog to digital converter. Voltage/current inputs to the A2D channels 2-7 are MUXed by 1:2 MUX. Therefore, two values can be read per channel. The MUX select signal is derived by the CPLD and can be controlled by SW and FW. The main 12V voltages and currents can be read directly from the AC/DC PS. The following table describes the voltage rail at each channel and the expected reading.

Acceptable range for each rail is +-10% from nominal value.

A/D Channel	MUX Select Signal	Monitored Voltage/Current		"B" Parameter
0	NA	DDR3 VTT 0.675V (VCC0P675_S)		0.008
1	NA VCC0P85_S (See Note 1)	***Rail level***	***[VID0,VID1]***	0.008
		0.675	[1,1]	
		0.725	[1,0]	
		0.8	[0,1]	
		0.9	[0,0]	
2	0	12V_AUX1 (12VSB1 current)		0.0586
	1	12V_AUX2 (12VSB2 current)		0.0586
3	0	3.3V (VCC3P3_A)		0.016
	1	5V (VCC5_A)		0.024
4	0	CPU 1.8V (VCC1P8_S)		0.008
	1	---		0.008
5	0	CPU/PCH 1.05V (VCC1P05_LAN)		0.008
	1	CPU 1.5V (VCC1P5_S)		0.008
6	0	CPU 1.05V (VCCP1P05_S)		0.008
	---	---		0.008
7	0	DDR3 1.35V (VCC1P35)		0.008
	---	---		0.008

Table 16: A/D Channels

Notes:

Parameter B is used to calculate the real voltage/current from I²C reading, using the following formula: V=B*R, where V is the real voltage and R is the I²C reading. This voltage is divided by voltage divider before it is sampled; therefore, its parameter B contains the divider.

6 JTAG

The system's MGMT board has standalone JTAG support; in addition, the system allows JTAG support via a special fan FRU, while the top cover is assembled.

In order to connect to the MGMT JTAG chain, a special extension card is required (no on-board connectors).

For lab CPLD burning, a few jumpers might have to be assembled, depending on the board and on the system structure. For CPLD JTAG chain, please refer t o **Error! Reference source not found.**.

The SwitchX chain and the MGMT board chain are described in the following figure.

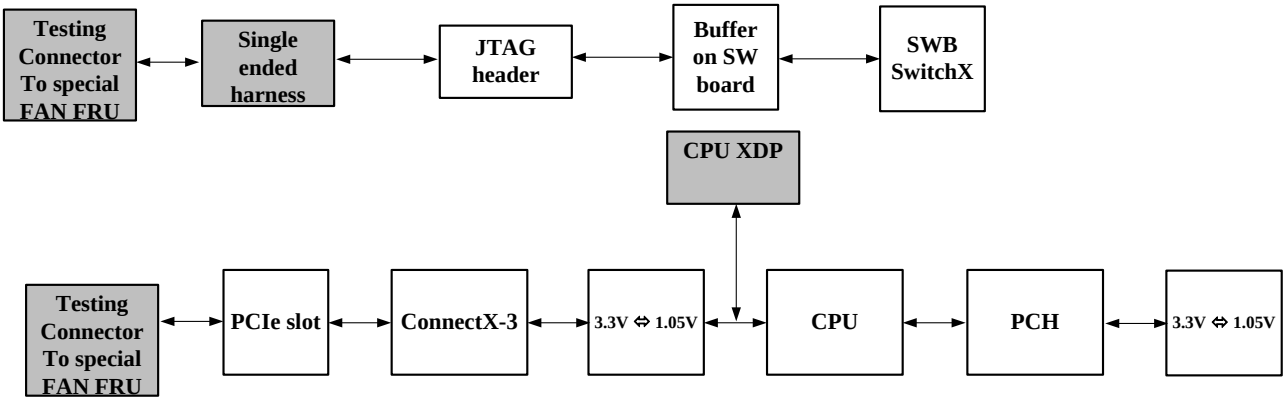

Figure 32: MGMT Board and Switch Board JTAF Chains

6.1 Testing I²C GPIO

Each of the two boards has a dedicated I²C GPIOs for testing. This GPIO is for use only during JTAG tests. The CPLD emulates I²C (using JTAG) to control this GPIO.

The GPIO allows control on critical signals during the JTAG test.

In addition, the I²C GPIO on the MGMT board is connected to the "Safe BIOS" mechanism. This will allow switching between the 2 sets of SPI flash (Know Good Image and Current Image) without the need of a jumper during the burning process.

7 Compliancy

- Shock & Vibration: ETSI EN 300 019-2-2: 1999-09
- Humidity operating: 5% - 95% non-condensing
- Operating temperature 0C-45C
- Safety:
    - US/Canada: cTUVus
    - EU: IEC60950
    - International: CB
    - CCC
- EMC (Emissions):
    - USA: FCC, Class A
    - Canada: ICES, Class A
    - EU: EN55022, Class A
    - EU: EN55024, Class A
    - EU: EN61000-3-2, Class A
    - EU: EN61000-3-3, Class A
    - Japan: VCCI, Class A
- Environmental
    - EU: IEC 60068-2-64: Random Vibration
    - EU: IEC 60068-2-29: Shocks Type I /II
    - EU: IEC 60068-2-32: Fall Test
- Acoustics:
    - ISO 7779
    - ETS 300 753