Facebook - Wedge

16x40GE Top of Rack Switch

OPEN
Compute Project

Author: Yuval Bachar

Date	Revision	Changed by	History
March 01 2015	Rev 0.1	Yuval Bachar	Initial version
March 09 2015	Rev 0.3	Yuval Bachar	OCP release version
October 17 2015	Rev 0.4	Zhiping Yao/Omar Baldonado	Misc cleanup, update license to OCPHW-R for OCP IC review

Table of Content

Table of Figures

Table of Tables

1. Overview

The Wedge platform is a ToR platform with 16xQSFP ports that can be used for any combination of 10/40GE interfaces using breakout cables for 10GE. The Wedge system has the available capacity to expand to 32x40GE ports and can be extended using either a cable or a modified PCB. The 16 unused ports of Trident2 will be in shut down mode during normal mode of operation and cannot be access via any network interface.

2. Licenses

Wedge (ToR) System Description

The Wedge platform is packaged into a 19” 1RU box. Special mechanical adapters are used for the open Rack environment; adapter are available for the open compute rack V1, and the open compute rack V2.

The Wedge Platform has the following building block:

- Main switching board with an integrated (in system pluggable) Micro Server
- Four fans cooling system with 2+2 redundancy
- Supports ports interface cards
    - GE Management ports (front and back)
    - uServer USB
    - BMC direct console
    - Facebook specific debug interface
- Power entry system – AC or DC modules with 1+1 redundancy
- Facebook specific ports in the system rear panel, which are not used for general switching platforms

The following picture shows a high level description of the Wedge configuration.

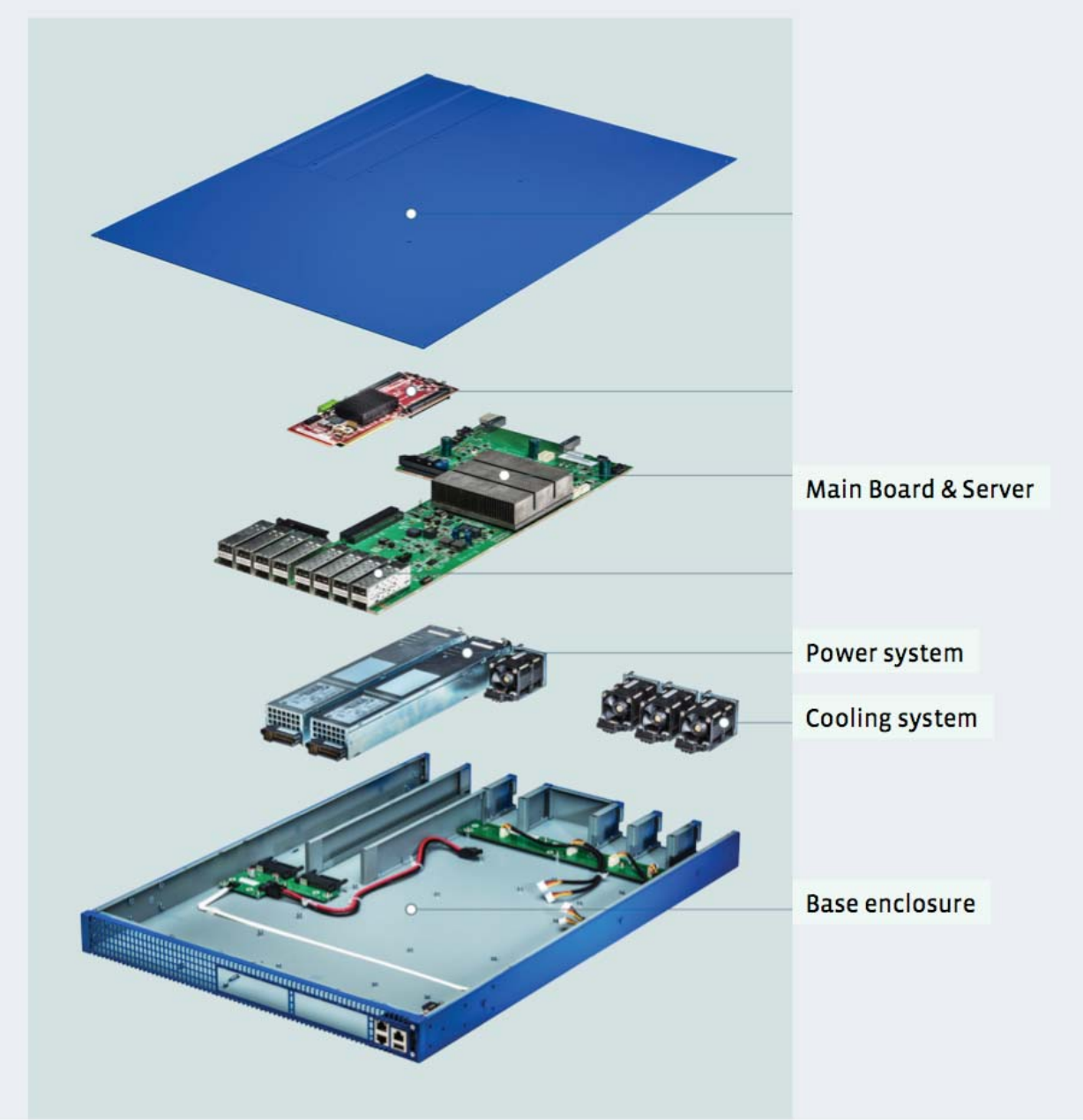


Figure 1: Wedge System Building Blocks

3. *Wedge System Level Connectivity Block Diagram*

The Wedge system is built out of the base building block of 16xQSFP configuration with the following additional modules (boards):

- Fan module – Bridge board to enable fan modules (FRU) termination
- Rear extension module - for Facebook Specific and control ports and GE management port
- Front extension module - Aux Ports front access module – Enables console, GE, USB, and Facebook specific debug ports
- Power adapter module (optional DC or AC) – Power supply termination module

See the board description for details about the baseboard and the additional boards

The following block diagram describes the connectivity of the modules:

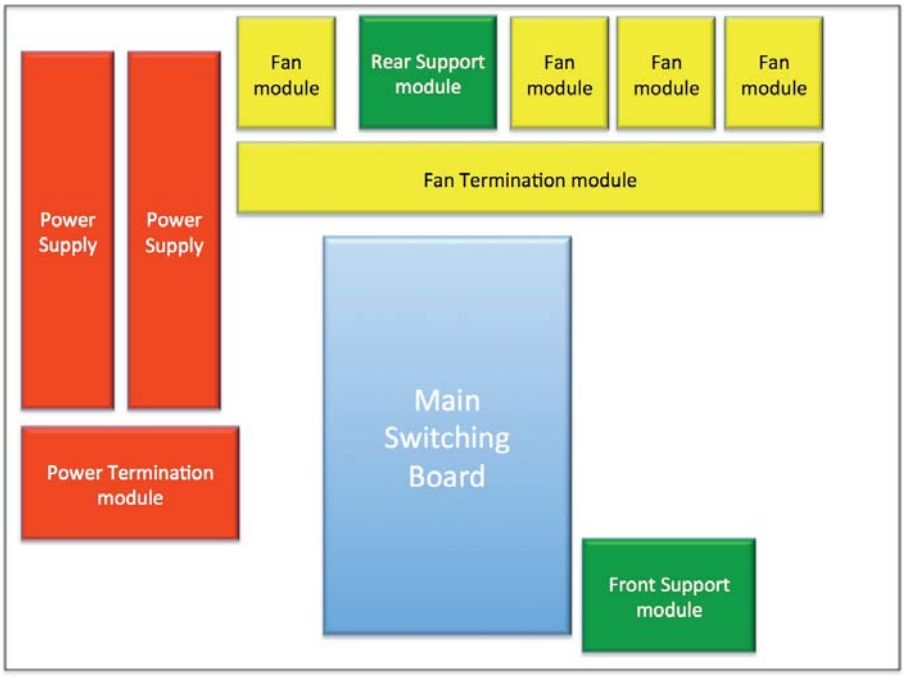


Figure 2: Wedge System Level Block Diagram

4. PCB Modules Specification

The following section describes the different Wedge system boards.

4.1. Main Switch Module - Base 16x40GE/QSFP

The switching modules are the base building blocks for the Wedge system. The switching module is based on the Broadcom Trident2 switching chip combined with Micro Server to handle the control plane and a BMC micro controller to manage all the low level board control as well as function as a console during boot time as well as support the “server like” management.

The switching module can handle 16x40GE worth of bandwidth using 16 QSFP front connectors, each connector can be broken up into four 10GE ports using breakout cable.

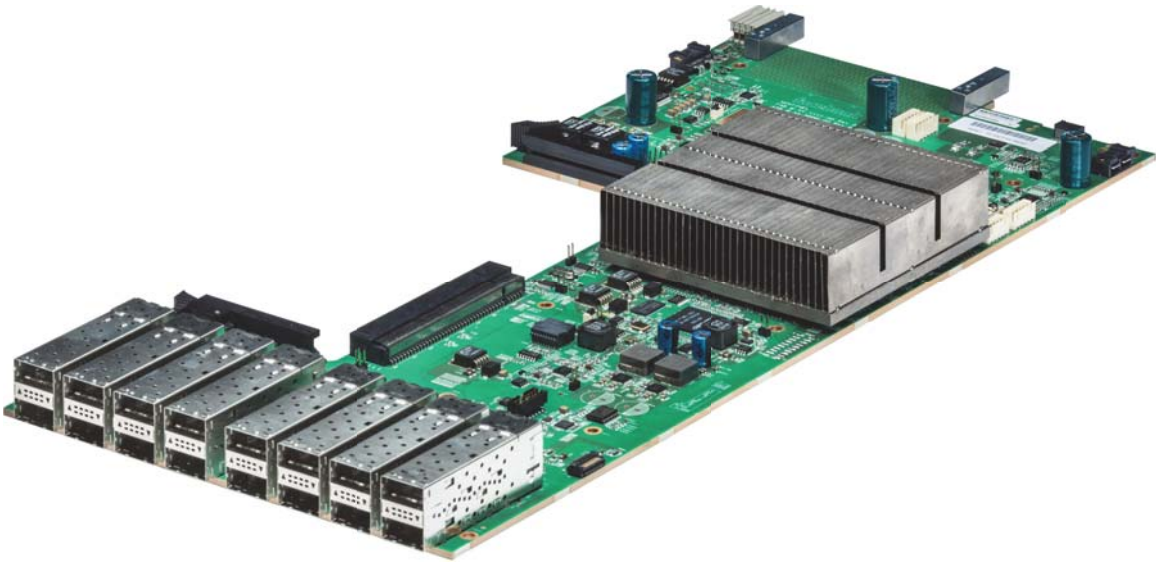


Figure 3: Base 16x40GE Switching Module

4.1. OCP Micro Server

The Wedge system has a pluggable (not in-line FRU) integrated Micro Sever. The Micro Server is the OCP released Micro Server (latest version when writing this document is 0.7). For a full specification of the Micro Server please refer to the Open Compute site under:

Micro Server card version 0.7 http://www.opencompute.org/wiki/Motherboard/SpecsAndDesigns - Micro-server

Note: Refer to the cards full design specification for the specific connectivity of the Micro Server to the rest of the system.

Figure 4: OCP Micro Server

4.2. Fan module

The fan module termination board is a bridge board to enable fan modules (FRU) termination, mechanically and electrically. It connects to the main base card using four cables that provide isolated power feed, control, and monitoring for the four pluggable (FRU) fan modules.

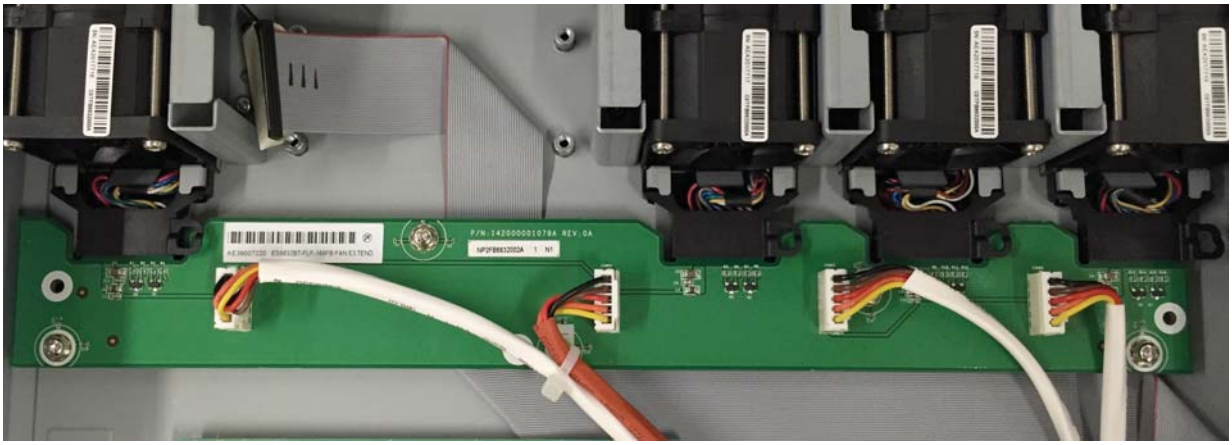


Figure 5: Fan Module

4.3. Rear Extension module

The rear extension module is a module to support specialty Facebook specific monitoring ports (labeled Rack Mon and GPIO) and an external access GE management port. This module is not required for the normal operations of the system. This module connects to the main card via a flat cable that is a mix of high speed and low speed signaling cable.

Figure 6: Rear Extension Module

4.4. Front Extension module

The Front extension module enables the connection to all the front axillary ports. The modules enable access to the BMC console, GE management port , USB port and the Facebook specific debug port. It also enables the LED selection button and LEDs.

This board connects to the main card using a board to board mount connector.

Figure 7: Front Extension Module

4.5. Power adapter module (optional DC or AC)

The power supply termination module function as a mechanical and electrical termination card for the plugin power modules (AC and DC), the module is identical for AC and DC power supplies and function as the main power feed into the main card using a cable to the power entry connector on the main card. The Power adapter also connects to the main card via a flat cable interface for all the management and monitoring interfaces to the power supplies (I^2C).

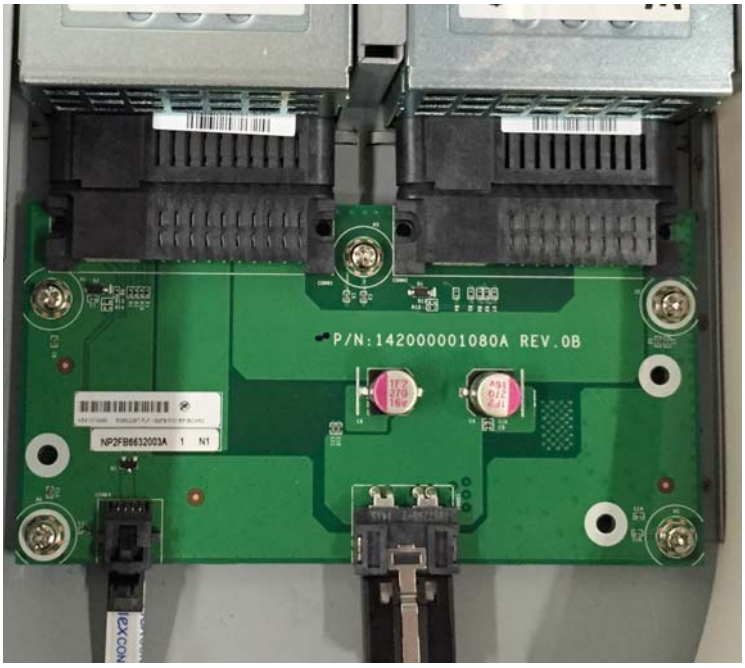


Figure 8: Power Adapter Module

5. Mechanical & Power Distribution

5.1. Wedge ToR System

The Wedge system is enclosed in a 19” rack compatible enclosure. The Wedge system requires a 21” adapter kit as today’s merchant switches to fit into the Open Rack solution. The system dimensions are as follows:

Width:17.5” Depth: 18” Height: 1.75” (1RU)

5.2. Power Distribution

The system has an internal 12v power distribution. The 12v feed will come from the power termination mini-card that terminates the two power supply connectors and enable the BMC management connection to the power supplies. The Mini-Power card is connected to the main card via a power feed cable and a flat cable for the BMC management.

5.3. Front Ports Cabling

5.3.1. Data Path cabling

All data path cabling is based a strip down version of a QSFP connector that will carry the 40GE as well as a 4x10GE option. The 40GE cabling is a 1:1 QSFP to QSFP cable while to enable 10GE connectivity to the data path a 1:4 breakout cable is required. The picture below shows a picture of the 1:4 cable. The plan of record for the in rack communication is a passive copper cable however an optional optical cable can be deployed in the same way using an active break configuration for 10GE

Figure 9: 1:4 Cable solution – Passive Copper

5.3.2. Out of Band management cable

The Wedge system has two RJ45 connectors carrying 1GE management out of band interface. A standard Cat5/6 cable should be connected between the Wedge system and the management centralized switch. The GE ports are on the front and the rear of the system.

Note: Do not connect both ports to the same uplink switch as it will create an undesired loop and potentially disruption of the management interface. It is completely OK to connect the two ports to two separate networks for redundancy. Both ports will be active by default at all times.

5.3.3. Console Cabling

The Wedge system has a front RJ45 connector for the local console port for the BMC processor.

Note: To access the physical Micro Server console port a Facebook Debug card is required, please contact Facebook for additional information. SoL solution is available on all GE ports.

6. Power Specifications

The Wedge system will have two power supplies that can be 42-72v DC and/or 85-264v AC in the same for factor, fully redundant power 1+1. The power feed will connect to the 12v input of the switching main module. Max system power is 275w.

6.1. Power modules

The power supply that will use the are the PowerOne PFE750 family and the PFE400, the system supports DC and AC modules. See the picture below for

the description of the modules:

Figure 10: Wedge Power modules

For more information about the power specification refer to the power specification under:

http://www.power-one.com/power-solutions/products/embedded-power/ac-dc-hot-swap-front-ends/12-volt-output/fnp-300-series/series?sid=169871

7. Detailed Hardware specification

7.1. Introduction

This section covers the detailed hardware specification of the main switching card. The document briefly explains the main switching card architecture and its interfaces. It also discusses the Power, Clock, Reset and I2C architecture of the card

Main switching board is designed to use in standalone mode to provide 16 QSFP port interfaces to provide 16x40GE ports, 64x10GE ports, or any combination in groups of 4. The width of the card allows placing 2 boards within the 19" racks, if needed.

7.2. Main Module Hardware Architecture

The main module hardware is built around the Trident2 device from Broadcom. The Hardware Architecture diagram is given figure given below. The major blocks in the architecture are:

- Networking Switch (Trident 2)
- Board management Controller (BMC)
- Micro Server
- USB hub
- USB to I2C bridge
- FPGA
- QSFP connectors
- Management local GE Switch (not shown in the figure below)

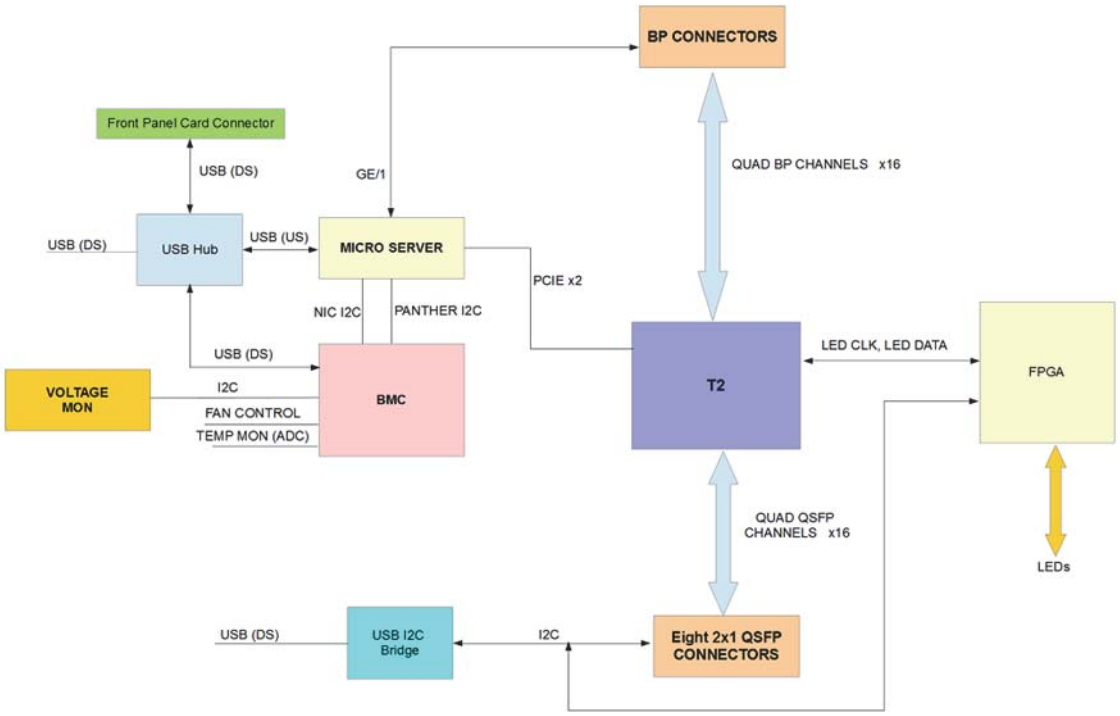


Figure 11: Main Switching Module Hardware Architecture

7.3. Main switching ASIC – Trident2

7.3.1. In Band Networking Switch

Broadcom's BCM56850 (Trident2) is used as the networking switch in this board. BCM56850 supports up to 104x 10GbE or up to 32x 40GbE switch ports. It features a maximum of 32 integrated Warpcores, each with four integrated 10G SerDes Transceivers. Its architecture delivers complete L2/L3 switching and routing capabilities at line rate (refer to line rate limitation in the Trident2 specification) and maximum port density.

A 2 lane Gen3 PCIe interface from the Micro Server is the interface for configuring and monitoring Trident2.

7.3.2. SerDes Interface

The Broadcom SerDes Warpcore is the versatile physical layer interface for the BCM56850, specifically designed to support up to 42Gbps. Each Warpcore consists of four SerDes lanes. Each lane can operate from 1.25Gbps to 10.9375Gbps. We can configure each Warpcore to operate up to a single 40Gbe link. There are 32 SerDes Warpcores available in BCM 56850. That provides a total of 128 links. Half the ports (64 ports) go to 16 QSFP front ports and remaining 64 are not used in the Wedge system.

7.3.3. High Speed Interface Connectivity

The traffic coming through one QSFP port is switched between other QSFP ports in the same board by the Trident2 switching ASIC. The data path diagram below shows the Wedge configuration.

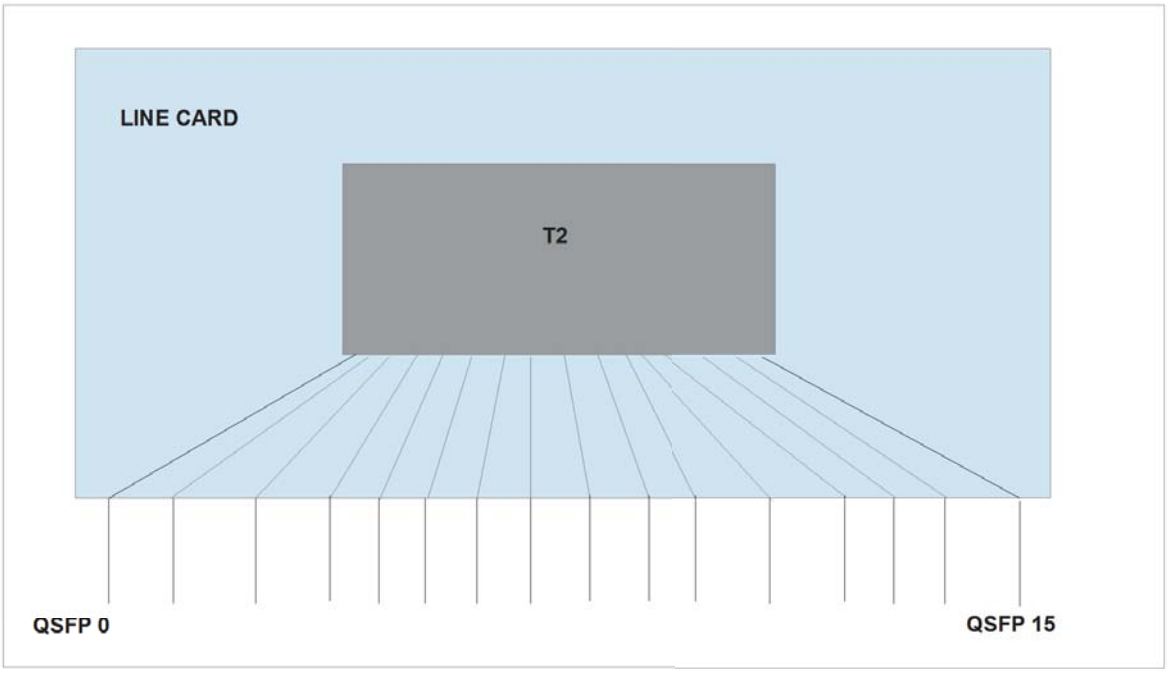


Figure 12: Trident2 to QSFP mapping

Figure 13 shows how QSFP ports are mapped.

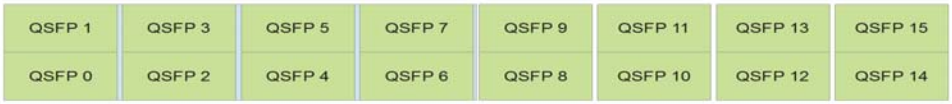


Figure 13: QSPF Ports

Table 1 shows the Trident2 to QSFP port connection.

Sl No.	Trident2 WC port	Trident2 Port/Lane	QSFP PORT	Port	QSFP Lane
1	8	TD0	0	Bottom	TX1
2	8	TD1	0	Bottom	TX4
3	8	TD2	0	Bottom	TX3
4	8	TD3	0	Bottom	TX2
5	8	RD0	0	Bottom	RX3
6	8	RD1	0	Bottom	RX2
7	8	RD2	0	Bottom	RX1
8	8	RD3	0	Bottom	RX4
9	9	TD0	1	Top	TX1
10	9	TD1	1	Top	TX2
11	9	TD2	1	Top	TX3
12	9	TD3	1	Top	TX4
13	9	RD0	1	Top	RX3
14	9	RD1	1	Top	RX2
15	9	RD2	1	Top	RX1
16	9	RD3	1	Top	RX4
17	10	TD0	2	Bottom	TX1
18	10	TD1	2	Bottom	TX4
19	10	TD2	2	Bottom	TX3
20	10	TD3	2	Bottom	TX2
21	10	RD0	2	Bottom	RX3
22	10	RD1	2	Bottom	RX2
23	10	RD2	2	Bottom	RX1
24	10	RD3	2	Bottom	RX4

25	11	TD0	3	Top	TX1
26	11	TD1	3	Top	TX2
27	11	TD2	3	Top	TX3
28	11	TD3	3	Top	TX4
29	11	RD0	3	Top	RX3
30	11	RD1	3	Top	RX2
31	11	RD2	3	Top	RX1
32	11	RD3	3	Top	RX4
33	12	TD0	4	Bottom	TX1
34	12	TD1	4	Bottom	TX4
35	12	TD2	4	Bottom	TX3
36	12	TD3	4	Bottom	TX2
37	12	RD0	4	Bottom	RX3
38	12	RD1	4	Bottom	RX2
39	12	RD2	4	Bottom	RX1
40	12	RD3	4	Bottom	RX4
41	13	TD0	5	Top	TX1
42	13	TD1	5	Top	TX2
43	13	TD2	5	Top	TX3
44	13	TD3	5	Top	TX4
45	13	RD0	5	Top	RX3
46	13	RD1	5	Top	RX2
47	13	RD2	5	Top	RX1
48	13	RD3	5	Top	RX4
49	14	TD0	6	Bottom	TX1
50	14	TD1	6	Bottom	TX4
51	14	TD2	6	Bottom	TX3
52	14	TD3	6	Bottom	TX2
53	14	RD0	6	Bottom	RX3
54	14	RD1	6	Bottom	RX2
55	14	RD2	6	Bottom	RX1
56	14	RD3	6	Bottom	RX4

57	15	TD0	7	Top	TX1
58	15	TD1	7	Top	TX2
59	15	TD2	7	Top	TX3
60	15	TD3	7	Top	TX4
61	15	RD0	7	Top	RX3
62	15	RD1	7	Top	RX2
63	15	RD2	7	Top	RX1
64	15	RD3	7	Top	RX4
65	16	TD0	8	Bottom	TX1
66	16	TD1	8	Bottom	TX2
67	16	TD2	8	Bottom	TX3
68	16	TD3	8	Bottom	TX4
69	16	RD0	8	Bottom	RX3
70	16	RD1	8	Bottom	RX2
71	16	RD2	8	Bottom	RX1
72	16	RD3	8	Bottom	RX4
73	17	TD0	9	Top	TX1
74	17	TD1	9	Top	TX2
75	17	TD2	9	Top	TX3
76	17	TD3	9	Top	TX4
77	17	RD0	9	Top	RX3
78	17	RD1	9	Top	RX2
79	17	RD2	9	Top	RX1
80	17	RD3	9	Top	RX4
81	18	TD0	10	Bottom	TX1
82	18	TD1	10	Bottom	TX4
83	18	TD2	10	Bottom	TX3
84	18	TD3	10	Bottom	TX2
85	18	RD0	10	Bottom	RX3
86	18	RD1	10	Bottom	RX2
87	18	RD2	10	Bottom	RX1
88	18	RD3	10	Bottom	RX4

89	19	TD0	11	Top	TX1
90	19	TD1	11	Top	TX2
91	19	TD2	11	Top	TX3
92	19	TD3	11	Top	TX4
93	19	RD0	11	Top	RX3
94	19	RD1	11	Top	RX2
95	19	RD2	11	Top	RX1
96	19	RD3	11	Top	RX4
97	20	TD0	12	Bottom	TX1
98	20	TD1	12	Bottom	TX4
99	20	TD2	12	Bottom	TX3
100	20	TD3	12	Bottom	TX2
101	20	RD0	12	Bottom	RX3
102	20	RD1	12	Bottom	RX2
103	20	RD2	12	Bottom	RX1
104	20	RD3	12	Bottom	RX4
105	21	TD0	13	Top	TX1
106	21	TD1	13	Top	TX2
107	21	TD2	13	Top	TX3
108	21	TD3	13	Top	TX4
109	21	RD0	13	Top	RX3
110	21	RD1	13	Top	RX2
111	21	RD2	13	Top	RX1
112	21	RD3	13	Top	RX4
113	22	TD0	14	Bottom	TX1
114	22	TD1	14	Bottom	TX4
115	22	TD2	14	Bottom	TX3
116	22	TD3	14	Bottom	TX2
117	22	RD0	14	Bottom	RX3
118	22	RD1	14	Bottom	RX2
119	22	RD2	14	Bottom	RX1
120	22	RD3	14	Bottom	RX4

121	23	TD0	15	Top	TX1
122	23	TD1	15	Top	TX2
123	23	TD2	15	Top	TX3
124	23	TD3	15	Top	TX4
125	23	RD0	15	Top	RX3
126	23	RD1	15	Top	RX2
127	23	RD2	15	Top	RX1
128	23	RD3	15	Top	RX4

Table 1: Trident2 to QSFP port mapping in Main switching

7.4. Control Interface

The Trident2 (BCM56850) has 2 lane (x2) Gen2 PCIe bus as the control Interface. It is directly interfaced with Micro Server PCIe bus.

7.5. Trident2 STRAPPING OPTIONS

The following table describes the Trident2 optional strapping configuration that sets the base startup configuration for the chip.

SIGNAL NAME	Trident2 PIN NO	In MAIN SWITCHING	Datasheet Description
GLOBAL_DISABLE*	V37	NOT CONNECTED	DO NOT CONNECT THIS PIN
DIRECT_SCAN	AR32	NOT CONNECTED	DO NOT CONNECT THIS PIN
DFT_GATING	AT32	NOT CONNECTED	DO NOT CONNECT THIS PIN
XGXS_SCAN_MODE	W37	NOT CONNECTED	DO NOT CONNECT THIS PIN
XGXS_TEST_EN	R37	NOT CONNECTED	DO NOT CONNECT THIS PIN
SERDES_TEST_EN	AY47	NOT CONNECTED	DO NOT CONNECT THIS PIN
EEPROM_LOAD_EN	N32	NOT CONNECTED	0:LOAD NOT ENABLE; 1:LOAD ENABLE
MONITOR	AG21	NOT CONNECTED	DO NOT CONNECT THIS PIN
BSC_DEBUG_MODE	AU27	NOT CONNECTED	0:PCIe IS ENABLED
CMIM_BSC_DEBUG_MODE	BD49	NOT CONNECTED	DO NOT CONNECT THIS PIN

TS_PLL_BYP	BF46	NOT CONNECTED	DO NOT CONNECT THIS PIN
PCIE_MODE	N35	NOT CONNECTED	CAN BE LEFT AS NO CONNECT OR TIED DIRECTLY TO GROUND
MCS_PLL_BYP	BD46	NOT CONNECTED	DO NOT CONNECT THIS PIN
XGXS_PLL_BYP	V36	NOT CONNECTED	DO NOT CONNECT THIS PIN
CORE_PLL_BYP	BE46	NOT CONNECTED	CAN BE LEFT AS NO CONNECT OR TIED DIRECTLY TO GROUND
EXT_UC_PRESENT	N34	EXTERNALLY PULLED DOWN TO GROUND THROUGH 4.7 RESISTOR	INDICATES THAT EXTERNAL U-CONTROLLER IS PRESENT
EXT_UC_IS_SPI	BC49	EXTERNALLY PULLED DOWN TO GROUND THROUGH 4.7 RESISTOR	INDICATES THAT THE EXTERNAL U-CONTROLLER PRESENT IS SPI
SPI_CHIPID2	BC48	NOT CONNECTED	CHIP ID TO SPI SLAVE
SPI_CHIPID1	BC47	NOT CONNECTED	CHIP ID TO SPI SLAVE
SPI_CHIPID2	BC46	NOT CONNECTED	CHIP ID TO SPI SLAVE
MDIO_22_SEL	Y36	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	WHEN PULLED HIGH ,THE MIIM INTERFACE SIGNALLING IS @2.5V OR 3.3V AND THE CL22 ELECTRICAL SIGNALLING IS IN USE .THE MDIO_VDDO PINS CAN BE DRIVEN WITH A 2.5 V OR 3.3V SUPPLY
JTAG_TCE	U36	EXTERNALLY PULLED DOWN TO GROUND	JTAG TEST ENABLE.MUST BE PULLED LOW DURING NORMAL

		THROUGH 4.7 RESISTOR	SWITCHING OPERATION
JTAG_TDI	T36	NOT CONNECTED	JTAG TEST DATA IN
JTAG_TDO	Y37	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	JTAG TEST DATA OUT
JTAG_TMS	T37	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	JTAG TEST MODE SELECT
JTAG_TCK	U37	NOT CONNECTED	JTAG TEST CLOCK
JTAG_TRST*	W36	EXTERNALLY PULLED DOWN TO GROUND THROUGH 4.7K RESISTOR	JTAG TEST RESET MUST BE PULLED LOW DURING NORMAL OPERATION

Table 2: Trident2 Strapping

7.6. Board Management Controller (BMC)

The Wedge System includes the AST1250, an Integrated Remote Management Processor from Aspeed Technologies is used as the Board Management Controller (BMC) for main switching board and all of the support mini-cards and sub-systems. The major functions of BMC are:

- Trident2 and Micro Server power management
- Temperature monitoring
- Voltage monitoring
- Fan control
- Reset Control
- Programming FPGA flash
- Read the Rx loss signals from the QSFP ports
- Micro Server boot up status
- ROV voltage and Board revision ID.
- I2C interfaces to Micro Server, USB, and Hot swap controller, temperature sensors, and voltage controllers.
- Monitoring detect signals
- GE direct access

- External Facebook specific debug port control and monitoring
- Facebook specific ports interface (RackMon)

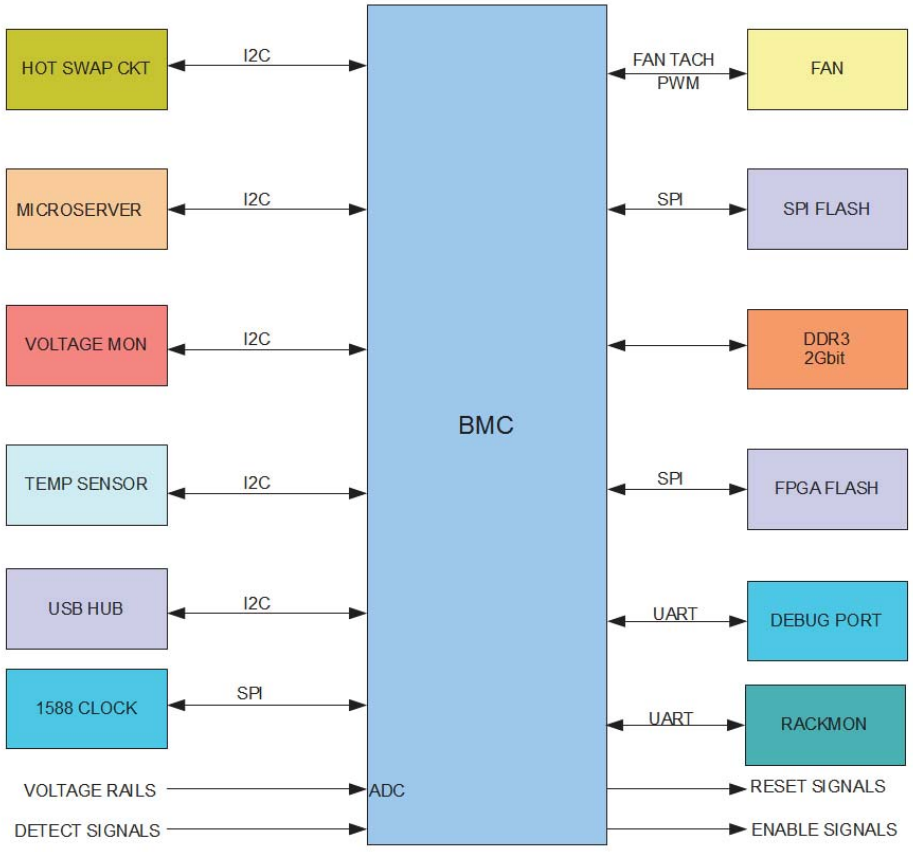


Figure 14: BMC block diagram

The above block diagram shows the various functionalities of BMC.

- BMC is configured to boot from 32 bit SPI flash by setting hardware strapping pins ROMA[3:2] =1:0.
- DDR3 memory used is K4B2G1646Q-BCK0 (96ball FBGA –x16), which is 2Gbit and having speed 1600.
- BMC can control the speed of FANS in the system through the PWM outputs. It can also read the FAN speed using the TACH inputs.
- UART signals from BMC is routed to the Facebook specific debug port (front panel).
- BMC monitors the Voltages rails in the board using I2C interface to a voltage monitor IC (MAX127)
- BMC can power cycle the board using the I2C interface of Hot Swap controller (ADM1278)
- I2C interface of BMC to various devices are explained in the I2C section.
- BMC also monitor voltages (5V, 3.3V, 2.5V, 1V_Analag, 1V_core) connected to its ADC.
- To reprogram FPGA flash, BMC will make FL_PRG_SEL high, with this the switches U14 and U15 connected to FPGA Flash will get connected to BMC. Then BMC can write into FPGA flash.

- The BMC processor has a direct GE port interface with MDC/MDIO it is using the RGMII interface to a PHY chip that connect to the on board local unmanaged GE switch
- Rack Monitor (Facebook specific) can read health of board in rack through BMC

7.6.1. BMC Hardware Strapping

The hardware strapping function is used default operation of BMC before the software runs. The hardware strapping table is given below

SIGNAL NAME	AST1250 IN MAIN SWITCHING	INTERNAL PULL UP/PULL DOWN	Description
BMC_JTAG_TRST_N	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	Reserved
BMC_UART_TX_5	DIRECTLY CONNECTED TO U25	INTERNAL PULL UP	Disable ICE reset mode.
SPI_IBMC_BT_CLK_R	NOT CONNECTED	INTERNAL PULL DOWN	010:DDR3 SDRAM configuration with CL=6,CWL=5
SPI_IBMC_BT_DO_R	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	
SPI_IBMC_BT_DI	NOT CONNECTED	INTERNAL PULL DOWN	
SPI_IBMC_BT_D2	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	01:128MByte DRAM size setting for VGA use
SAS_FLT_0_R	NOT CONNECTED	INTERNAL PULL DOWN	
SAS_FLT_1_R	NOT CONNECTED	INTERNAL PULL DOWN	Software designed strapping registers.
SAS_FLT_2_R	NOT CONNECTED	INTERNAL PULL DOWN	
ROMD7	NOT CONNECTED	INTERNAL PULL DOWN	
ROMA0	NOT CONNECTED	INTERNAL PULL	10:Boot from SPI flash memory

		DOWN	
ROMA1	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	
ROMA2	NOT CONNECTED	INTERNAL PULL DOWN	
ROMA3	NOT CONNECTED	INTERNAL PULL DOWN	00: Select 8mb VGA memory
ROMA4	NOT CONNECTED	INTERNAL PULL DOWN	0:select 8bit interface for NOR type flash 0:boot with 24bit address mode for SPI type flash
ROMA5	NOT CONNECTED	INTERNAL PULL DOWN	0:No VGA BIOS ROM for on-board applications
ROMA6	NOT CONNECTED	INTERNAL PULL DOWN	0:MAC1 interface RMII/NCSI
ROMA7	NOT CONNECTED	INTERNAL PULL DOWN	0:MAC2 interface RMII/NCSI
ROMA8	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	
ROMA9	NOT CONNECTED	INTERNAL PULL DOWN	01:H-PLL clock frequency 360Mhz
ROMA10	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	
ROMA11	NOT CONNECTED	INTERNAL PULL DOWN	01:CPU/AHB clock frequency ratio =2:1
ROMA12	NOT CONNECTED	INTERNAL PULL DOWN	
ROMA13	NOT CONNECTED	INTERNAL PULL DOWN	00:disable SPI interface
ROMA14	NOT CONNECTED	INTERNAL PULL DOWN	0:LPC reset is shared withPCI reset

ROMA15	NOT CONNECTED	INTERNAL PULL DOWN	0:select the class code for video device
ROMA16	NOT CONNECTED	INTERNAL PULL DOWN	0:superIO configuration address selection(decode0x2E)
ROMA17	NOT CONNECTED	INTERNAL PULL DOWN	0:Disable BMC 2nd watch dog timer
ROMA18	NOT CONNECTED	INTERNAL PULL DOWN	[ROMA23,ROMA18]00: 24 mhz feed clkin
ROMA19	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	1:Disable ACPI function
ROMA20	NOT CONNECTED	INTERNAL PULL DOWN	0:Enable LPC to decode superIO 0x2E/0x4E address
ROMA21	EXTERNALLY PULLED HIGH TO 3.3V THROUGH 4.7K RESISTOR	INTERNAL PULL DOWN	1:Enable GPIOD pass through mode
ROMA22	NOT CONNECTED	INTERNAL PULL DOWN	0:Disable GPIOE pass through mode
ROMA23	NOT CONNECTED	INTERNAL PULL DOWN	0 : clkin is 24/48 mhz

Table 3: System BMC (AST2400) Strapping

7.7. Microserver

A micro-server is a PCI-like card that hosts the system management SOC, dynamic memory modules for the SOC (DDR) , and a storage device (mSATA). The main switching board has a Micro Server slot in it, to which we can plug-in a Micro Server card.

Note: Please not the Micro Server is not an online FRU and requires opening the box to be serviced or replaced.

7.7.1. MicroServer PCIE

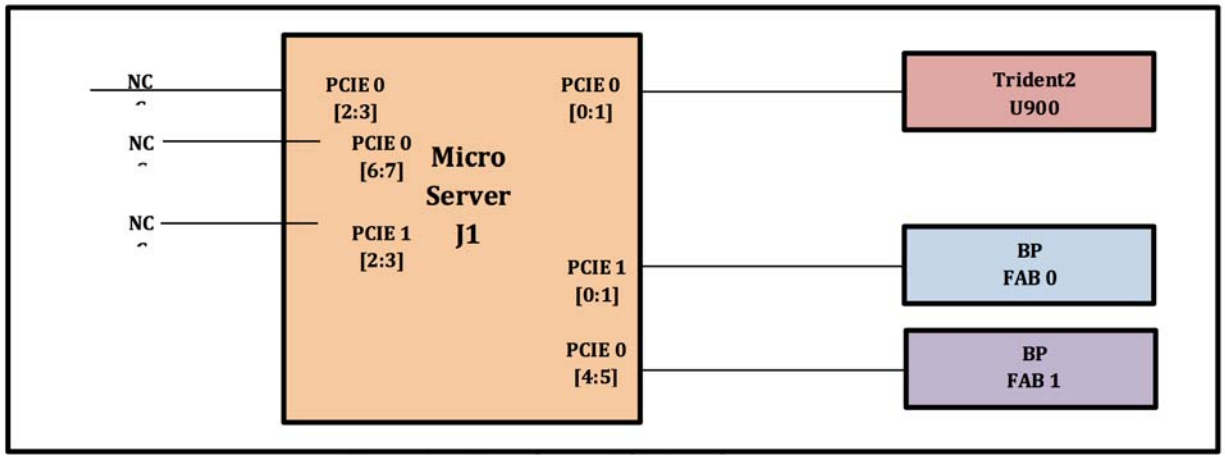


Figure 15: Micro Server PCIe Interface

One Micro Server PCIE x1 ports is routed to Trident2 and two other Micro Server PCIE x1 ports are routed to FABRIC 0 and FABRIC 1 respectively through the backplane connectors.

7.7.2. Gigabit Ethernet

Micro Server has a Gigabit Ethernet port. In the Wedge system the Micro Server GE port is routed to the local GE switch that is shared between the Micro Sever the BMC and the external ports (front and back panels).

7.8. USB HUB

In Main switching we are using USB2513B-AEZG from Microchip, which is having 3 downstream ports. Here the downstream ports are connected to the USB I2C bridge (U70), BMC (U1) and front port JP7 respectively. The upstream USB interface is connected to the USB port of Microserver.

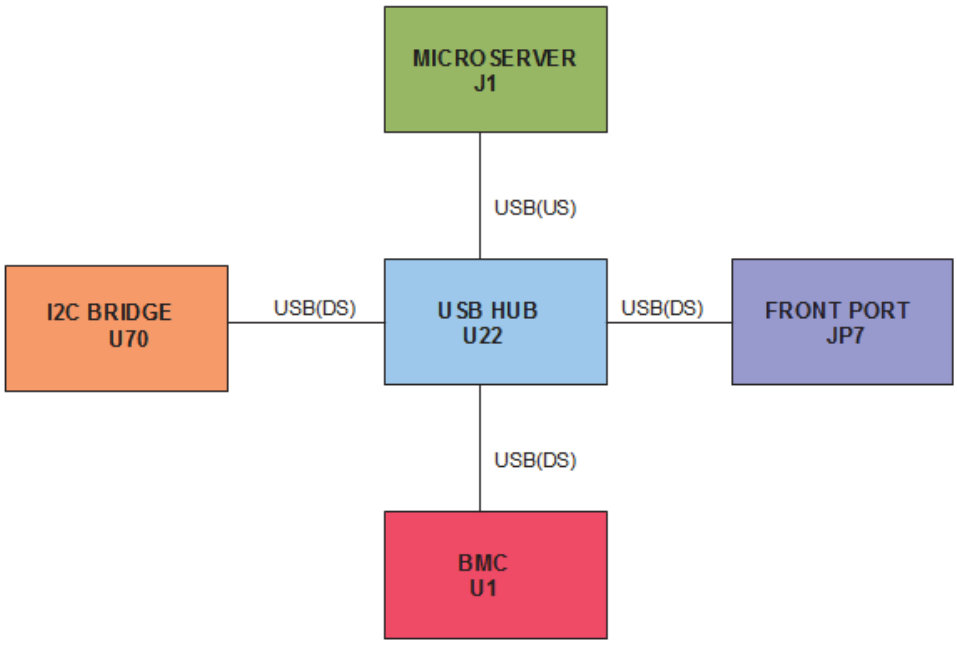


Figure 16:USB HUB

7.9. USB- I2C BRIDGE

Micro Server needs interface with QSFP connectors to monitor the status of each connector. It is done using the USB-I2C bridge CP2112. The USB hub is connected to U70 (CP2112) through one of its downstream port. The Micro Server can read the QSFPs through the USB hub and I2C Bridge.

I2C bus from USB Bridge is also connected to the FPGA. Micro Server can write into an 8bit register in FPGA through I2C to indicate which QSFP port is configured for 10G /40G.

Update: Change in the Trident2 LED stream will eliminate the need to write to the register.

7.10. Local GE Switch

The Wedge platform have a local GE switch that is connected to the Micro Server, the BMC, the Front and Back connectors, and has two spare unused ports.

All Out Of Band management traffic is rounded through the GE switch with basic L2 learning and forward via the GE switch.

Note: Do not connect both ports to the same uplink switch as it will create an undesired loop and potentially disruption of the management interface. It is completely OK to connect the two ports to two separate networks for redundancy. Both ports will be active by default at all times.

7.11. VOLTAGE MONITOR

In Main switching board MAX127 from Maxim is used as the voltage monitor. It has eight analog input channels. Below table shows the voltages monitored by U3304.

Channel	Voltage	Value	Tolerance

CH0	VDD CORE	1V	+/-2%
CH1	VDD2_5V	2.5V	+/-5%
CH2	Not Used	Not Used	Not Used
CH3	VDD_ANLG	1.03V	+/-2%
CH4	Not Used	Not Used	Not Used
CH5	VDD3_3V	3.3V	+/-5%
CH6	VDD5V	5V	+/-5%
CH7	Not Used	Not Used	Not Used

Table 4: Voltage rails monitored

7.12. FPGA

7.12.1. Description

We are using XC3S50AFT256, which belongs to Spartan-3A family from Xilinx. It is a 256 pin BGA.

The functions of FPGA are

- Convert the serial LED data into parallel and drive the LEDs to show the status on 10G and 40G ports.
- Act as an I2C slave to USB-I2C Bridge. BMC can write into an eight bit FPGA register through I2C to configure each QSFP port into 10G or 40G ports.

Update: Change in the Trident2 LED stream will eliminate the need to write to the register.

- Read the ROV voltage from Trident2 and drive corresponding VID codes to the Voltage controller.

FPGA Configuration mode set in Main switching board is master SPI mode. This is done by hardware strapping (M[2:0] = <0:0:1>). The size of FPGA Flash is 1Mega bit. JTAG header can be used to program the FPGA or upgrade the configurations stored in the EEPROM .

7.13. QSFP LED

7.13.1. Block Diagram

Below figure shows the LED arrangement for QSFP PORT 0 /1

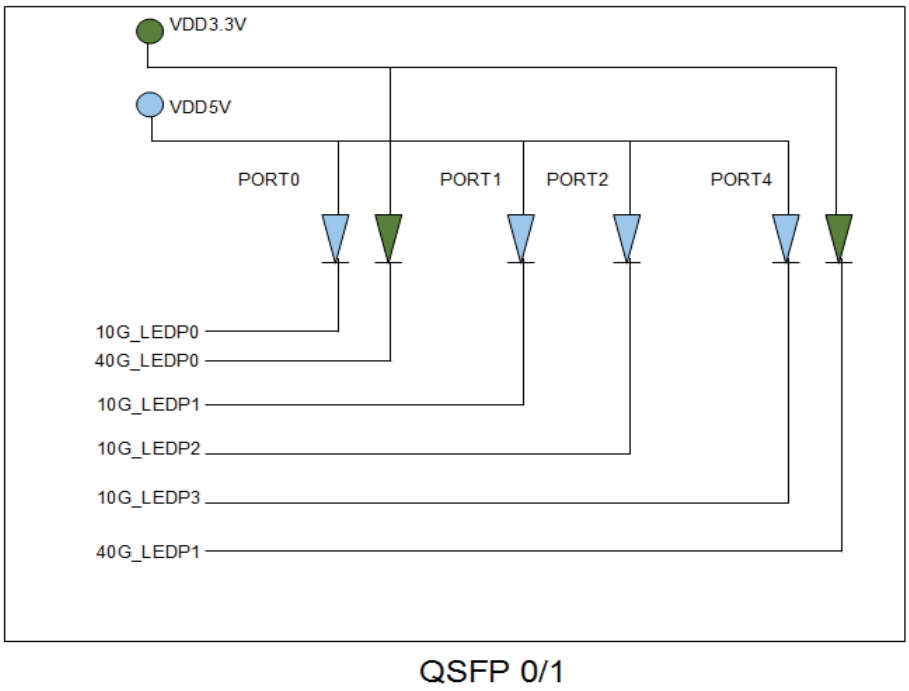


Figure 17: QSFP LED

7.13.2. Description

The FPGA drives the 48 LEDs after converting the serial LED data into parallel to indicate the line conditions on the QSFP ports. Corresponding to two QSFP ports there are two bicolor LEDs and two single colour Blue LEDs. The Blue LEDs are connected to 5V and green LEDs are connected to 3.3V.

7.13.3. ON Panel LED function

Each connector in the Wedge and Six Pack can accommodate 2 QSFP ports; the ports can be configured to be 4x10G or 1x40G.

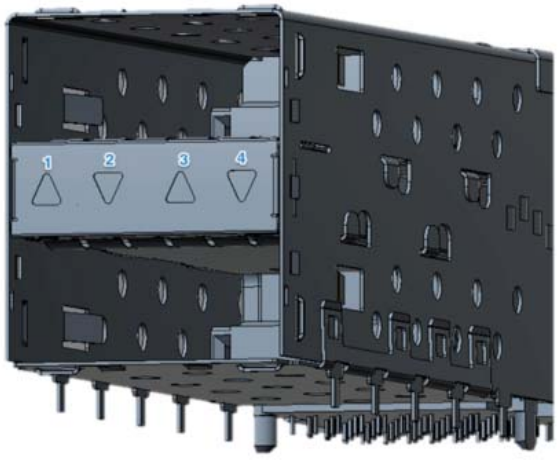


Figure 18: QSFP LED configuration

The System indication will be a combination of color-coded LED light combined with LED state.

If a port is set to be in 4x10G port, all 4 LEDs (1-4) will be active if a port is set to be in 40G mode only the LEDs on the edge (1 and 4) will be active.

Each Wedge **system** will have section push button that will control if the LEDs reflect the top or bottom port and associated LEDs to indicate if the LEDs viewed are top or bottom.

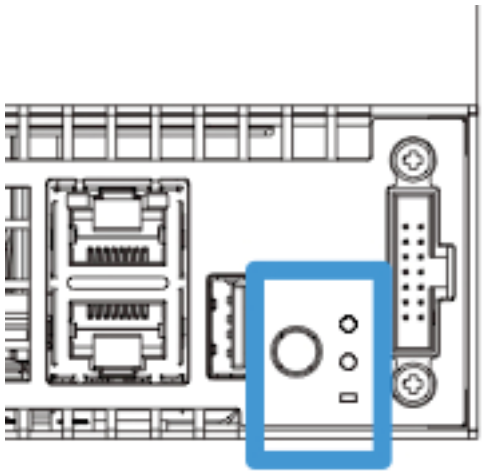


Figure 19: Up/Down Push Button

The following cases are supported:

7.13.4. Uniform Ports LED Configuration

The following tables describe the functionality of the LEDs in case of 40/10G or a mix of the two speeds in one cage. Please note that that 10G LEDs in the middle will be disabled when operating in 40G mode.

The Selection button will be functional when either both ports have 10G ports associated with it or in case that one port will have 10G and the other 40G. The selection push button will not have any function when both ports in a cage are 40G and the Green LEDs will be on at all time reflecting the top and bottom parts of the cage.

Switch State	Two ports 40G	Two ports 10G
Top	LED 1 – Reflects top QSFP 40G port - Green LED 2,3 – OFF Optional – to enable LED 4 to also reflect concurrently	LED 1,2,3,4 – Reflects the 4x10G sub-ports on the top QSFP - Blue
Bottom	LED 4 – Reflects top 40G port - Green LED 2,3 - OFF	LED 1,2,3,4 – Reflects the 4x10G sub-ports on the top QSFP - Blue

	Optional – to enable LED 1 to also reflect concurrently	

Table 5: Uniform Ports LED Configuration

7.13.5. Mixed Ports LED Configuration

Switch State	Top QSFP 40G Bottom QSFP 4x10G	Top QSFP 4x10G Bottom QSFP 40G
Top	LED 1 – Reflects top QSFP 40G port - Green LED 2,3,4 - OFF	LED 1,2,3,4 – Reflects the 4x10G sub-ports on the top QSFP - Blue
Bottom	LED 1,2,3,4 – Reflects the 4x10G sub-ports on the bottom QSFP - Blue	LED 4 – Reflects bottom QSFP 40G port - Green LED 1,2,3 - OFF

Table 6: Mixed Ports LED Configuration

7.13.6. Miscellaneous LEDS

BMC drives the Heart beat LED output to indicate whether the system is working fine. Heart bead LED will glow when when BMC_HEARTBEAT_N is low.
BMC drives the Power LED blue ON by making LED_PWR_BLUE signal high to indicate that the BMC power is good.
The BMC also drives the Fan Module LEDs indicating fault condition in any fan module, the fan module LEDs are on the rear support module exposed to the rear view of the system.

7.13.7. Communication ports and Debug ports

The BMC support multiple debug and support port like describe in the figure below:

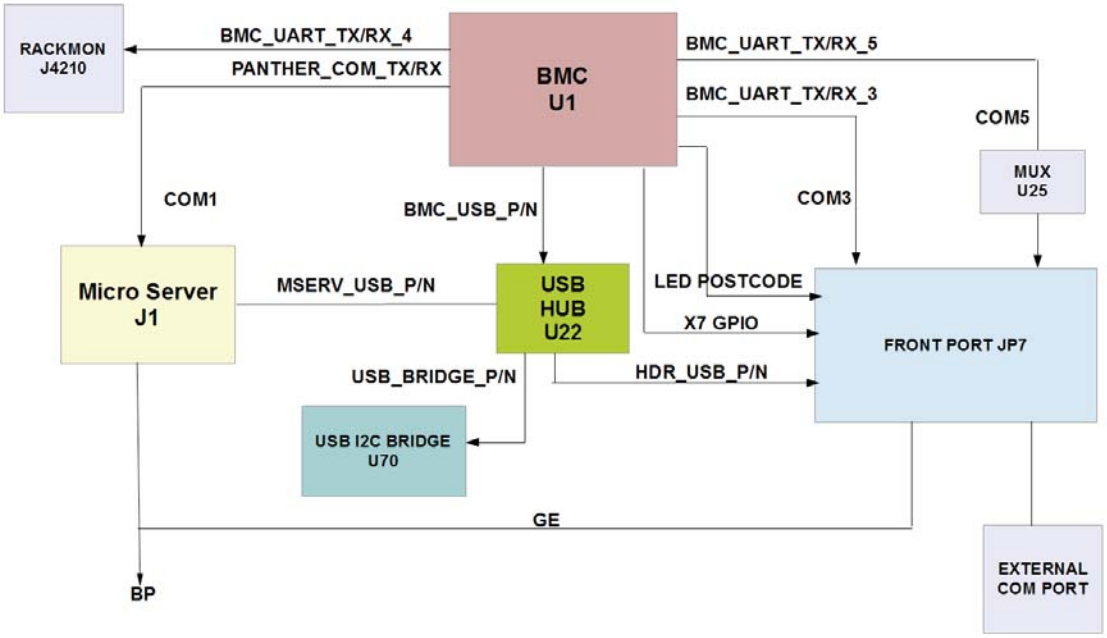


Figure 20: Communication ports

The communication port block diagram is given above. Here the com port 1 of BMC is connected to microserver. Com port 3 and com port 5 are connected to the front port JP7. Com port 4 of BMC is connected to Facebook Specialty connector which is used for rack monitoring using rack manager. The GE port of microserver is connected to front panel as well as to Back plane connectors. BMC can get the microserver bootup status through the LED post code GPIOs connected to front port. The USB Hub is connected to the BMC, microserver, header and I2C bridge. The debug port UART selection logic is given below.

DEBUG_PORT_UART_SEL_N is connected to a push button in the front panel. Whenever Debug port select button signal sees a high to low transition, BMC toggles the DEBUG_UART_SEL_0.

If DEBUG_UART_SEL_0 is low then Panther COM port gets connected to debug header.

If DEBUG_UART_SEL_0 is high then BMC com 5 connects to debug header and Panther com port connects to BMC com 1.

Detailed block diagram of the UART port switching is given below.

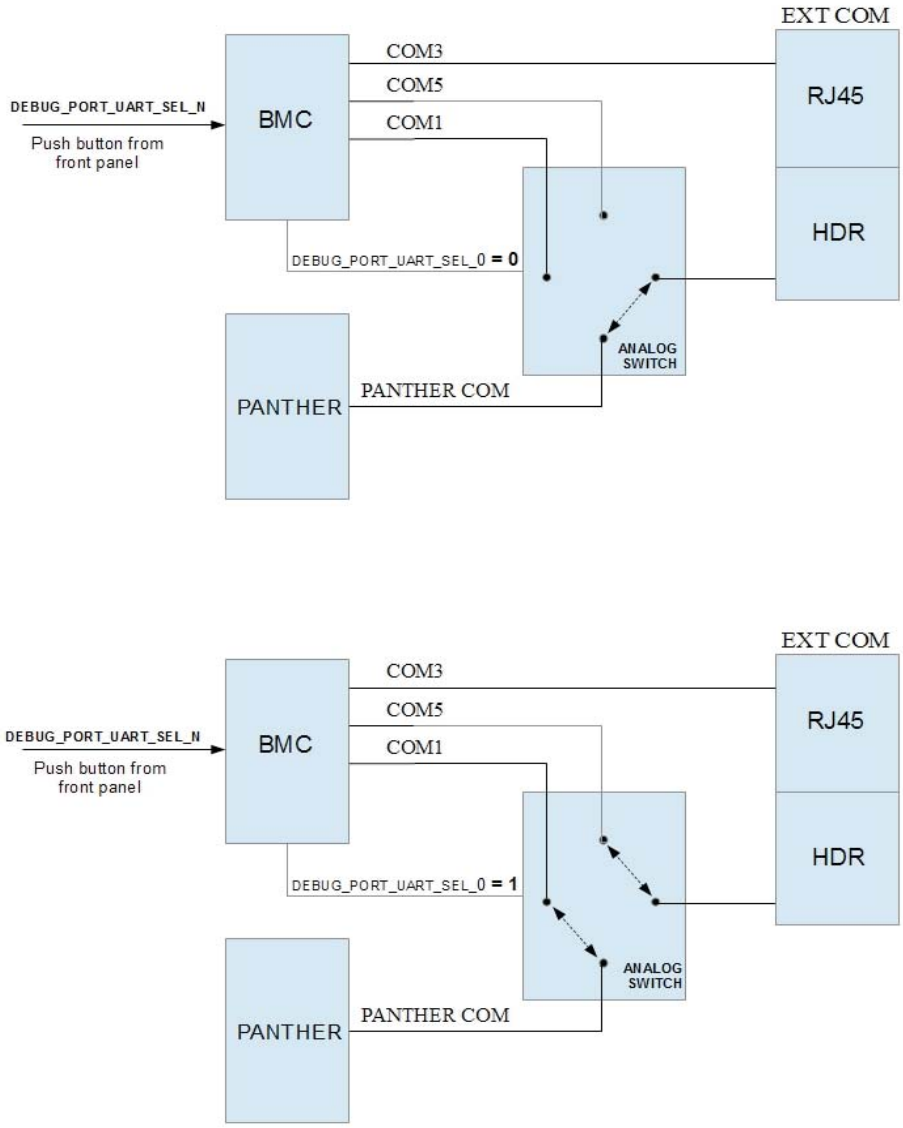


Figure 21: Debug port switching

7.14. Clock Architecture

7.14.1. Block Diagram

The clock Architecture is given in the figure below.

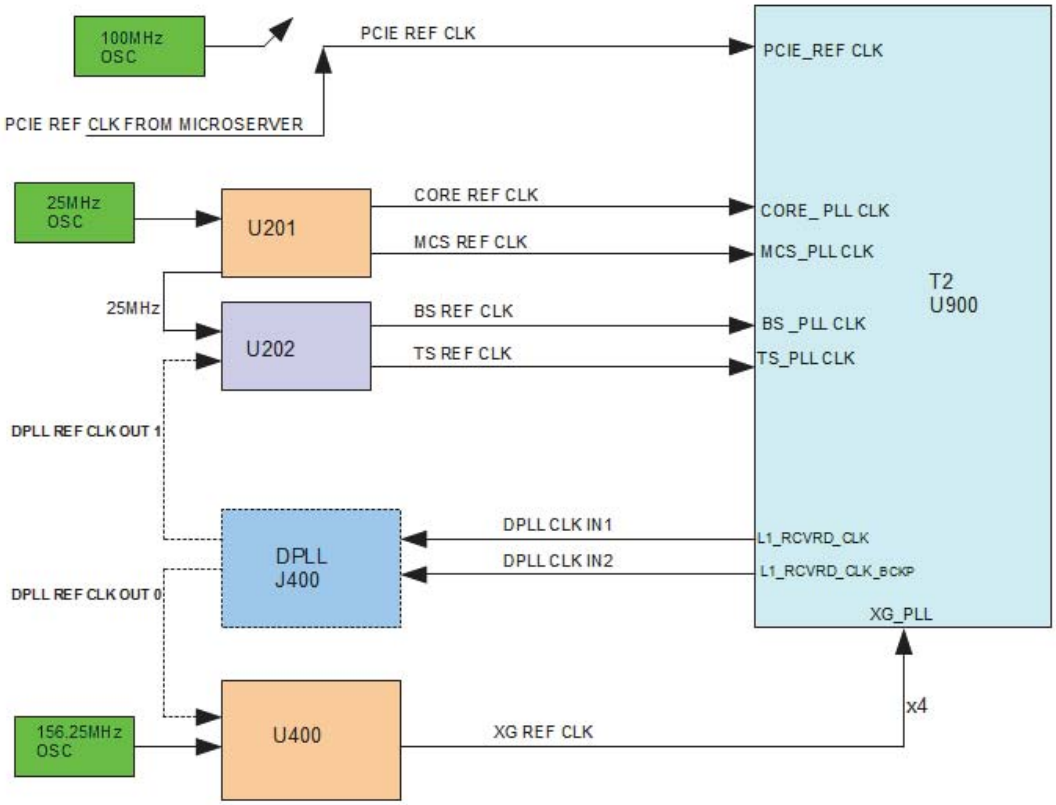


Figure 22: Clock Architecture

The PCIE reference clock for Trident2 is taken from the Micro Server PCIE ref clock. A backup option to feed 100MHz ref clock to PCIE ref clock is also provided. U201 is a 1-to-4 crystal oscillator fan-out buffer. 25MHz clock is fed to CORE_PLL clock and MCS_PLL clock.

U202 is a1-to-5 fan-out buffer which has two selectable inputs. Here one input is from the DPLL and the other input is from the from U201. Default input for U202 is Output of U201. Output from U202 is fed to BS_PLLclk (required to drive the broad sync interface) and TS_PLL clock (Reference clock required for time sync). U400 is a 1-to-4 crystal oscillator fan-out buffer which takes 156.25Mhz input and feeds four XG_PLL clock of Trident2.(Differential clock to drive 4 Warpcores of Trident2)

7.15. Power Architecture

7.15.1. Block Diagram

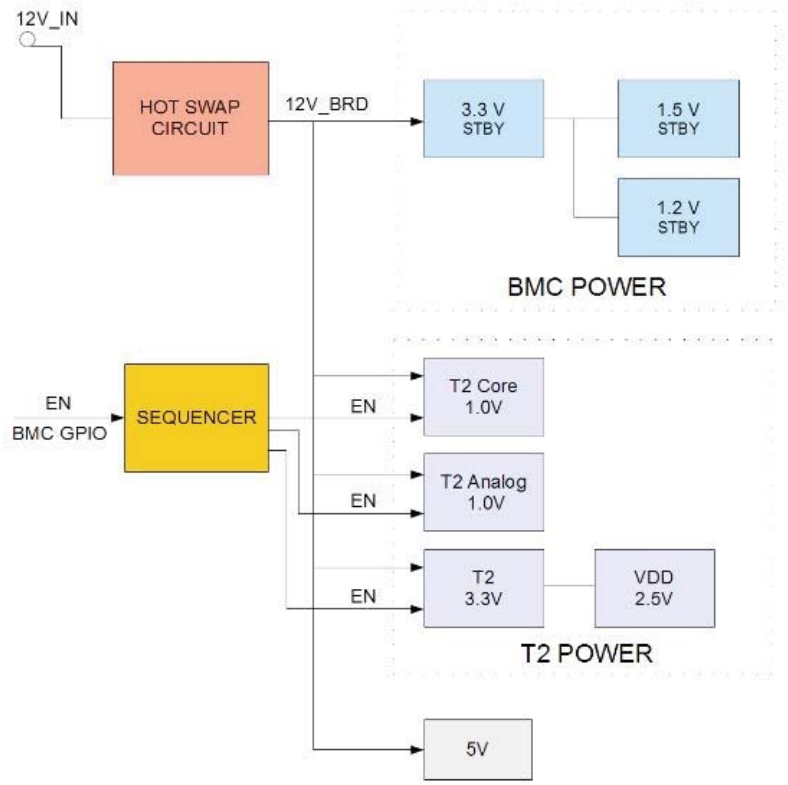


Figure 23: Power Architecture

7.15.2. Description

7.15.2.1. HOT SWAP CONTROLLER

The ADM1278 is a hot swap controller that allows a circuit board to be removed from or inserted into a backplane. It carefully controls the inrush current to ensure a safe power-up interval. The hot-swap controller will also continually monitor the supply current after power-up for protection against short circuits and overcurrent conditions during normal operation.It also allows voltage, and power read back. The hot-swap controller will be enabled by a low on the CARD_EN signal. Since CARD_EN is a short pin in Backplane connector, it will become low, only when the Card is fully inserted. The current limit is set to 30A. The BMC can read the power good signal from Hot swap controller to know whether the 12V input is within the tolerance level. BMC can read whether there is any fault in the hot swap by reading a low on the fault pin. BMC can access information regarding fault, over current sense from hot swap circuit though it's I2C interface. I2C Device address of hot swap controller is 00010000.

7.15.2.2. POWER SEQUENCER

Power sequencing is done using MAX16050, which can sequences up to 4 voltages. The power sequencer is turned ON when the 12V reaches a threshold value. (Power sequencer is turned on 750msec after 12V

reaches 8.5V). MAX16050 can also monitor each power-supply voltage. If any voltage falls below its threshold, the reset output asserts and all voltage supplies gets turned off. The sequencing of voltages is in the order VDD3_3V, VDD_ANLG, and VDD_CORE respectively.

7.15.2.3. VOLTAGE CONTROLLERS

Trident2 Voltages are generated using NCP4200, it is an integrated power control IC with an I2C interface. NCP4200 uses an internal 8−bit DAC to read a Voltage Identification (VID) code, which is used to set the output voltage. BMC can use the I2C interface to program system set points such as voltage offset, load−line and phase balance and output voltage. Performance data, such as current, voltage, power and fault conditions can be read back by the BMC through the I2C. 2.5V is generated from 3.3V using LT1764.

The VID code and device address for different voltage controllers are given below. NOTE: VID code in the table represent voltages after taking into account the voltage drop due to Iref flowing through the feedback resistors. For Example, for the Core voltage, VID is set for 1.01875. But the actual core voltage will be 1V.

Voltage Rail	Refdes	Device ID	VID code
VDD3.3V	U3900	11000A1A0	1010010
VDD_ANLG	U3700	11000A1A0	1011111
VDD_CORE	U3500	11000A1A0	1011111

Table 7: VID codes for NCP4200

Voltage rail	current rating
VDD3_3V	70A
VDD_ANLG	70A
VDD_CORE	140A

Table 8: Current rating of Trident2 voltage rails

3.3V_stby power for BMC is generated from the 12V input supply using MP38872 and the current rating of 3.3V_stby voltage rail is 6A. 1.5V_stby, 1.2V_stby power for BMC is generated from the 3.3V_stby using LT1764 regulator, and this device is capable of supplying 3A of output current. 5V power is also generated from 12V using MP38872 and the current rating 5V voltage rail is 6A.

Device	12V	3.3V	2.5V	5V	VDD CORE(1V)	VDD ANLG(1V)	3.3V_STBY	1.5V_STBY	1.2V_STBY
Microserver	Y								
T2		Y	Y		Y	Y			
DDR3								Y	
BMC							Y	Y	Y
GE crossconnect							Y		
I2C Bridge							Y		
Voltage monitor				Y					
USB HUB							Y		
FAN	Y								
FPGA		Y							
LED(BLUE)				Y					
LED(GREEN)		Y							

Table 9: Voltage rail table

Power up sequence

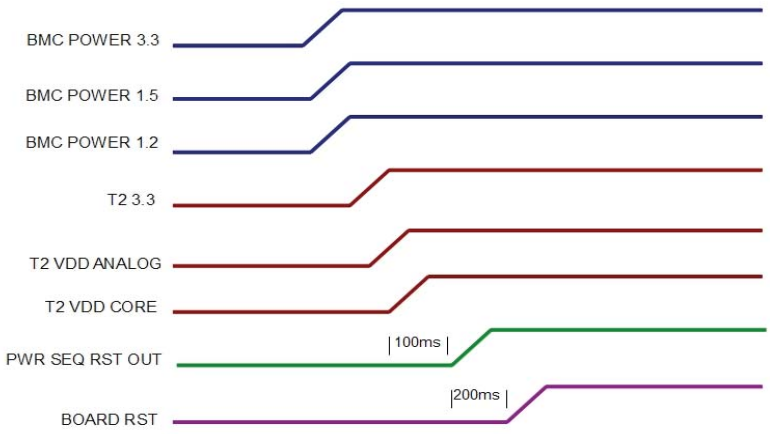


Figure 24: Power up sequence

7.16. I2C Architecture

7.16.1. Block Diagram

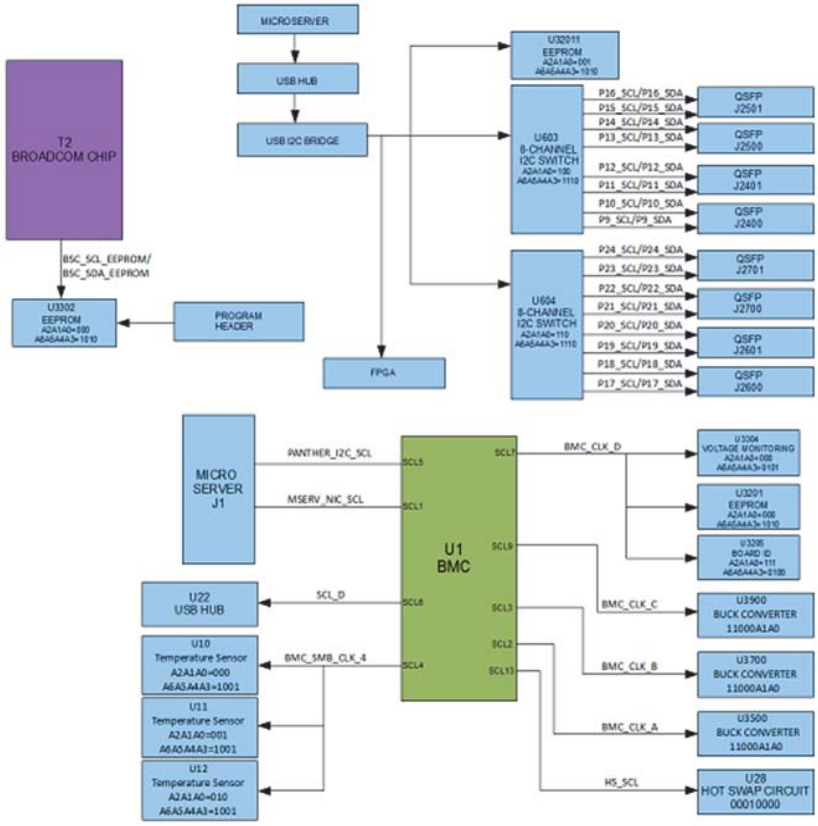


Figure 25: I2C architecture

7.16.2. Description

I2C communication block diagram is shown above. The Micro Server can communicate to the QSFP's through the I2C Bridge and USB hub. BMC gathers information from different circuits onboard using the I2C communication

1. BMC can access information regarding fault, over current sense from hot swap circuit through I2C13. BMC can also power cycle the whole board using this I2C interface.
2. BMC is using I2C 2,3,9 respectively to interface to 1V Core supply controller, 1V Analog supply controller and 3.3V supply. BMC can use the I2C interface to program system set points such as voltage offset, load–line and phase balance and output voltage. Performance data, such as current, voltage, power and fault conditions can be read back by the BMC through the I2C.
3. BMC uses I2C 4 interface to temperature sensors. BMC can read on-board temperature by accessing these sensors(inlet, outlet and Trident2 temp).
4. BMC uses I2C 7 to interface with voltage monitoring IC, General Purpose EPROM, Board ID.
5. BMC I2C 1 is connected to microserver NIC for IPMI interface. (MSERV_NIC_SMBUS_ALRT will be high to indicate an alert event in that particular I2C)
6. BMC I2C 5 provide I2C interface to Panther. (PANTHER_I2C_ALERT_N will go low to indicate an alert event in that particular I2C)

7. BMC is using I2C 6 to interface to USB HUB.

7.17. Reset Architecture

The following figure describes the system level reset architecture:

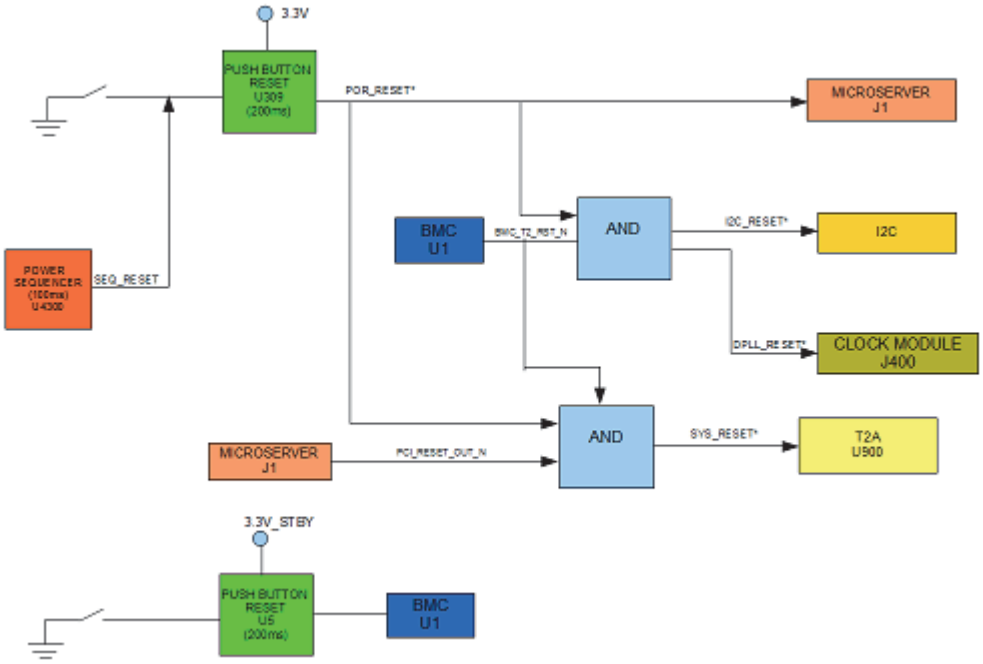


Figure 26: Reset Architecture

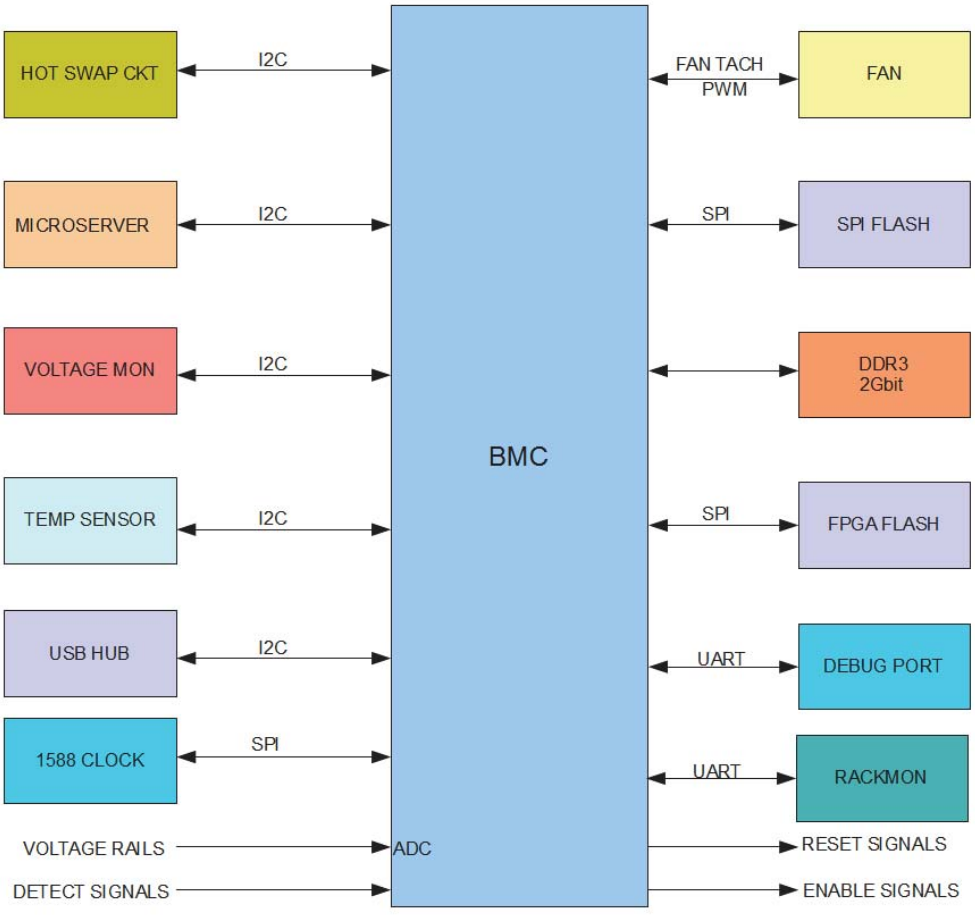
The reset block diagram is given above. The U309 (MAX708RCSA) is a supervisory circuit. In the power sequencer when any of the monitored voltages falls below its threshold, the SEQ_RESET is asserted and U309 in turn resets Micro Server, Trident2, I2C and DPLL Clock module .U309 can be manually reset by using the push button. BMC can also reset Trident2, I2C and DPLL clock module by driving BMC_Trident2_RESET_N low. PCI_RESET_OUT_N can also reset Trident2. BMC can be reset by triggering the push-button of power monitor U5.

8. Detailed Software interface specification

8.1. Introduction

This section covers the Software Interface Specification of the networking switch, Wedge. The document briefly describes functions of the BMC on the main switching module

8.2. BMC Architecture



8.3. BMC Functions

8.3.1. Power up Trident2 and enable BMC interface.

To power up Trident2 and enable interfaces.

BMC pin number	BMC pin name	Signal name
D9	RGMII2TXCK_RMII2TXEN_GPIOT6	Trident2_POWER_UP
A2	GPIOC2_SD1DAT0_SCL11	BUF_EN
E5	GPIOC3_SD1DAT1_SDA11	CARD_EN
B19	GPIOF1_NDCD4_SIOPBI_N	SEQ_RESET

Table 10: Trident 2 Power up GPIO list

8.3.2. Sequence

1. Enable main board power by making Trident2_POWER_UP high.
2. Check if system is powered by monitoring the SEQ_RESET signal.
3. If SEQ_RESET is high, Enable isolation buffer of BMC by making BUF_EN high.

Signal. It should proceed with power up, only when CARD_EN is low

8.3.3. Trident2 Reset.

The objective of this function is to set the appropriate reset sequence of Trident2 and reset Trident2. The state of RESET_SEQ pins determines the mode in which Trident2 reset should work. RESET_SEQ function table is given below.

RESET_SEQ[1,0]	Sequence
11	MCS only
10	MCS is out of reset first and then PCIE
01	PCIE is out of reset first and then MCS
00	PCIE is out of reset only

Table 11 :RESET_SEQ function table

BMC pin number	BMC pin name	Signal name
A20	GPIOF2_NDSR4_SIOPWRGD	RESET_SEQ1
D17	GPIOF3_NRI4_SIOPBO_N	RESET_SEQ0
C4	GPIOC0_SD1CLK_SCL10	BMC_Trident2_RESET_N

Table 12:Trident2 Reset GPIO list

8.3.4. Sequence

1. SW should drive appropriate signals to RESET_SEQ pins from BMC.
2. Once appropriate reset sequence is set drive BMC_Trident2_reset_N low for 2µs to reset Trident2.

8.3.5. Microserver detection

A low on the PANTHER_PRSNT_N signal indicates that Micro Server is present.

BMC Pin Number	Symbol Pin Name	Signal Name
E14	GPIOE4_NDTR3	PANTHER_PRSNT_N

Table 13:Micro Server detect GPIO list

8.3.6. Micro Server Reset.

To reset Micro Server

BMC pin number	BMC pin name	Signal name
B3	GPIOC1_SD1CMD_SDA10	MSRVR_SYS_RST

Table 14:Micro Server Reset GPIO list

8.3.7. Sequence

1. To reset Micro Server, MSRVR_SYS_RST signal should be pulled low for xx µS.

8.3.8. Micro Server Power button

The objective of this function is to emulate Microserver power button function using BMC.

BMC pin number	BMC pin name	Signal name
A18	GPIOD0_SD2CLK	BMC_PWR_BTN_IN_N
D16	GPIOD1_SD2CMD	BMC_PWR_BTN_OUT_N
AB7	GPIOP7_TACH15_FLACK	BMC_READY_N

Table 15:Power Button GPIO list

8.3.9. Sequence

1. Once the BMC software is up, it should make BMC_READY signal low.
2. When BMC_READY is low, BMC can put microserver in sleep mode or power off

- When Micro Server is power on state, BMC can put it in sleep mode by making BMC__PWR_BTN_OUT_N low for xx µs (short press)
- To wake up from sleep mode make BMC__PWR_BTN_OUT_N low for xx µs (short press)
- When Micro Server is power on state, BMC can power off Micro Server by making BMC__PWR_BTN_OUT_N low for xx µs (long press)
- To power on from power off state make BMC__PWR_BTN_OUT_N low for xx µs (long press)

8.3.10. Microserver I2C alert

I2C alert signals that alert the BMC that an event has occurred that needs to be taken care off. The signal PANTHER_I2C_ALERT_N will go low to indicate an alert event in that particular I2C.MSERV_NIC_SMBUS_ALRT will be high to indicate an alert event in that particular I2C .

BMC Pin Number	Symbol Pin Name	Signal Name
J21	GPIOB0_SALT1	PANTHER_I2C_ALERT_N
J20	GPIOB1_SALT2	MSERV_NIC_SMBUS_ALRT

Table 16:I2C alert GPIO list

8.3.11. Microserver Post code read

BMC can read Microserver boot up status by reading GPIOS shown in below table.

BMC pin number	BMC pin name	Signal name
A14	GPIOG0_SGPSCK	LED_POSTCODE_0
E13	GPIOG1_SGPSLD	LED_POSTCODE_1

D13	GPIOG2_SGPSI0	LED_POSTCODE_2
C13	GPIOG3_SGPSI1	LED_POSTCODE_3
E19	GPIOB4_LPCRST_N	LED_POSTCODE_4
H19	GPIOB5_LPCPD_N_LPCSMI_N	LED_POSTCODE_5
H20	GPIOB6_LPCPME_N	LED_POSTCODE_6
E18	GPIOB7_EXTRST_N_SPICS1_N	LED_POSTCODE_7

Table 17:Post Code GPIO list

8.3.12. BMC I2C interface to Hot swap circuit.

BMC can access information regarding fault, over current sense from hot swap circuit

Though it's I2C13. For programing details refer to ADM1278 data sheet.

I2C Device address of Hot swap controller is 00010000

BMC pin number	BMC pin name	Signal name
B2	GPIOC6_SD1CD_N_SCL13	HS_SCL
A1	GPIOC7_SD1WP_N_SDA13	HS_SDA

Table 18:I2C to Hot Swap GPIO list

8.3.13. BMC-Microserver NIC I2C interface.

BMC I2C 1 is connected to microserver NIC for IPMI interface

BMC pin number	BMC pin name	Signal name
K21	SCL1	MSERV_NIC_SCL
K22	SDA1	MSERV_NIC_SDA

Table 19: I2C NIC GPIO list

8.3.14. BMC- Panther I2C interface.

BMC I2C 5 provide I2C interface to Panther **.**

BMC pin number	BMC pin name	Signal name
E3	GPIOK0_SCL5	PANTHER_I2C_SCL
D2	GPIOK1_SDA5	PANTHER_I2C_SDA

Table 20: I2C to Micro Server GPIO list

8.3.15. BMC I2C interface to 1V Core supply controller.

BMC is using I2C 2 to interface to 1V Core supply controller. BMC can use the I2C interface to program system set points such as voltage offset, load−line and phase balance and output voltage. Performance data,

such as current, voltage, power and fault conditions can be read back by the BMC through the I2C.For programing details refer to NCP4200 data sheet.

Device address is: 11000A1A0.

BMC pin number	BMC pin name	Signal name
J19	SCL2	BMC_CLK_A
J18	SDA2	BMC_DATA_A

Table 21:I2C to Core voltage GPIO list

8.3.16. BMC I2C interface to 1V Analog supply controller.

BMC is using I2C 3 to interface to 1V Analog supply controller. BMC can use the I2C interface to program system set points such as voltage offset, load−line and phase balance and output voltage. Performance data, such as current, voltage, power and fault conditions can be read back by the BMC through the I2C. For programing details refer to

NCP4200 data sheet.

Device address is: 11000A1A0.

BMC pin number	BMC pin name	Signal name
D3	GPIOQ0_SCL3	BMC_CLK_B
C2	GPIOQ1_SDA3	BMC_DATA_B

Table 22:I2C to 1v Power GPIO list

8.3.17. BMC I2C interface to 3.3V supply controller.

BMC is using I2C 9 to interface to 3.3V supply controller. BMC can use the I2C interface to program system set points such as voltage offset, load−line and phase balance and output voltage. Performance data, such as current, voltage, power and fault conditions can be read back by the BMC through the I2C.For programing details refer to

NCP4200 data sheet.

Device address is: 11000A1A0.

BMC pin number	BMC pin name	Signal name
C5	GPIOA4_TIMER5_SCL9	BMC_CLK_C
B4	GPIOA5_TIMER6_SDA9	BMC_DATA_C

Table 23:I2C to power controller GPIO list

8.3.18. BMC I2C interface to USB HUB.

BMC is using I2C 6 to interface to USB HUB.

BMC pin number	BMC pin name	Signal name
C1	GPIOK2_SCL6	SCL_D
F4	GPIOK3_SDA6	SDA_D

Table 24:I2C to USB Hub GPIO list

8.3.19. BMC I2C interface to temperature sensors.

BMC is using I2C 4 interface to temperature sensors. BMC can read on-board temperature by accessing these sensors. Details of each Temperature sensor is given in the table below.

Ref des	Device ID	Location
U10	1001000	outlet
U11	1001001	Trident2 temperature
U12	1001010	inlet

Table 25: Temperature sensors

BMC pin number	BMC pin name	Signal name
B1	GPIOQ2_SCL4	BMC_SMB_CLK_4
F5	GPIOQ3_SDA4	BMC_SMB_DATA_4

Table 26:Temp sensors GPIO list

8.3.20. BMC I2C interface to voltage monitoring IC, General purpose EEPROM, Board ID.

BMC is using I2C 7 to interface with voltage monitoring IC, General Purpose EPROM, Board ID.

Board ID device id: 0101000 .

General purpose EPROM; device id: 1010000 .

Voltage monitoring IC(U3304); device id: 0101000 .

BMC pin number	BMC pin name	Signal name
E2	GPIOK4_SCL7	BMC_CLK_D
D1	GPIOK5_SDA7	BMC_DATA_D

Table 27:BMC I2C GPIO list

Below table shows the voltages monitored by U3304.

Channel	Voltage	Value	Tolerance
CH0	VDD CORE	1V	+/-2%
CH1	VDD2_5V	2.5V	+/-5%
CH2	Not Used	Not Used	Not Used
CH3	VDD_ANLG	1.03V	+/-2%
CH4	Not Used	Not Used	Not Used

CH5	VDD3_3V	3.3V	+/-5%
CH6	VDD5V	5V	+/-5%
CH7	Not Used	Not Used	Not Used

Table 28: Voltages monitored and tolerance value

8.3.21. BMC Power led blue

Power LED blue will be ON when LED_PWR_BLUE is high.

Power LED blue will be off when LED_PWR_BLUE is low.

BMC pin number	BMC pin name	Signal name
D14	GPIOE5_NRTS3	LED_PWR_BLUE

Table 29: BMC Power LED GPIO list

8.3.22. HEART Beat LED.

Heart beat LED output indicates whether system is working fine.

Heart beat LED will be off when BMC_HEARTBEAT_N is high

Heart beat LED will be ON when BMC_HEARTBEAT_N is low

BMC pin number	BMC pin name	Signal name
H1	GPIOQ7	BMC_HEARTBEAT_N

Table 30 : Heartbeat LED GPIO list

8.3.23. FAN control and monitoring

BMC can control the speed of FANS in the system through PWM outputs. It can read the FAN speed using the TACH inputs.

Below list shows how each FAN is connected to BMC.

1. PWM0, TACH0, TACH1 for FAN1
2. PWM1, TACH2, TACH3 for FAN2
3. PWM6, TACH4, TACH5 for FAN3
4. PWM7, TACH6, TACH7 for FAN4

BMC pin number	BMC pin name	Signal name
W4	GPION0_PWM0_VPIG0	PWM_OUT1
Y3	GPION1_PWM1_VPIG1	PWM_OUT2
AA3	GPION6_PWM6_VPIG6	PWM_OUT3
AB2	GPION7_PWM7_VPIG7	PWM_OUT4

V6	GPIOO0_TACH0_VPIG8	FAN1_TACH
Y5	GPIOO1_TACH1_VPIG9	FAN2_TACH
AA4	GPIOO2_TACH2_VPIR0	FAN3_TACH
AB3	GPIOO3_TACH3_VPIR1	FAN4_TACH
W6	GPIOO4_TACH4_VPIR2	FAN1A_TACH
AA5	GPIOO5_TACH5_VPIR3	FAN2A_TACH
AB4	GPIOO6_TACH6_VPIR4	FAN3A_TACH
V7	GPIOO7_TACH7_VPIR5	FAN4A_TACH

Table 31: Fan Control GPIO list

8.3.24. Trident2 ROV read

The objective of this function is to read the ROV value from Trident2. BMC can read ROV from Trident2 through GPIOs given in table below.

BMC pin number	BMC pin name	Signal name
A8	GPIOH0_ROMD8_NCTS6	ROV1
C7	GPIOH1_ROMD9_NDCD6	ROV2
B7	GPIOH2_ROMD10_NDSR6	ROV0

Table 32: Trident2 ROV GPIO list

ROV 3 bit coding is given below

010=0.95V

001=1.025V/1.0V

8.3.25. ADC voltage monitoring

BMC can monitor voltages (5V, 3.3V, 2.5V,1V_Analag, 1V_core) connected to ADCs. This is in addition to the

External voltage monitoring circuit connected through I2C.

BMC pin number	BMC pin name	Signal name
M5	ADC5_GPIW5	VDD_CORE
M4	ADC6_GPIW6	VDD_ANLG
M3	ADC7_GPIW7	V5_SNS
M2	ADC8_GPIX0	V3P3_SNS
M1	ADC9_GPIX1	V2P5_SNS

Table 33:ADC Voltage monitor GPIO list

ADC voltage mapping

ADC	Voltage rail Monitored	Remarks
5	1V_CORE	Division factor =1
6	1V_ANALOG	Division factor =1
7	5V	Division factor =2
8	3.3V	Division factor =2
9	2.5V	Division factor =2

Table 34:ADC voltage mapping

8.3.26. Monitoring Hot Swap Controller

HOTSWAP_PG will be high when the12V input is within the tolerance level.

If the HS_FAULT_N signal is low then it indicates that the hot swap went to shut down.

BMC Pin Number	Symbol Pin Name	Signal Name
V1	GPIOL3_NRI1_VPIHS	HOTSWAP_PG
E9	RGMII2TXCTL_GPIOT7	HS_FAULT_N

Table 35:Hot Swap Monitor GPIO list

8.3.27. BP SLOT ID read

BMC can read the below mentioned GPIOs to understand to which slot the wedge is

Inserted in the back plane.

BMC pin number	BMC pin name	Signal name
A10	RGMII2TXD0_RMII2TXD0_GPIOU0	BP_SLOT_ID0
B10	RGMII2TXD1_RMII2TXD1_GPIOU1	BP_SLOT_ID1
C10	RGMII2TXD2_GPIOU2	BP_SLOT_ID2
D10	RGMII2TXD3_GPIOU3	BP_SLOT_ID3

Table 36 :Slot ID GPIO list

If BP_SLOT_ID3 is low It represents the slots in left back plane .

If BP_SLOT_ID3 is high it represents the slots in right back plane.

BP_SLOT_ID[2:0]	Slot
000	0
001	1
010	2
011	3

Table 37:Slot ID

8.3.28. Board rev ID read

BMC can read board revision ID from below mentioned GPIOs.

For rev 0 the strapped value is 000.

For rev 1 the strapped value is 001.

BMC pin number	BMC pin name	Signal name
C21	GPIOY0_SIOS3_N	BOARD_REV_ID0
F20	GPIOY1_SIOS5_N	BOARD_REV_ID1
G20	GPIOY2_SIOPWREQ_N	BOARD_REV_ID2

Table 38:Board Rev ID GPIO list

8.3.29. BMC Programming FPGA Flash.

The objective of this function is to update the FPGA Flash image into the EEPROM

BMC Pin Number	Symbol Pin Name	Signal Name
B17	GPIOD2_SD2DAT0	FLASH_CS
A17	GPIOD3_SD2DAT1	FLASH_SCK
C16	GPIOD4_SD2DAT2	FLASH_SI
B16	GPIOD5_SD2DAT3	FLASH_SO
A16	GPIOD6_SD2CD_N	FL_PRG_SEL
H4	GPIOQ4_SCL14	FLASH_WP
H3	GPIOQ5_SDA14	FLASH_HOLD

Table 39 :FPGA Programing GPIO list

8.3.30. Sequence

1. Make FL_PRG_SEL high, when FPGA flash needs to be reprogrammed. With this, bus Switch (U15, U14) connected to FPGA EEPROM will connect the EEPROM signals to BMC.
2. Make FLASH_WP high to disable Write protect of FPGA PROM
3. Initiate FLASH_EEPROM memory write by making FLASH_CS low.

4. Execute FPGA_EEPROM programming function.

8.3.31. USB HUB Reset

The objective of this function is to reset the USB HUB. BMC can reset USB HUB by pulling BMC_USB_RESET_N low.

BMC pin number	BMC pin name	Signal name
E15	GPIOD7_SD2WP_N	BMC_USB_RESET_N

Table 40: USB Hub Reset GPIO list

8.3.32. Monitoring USB over current event

A low on the USB_OSC_N1 indicates that the resettable fuse has tripped.

BMC pin number	BMC pin name	Signal name
H2	GPIOQ6	USB_OCS_N1

Table 41: Monitoring USB GPIO list

8.3.33. RX loss signal read

The objective of this function is to read RX loss signals from QSFP ports through the GPIOs

RX loss port map is given in below table.

BMC Pin Number	Symbol Pin Name	Signal Name	QSFP ports
V3	GPIOM0_NCTS2_VPIB2	P24_RX_LOSS	Port 7T
W2	GPIOM1_NDCD2_VPIB3	P23_RX_LOSS	Port 7B
Y1	GPIOM2_NDSR2_VPIB4	P22_RX_LOSS	Port 6T
V4	GPIOM3_NRI2_VPIB5	P21_RX_LOSS	Port 6B
W3	GPIOM4_NDTR2_VPIB6	P20_RX_LOSS	Port 5T
Y2	GPIOM5_NRTS2_VPIB7	P19_RX_LOSS	Port 5B
AA1	GPIOM6_TXD2_VPIB8	P18_RX_LOSS	Port 4T
V5	GPIOM7_RXD2_VPIB9	P17_RX_LOSS	Port 4B
A12	RGMII1TXCK_RMII1TXEN_GPIOT0	P16_RX_LOSS	Port 3T
B12	RGMII1TXCTL_GPIOT1	P15_RX_LOSS	Port 3B
C12	RGMII1TXD0_RMII1TXD0_GPIOT2	P14_RX_LOSS	Port 2T
D12	RGMII1TXD1_RMII1TXD1_GPIOT3	P13_RX_LOSS	Port 2B
E12	RGMII1TXD2_GPIOT4	P12_RX_LOSS	Port 1T
A13	RGMII1TXD3_GPIOT5	P11_RX_LOSS	Port 1B
C6	GPIOR6_MDC1	P10_RX_LOSS	Port 0T

A5	GPIOR7_MDIO1	P9_RX_LOSS	Port 0B

Table 42: RX Loss GPIO list

Note: T denotes Top, And B denotes bottom

8.3.34. SPI to clock module

The objective of this function is to provide SPI interface from BMC to the clock module.

BMC Pin Number	Symbol Pin Name	Signal Name
B22	GPIOI4_SPICS0_N_VBCS_N	BMC_SPI_SS
G19	GPIOI5_SPICK_VBCK	BMC_SPI_SCK
C18	GPIOI6_SPIDO_VBDO	BMC_SPI_MOSI
E20	GPIOI7_SPIDI_VBDI	BMC_SPI_MISO

Table 43:SPI to clock GPIO list

8.3.35. Debug port reset signal.

When Debug Reset button in Front panel is pressed, DEBUG_RST_BTN_N will become low. BMC can detect this by reading this GPIO.

BMC pin number	BMC pin name	Signal name
Y22	GPIOR2_ROMCS3_N	DEBUG_RST_BTN_N

Table 44:Debug Port reset GPIO list

8.3.36. Debug port UART selection

This function selects which COM port will go to Debug Header. The DEBUG_PORT_UART_SEL_N is connected to a push button in the front panel. When Debug port select button in Front panel is pressed DEBUG_PORT_UART_SEL_N becomes low.

BMC pin number	BMC pin name	Signal name
H18	GPIOB2_SALT3	DEBUG_PORT_UART_SEL_N
D15	GPIOE0_NCTS3	DEBUG_UART_SEL_0
H5	TXD5	BMC_UART_TX_5
G1	RXD5	BMC_UART_RX_5

Table 45:Debug Port UART selection GPIO list

- When the Software sees a low on DEBUG_PORT_UART_SEL_N, it toggles the DEBUG_UART_SEL_0.
- If DEBUG_UART_SEL_0 is low then Panther COM port gets connected to debug header .
- If DEBUG_UART_SEL_0 is high then BMC com 5 connects to debug header and Panther com port connects to BMC com 1 .

8.3.37. RackMon data read

BMC can read RackMon data from the front panel through these GPIOs.

BMC Pin Number	Symbol Pin Name	Signal Name
D7	GPIOH4_ROMD12_NDTR6	RACK_PORT_D0
B6	GPIOH5_ROMD13_NRTS6	RACK_PORT_D1
A6	GPIOH6_ROMD14_TXD6	RACK_PORT_D2
E7	GPIOH7_ROMD15_RXD6	RACK_PORT_D3
U1	GPIOL0_NCTS1	RACK_PORT_D4
T5	GPIOL1_NDCD1_VPIDE	RACK_PORT_D5
U3	GPIOL2_NDSR1_VPIODD	RACK_PORT_D6

Table 46: RackMon GPIO list

8.3.38. Debug port Spare

This is a spare signal from front panel to BMC.

BMC Pin Number	Symbol Pin Name	Signal Name
A7	GPIOH3_ROMD11_NRI6	DEBUG_PORT_SPARE

Table 47: Debug port Spare GPIO list

8.3.39. Serial UART signals

Signals from BMC to UART port in the front panel

BMC Pin Number	Symbol Pin Name	Signal Name
C14	GPIOE6_TXD3	BMC_UART_TX_3
B14	GPIOE7_RXD3	BMC_UART_RX_3

Table 48: Serial UART signal GPIO list

8.3.40. I2C from front panel

I2C signals from the power module of front panel to the BMC.

BMC Pin Number	Symbol Pin Name	Signal Name
G5	GPIOK6_SCL8	RFID_SMB_CLK_2
F3	GPIOK7_SDA8	RFID_SMB_DATA_2

Table 49: I2C from front panel GPIO list

8.3.41. Monitoring health of board in rack

Rack manager can read health of board in rack through these GPIOs going to Facebook Specify connector from BMC.

BMC Pin Number	Symbol Pin Name	Signal Name
Y6	GPIOP0_TACH8_VPIR6	H_GPIOP0
AB5	GPIOP1_TACH9_VPIR7	H_GPIOP1
W7	GPIOP2_TACH10_VPIR8	H_GPIOP2
AA6	GPIOP3_TACH11_VPIR9	H_GPIOP3
AB6	GPIOP4_TACH12	H_GPIOP4
Y7	GPIOP5_TACH13	H_GPIOP5
AA7	GPIOP6_TACH14_BMCINT	H_GPIOP6
B13	GPIOG4_WDTRST1_OSCCLK	H_GPIOG4
Y21	GPIOG5_WDTRST2_USBCKI	H_GPIOG5
AA22	GPIOG6_FLBUSY_N	H_GPIOG6
U18	GPIOG7_FLWP_N	H_GPIOG7
K20	GPIOY3_SIOONCTRL_N	H_GPIOY3
J5	GPIOJ0_SGPMCK	H_GPIOJ0
J4	GPIOJ1_SGPMLD	H_GPIOJ1
K5	GPIOJ2_SGPMO	H_GPIOJ2
J3	GPIOJ3_SGPMI	H_GPIOJ3

Table 50: Monitoring health of board in rack GPIO list

8.3.42. UART for rack monitoring

UART signals to Facebook Specify connector for rack monitoring using rack manager.

BMC Pin Number	Symbol Pin Name	Signal Name
E16	GPIOF6_TXD4	BMC_UART_TX_4
C17	GPIOF7_RXD4	BMC_UART_RX_4

Table 51: UART for rack monitoring GPIO list

8.4. Microserver I2C access to QSFP

Micro Server needs interface with QSFP connectors to monitor the status of each port. Each QSFP connector has I2C bus through which we can read the status of each QSFP port. The micro server can read the QSFPs through the USB hub, USB-I2C bridge and I2C switch. CP2112 is used as the USB to I2C bridge. The USB hub is connected to U70 (CP2112) through one of its downstream port(Port3). There are 16 QSFP ports. The interface between USB I2C bridge and each QSFP connector I2C is through two 8-channel I2C switches (PCA9548). To access QSFP 1, first Micro server has to enable Downstream port3 (which is connected to USB to I2C bridge) of the USB hub. Through this USB port, I2C bus connected to two I2C switches, FPGA and EEPROM (USB_Trident2_I2C) can be accessed. We can select a particular device by writing the slave address of that particular device in the USB-I2C bridge. I2C slave address is given below.

FPGA- 7’b 101 0101

U603- 7’b100 1110

U604- 7’b110 1110

QSFP1 is connected to I2C channel 7 of the U603. If Micro Server wants to access QSPF1, it has to enable channel7 of the 8- channel I2C switch, U603. This can be done by accessing device address of U603 and writing into the control register of 8-channel I2C switch. Once that particular channel is enabled, all the I2C transactions addressed to U603, will be routed to QSFP1. The I2C architecture for the QSFP port is given below.

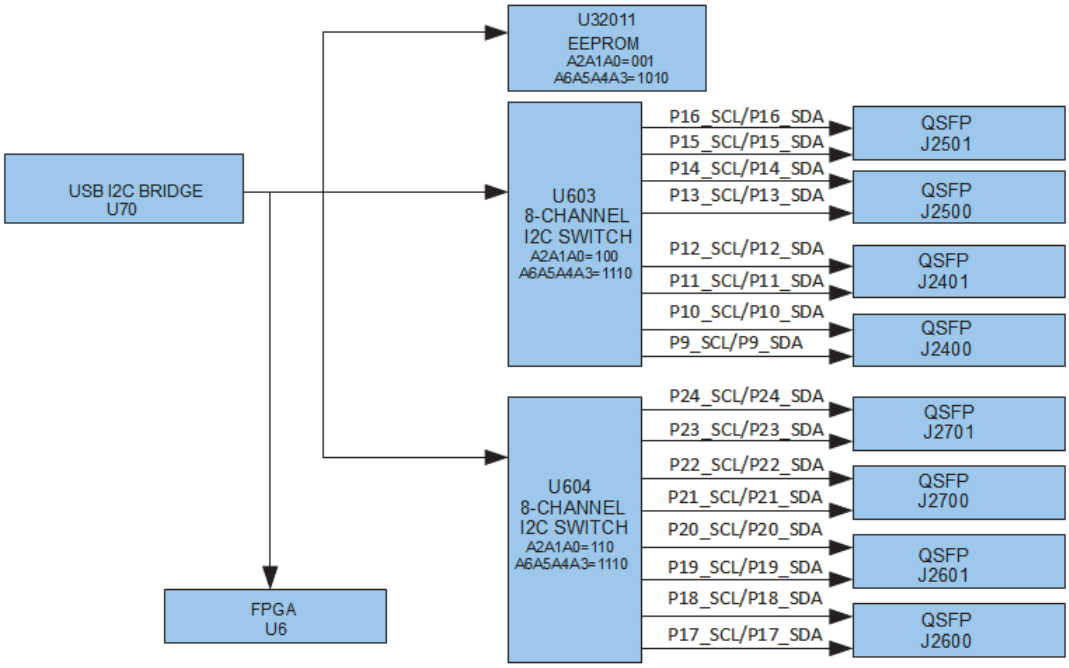


Figure 27: Micro Server I2C access to QSFP

8.5. Board power cycle using BMC.

BMC can power cycle the hot swap circuit by doing a write cycle into address 0xD9 through its I2C13. There by it will power cycle the whole board.

I2C Device address of Hot swap controller is 00010000

BMC pin number	BMC pin name	Signal name
B2	GPIOC6_SD1CD_N_SCL13	HS_SCL
A1	GPIOC7_SD1WP_N_SDA13	HS_SDA

Table 45: Board power cycle using BMC GPIO list

9. Software compatibility and support

Micro Server

The Micro Server supports any network operating system that can be loaded using PXE. For example, Facebook packages a Linux-based distribution (CentOS 6), and then loads FBOSS applications (some of which has already been contributed to open source) along Broadcom's OpenNSL

BMC - OpenBMC

The BMC (based on the Aspeed 1250 currently) runs the OpenBMC code to the open source community,

the open BMC has been released by Facebook and can be found on the public distributions.

10. Physical Specifications

Power Consumption

The total estimated system power consumption of the AS5410-54X is ~360 Watts. This is based upon worst case power assumptions for traffic, optics used, and environmental conditions. Typical power consumption is ~282 Watts

Environmental

- Weight 19.56lbs / 8.8kg
- 0 to 40 Degrees C operating range
- -40 to 40 Degrees C storage temperate range
- Humidity 5% to 95% non-condensing (operational and storage)
- Vibration – IEC 68-2-36, IEC 68-2-6
- Shock – IEC 68-2-29
- Acoustic Noise Level – Under 60dB in 40 degree C
- Altitude - 15,000 (4572 meters) tested operational altitude

Safety

- UL/ Canada
- CB (Issued by TUV/RH)
- China CCC

Electromagnetic Compatibility

- CE
- EN55022 Class A
- EN55024
- EN61000-3-2
- EN61000-3-3
- FCC Title 47, Part 15, Subpart B Class A
- VCCI Class A
- CCC

ROHS

Restriction of Hazardous Substances (6/6)

Compliance with Environmental procedure 020499-00 primarily focused on Restriction of Hazardous Substances (ROHS Directive 2002/95/EC) and Waste and Electrical and Electronic Equipment (WEEE Directive 2002/96/EC)

MTBF

The expected calculated MTFB for the Wedge system is as described in the table below:

	Temperature

Predicted MTBF	25C	45C
System Failure Rate (FIT)	2717.807230	5851.596668
System MTBF (Hours)	**367,944**	**170,894**
MAIN BOARD	855,811	377,874
MINI BOARD	8,469,937	3,478,319
FAN EXTENDED	88,531,306	55,415,353
POWER BOARD	61,210,709	24,584,452

Table 52: Wedge MTBF