OPEN

Compute Project

Project Olympus
FX-16 Flash Expansion Specification

Author:

Mark A. Shaw, Principal Hardware Engineering Manager, Microsoft

Engineering Version History

Rev	Version/ Date	Notes
A	1.0	Initial Release Revision

Table of Contents

List of Figures

List of Tables

1 Overview

The FX-16 is 1U PCIe based flash storage expansion platform. It supports up to 16 SFF-TA-1007 EDSFF 1U Long (E1.L) 18mm SSD form factor storage devices. Each device is hot pluggable at the front of the server (cold aisle). The server supports two PCIe x16 interfaces to an external host using MiniSAS HD cabling. Bifurcation of the PCIe is determined by the host server.

A 3D drawing of the FX-16 is shown in Figure 1.

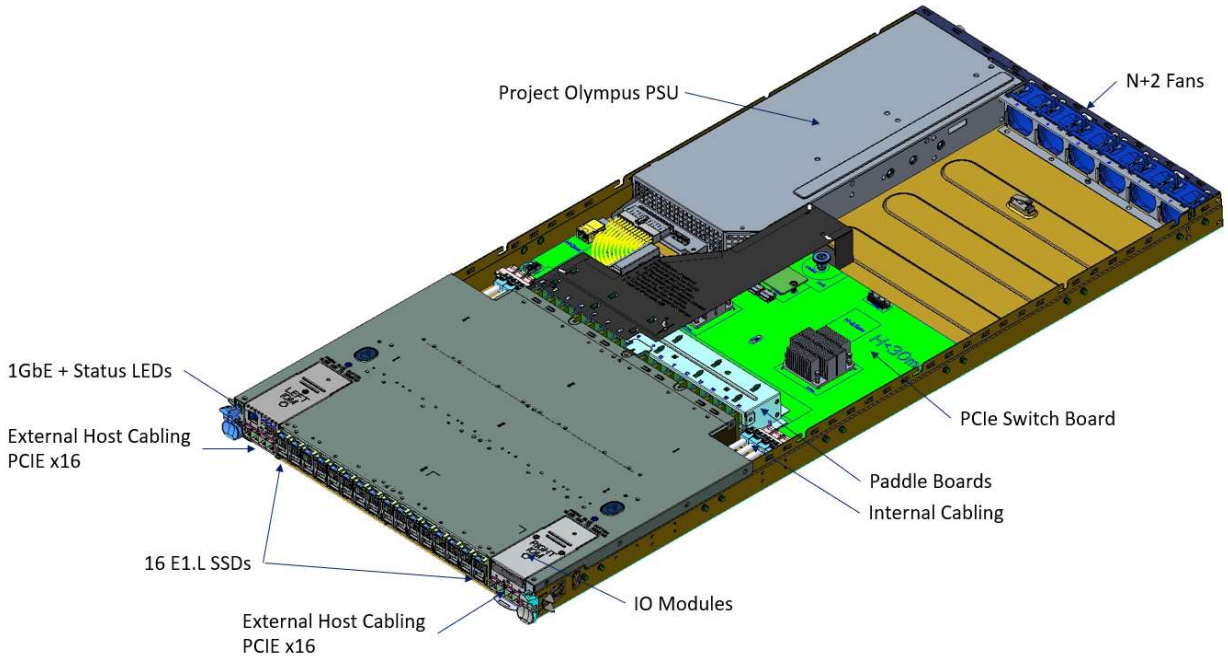


Figure 1. FX-16 3D

The FX-16 supports the following primary functional features:

- 16 hot pluggable EDSFF 1U Long (E1.L) SSDs
  - Power and signal isolation between SSDs
- Two PCIe x16 external front external cable interfaces (MiniSAS HD SFF-8644))
- AST2520 BMC for FX-16 management
  - Cerberus Security
  - Dual flash architecture
  - Open BMC firmware
- 1GbE Management front RJ45 connector
- Support for Project Olympus PSU and server management
- N+2 fan support

- PCIe switch board with 2 PCIe switches
    - PCIe x2, x4, x8, SSD bifurcation support
    - Supports SSD hot install, hot removal, and surprise removal

2 Features

2.1 PCIe

The design supports an interface to two PCIe x16 root ports (Host). An individual x16 is fanned out to up to 8 SSD slots through a single Microsemi PSX8535 PCIe switch. The design incorporates two x16 cable interfaces to two PCIe switches and 16 SSDs. Key features of the PSX8535 are:

- 80 lanes
- X16 Host bifurcation
- 4x2, 2x4, or 1x8 SSD bifurcation
- Up to 48 ports and 48 NTBs

A block diagram of the PCIe topology is shown in Figure 2. The figure shows the complete PCIe data path from the host. The path assumes a specific PCIe retimer and custom MiniSAS HD cable that supports common PCIe reference clock from Host as well as other sideband signals including PERST#, USB, and I2C.

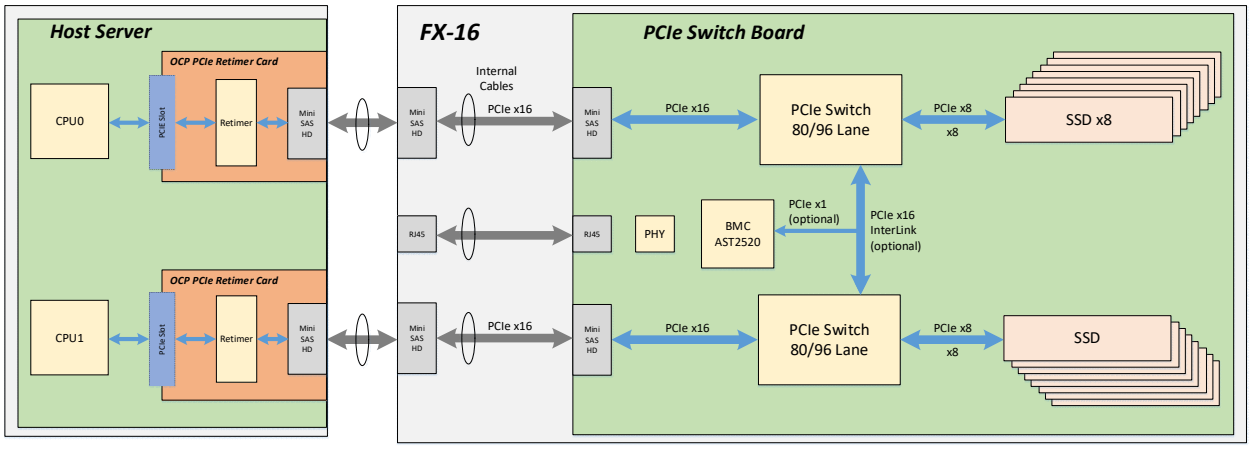


Figure 2. PCIe Block Diagram

Note that the design shall include support for the following BOM load options:

- PFX8536 96 lane switch with a PCIe x16 crossover connection between the two switches.
- PCIe x1 connection to the BMC.

2.2 PCIe Clocks

The PCIE clock tree can be sourced from one of two clock modes:

- Common Clock – In this mode, the base of the clock tree is driven from the Host to the individual PCIe switches. This mode supports spread spectrum and non-spread spectrum clocking. This is expected to be the primary clock mode.
- Local Clock – In this mode, the base of the clock tree is driven from a local clock generator to the individual PCIe switches. This mode does not support spread spectrum clocking. This is expected to be an alternative clock mode but is not expected to be the primary clock mode.

A block diagram of the clock topology is shown in Figure 3. The PCIe Switch receives the reference 100Mhz clock from the Host or CLK generator. It then sources a single clock for each of its PCIe x16 stacks. Since the design bifurcates each x16 stack to a 2x8 or 4x4 topology, a 1:2 clock buffer is required to drive clocks to the SSDs.

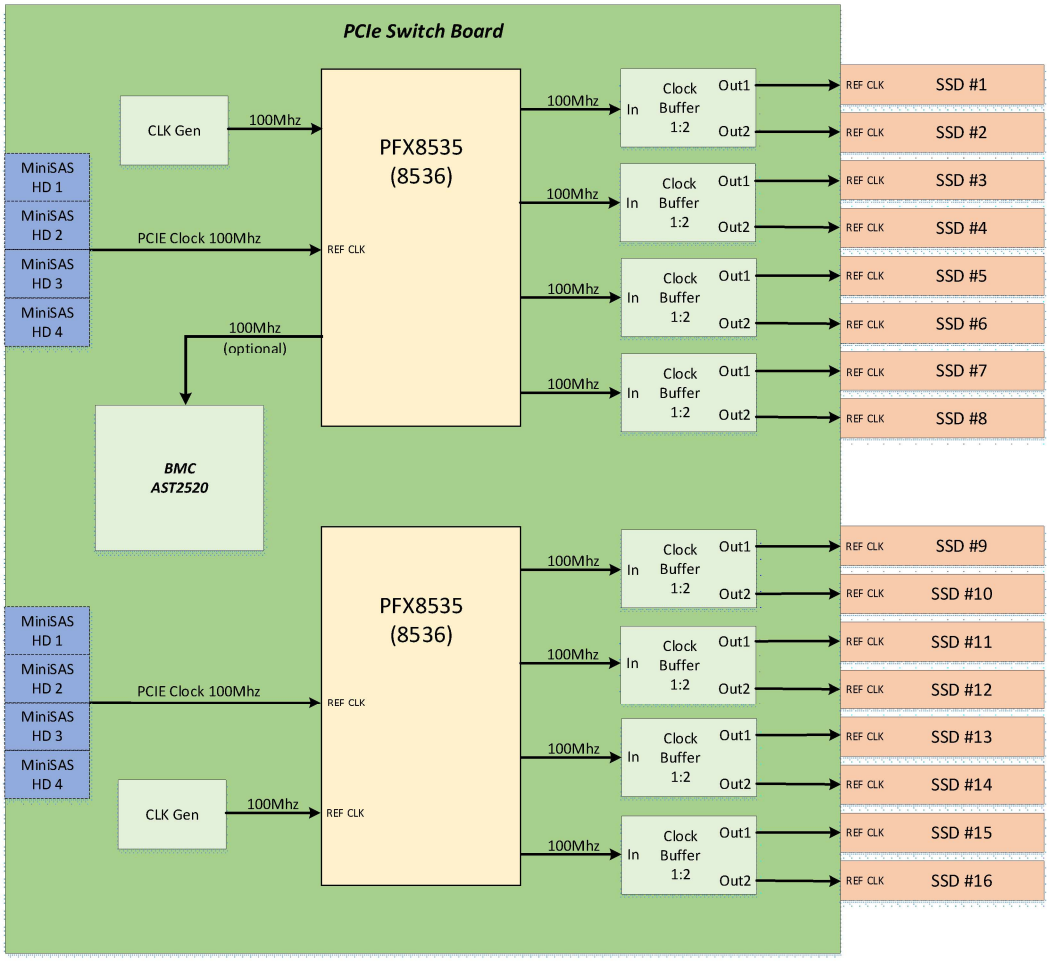


Figure 3. PCIe Clock Block Diagram

2.3 PCIe Reset

PCIe reset shall be sourced from the Host and fanned out to the end points through the PCIe switch as part of the hot swap control logic. The BMC has the capability of asserting PERST# to the SSDs but cannot de-assert PERST# if it is asserted by the PCIe switch. A block diagram is shown in Figure 4.

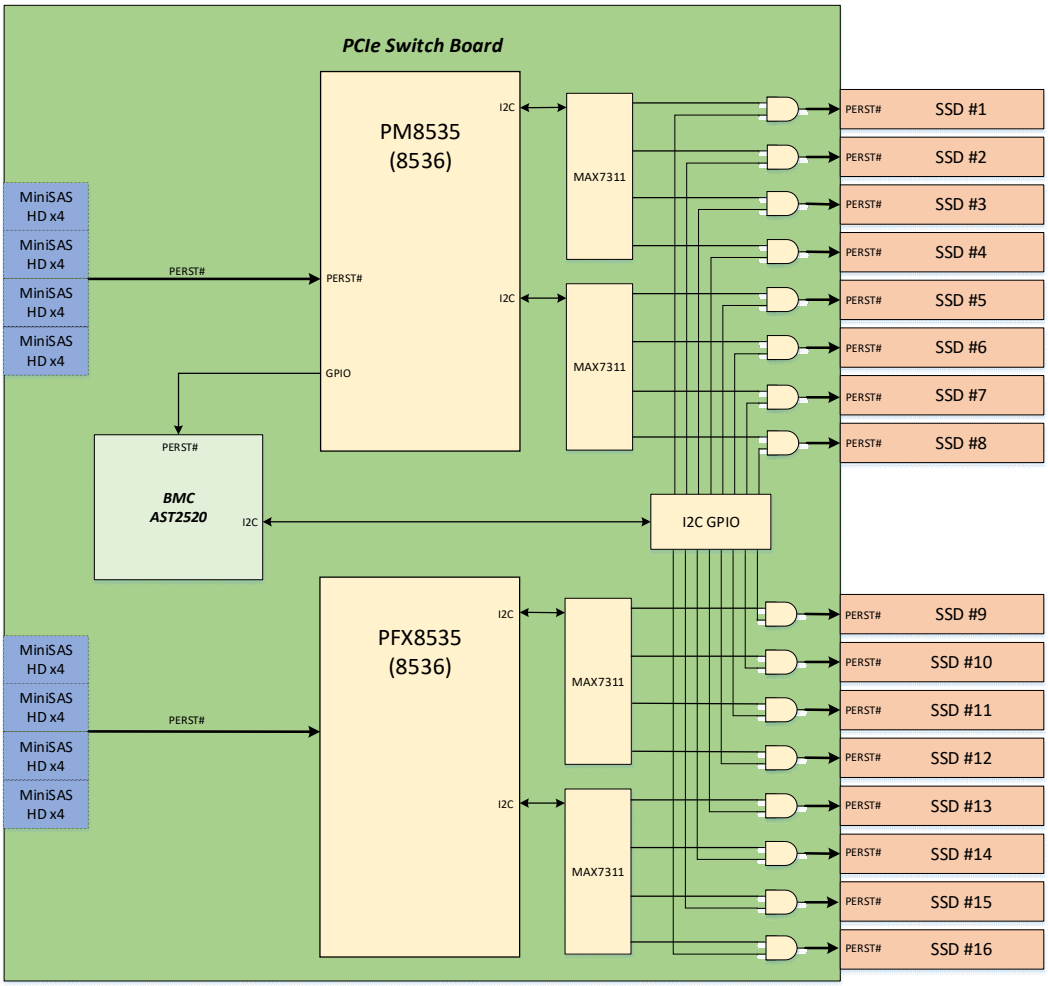


Figure 4. PCIe Reset Block Diagram

2.4 PCIe Hot Swap

The design shall support hot swap of up to 16 SSDs. A block diagram of the hot swap circuitry is shown in Figure 5. Upon insertion of a device, 3.3V Aux power shall be immediately available to the device. The PCIe switch shall be responsible for detecting physical presence of the SSD and managing power enable and PCIe reset as part of its hot swap control logic. The BMC shall track presence and have the capability of asserting PERST# to the device.

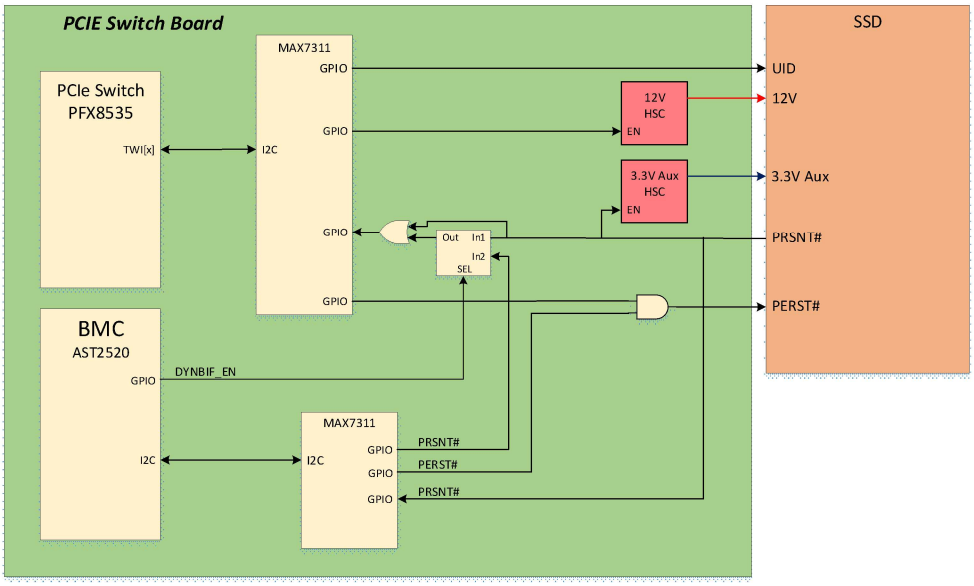


Figure 5. Hot Swap Block Diagram

2.4.1 Dynamic Bifurcation

The design shall optionally support dynamic bifurcation of the PCIe x8 slots enabling the system to detect the PCIe bifurcation of the endpoint(s) and alter the PCIe switch setup to support either PCIe 1x8, 2x4, or 4x2 configurations. The design shall contain the following features to support dynamic bifurcation

- Selectable feature controlled by the BMC to determine the direct source (BMC or SSD) of presence detection by the PCIe switch. The default state of the features shall be “off” meaning presence is driven directly from the SSD to the PCIe Switch.
- When enabled, the BMC shall be responsible for detecting SSD presence, setting the bifurcation registers of the PCIe switch, and forwarding the SSD presence to the switch.
- Removal of the SSD is detected immediately by the PCIe switch and is not dependent on the BMC.

The BMC shall determine the bifurcation requirement by either reading an I2C GPIO device on the SSD or by reading a NVME specified address on the SSD FRUID EEPROM or VPD EEPROM

2.5 PCIe Mapping

Table 1 shows the mapping of PCIe lanes from the input cable ports through the PCIe switches to the SSD slots numbers. The description of the columns is listed below.

- Cable Port – Refers to the PCIe input cable connectors labelled 1-8 at the front of the chassis
- PCIe Switch – Refers to the two PCIe switches on the PCIe Switch Board.

- SB Connector – Refers to the 16 EDSFF connectors on the PCIe Switch Board. The numbering matches the silkscreen identifiers for each connector.
- Port Number – Refers to the internal port numbering of the PCIe switch. In the device datasheet, these are also referred to as a stack.
- Lanes – Refers to the lanes of the PCIe switch based on the Main Board schematics which utilizes the schematic symbol for the 96-lane switch.
- Slot – Refers to the 16 SSD hot plug slot. The designations for the slots match the MB Connector designations.
- EDSFF Map – Refers to the PCIe mapping on the EDSFF connector.
    - x4 Base – Indicates a connection to the x4 base connections of the connector or lanes [3:0]. These lanes are used by any SSD that supports a x4 interface.
    - x8 Expansion – Indicates a connection to the x8 expansion of the connector or lanes [7:4]. These lanes are only used by SSDs that support greater than a x4 interface.

Table 1. PCIE Mapping

Cable Port	PCIe Switch	SB Connector	Port	Lanes	Slot	EDSFF Map
1-4	1	1	2	35:32	1	x8 Exp
				39:36	1	x4 Base
		2		43:40	2	x8 Exp
				47:44	2	x4 Base
		3	3	51:48	3	x8 Exp
				55:52	3	x4 Base
		4		59:56	4	x8 Exp
				63:60	4	x4 Base
		5	4	67:64	5	x8 Exp
				71:68	5	x4 Base
		6		75:72	6	x8 Exp
				79:76	6	x4 Base
		7	5	83:80	7	x8 Exp
				87:84	7	x4 Base
		8		91:88	8	x8 Exp
				95:92	8	x4 Base
5-8	2	9	1	19:16	9	x8 Exp
				23:20	9	x4 Base
		10		27:24	10	x8 Exp
				31:28	10	x4 Base
		11	2	35:32	11	x8 Exp
				39:36	11	x4 Base
		12		43:40	12	x8 Exp
				47:44	12	x4 Base
		13	3	51:48	13	x8 Exp
				55:52	13	x4 Base

		14		59:56	14	x8 Exp
				63:60	14	x4 Base
		15	4	67:64	15	x8 Exp
				71:68	15	x4 Base
		16		75:72	16	x8 Exp
				79:76	16	x4 Base

2.6 Management Subsystem

The management circuitry for the server shall use the ASPEED AST2500/2520 Baseboard Management Controller (BMC). This section describes requirement for the server management. Primary features include.

- BMC (ASPEED AST2520)
- BMC dedicated 1GbE LAN
- I2C support
- Thermal sensor for exhaust temperature monitoring
- Power monitoring
- Service and Debug LEDs

Figure 6 shows the management block diagram.

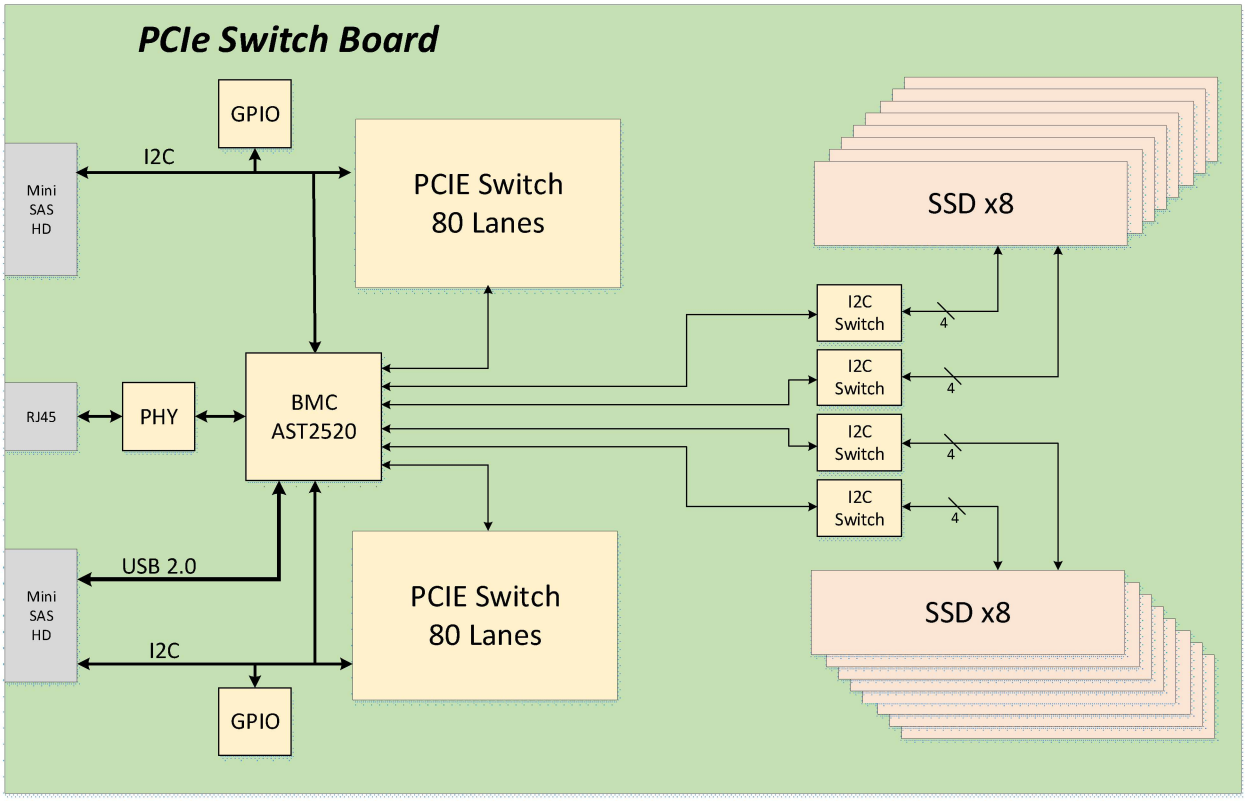


Figure 6. Management Block Diagram

2.6.1 BMC

The design for the BMC is based on the Aspeed AST2500/2520. Primary features include:

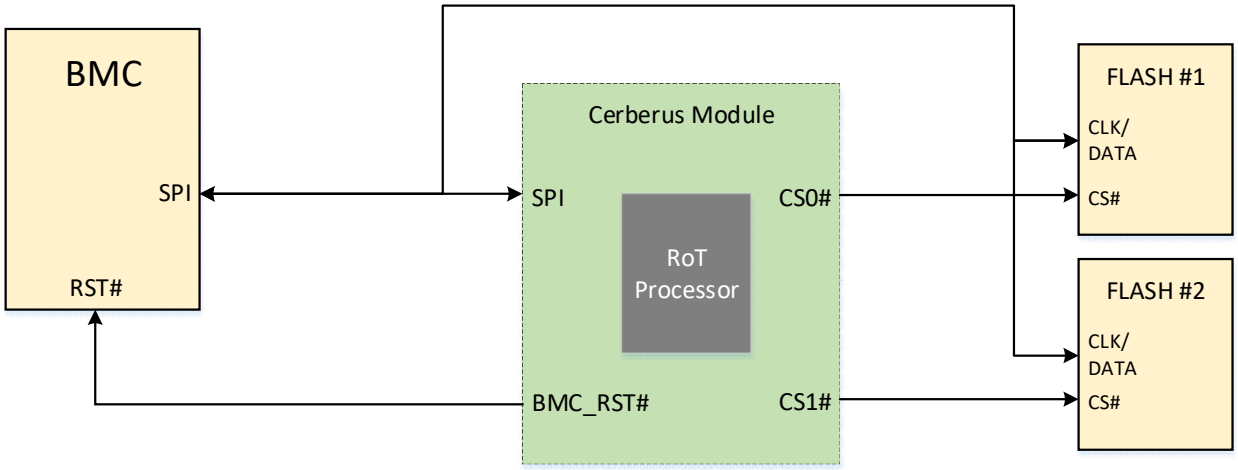
- Embedded ARM176JZF-S 32-bit RISC CPU, maximum 800MHz, 16KB Instruction cache and 16KB data cache
- Embedded 36KB SRAM buffer
- 16-bit DDR4 Memory up to 1024MB
- Two SPI flash memory clock speed up to 100MHz
- Two 10/100/1000Mbps Ethernet MAC
- Fourteen I2C/SMBus
- Six UART
- Eight PWM outputs
- Sixteen fan tachometer inputs

2.6.2 DRAM

The BMC requires a minimum of 512MB of DDR4 memory.

2.6.3 Cerberus Root of Trust Security

The management circuitry shall utilize a root of trust security module to ensure secure boot of the BMC and enable secure update of BMC firmware. A block diagram is shown below in



2.6.3.1 Cerberus Module

The design shall incorporate support for a root of trust security module. The module is based on a LPC540 ARM processor. The processor is responsible for ensuring secure BMC access to the BMC firmware and ensuring that the boot flash image is validated prior to de-assertion of reset to the BMC. The connector interface is defined in Section 3.4.

2.6.3.2 BMC Local Boot Flash

The design shall support a local boot flash made available to the BMC through a 1:2 mux. The device shall be 256Mb (32MB) minimum. Recommend using Windbond W25Q256 or equivalent. A secondary device shall be supported to provide BMC recovery. The purpose of the device is to provide boot capability during early design testing without the presence of the Security Module. This device shall be removed for production. The design shall support a socket for the device during preproduction.

The server shall support external management through a 1GbE PHY connected to the BMC. A RJ45 connector shall be located at the front of the server to provide Gbe connectivity to an external management switch. The PHY shall be Broadcom BCM54612E.

2.6.4 UARTS

The design shall support four UART connections to the BMC for managing the PCIe switches and debugging the PCIe. A block diagram is shown in Figure 7.

- UART3 – PCIe Switch #1 programming port
- UART4 – PCIe Switch #2 programming port
- UART5 – BMC Debug Console.
- UART 6 – BMC Debug Console (Debug Card Access)

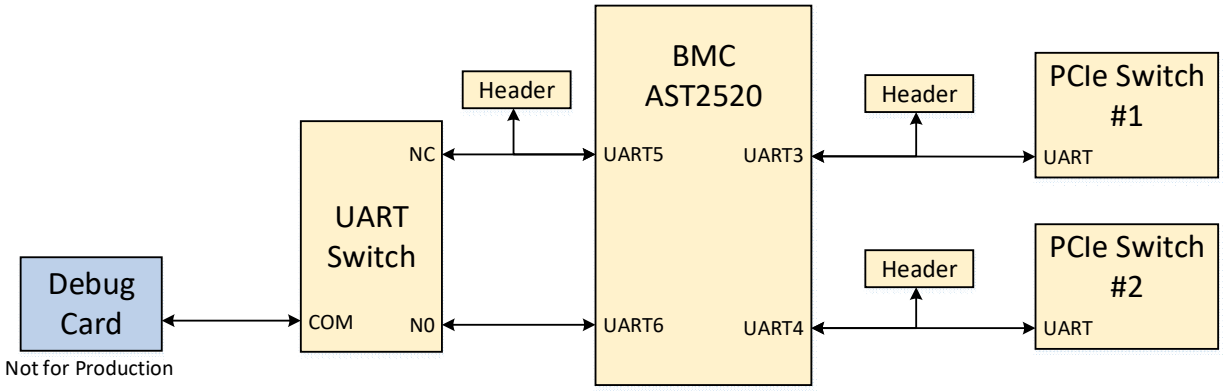


Figure 7. UART Block Diagram

2.6.5 I2C

The BMC shall support I2C communication to all SSDs as well as other local required components as shown in Figure 8. I2C switches are utilized to electrically isolate devices and enable deconfiguration of devices in the event of a device failure. All I2C switches shall support reset from the BMC to enable recovery. Note that addresses shown are 8-bit address with the R/W bit as the LSB set to 0 (0x48=0100100x). All I2C devices shall support 100Khz operation.

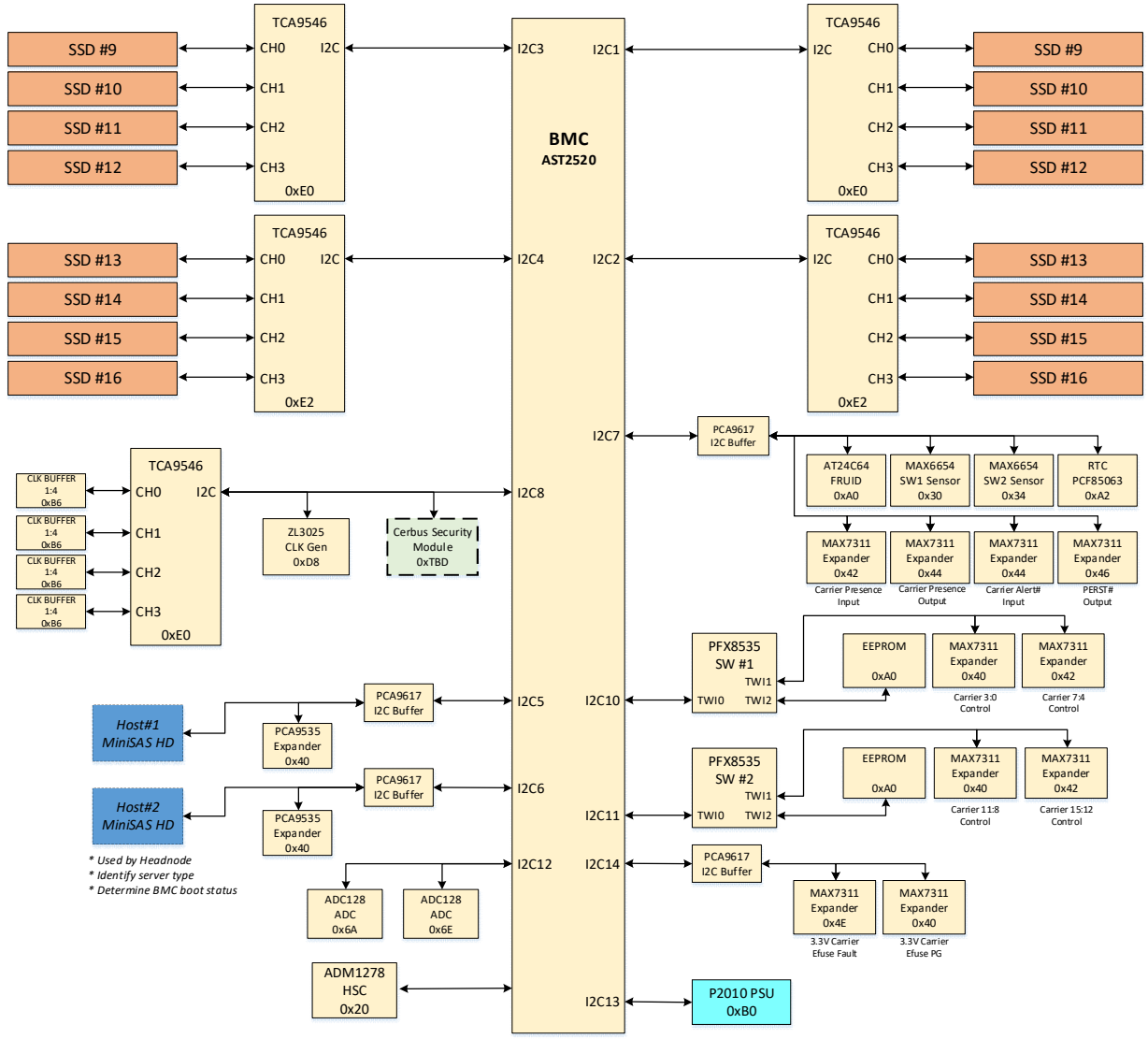


Figure 8. I2C Block Diagram

2.6.5.1 SSDs

The design shall support I2C communication with all hot pluggable SSDs. The devices shall be electrically isolated using I2C switches so that in the event of an I2C failure of an individual device, the device can be de-configured from the I2C path and I2C functionality recovered.

2.6.5.2 FRU EEPROM

The Rack Manager shall support a 64Kb FRUID EEPROM for storage of manufacturing data. The device shall be M24128-BWMN6TP or equivalent. Data in the EEPROM will be configured as an IPMI FRU device and fields are defined in the FRU ID Specification for the FX-16 Switch Board.

2.6.5.3 Thermal Sensor

The design shall include I2C support for temperature sensors, MPN TMP421 or equivalent, for

monitoring outlet temperatures of the board. The outlet temperature sensor is expected to be located at the back of the PCIe Switch PCBA closest to the fans.

2.6.5.4 I2C GPIO

The design shall support two I2C GPIO devices for communicating SKU ID and BMC boot status to the individual Hosts. The pinout is shown in Table 2.

Table 2. I2C GPIO

0x40	PCA9535C		
GPIO	**I/O**	**Signal**	**Description**
7	I	N/A	N/A
6	I	N/A	N/A
5	I	N/A	N/A
4	I	BOOT_COMPLETE#	0 = BMC Boot Complete
3	I	SKU_ID[3]	SKU_ID Bit 3
2	I	SKU_ID[2]	SKU_ID Bit 2
1	I	SKU_ID[1]	SKU_ID Bit 1
0	I	SKU_ID[0]	SKU_ID Bit 0

2.6.5.4.1 SKU ID

The GPIO includes SKU_ID setting to enable the Host to determine the type of appliance that it is communicating with. To identify itself as a FX-16 Flash Storage SKU, the SKU_ID should be set to 4'b0001.

2.6.5.4.2 BMC Boot Status

The GPIO includes a BOOT_COMPLETE# signal to enable the Host to determine the BMC boot status. The signal shall be pulled up on the PCIe Switch Board and driven low by the BMC once boot is completed.

2.6.5.5 PCIE Switch

The design shall support I2C communication with the PCIe switches to enable retrieval of PCIe switch status information and setting of switch configuration registers.

2.6.5.6 HSC Support

The design shall support I2C communication with all PMBus compliant HSC to enable retrieval of system power telemetry data.

2.6.5.7 Current Sense ADC Support

The design shall support A/D converters enabling current monitoring of 12V input to each of the SSDs.

2.6.6 Fan Control

The design shall include BMC fan control support for up to six fans located in the back of the server assembly. Two fans can fail and not affect system operation for N+2 support.

2.7 Power Capping and Throttling

The design shall support the PCIe PWRBRK# function enabling BMC controlled power capping and emergency throttling of each SSD. A block diagram of the logic is shown below in Figure 9. The circuit shall support throttling from two external sources:

- Local PSU PMBUs Alert
- External Throttle command from a PMDU (Power and Monitoring Distribution Unit)

The design shall enable the BMC to enable and disable each function as well as force a throttle event. Any throttle even shall be buffered and fanned out to all SSDs. At the time of this publication, this feature is unsupported by SSDs and is disabled by default.

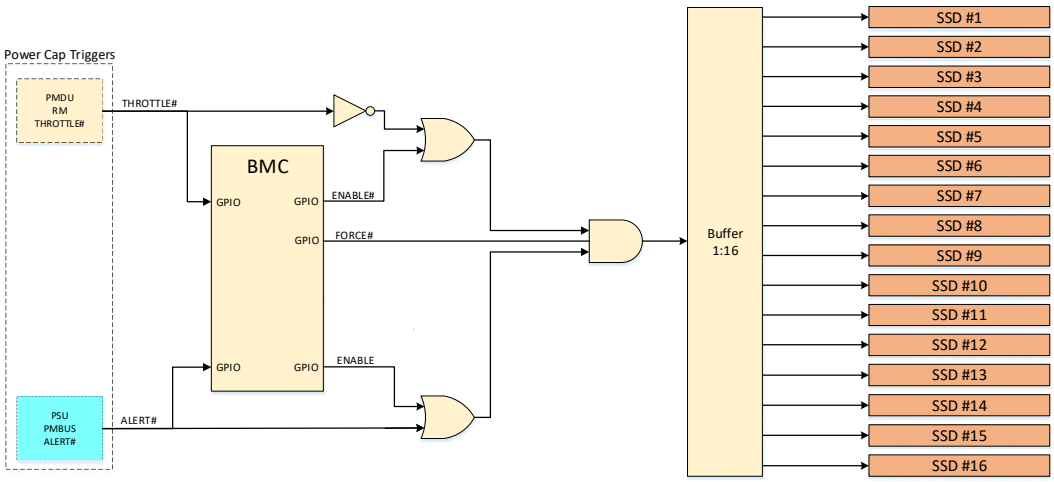


Figure 9. Power Capping Block Diagram

2.8 LEDs

2.8.1 Front Panel LEDs

The server shall support the following LEDs located on the front panel of the server.

Table 3. LEDs

LED Name	REF Des	Color	Description

UID LED	LED1	Blue 428-450nm d	Unit Identification LED – Location LED used to indicate that assembly requires servicing. When enabled this LED shall have a blink rate of 0.5 Hz at 50% duty cycle. • On = UID Enabled • Off = UID Disabled
Power Status LED	LED2	Green	Indicates Power Status of Blade • On = Switch Board Power is Good • Off = Switch Board Power is not Good
Attention LED	LED3	Red 618-635nm d	Attention LED – Indicates that an error has occurred • On = Error has occurred • Off = Normal Operation
1GbE Mgmt. Status	LED4	Green ~525nm λD	Indicates PCIe x16 Link 0 has connected to the PCIe Switch • On = 1GbE Mgmt port is up • Blinking = 1GbE Mgmt Port is active
1GbE Mgmt Speed	LED5	Green ~525nm λD	Indicates link speed of Management Port (located on RJ45) • On = 1Gbps • Off = 10 or 100Mbps

2.8.1.1 UID LED

Blue LED used to help visually locate a specific blade within a datacenter. This LED should be controlled by the BMC. A pictogram as referenced below shall be printed or metal-stamped adjacent to the unit identification LEDs. The text “UID” may accompany the pictogram.

Figure 10. Attention LED Symbol

2.8.1.2 Power Status LED

Green LED that indicates the power status of the server. LED should illuminate if Zone 0 power is good. A pictogram (IEC60417-2 symbol 5009) shall be printed or metal-stamped adjacent to the LED with no accompanying text.

Figure 11. Power Status LED

2.8.1.3 Attention LED

Red LED that indicates that a system error has occurred. This LED should be controlled by the BMC. A pictogram (ISO 7000, symbol 0434) referenced below shall be printed or metal-stamped adjacent to the LED with no accompanying text.

Figure 12. Attention LED Symbol

2.8.2 Debug LEDs

The design shall support the following LEDs visible on the PCIe Switch Board. These LEDs enable debug and are not required for production shipment.

Table 4. Debug LEDs

LED Name	REF Des	Color	Description
PCIe Attention LED	TBD	Green/Orange	Indicates PCIe connection status for each PCIe end point (SSD)
BMC Heartbeat	TBD	Green	Blinking LED that indicates that the BMC is running
PSU Status LED1	TBD	Green	Status LED for PSU1 • Solid Green = AC and DC Power Good • Blinking Green = Battery Power Good
PSU Status LED2	TBD	Amber	Status LED for PSU1 • Solid Amber = Failure of PSU Phase • Blinking Amber = Failure of 2 PSU Phases
PCIe Link 0 Status	LED6	Green ~525nm λD	Indicates PCIe x16 Link 0 has connected to the PCIe Switch
PCIe Link 1 Status	LED7	Green ~525nm λD	Indicates PCIe x16 Port 1 has connected to the PCIe Switch

2.8.2.1 PCIe Attention LED

Red LED driven by the PCIe Switch that indicates that an error has been detected to the corresponding SSD. The design shall support one Attention LED per SSD. The LED shall be placed uniformly near the corresponding SSD's connector.

2.8.2.2 BMC Heartbeat LED

Green LED that blinks to indicate that the BMC is alive.

2.8.2.3 PSU Status LEDs

Green and Amber LEDs that indicate the status of the Project Olympus Power Supply.

2.8.2.4 PCIe Link 0 Status

Green LED that indicates that PCIe Link 0 has connected to the PCIe Switch 0. This LED should be controlled by the PCIe Switch. This LED is for bring-up and test and is not expected to be required for production.

2.8.2.5 PCIe Link 1 Status

Green LED that indicates that PCIe Link 1 has connected to the PCIe Switch 0. This LED should be controlled by the PCIe Switch. This LED is for bring-up and test and is not expected to be required for production.

2.8.3 SSD UID LED

The design shall support control of a single LED through each of the SSD connector interfaces. The LED be controlled from the PCIe switch enabling both the BMC and the Operating System to control the LED. The UID LED on the SSD must be powered by the Switch Board so that UID can be visible when the SSD power is disabled.

2.9 PSU Interface

The design shall support Project Olympus rack and PSU management communication through the management cable connecting to the P2010 PSU. A block diagram is shown in Figure 13.

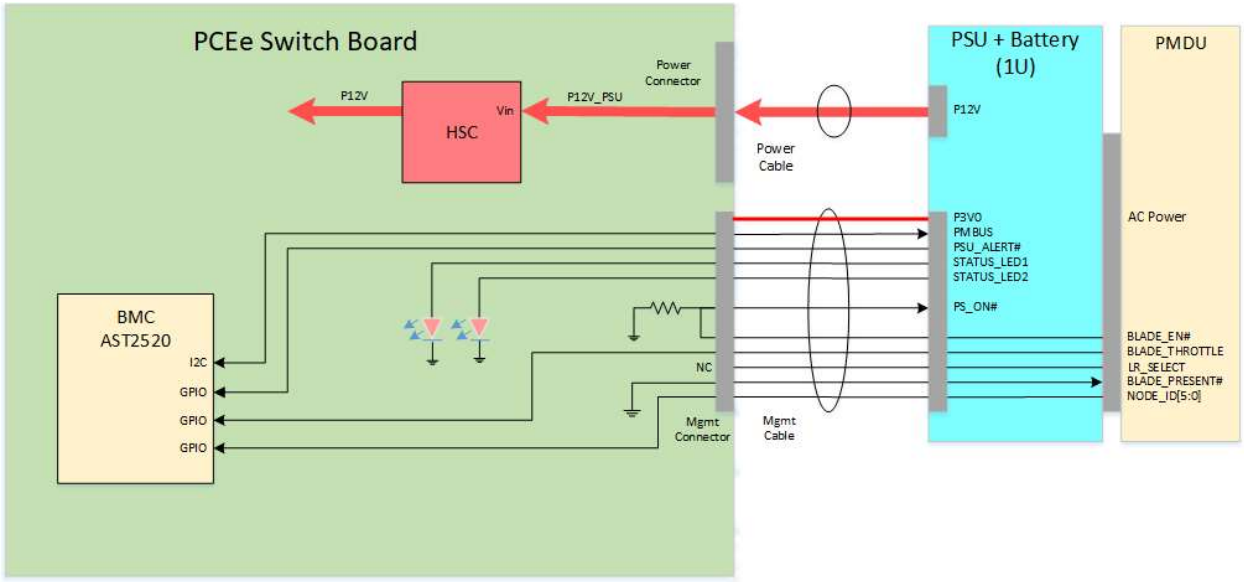


Figure 13. Rack Management Block Diagram

Below is a description of the signals.

- BLADE_EN# - Active low signal used to enable/disable power to the PSU. A pulldown resistor is used on the PCIe Switch PCBA to ensure a default low state if the Rack Manager is removed. This signal connects to the PS_ON# signal of the PSU. This signal is also connected to the BMC so that the BMC can disable error logging during shutdown.
- BLADE_PRESENT# - Active low signal used to communicate physical presence of the server to the Rack Manager. This signal should be tied to GND on the PCIe Switch PCBA.

- BLADE_THROTTLE - Active high signal used to put the server into a low power (power cap) state. This signal should default low (inactive) if the Rack Manager is not present. The Rack Manager drives this signal to multiple server slots. Therefore, the server must take care to electrically isolate this signal to avoid power leakage between servers. This signal should be connected to the BMC, but there is no current use case for this signal.
- SLOT_ID[5:0] – Identifies the PMDU physical slot that the server is plugged into. Each slot on the PMDU will report a different address. These signals should be connected to the BMC.
- PS_ON# - Active low signal used to enable/disable power to the PSU. This signal is driven by the Rack Manager. A pulldown resistor is used on the board to ensure a default low state if the RM is removed.
- PSU_ALERT# - Active low signal used to alert the BMC that a fault has occurred in the PSU. This signal is connected to the BMC for monitoring PSU status.
- PMBUS – I2C interface to the PSU. This is used by the BMC to read status of the PSU.
- STATUS_LED – Controls LED blade to provide visual indication of PSU status.

2.10 PCB Revision

The design shall support reading of a 3-bit PCB revision using GPIOs on the AST2520 processor. Revision shall be set using strapping resistors. The purpose of the ID bits is to communicate to firmware the PCBA assembly revision if different PCBA revisions require different firmware features to be enabled or disabled. At a minimum, the ID should be incremented with each revision of the PCBA.

Table 5. PCB Revision ID

ID	Description
000	EV
001	DV
010	PV
011	TBD
100	TBD
101	TBD
110	TBD
111	TBD

2.11 Power

2.11.1 Power Subsystem

The PCIe Switch PCBA shall be powered through a 12V connector. The design supports cabling to the Project Olympus PSU and can alternatively support 12V cabling to OCP Open Rack. Each SSD is isolated using an Efuse/HSC for both the 12V and 3.3V Aux power. The BMC and PCIe switches are isolated from

all the hot pluggable SSD through a separate Efuse/HSC. A block diagram of the power delivery is shown in Figure 14.

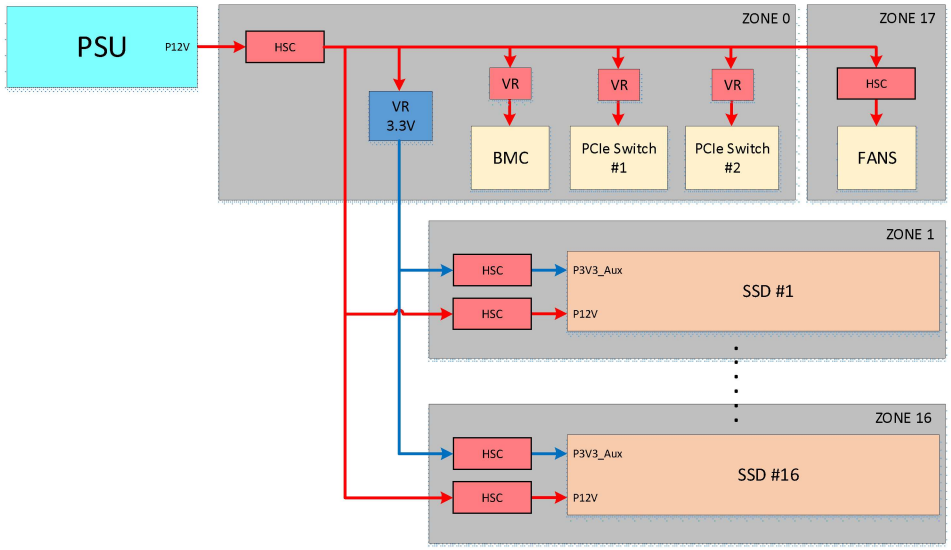


Figure 14. Power Block Diagram

2.11.2 Power-up/down Sequence

The power behavior of the FX-16 when paired with a head node shall conform to the requirement below

- FX-16 must be installed or powered prior to or coincident with install or power of the head node so that the head node BMC can detect the presence of and identify the FX-16.
- Powering the F2010 after the head node can result in failure of the head node to detect the PCIe endpoint devices (SSDs).

2.11.3 Grounding and Return

The chassis grounding/return is provided to the board from the front bezel assembly through the alignment and mounting holes that secure the board to PSU. The Rack Manager board ground is also tied to the PSU ground through the edge finger connector. Chassis ground and Logic ground are tied together on the board.

2.11.4 Power Monitoring

The design shall support monitoring the input power to each of the power zones. This includes the following:

- 12V input HSC enabling monitoring of total power of the chassis.
- 12V Fan HSC enabling monitoring of fan power.
- ADC monitoring of each 12V input to the SSDs.

2.12 PCIe Switch Board Stackup

The PCB stackup is shown in Figure 15. To enable visual tracking, the PCB color shall be different for EVT (red), DVT (blue) and PVT (green) builds. PVT build (green) shall be the production PCB color.

0.071", 8 Layers, TU863/IT170GRA/NPG171 Lead-Free / OSP / Tg>170 All dimensions in mils				Single Ended		Differential							
			Signals ->	Misc I/O; SE clocks		PCIe Diff Clocks; USB2/3;			LAN; MIDI			Signal Integrity	
			Z0->	50 Ohms		85 Ohms			100 Ohms			Insertion Loss	
Layer Name	Plane Description	Copper Weight (oz)	Layer Thickness (mil)	Width	Z0 tol.	Width	Space	Z0 tol.	Width	Space	Z0 tol.	Max SDD21 4GHz	Max SDD21 8GHz
	solder mask		0.5										
L1	SIGNAL	0.5 oz.	1.9	4.3	+/-10%	5.0	7.0	+/-10%	4.0	15.0	+/-10%	0.69	1.33
	prepreg		2.7										
L2	GND	1.0 oz.	1.2										
	core		4.0										
L3	SIGNAL / GND	1.0 oz.	1.2	5.2	+/-10%	5.5	6.0	+/-10%	4.8	12.8	+/-10%	0.65	1.25
	prepreg		15.0										
L4	POWER	2.0 oz.	2.4										
	Core		13.2										
L5	POWER	2.0 oz.	2.4										
	prepreg		15.0										
L6	SIGNAL / GND	1.0 oz.	1.2	5.2	+/-10%	5.5	6.0	+/-10%	4.8	12.8	+/-10%	0.65	1.25
	core		4.0										
L7	GND	1.0 oz.	1.2										
	prepreg		2.7										
L8	SIGNAL	0.5 oz.	1.9	4.3	+/-10%	5.0	7.0	+/-10%	4.0	14.0	+/-10%	0.69	1.33
	solder mask		0.5										
		9.0 oz.	71.0	mils									
			1.80	mm									

Figure 15. PCB Stackup

3 Connectors

This section defines the primary connectors and their signal interface requirements. The I/O direction of the signals is defined from the perspective of the Main Board.

3.1 PCIe x8 Card Edge

The design shall support 16 PCIe x8 card-edge connectors for supporting paddle card interfaces to the SSDs. The connector shall be compliant with the SFF-TA-1002 connector specification. The pinout is shown in Table 6. The pinout is compatible with the SFF-TA-1009 Enterprise and Datacenter SSD Pin and Signal Specification.

Table 6. PCIe x8 Connector Pinout

Pin	Signal	Signal	Pin
A42	RFU	PRSNT1#	B42
A41	GND	GND	B41

A40	PETp7	PERp7	B40
A39	PETn7	PERn7	B39
A38	GND	GND	B38
A37	PETp6	PERp6	B37
A36	PETn6	PERn6	B36
A35	GND	GND	B35
A34	PETp5	PERp5	B34
A33	PETn5	PERn5	B33
A32	GND	GND	B32
A31	PETp4	PERp4	B31
A30	PETn4	PERn4	B30
A29	GND	GND	B29
Key			
A28	GND	GND	B28
A27	PETp3	PERp3	B27
A26	PETn3	PERn3	B26
A25	GND	GND	B25
A24	PETp2	PERp2	B24
A23	PETn2	PERn2	B23
A22	GND	GND	B22
A21	PETp1	PERp1	B21
A20	PETn1	PERn1	B20
A19	GND	GND	B19
A18	PETp0	PERp0	B18
A17	PETn0	PERn0	B17
A16	GND	GND	B16
A15	REFCLKp1	REFCLKp0	B15
A14	REFCLKn1	REFCLKn0	B14
A13	GND	GND	B13
A12	PRSNT0#	PWRDIS	B12
A11	PERST1#	3.3VAux	B11
A10	LED	PERST0#	B10
A9	SMBRST#	DualPortEn#	B9
A8	SMBDAT	PWRBRK#	B8
A7	SMBCLK	MFG	B7
A6	GND	12V	B6

A5	GND	12V	B5
A4	GND	12V	B4
A3	GND	12V	B3
A2	GND	12V	B2
A1	GND	12V	B1

Table 7 provides additional descriptions of the signals and their use in this design.

Table 7. PCIe x8 Signal Descriptions

Signal Name	I/O	V	Notes
12V	O	12V	12V Power Output from HSC
3.3V Aux	O	3.3V	3.3V Aux Power Output from HSC
GND	O	0V	Ground
PET[7:0]	O	CML	PCIe Transmit from the PCIe Switch to the SSD
PER[7:0]	I	CML	PCIE Receive driven from the SSD to the PCIe Switch
REFCLK0	O	CML	PCIe Reference Clock to the SSD
REFCLK1	O	CML	Not used by this Design. This should be NC.
PERST0#	O	3.3V	PCIe Reset to the SSD
PERST1#	O	3.3V	Not used by this Design. This should be NC.
PRSNT0#	O	3.3V	This is connected to GND on the Switch Board.
PRSNT1#	I	3.3V	Active low signal. Indicates physical connection of the SSD to PRSNT0# and PRSNT1#. This signal is pulled high on the Switch Board.
DualPortEn#	O	3.3V	This should be pulled to 3.3V on the MB.
LED	O	3.3V	Driven by the PCIe Switch through a I2C GPIO. Drives UID on the SSD.
SMB_CLK	O	3.3V	I2C Clock to the SSD
SMB_DATA	I/O	3.3V	I2C Data to the SSD
SMB_RST#	O	3.3V	Management reset to the SSD
PWRDIS	O	3.3V	This should be pulled to GND on the MB.
RFU	I	N/A	Reserved for Future Use. Not used by this Design
PWRBRK#	O	3.3V	This function is not supported by vendor products and shall be buffered with a 3-state buffer so that this function can be disabled.

3.2 PCIe MiniSAS HD Connectors

PCIe and sideband signals are cabled to the assembly through two quad MiniSAS HD connector/cable assemblies. Each assembly consists of four PCIe x4 MiniSAS HD ports. The pinout for an individual port is shown in Table 8.

Table 8. MiniSAS HD Cable Connector Pinout

Pin	Signal		Signal	Pin
D9	GND	→	GND	B9
D8	TX2-	→	RX2-	B8
D7	TX2+	→	RX2+	B7
D6	GND	→	GND	B6
D5	TX1-	→	RX1-	B5
D4	TX1+	→	RX1+	B4
D3	GND	→	GND	B3
D2	USB_DP	↔	SB6	D2
D2	USB_DN	↔	SB5	D1
C1	SCL	→	SCL	C1
C2	SDA	↔	SDA	C2
C3	GND	→	GND	A3
C4	TX0+	→	RX0+	A4
C5	TX0-	→	RX0-	A5
C6	GND	→	GND	A6
C7	TX3+	→	TX3+	A7
C8	TX3-	→	TX3-	A8
C9	GND	→	GND	A9
B9	GND	→	GND	D9
B8	RX2-	←	TX2-	D8
B7	RX2+	←	TX2+	D7
B6	GND	→	GND	D6
B5	RX1-	←	TX1-	D5
B4	RX1+	←	TX1+	D4
B3	GND	→	GND	D3
B2	CABLE_PRSNT#	→	CABLE_PRSNT#	B2
B1	PERST#	→	PERST#	B1
A1	PCIE_CLKP	→	PCIE_CLKP	A1
A2	PCIE_CLKN	→	PCIE_CLKN	A2
A3	GND	→	GND	C3
A4	RX0+	→	TX0+	C4
A5	RX0-	→	TX0-	C5
A6	GND	→	GND	C6
A7	RX3+	→	TX3+	C7
A8	RX3-	→	TX3-	C8
A9	GND	→	GND	C9

3.3 RJ45 GbE Connectors

The design shall support a single RJ45 connector located on the bulkhead of the assembly for cabling 1GbE. The pinout for the connector is shown in Figure 5.

Table 9. RJ45 GbE Pinout

Pin #	Signal	I/O	Description
1	TX_D1+	O	Transmit Data +
2	TX_D1-	O	Transmit Data -
3	RX_D2+	I	Receive Data +
4	BI_D3+	I/O	Bi-directional Data +
5	BI_D3-	I/O	Bi-directional Data -
6	RX_D2-	I	Receive Data -
7	BI_D4+	I/O	Bi-directional Data +
8	BI_D4-	I/O	Bi-directional Data -

3.4 Cerberus Module Connector

The board shall support a single mezzanine connector FCI/Amphenol MPN- 61082-04 or equivalent. Figure 16 shows the physical layout of the connector. The pinout is shown in Table 10.

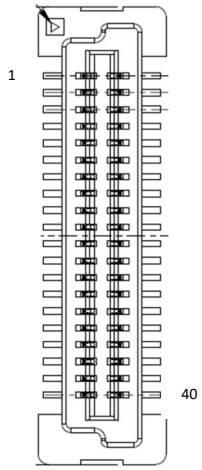


Figure 16. Mezzanine Connector Pin Numbering

Table 10. Mezzanine Connector Pinout

Left side-

Pin #	Signal name	I/O	Description
1	PORT0_SPI_SCK	O	Port 0 should be connected to the BMC
3	GND		Ground
5	PORT0_SPI_MOSI	O	Port 0 should be connected to the BMC
7	PORT0_SPI_CS0#	O	Port 0 should be connected to the BMC

9	PORT0_SPI_MISO	I	Port 0 should be connected to the BMC
11	PORT0_SPI_CS1#	O	Port 0 should be connected to the BMC
13	PORT0_RST_IND#	O	Should be AND logic of PORT0_RST_REQ0# (BMC reset request) and 1# (WDT timeout).
15	PORT0_RST#	I	Should be connected to BMC reset control logic
17	I2C0_ALERT#	I	Should be connected to BMC I2C alert pin
19	PORT1_RST_AUTH#	O	Should be connected to I2C Alert from CPLD
21	PON_RESET#	O	Should be connected to main board PON Reset/ 3.3V power good
23	PORT1_RST_REQ#	O	Should be connected to PLTRST# gated via CPLD
25	PORT1_RST#	I	Should be connected to CPLD RSMRST# logic
27	PORT1_SPI_QSPI3	I/O	Should be connected to PCH
29	PORT1_SPI_QSPI2	I/O	Should be connected to the PCH
31	PORT1_SPI_QSPI1	I/O	Should be connected to the PCH
33	PORT1_SPI_QSPI0	I/O	Should be connected to the PCH
35	PORT1_SPI_CS0#	O	Should be connected to the PCH
37	GND		Ground
39	PORT1_SPI_SCK	O	Should be connected to the PCH

Right side-

Pin #	Signal name	I/O	Description
2	PORT3_SPI_SCK	I	Should be connected to BMC Flashes
4	GND		Ground
6	PORT3_SPI_MOSI	I	Should be connected to BMC Flashes
8	PORT3_SPI_CS0#	I	Should be connected to BMC Flash #1
10	PORT3_SPI_MISO	O	Should be connected to BMC Flashes
12	PORT3_SPI_CS1#	I	Should be connected to BMC Flash #2
14	I2C0_SCL (BMC Master)	O	Should be connected to the BMC and debug header
16	I2C0_SDA (BMC Master)	I/O	Should be connected to the BMC and debug header
18	P3V3		Should be connected to 3.3V from main board
20	P3V3		Should be connected to 3.3V from main board
22	I2C1_SCL (RoT Master)	I	Should be connected to the CPLD and debug header
24	I2C1_SDA (RoT Master)	I/O	Should be connected to the CPLD and debug header
26	PORT2_QSPI_IO3	I/O	Should be connected to BIOS Flash (QSPI)
28	PORT2_QSPI_IO2	I/O	Should be connected to BIOS Flash (QSPI)
30	PORT2_QSPI_IO1	I/O	Should be connected to BIOS Flash (QSPI)
32	PORT2_QSPI_IO0	I/O	Should be connected to BIOS Flash (QSPI)
34	PORT2_QSPI_CS1#	I	Should be connected to BIOS Flash #1 (QSPI)
36	PORT2_QSPI_CS0#	I	Should be connected to BIOS Flash #2 (QSPI)
38	GND		Ground
40	PORT2_QSPI_SCK	I	Should be connected to BIOS Flash (QSPI)

3.5 I2C Debug Header

The board shall include a 3 pin I2C header on all I2C busses. This is a standard 100mil pitch 3-pin header. The pinout should be compatible with standard I2C debug solutions such as Beagle and Aardvark.

4 Component Labelling

4.1 Front View

Figure 17 shows a front view of the assembly demonstrating the label assignments for each feature.

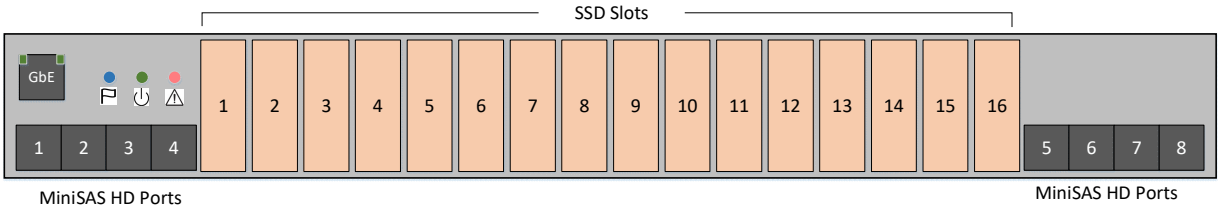


Figure 17. Component Labelling - Front View

4.2 Top View

Figure 18 shows a top view of the assembly demonstrating the label assignment for each feature.

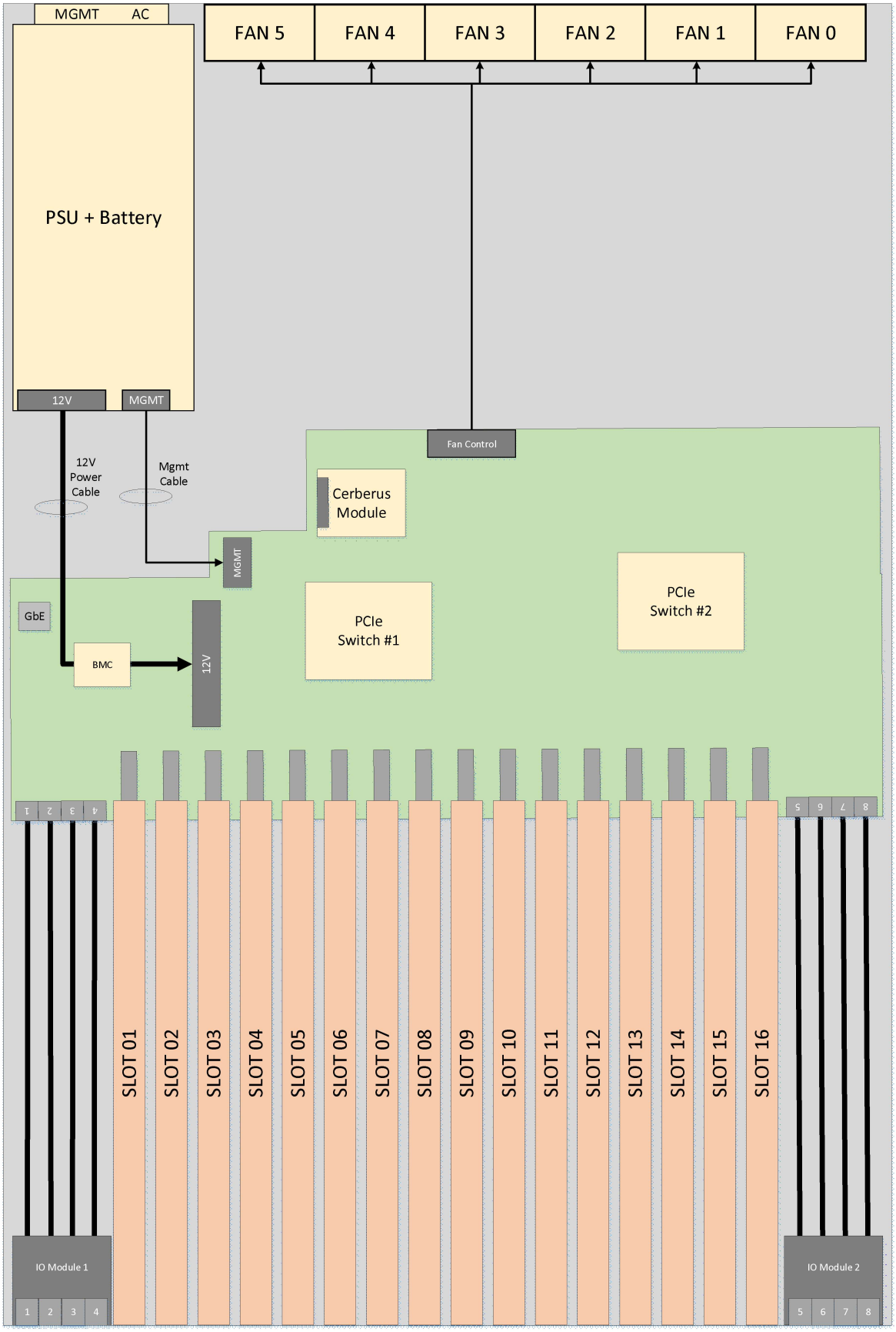

MGMT
AC
FAN 5
FAN 4
FAN 3
FAN 2
FAN 1
FAN 0
PSU + Battery
12V
MGMT
Fan Control
12V Power Cable
Mgmt Cable
Cerberus Module
MGMT
GbE
BMC
12V
PCIe Switch #1
PCIe Switch #2
SLOT 01
SLOT 02
SLOT 03
SLOT 04
SLOT 05
SLOT 06
SLOT 07
SLOT 08
SLOT 09
SLOT 10
SLOT 11
SLOT 12
SLOT 13
SLOT 14
SLOT 15
SLOT 16
IO Module 1
IO Module 2

Figure 18. Component Labelling - Top View

5 Electrical Specifications

The following sections provide specifications for the input voltage and current, as well as the primary blade signals.

5.1 Input Voltage, Power, and Current

Table 11 lists the nominal, maximum, and minimum values for the blade input voltage. The maximum and minimum voltages include the effects of connector temperature, age, noise/ripple, and dynamic loading.

Table 11. Input Power Requirements

	Minimum	Nominal	Max
Input voltage	11V DC	12.3V DC	14V DC
Input Power	<50W	430W	680W
Inrush Rise Time	5ms	n/a	200ms

12V power to the PCIe Switch Board is supplied through a 24 pin Mini-Fit Jr or equivalent connector. The server provides inrush current control through the 12V bus rail; return-side inrush control is not used. The inrush current rises linearly from 0A to the load current over a minimum 5 millisecond (ms) period (this time must be no longer than 200ms).

Assumptions for input power numbers are as shown below. This is provided for guidance purposes only as the actual numbers will be dependent on the SSD technology supported.

- Minimum assumes a base configuration of two SSDs with SSDs at idle power.
- Nominal power is based on a fully loaded configuration with 16 SSDs.
- Max power is based on the supported power of the P2010 PSU running at N.

5.2 Current Interrupt Protection and Power, Voltage, and Current Monitoring

The server provides a cost-effective way to measure and report blade voltage and current consumption, and to make instance reporting available to the system. The blade shall include power consumption measurements at the inrush controller. (Accuracy of the measurement shall be +/- 2%)

The blade also provides a way to interrupt current flow within 1 microsecond (µs) of exceeding the maximum current load at each hot swap control instance.

5.3 Filtering and Capacitance

The blade should provide sufficient input filtering and capacitance to support operation with the power supply as specified in the Project Olympus PSU Specification.

5.4 Grounding and Return

The server chassis grounding/return is provided to the PCIe Switch Board from the tray assembly through the alignment and mounting holes that secure the board to the tray. The board is also tied to the PSU ground through the 12V power connector. Chassis ground and Logic ground are tied together on the board.

6 Mechanical Specifications

This specification describes key mechanical elements of the assembly.

6.1 Server 2D mechanical control outline

Figure 19 shows the dimensions for the 1U assembly. The FX-16 mechanical dimensions are consistent with all Project Olympus 1U Severs.

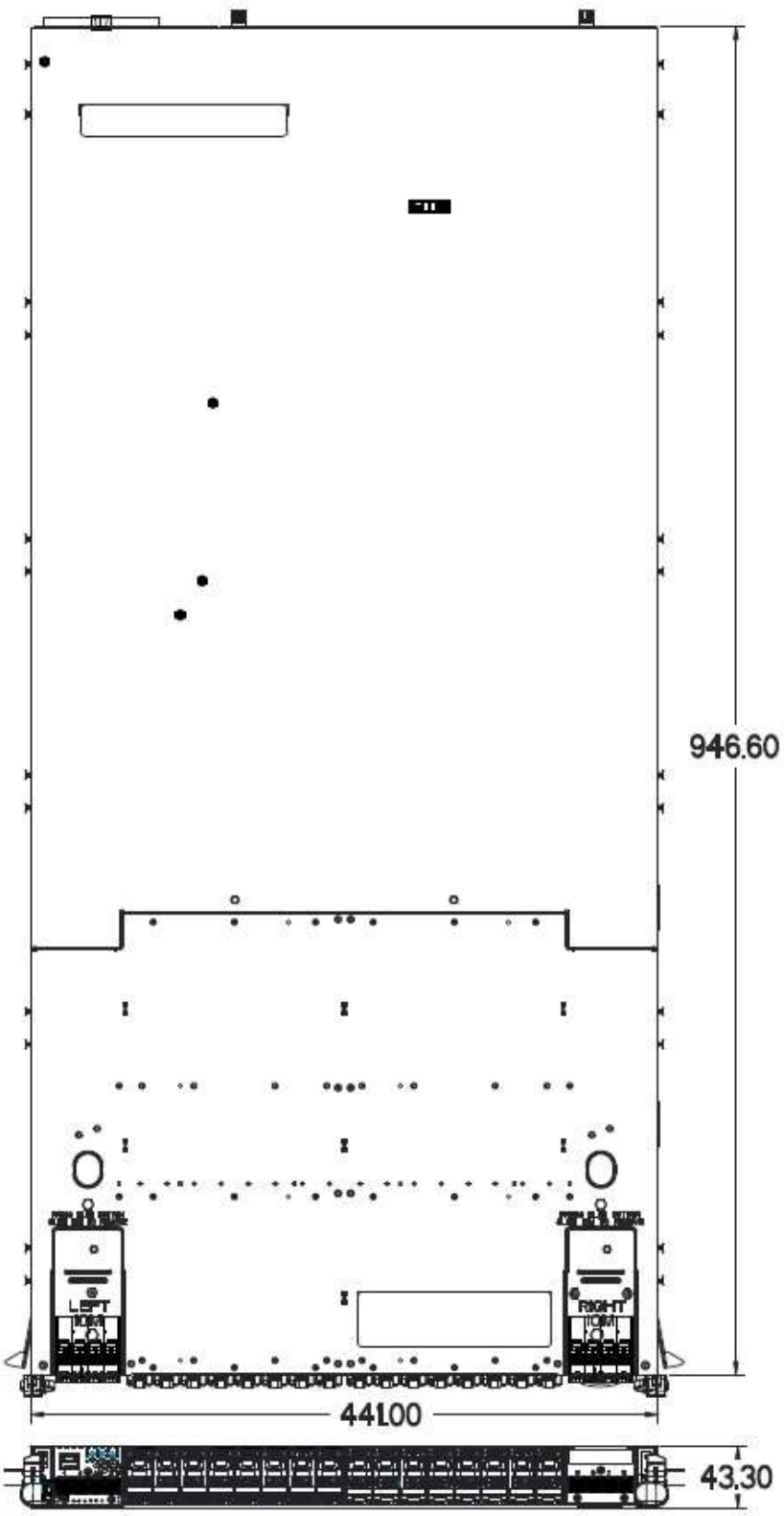


Figure 19. 1U Mechanical Dimensions

6.2 PCIe Switch Board 2D mechanical control outline

Figure 20 shows the dimensions for the PCIe Switch Board.

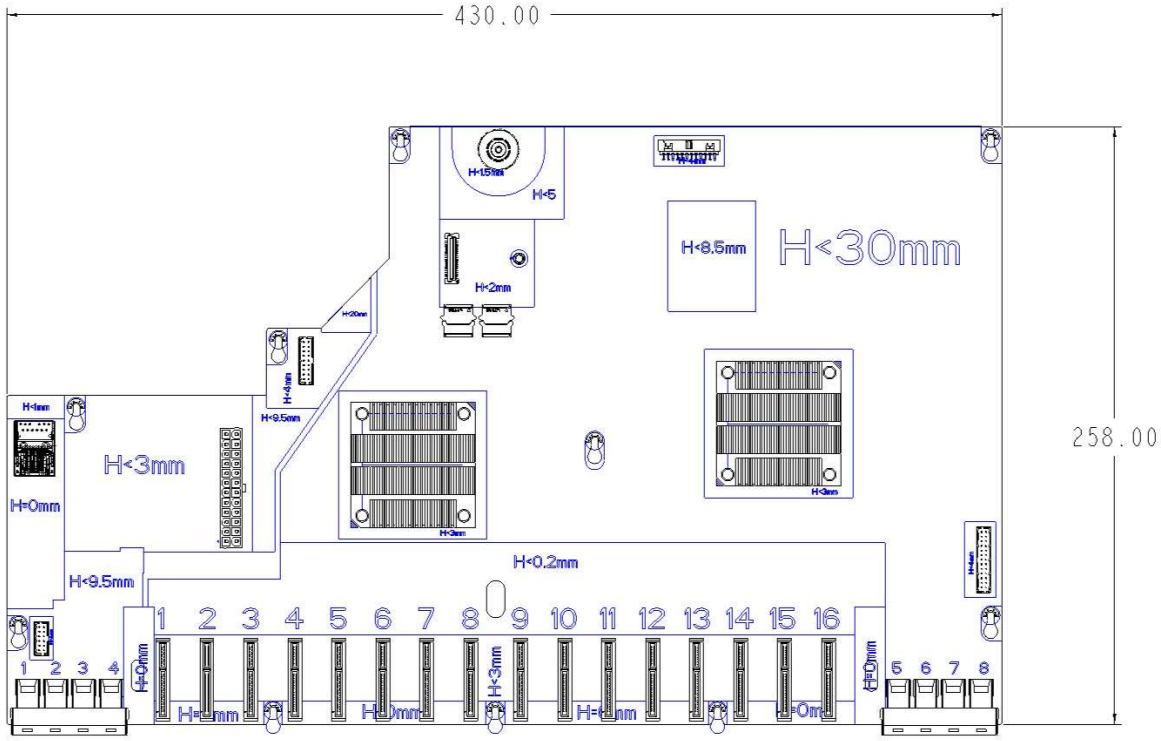


Figure 20. PCIe Switch PCB 2D Mechanical Outline

6.3 Paddle Board 2D mechanical control outline

Figure 21 shows the dimensions for the Paddle Board.

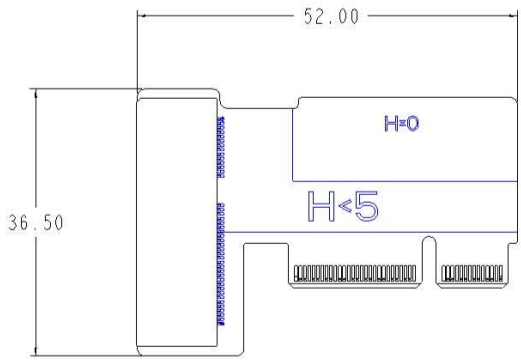


Figure 21. Paddle Board 2D Mechanical Outline

7 Environmental

7.1 Thermal

The server is to be deployed in an environmentally controlled location. The inlet to the server will be exposed to the environment described in Table 12. The server must have the capability to provide full functional operation under the conditions given.

Table 12. Thermal Requirement

Specification		Requirement
Inlet temperature	Operating	• 50°F to 95°F (10°C to 35°C) • Maximum rate of change: 18°F (10°C)/hour • Allowable derating guideline of 1.6°F/1000ft (0.9°C/304m) above 3000 ft.
	Non-operating	• -40°F to 140°F (-40°C to 60°C) • Rate of change less than 36°F (20°C)/hour

The server is required to use on-board fans to cool the electrical components. A maximum of six 40mm fans may be used to cool the components in 1U. The maximum airflow allowed in a server must not exceed 158 CFM/kW. The server must also operate at its maximum power configuration without performance degradation with 2 failed fans.

Variable fan speed capability allows the rack to minimize energy consumption of the air movers and facilities in conditions that permit it. The speed of airflow is based on component temperature requirements within the server. The maximum airflow available to a single EDSFF module is 2.2 cfm. The server has provisions for one airflow zone.

7.2 Shock and Vibration

The FX-16 shall meet the shock and vibrations requirements specified at the system level.

7.3 Regulatory

The FX-16 shall meet Safety (UL, IEC), EMC (EN and FCC Class A), Ecodesign, WEEE, and ROHS requirements.

8 Labels and Markings

The FX-16 shall have the following barcoded labels in visible locations where they can be easily scanned during integration.

- Vendor P/N, S/N, Rev
- Date Code (industry standard: WEEK/YEAR)
- Regulatory information and certification marks

9 Electromagnetic Interference Mitigation

For electromagnetic interference (EMI) containment, EMI shielding, and grounding must be accounted for at the unit assembly level. A top cover shall be installed to prevent leakage of electromagnetic fields.