OPEN
Compute Project

Facebook
Lightning Hardware System Specification
V1.0

Authors:
Mike Yan, Hardware Engineer
Clark Shao, Hardware Engineer
Chris Petersen, Hardware System Engineer

1 Contents

2 Scope

This document describes the hardware specification used in the design of Facebook's PCIe based storage system, code name "Lightning".

3 Overview

Lightning is a PCIe version of the Open Vault Storage platform of Open Compute Project. It supports up to 60 Solid State Drives (SSDs) connected via Gen3 PCIe links from an external host or hosts. Lightning is a PCIe JBOF (Just a Bunch Of Flash).

The Lightning board set resides in the same 2U chassis of OCP Honey Badger storage server. The Honey Badger Drive Plane Board is replaced with a PCIe Drive Plane Board (PDPB) that supports PCIe SSDs. The Honey Badger Baseboard and Micro Server Card are replaced with the Lightning PCIe Expansion Board (PEB). This PEB connects one or two sets of 16 lanes of PCIe Gen3 connections to one, two or four server hosts and uses PCIe Gen3 switches to fan out to the SSDs on the PDPB.

Outside of the PDPB and PEB, the rest of the storage server is unchanged. This includes the six fan modules and fan control board (FCB), the power transition board (PTB), the bus bar clip, and the tray cable arms, etc. The tray latches remain unchanged. The PEB will use the same card guides as Honey Badger Baseboard, while the card latches will be changed to accommodate the increased insertion force by the new connectors between PEB and PDPB.

The mechanical design of the chassis is mostly unchanged, so the 15 bays for 3.5-inch Hard Disk Drives (HDDs) are utilized to connect SSD Carriers that contain either a single U.2 SSD, or two M.2 modules. The SSDs are small form factor (SFF) 2.5-inch drives, and a carrier is used to hold the one U.2 SSD, or includes an adapter card to carry the two M.2 modules. The SSD carrier is mounted and retained in the chassis in a similar way as Honey Badger / Knox.

3.1 Reference Document

Below reference documents about Open Vault Storage and Honey Badger can be found on Open Compute Project website, storage track:
http://www.opencompute.org/wiki/Specs_and_Designs_Page#Storage

- Open Compute Project, Facebook, Open Vault Storage Hardware Specification, v0.8, April 16, 2014.
- Web link:
  http://files.opencompute.org/oc/public.php?service=files&t=3834b65de2b772fa4217c288ad5d1af1
- Open Compute Project, Facebook, Honey Badger – Light Weight Compute Module in Open Vault Storage, v0.8, Jan 8, 2015
  Web link:
  http://files.opencompute.org/oc/public.php?service=files&t=5dc72e32ba081bafoa8a99b46183888c

4 Lightning Hardware System Overview

4.1 Lightning Key Features

As a PCIe JBOF, Lightning system has below key features list:

- 16 lanes of PCIe Gen3 cabled to a single host (POR)
  - ✓ Mini-SAS HD (SFF-8644) cable, x4 from each connector, cable could be single x4 or "x8" (two x4 bound together)
  - ✓ A custom version cable is needed to carry PCIe Clock, Reset and other sideband signals (I2C, USB, etc.)
  - ✓ Host or hosts connection use a x16 PCIe expansion adapter (re-timer card)
  - ✓ Stretch goal: the design may support two hosts(8 lanes each), four hosts(4 lanes each) or up to eight hosts (32 lanes, x4 each)
- One or two SSDs per existing drive bay
  - ✓ NVMe compliant
  - ✓ Target power less than 14W for each drive bay
  - ✓ Only single port is supported per SSD (no dual-port support)
- SSD Carrier 1:
  - ✓ Single SSD, x4 PCIe Gen3
  - ✓ 2.5" form factor, 15 mm thickness
- SSD Carrier 2:
  - ✓ Single SSD, x4 PCIe Gen3
  - ✓ 2.5" form factor, 7 mm thickness
- SSD Carrier 3:
  - ✓ Two M.2 flash modules with adapter card, x2 PCIe Gen3 to each
  - ✓ Support 22110 M.2 form factor
  - ✓ This dual-M.2 Carrier is implemented as a single FRU for PCIe hot-plug
- PCIe switch supports:
  - ✓ PCIe hot-plug: add, remove, replace, and surprise add/removal
  - ✓ PCIe Downstream Port Containment (DPC)
  - ✓ PCIe Extended Downstream Port Containment (eDPC)
  - ✓ One host (x16) as the upstream of the switch
  - ✓ Stretch goal: two hosts (x16 each) or four (x4) hosts as the upstream of the switch
  - ✓ Up to 30 SSD devices are downstream of the switch, x2 to each device; Or up to 15 SSD Carriers, x4 to each drive bay
  - ✓ Each SSD Carrier is assigned and mapped to only one host (i.e. there is no "pooling" of storage)
  - ✓ There is no host to host communication through the switch
- A BMC running OpenBMC firmware on the PEB will provide the enclosure management functionality on each Lightning tray and the chassis
  - ✓ One BMC for each PEB, is responsible for tray level management functionality
  - ✓ Two BMCs on two trays work together for chassis level management functionality, in the same way as Open Vault Storage (Knox)
  - ✓ Each BMC will be accessible by the host CPU subsystem (in-band management), through a USB bus routed on the customized version of Mini-SAS HD cable; for current OCP server, there will be a USB cable between the rear USB connector on server motherboard and a USB connector on the PCIe re-timer card

- ✓ ASPEED AST2400 BMC, with a PCIe x1 from BMC to PCIe switch. The BMC is an endpoint device only
- ✓ A 256MB DRAM is used for compatibility with other Facebook projects that implement OpenBMC
- ✓ Stretch goal: support host I2C to reset Lightning BMC MCTP protocol as an in-band management approach between host / hosts and BMC
- ✓ Stretch goal: Each BMC could support up to 4x host connections using PCIe (MCTP) connections later

- Mechanically, Lightning tray is compatible with Honey Badger tray that accepts Baseboard, which also supports Knox tray configuration with single SEB on A side. Same sheet metal shall work for all 3x designs, except:
  - ✓ Due to the increased insertion force by the new connectors between PEB and PDPB, the card latches and jack screws in the front will be changed accordingly
  - ✓ HDD latches and "Z" bracket have been removed as they are not necessary
  - ✓ Most of the screw holes on PDPB are changed to T-slots to improve serviceability in data centers
  - ✓ Separate carriers have been re-designed for 7mm U.2 and M.2 22110 form factors
- System airflow "CFM per Watt" shall not to exceed the previous generation (Honey Badger light weight Storage Server)

Figure 1 and Figure 2 shows an overview of the Lightning hardware system.

Figure 1 Lightning Hardware System Overview

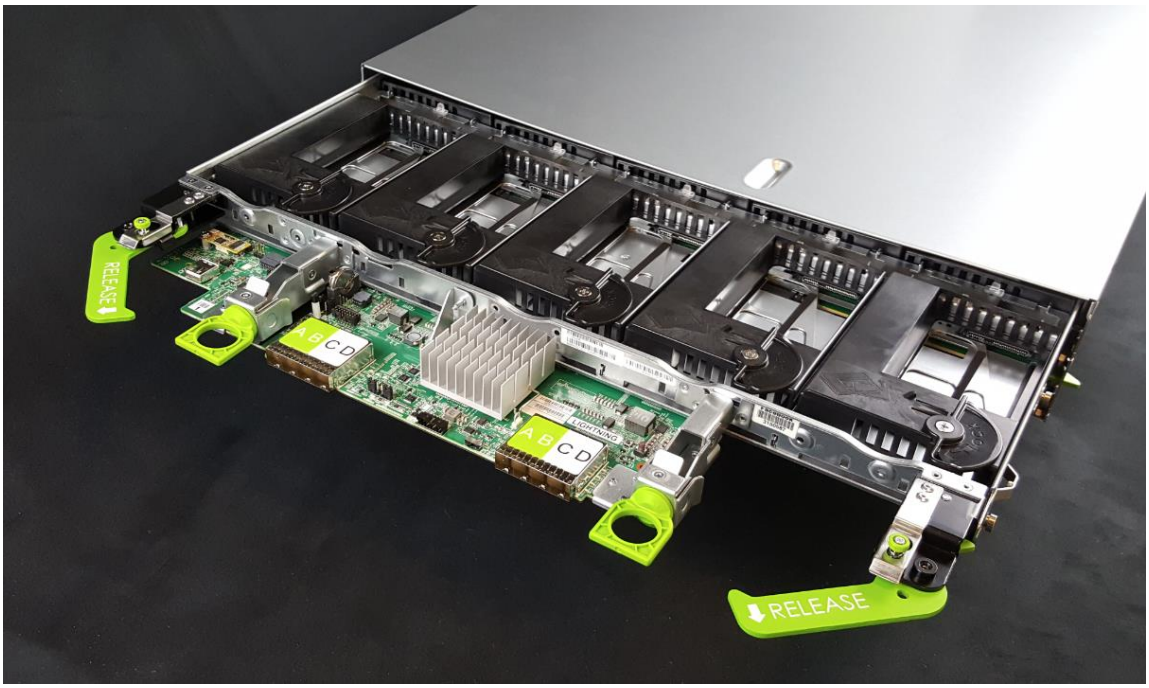


Figure 2 Lightning System with One Tray Pulled Out

4.2 Key Accessible Items

List of front accessible Lightning Hardware System components:

- Two SSD Trays
- Two PCIe Expansion Boards, one for each tray

The rest of accessible equipment resides on the PEB, and for each PEB it includes below items:

- Eight Mini-SAS HD Connectors, SFF-8644 form factor
- Two Status LEDs for system information
- Two heart beat LEDs for indication of BMC and PCIe switch functionality
- Eight LEDs for uplink indication
- Stretch goal: Six LEDs for switch zoning indication
- One OCP Debug Header
- One RJ-45 connector for 1GbE port, not populated
- One USB 1.0 port, reserved for BMC debug during development
- Two UART headers to access BMC and PCIe switch separately
- One EJTAG header for PCIe switch debug during development

Figure 3 Lightning PEB with Accessible Items

4.3 System Component Layout

Figure 4 shows the major system components layout from top-down view of the Lightning hardware system.

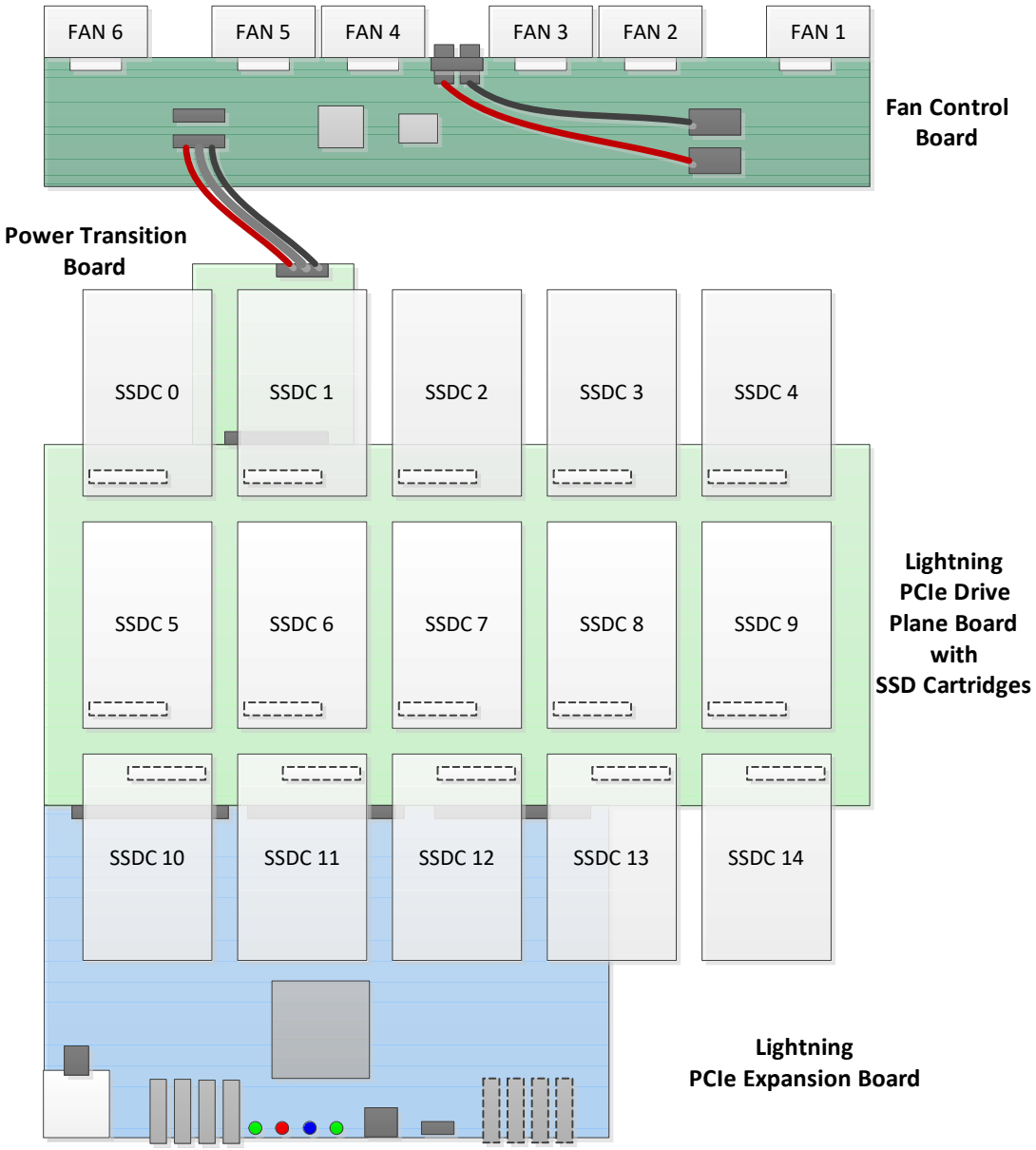


Figure 4 Lightning System Components Layout

4.4 Host Connection Block Diagram

Figure 5 shows the overall concept of the Lightning host connections. Each of the custom Mini-SAS HD cables carries the following signals:

- PCIe Gen3 x4 lanes
- PCIe sideband signals: Clock and Reset
- USB bus for BMC management
- I2C bus as backup connection (not POR)

Mini-SAS HD cable length in planning:

- 1.5 meters

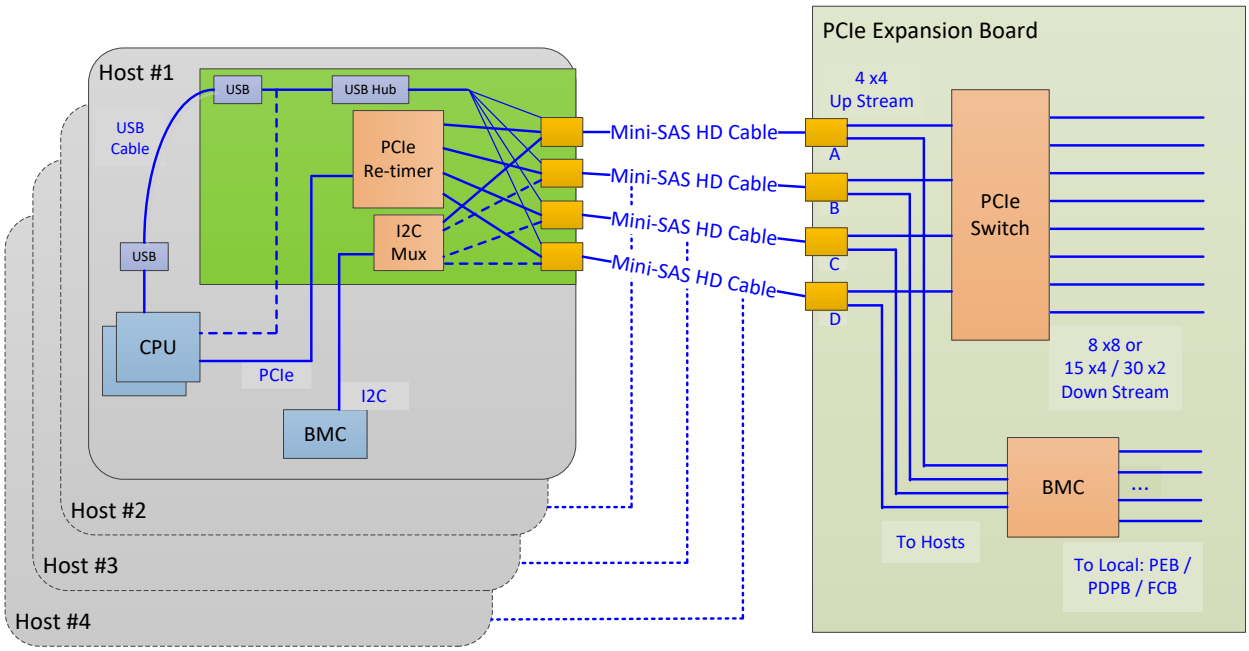


Figure 5 Lightning to Host Connection Block Diagram

4.5 PCIe Block Diagram

The Lightning PEB which will use a 96 lanes Microsemi PCIe switch.

Figure 6 shows the PCIe block diagram, which uses the Microsemi PM8536 PCIe switch. It's a single chip solution that supports 16 lanes or 32 lanes upstream to the host, and 60 lanes downstream to the SSDs. On the downstream side, it can be configured as x4 lane ports to support 15x U.2 SSDs, or configured as x2 lane ports to support up to 30x M.2 SSDs. An extra x1 downstream PCIe lane is used to connect to BMC for in-band programming. Figure 6 only shows the concept of x4 PCIe lanes for U.2 and x2 PCIe lanes for M.2. In practice, only one form factor will be supported per tray.

On the upstream side, an extra group of four Mini-SAS HD connectors is reserved on PEB to support an additional x16 uplinks, up to x32 uplinks. If there's no relevant production needs, the PM8535 may be substituted as the 80-lane chip provides enough lanes for one x16 uplink, and the 2nd four ports of SFF-8644 connectors can be de-populated. PCIe signal integrity measurements demonstrate that a PCIe re-driver is not needed on PDPB.

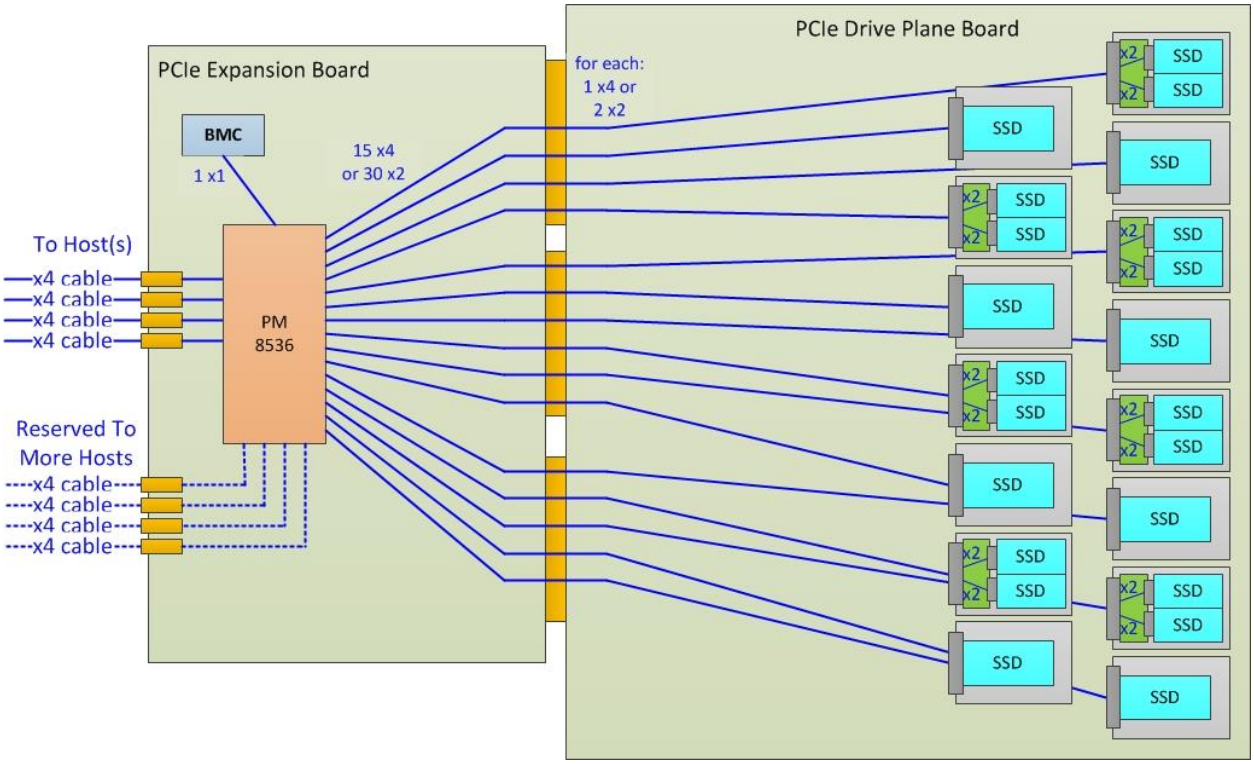


Figure 6 Lightning PCIe Block Diagram

4.6 System I2C Topology

Figure 7 shows the system I2C topology of Lightning. This mainly reflects the enclosure management structure of the Lightning hardware system. The BMC chip has a total of 14 I2C buses. A summary of the I2C bus segments implemented on the Lightning PEB and system is as follows:

Common portion:

- I2C_A: Combined with I2C_D, both are legacy from Honey Badger. It is used for PEB local sensors and components.
- I2C_B: Legacy from Honey Badger. It is for PDPB temperature and voltage sensors, and FRU EEPROM.
- I2C_C: Legacy from Honey Badger. It is used for shared control of the FCB by both BMCs in the chassis.
- I2C_Mini-SAS_1/2/3/4: These are used for host connections, and 1 per MiniSAS-HD connector.
- I2C_8 and I2C_9: they are used for all of the SMBus connections with each SSD(M.2 or U.2), and also used for clock buffers on the PDPB.
- I2C_10: It is used for the connection with the PCIe switch as a slave device.
- I2C_Mini-SAS_11/12/13/14: These are used for host connections and 1 per Mini-SAS HD connector.

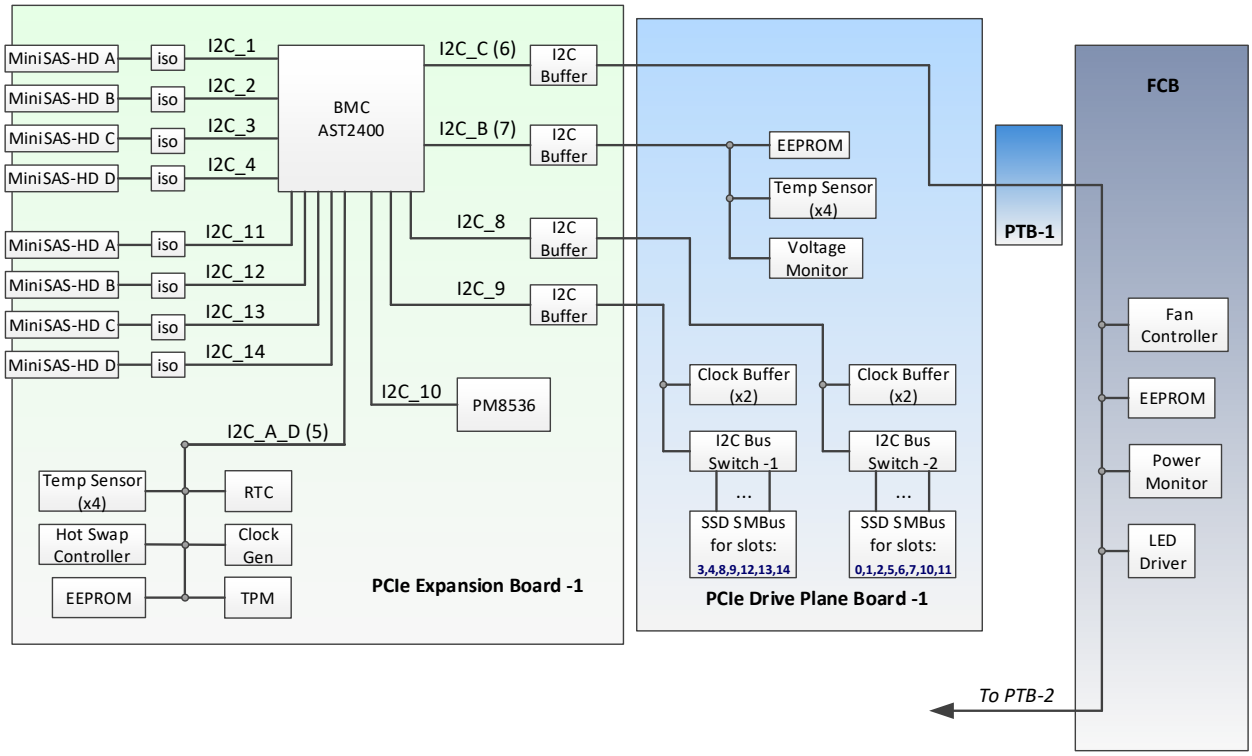


Figure 7 Lightning System I2C Topology

4.7 PCIe Clock Tree

Figure 8 shows the PCIe clock topology. A clock generator is provided on the PEB to enable a non-common clock solution. The clocks going to the PDPB for all SSDs are driven directly by the PCIe switch.

- Four clock buffers on the PDPB will be needed to fan out from 4x clocks to 15x / 30x clocks.
- A non-SSC clock mode is supported but default is a SSC common clock provide by the host.
- Due to the limitation of PM8536 that a single stack can only be connected with up to 3 pairs of host reference clock, for the scenarios of either 4 hosts or 8 hosts, only local reference clock on PEB will be used to provide reference clock to PM8536 as SRIS or SRNS.
- The re-timer chip on the host side currently only supports common clock (with or without SSC), and separate clocks without SSC (SRNS). As a result, SRIS is not supported as of 2017.

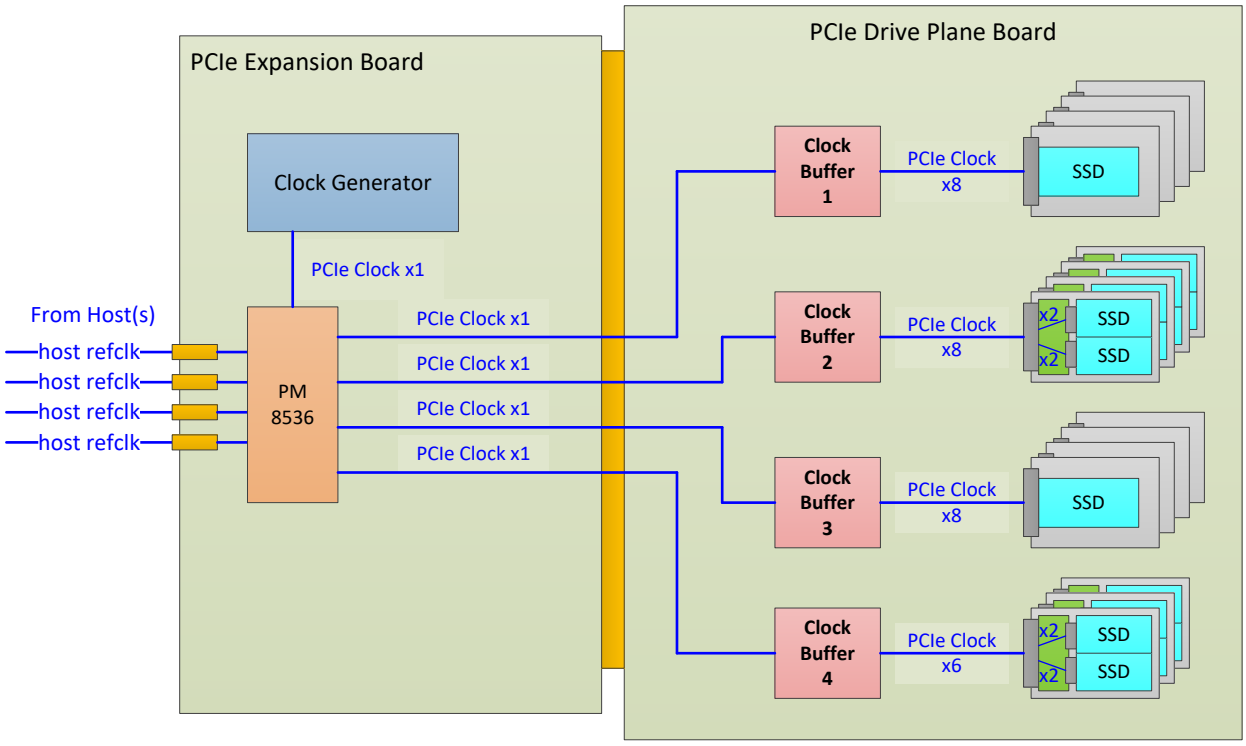


Figure 8 Lightning PCIe Clock Tree

4.8 PCIe Sideband Signals

PCIe sideband signals include PERST#(PCIe reset), IFDET#(Presence), ATNLED#(Attention LED) and PWREN(Power Control). All sideband signals should be connected to the PCIe switch.

Figure 9 shows the sideband signal connections. Since the GPIO expander pin assignment is configurable, 4 GPIO expanders are enough for a 15x SSD solution, or for FRU (slot) level hot-plug of dual SSD solutions.

More considerations:

- To create a switch-agnostic solution and keep a common PDPB design, all of the GPIO expanders must be placed on the PEB.
- The switch needs to properly handle the dual-M.2 adapter where the two x2 M.2 drives share the same ACTLED and IFDET# signals using “OR” gates. PWREN is also shared by the two M.2 drives on the carrier.

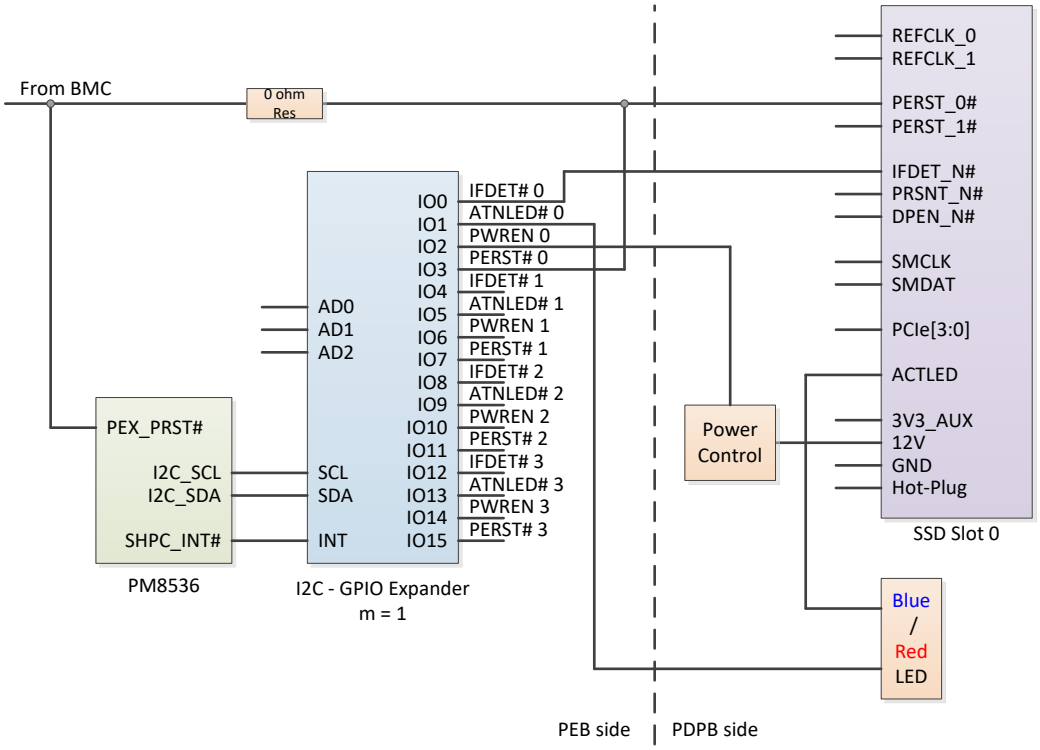


Figure 9 SSD Sideband Signals

4.9 PCIe Reset Signals

The PCIe reset signals to each SSD are shown above in 0. There are three approaches to driving PCIe reset to each SSD slot: from hardware power on reset(optional), from BMC GPIO(optional) and from host via the PCIe switch(by default). “AND” logic is used to connect those three approaches together and by default hardware power on reset and BMC GPIO are disabled by de-populating oohm resistors.

The hardware power on reset is added as an option in case a local PERST# is needed by switch.

The BMC GPIO PERST# is provided for debug purposes during development.

4.10 PCIe Lanes and Polarity Reversal

PCIe lane reversal and differential pair polarity reversal must be supported, but should be implemented carefully. For example, if lane reversal is implemented it must be done in both TX and RX directions for the same PCIe lane per port.

4.11 SSD Carrier Power Control

Power control capability needs to be implemented on a x4 slot basis. The switch needs to be able to power on or off each U.2 SSD or each SSD carrier holding two M.2 SSDs. This allows the user to power cycle each SSD(or pair of SSDs) individually if needed. When the system powers on or an SSD is inserted into a powered system, the SSD must be powered up automatically and the PCIe link between the SSD and the host must be trained successfully in both cases.

5 Lightning PCIe Expansion Board

5.1 Block Diagram

Figure 10 illustrates the functional block diagram of the Lightning PCIe Expansion Board (PEB), using the Microsemi PM8536 PCIe Gen3 switch.

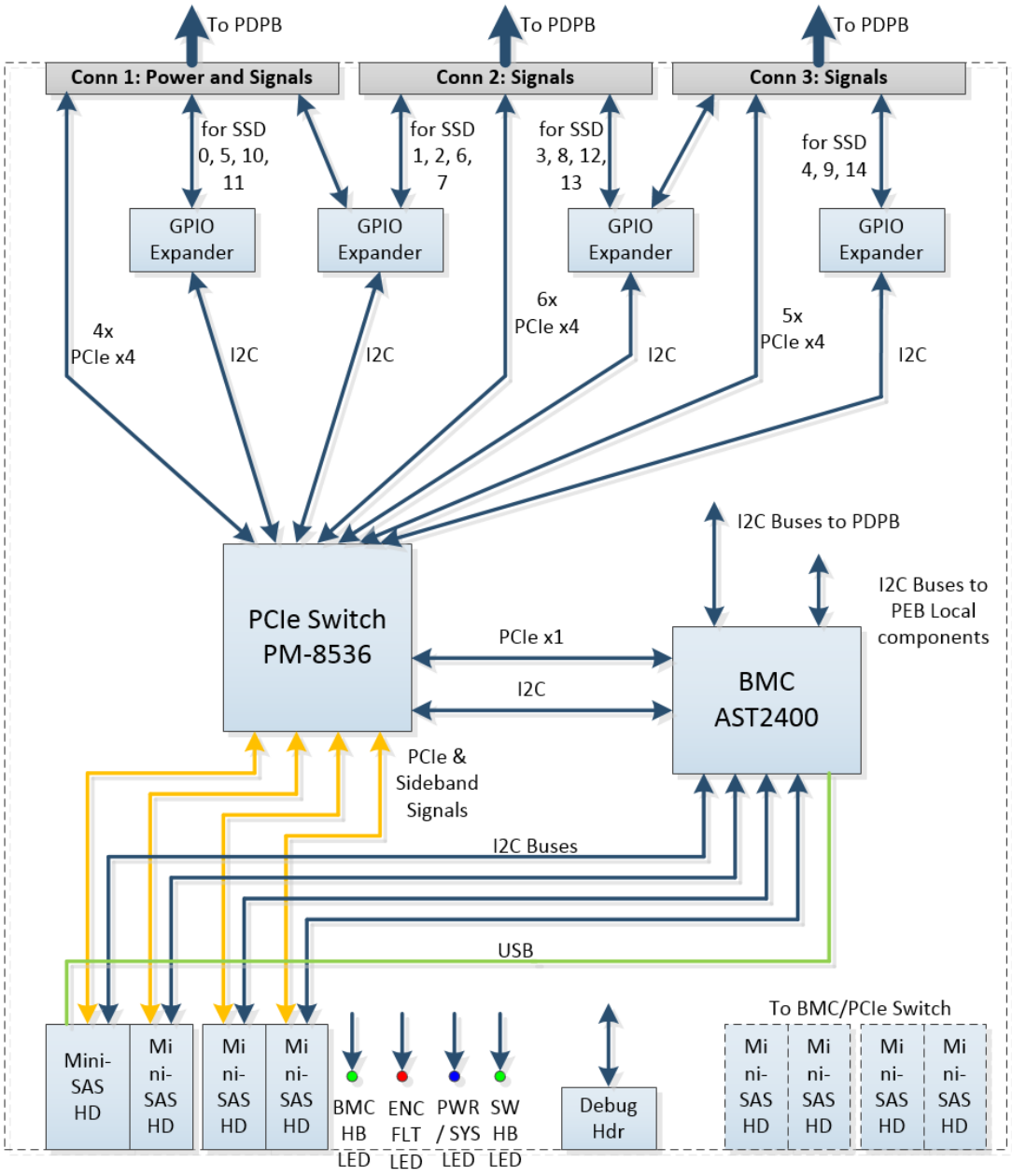


Figure 10 Lightning PEB Block Diagram

Notes:

- By default, the 4 Mini-SAS HD connectors on the left are used to connect the Lightning chassis to the host;
- In addition to the PCIe connection between switch and the host/SSDs, there's a 1-lane PCIe connection between the switch and the BMC for BMC FW in-band update;
- 4 GPIO expanders are used for sideband signal control from the switch to the SSDs
- USB and I2C interfaces are used for communication with Lightning BMC from host.

5.2 PEB Form Factor and Placement

The PCIe Expansion Board form factor is a 355.6 mm x 233 mm board. Figure 11 illustrates the outline and keep out area, etc. Note that the goal is to minimize the depth of the board since it is not anticipated that the depth required of Honey Badger will be required for the Lightning PEB.

Placement and routing is heavily restricted by the compatibility requirements of the main Lightning system. The PEB will need to slide into the chassis using the same guide rails as the Honey Badger Baseboard. In addition, the space available on the board for external connectors, PCIe switch, and other larger components is restricted to the section of the board that clears the front of the chassis, since the board space that sits under the chassis sheet metal is severely restricted in terms of component height and airflow, etc.

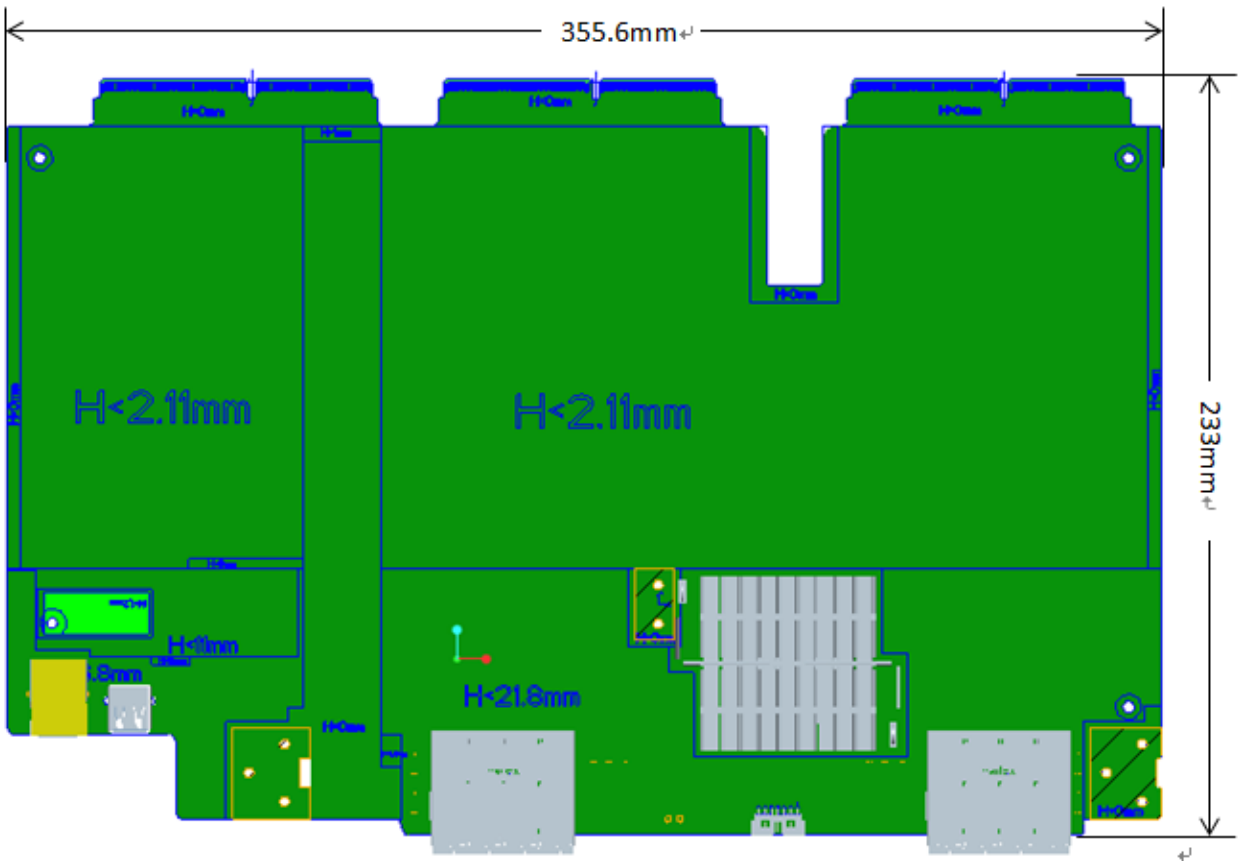


Figure 11 Lightning PEB PCB Dimension

Figure 12 shows the key components placement and layout analysis for PCIe port assignments, lane directions, etc.

Main considerations:

- The PM8536 location and orientation should be considered to: (1), Minimize PCIe trace length to edge connectors, so that the total PCIe trace length together with the portion on PDPB will be as short as possible; (2), Optimize PCIe trace length to the Mini-SAS HD connectors to Host(s).
- The primary four Mini-SAS HD connectors are placed to the left side of the PEB, while the secondary four Mini-SAS HD connectors are placed to the right side of PEB to ensure optimal external cable routing.
- The debug header and LEDs are placed in the middle portion of the PEB
- The RJ-45 connector is placed on the left portion of the PEB(depopulated by default)
- The USB connector is placed on the left portion of the PEB

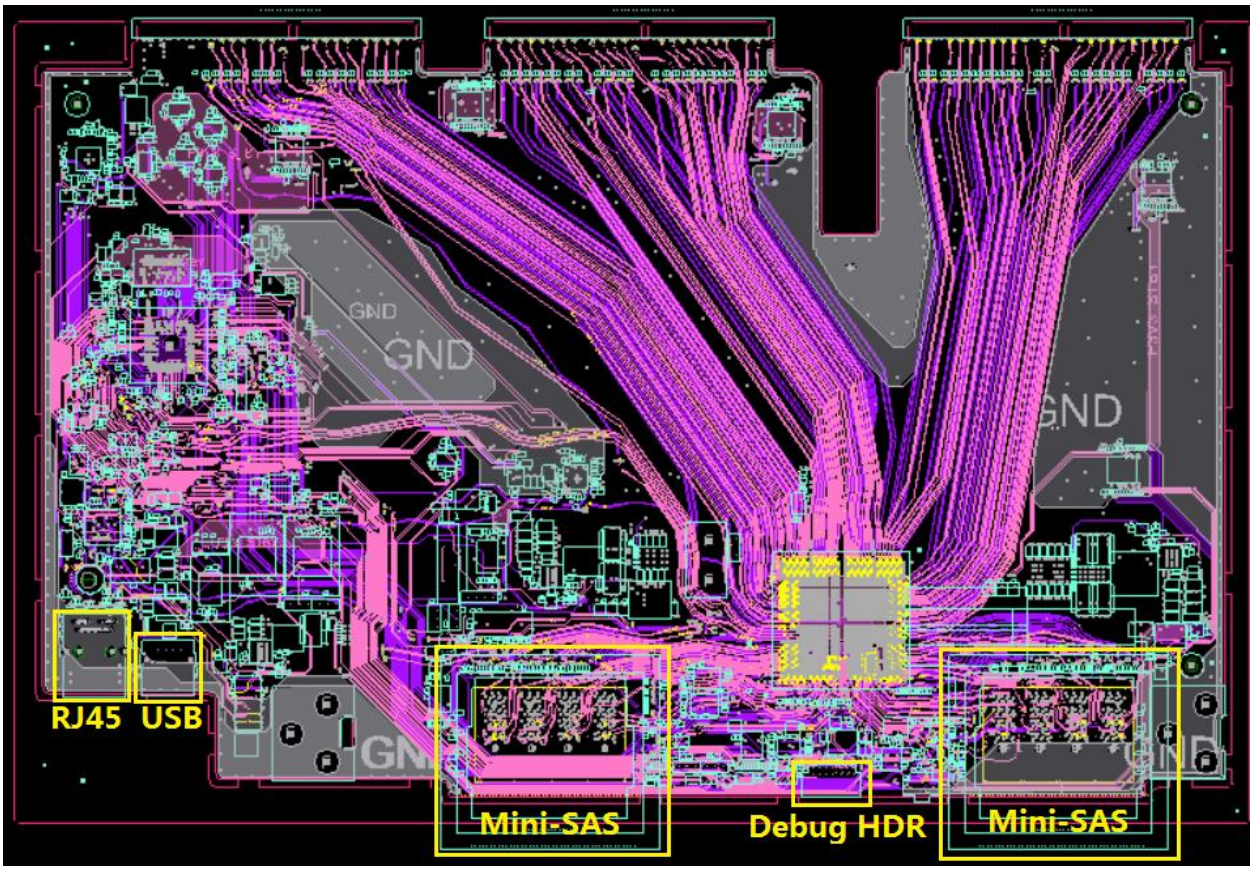


Figure 12 Lightning PEB Placement

5.3 PCIe Switch IC

Lightning is using PM8536 as the PCIe switch on the PCIe Expansion Board. The PM8536 PFX 96xG3 PCIe Gen3 Fanout Switch is a 96-lane, 48-port, resilient, PCIe Base Specification 3.1-compliant switch supporting up to 24 transparent switch partitions, 48

Non-Transparent Bridges (NTBs), hot- and surprise-plug controllers per port, advanced error containment, and comprehensive diagnostics and debug capabilities.

Applications for the PFX series include PCIe SSD enclosures, flash arrays, multi-host architectures, high-density servers, blade servers, and applications that require resilient PCIe switching. Highlights of this switch IC are as below:

- High performance: 174 GBps switching capacity; <170 ns latency non-blocking switch (transparent and through NTB)
- Virtual switch partition configuration can be reconfigured or reset without impacting operation of other partitions
- Downstream port bifurcation can be dynamically reconfigured or reset without impacting operation of adjacent ports
- Error containment: AER, DPC, CTS, Hot-plug controllers per port
- Diagnostics and Debug
- ChipLink Diagnostics Tools
- Common ref clock with or without SSC
- Up to 106 parallel GPIO pins
- 37.5mm x 37.5mm, 1311-pin FCBGA package, 1mm ball pitch

5.4 Voltage Monitor

BMC ADC ports are used to monitor all power rails for the Lightning PCIe Expansion Board in order to ensure proper operation of all power rails at all times. The voltages will be reported as part of the enclosure status as described in Chapter 9 (Baseboard Management Controller). The voltage rails to be monitored are shown in table 1.

Power Rail	Voltage
VCC_0V9	0.9V
VCC_PCIe_0V9	0.9V
P1V26_STBY	1.26V
P1V53V	1.53V
P1V8_STBY	1.8V
P3V3_STBY	3.3V
P5V_STBY	5.0V
P12V	12V

Table 1 Monitored Voltage Rails on PEB

5.5 Connectors

The sections below describe the connectors that reside on the Lightning PCIe Expansion Board.

5.5.1 PEB to DPB Connector and Pinout

The card edge connectors that connect the PEB and the PDPB are different for Lightning, and allow the routing of 60 lanes of PCIe to the PDPB in order to support x4 lanes per

SSD bay. A single U.2 SSD connects to 4 lanes of PCIe, while a dual-M.2 Carrier connects 2 lanes of PCIe per drive.

The PEB to PDPB connectors are Amphenol 200-pin 0.8mm straddle-mount connectors (G639f200221431HR). The design uses 3 of these connectors for both signal and power connections.

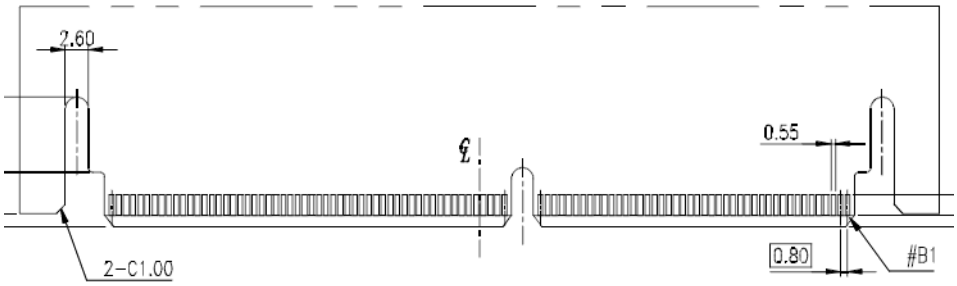


Figure 13 PCI-E 0.8mm Edge Golden Fingers

Due to the large number of signals, the full connector pin definition is provided in Section 15.1, part I, part II and part III.

5.5.2 External Mini-SAS HD Connector

The Lightning PCIe Expansion Board interfaces with a server head node via external Mini-SAS HD connectors. It's SFF-8644 standard form factor. The total quantity of host connectors is up to eight. To maximize space utilization, a quad connector is used for each group of 4x connectors. o shows the concept with either a single x4 cable connector, or a dual x4 (aka x8) cable connectors.

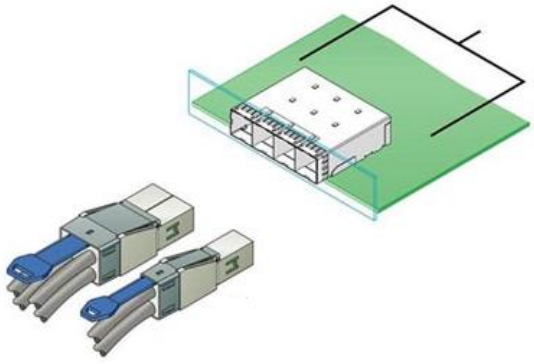


Figure 14 Mini-SAS HD Connector

The pinout of host connectors has been modified from a standard Mini-SAS HD connection to allow for a sideband signaling. Each of the x4 connectors has exactly the same pinout. Please refer to Figure 15 for the customized pinout definition and table 2 for the PCI-SIG pinout definition for the Mini-SAS HD connector. Both PCIe re-timer card and PEB designs shall have BOM options to support both the customized pinout and the PCI-SIG pinout.

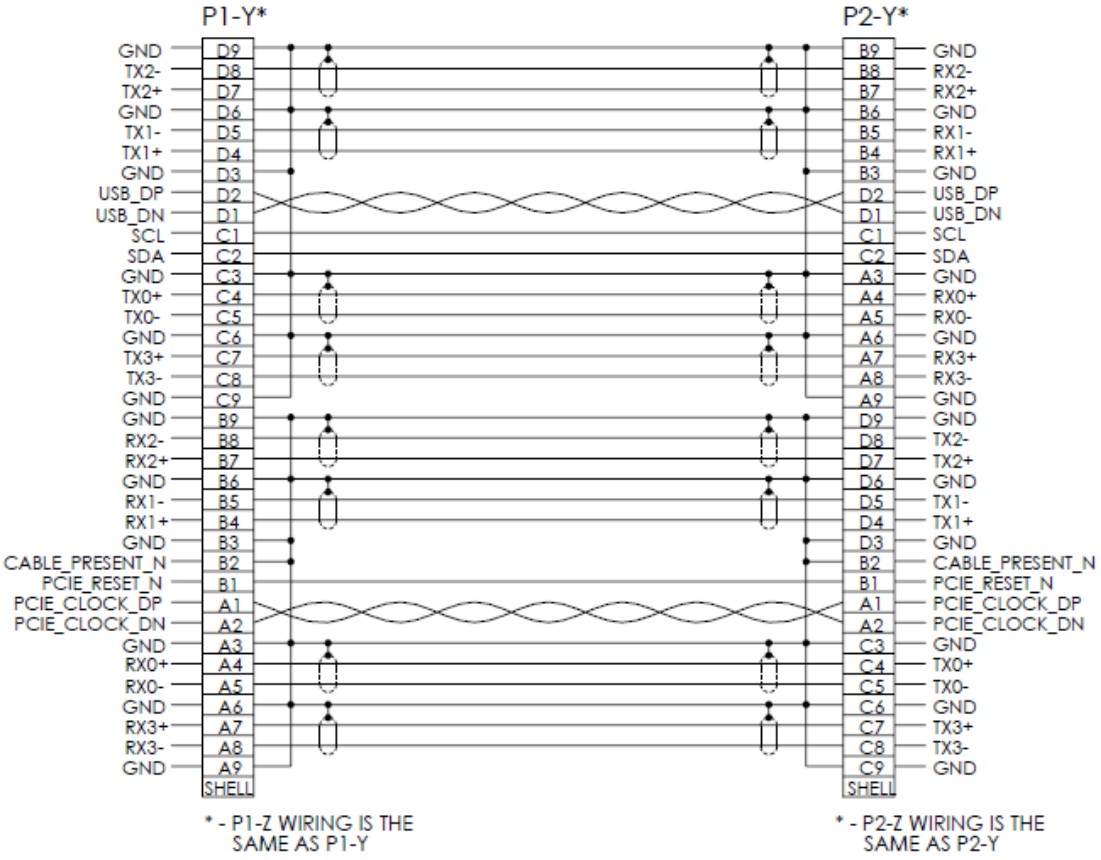


Figure 15 Mini-SAS HD Connector and Cable Pinout

	Column								
Row	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
D	GND	PETn2	PETp2	GND	PETn1	PETp1	GND	MGTPWR	PWR
C	GND	PETn3	PETp3	GND	PETn0	PETp0	GND	CMISDA	CMISCL
B	GND	PERn2	PERp2	GND	PERn1	PERp1	GND	CBLPRSNT#	PWR
A	GND	PERn3	PERp3	GND	PERn0	PERp0	GND	CINT#	CADDR

Table 2 Standard Mini-SAS HD Connector and Cable Pinout by PCI-SIG

The pinout definition difference between the customized definition (Lightning POR) and the PCI-SIG definition is shown in table 3.

	PCI-SIG Cabling Spec (0.7 Draft)	Lightning POR
A1	CAddr	REF_CLK_P
A2	CINT#	REF_CLK_N
B1	PWR	PERST#
D1	PWR	USB_DN
D2	MGTPWR	USB_DP

Table 3 Differences for the Mini-SAS HD Connector and Cable Pinout

BOM options that have been added onto the PEB are shown in Figure 16.

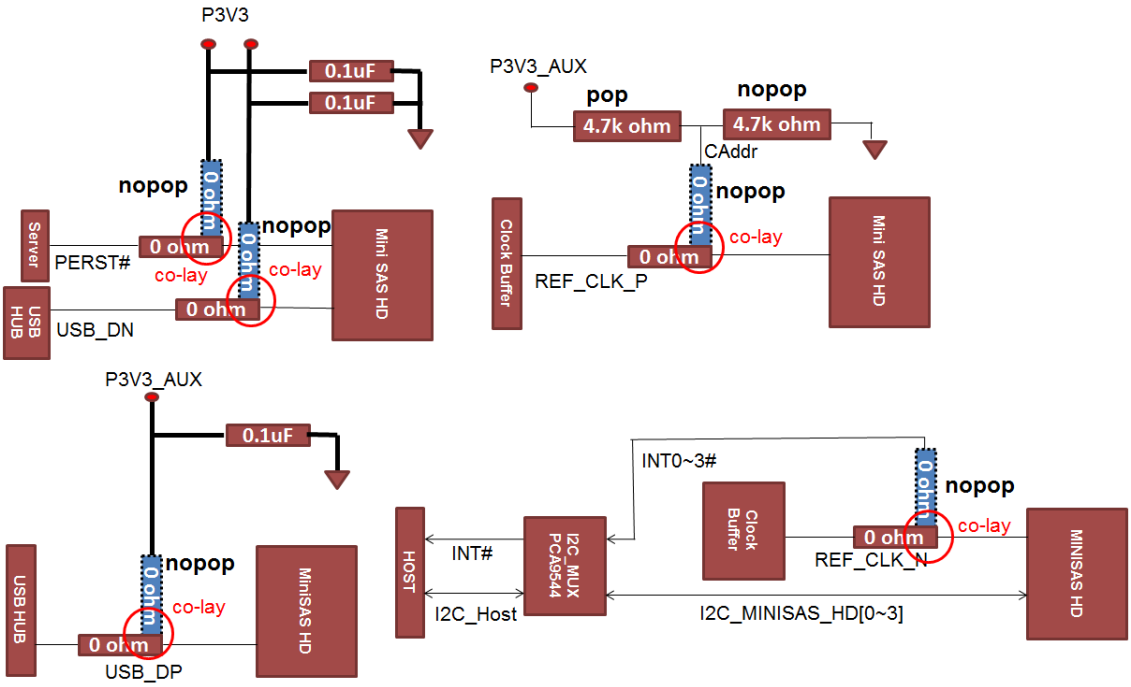


Figure 16 BOM options to support PCI-SIG pinout definition

5.5.3 Debug Header

The PCIe Expansion Board includes a debug header on the front edge. It supports hot plugging of an OCP debug card. The card supports the following functions:

- Two 7-segment LED displays: Show BMC firmware POST information and system error code defined in Sec 9.9.
- One RS-232 serial connection: Provides BMC or PCIe switch UART console connection.

The connector for the debug header is a 14-pin, shrouded, right-angled, 2mm pitch connector. Figure 17 shows an illustration. The debug card has a key to match with the notch to avoid pin shift when plugging it in.

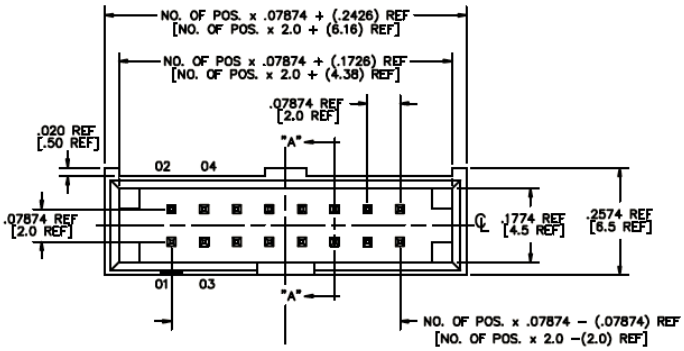


Figure 17 Debug Header Illustration

Table 4 lists the pin definition of the debug header:

Pin (CKT)	Function
1	Low HEX character [0] least significant bit
2	Low HEX character [1]
3	Low HEX character [2]
4	Low HEX character [3] most significant bit
5	High HEX character [0] least significant bit
6	High HEX character [1]
7	High HEX character [2]
8	High HEX character [3] most significant bit
9	Serial Transmit
10	Serial Receive
11	Reset. Slide switch determines if its BMC or PCIe switch reset
12	Serial Port Select (1=Switch; 0=BMC)
13	GND
14	VCC (+5VDC)

Table 4 Debug Header Pin-out

5.6 Switch and Buttons

The PCIe Expansion Board will have one push button together with a slide switch, to trigger a reset signal to either the BMC, or the PCIe switch.

The default setting of the slide switch allows the reset button to trigger a reset of the BMC.

There is also an option on the debug card through which the user can reset either the BMC or the PCIe switch depending on the slide switch position.

A yellow LED is located close to the slide switch, to indicate which UART console is connected (On: PCIe switch, Off: BMC).

5.7 LEDs

The PCIe Expansion Board will have several LEDs on its front edge to display various information as below:

- One blue single color LED, for power and system identification.It is controlled by BMC GPIO (refer to Table 5).
- One red single-color LED for enclosure fault status and controlled by BMC GPIO (refer to Table 6).
- Four green single-color LEDs per x16 uplink, for uplink status and controlled by the PCIe switch (refer to Table 7).
- Three green single-color LEDs per x16 uplink for switch zoning indication and controlled by the PCIe switch (refer to Table 8).
- One green single-color LED, for the PCIe switch heartbeat and controlled by the PCIe switch (blinking at 1s frequency if switch is running normally)
- One green single-color LED, for the BMC heartbeat (blinking at 500ms frequency if BMC is running normally)

The following tables summarize the LED behaviors.

Power and System Identification	Blue LED
Power on, System Identify off	Consistently on
Power on, System Identify on	On 0.9sec, off 0.1sec, and loop

Table 5 Power and System Identify LED

Enclosure Fault Status	Red LED
Normal system operation	Off
Any fault in whole enclosure	On
Reserved for future use	Blinking

Table 6 Enclosure Fault Status LED

Uplink Status	Green LED 1	Green LED 2	Green LED 3	Green LED 4
X16	ON	OFF	OFF	OFF
X8	ON	ON	OFF	OFF
X4	ON	ON	ON	ON

Table 7 LED Indicator for Uplink

Mini-SAS Port Link Status	Green LED 1	Green LED 2	Green LED 3
1x16	ON	OFF	OFF
2x8	OFF	ON	OFF
4x4	OFF	OFF	ON

Table 8 LED Indicator for Switch Zoning

5.7.1 Implementation of enclosure status LED

The enclosure fault LED (red) should be designed to meet the following scenarios:

- If BMC firmware hangs, the enclosure fault LED will be turned ON.
- If BMC firmware runs normally, it will turn on the enclosure fault LED for any fault / warning within the whole system. The list of sensors is defined in Section 9.11.

5.8 PCB Stack-up

Figure 18 shows an example of the PEB stack-up, where medium-loss PCB material is chosen to ensure that the PCIe switch can drive the PCIe signals at 8GT/s without requiring re-drivers. Also, the detailed information of impedance control is listed for each layer in Figure 19.

Board Number:	15105-1
Project name:	Lightning_PMC
Model Name:	PEB
Layer Count:	8 Layer
Date:	10/19/2016
Material:	TU863+VLP / NPG171+VLP
Gold Finger(Y/N)	Y
Customer:	XXXXX
EE engineer	Leon Du
SI Engineer	Eason YS Chen

Version
02

	Layer		Thickness		Glass/Copper Style	Er		Df(1G)
		Cu oz	Wiwynn			Wiwynn		
	Mask		0.5			3.4		0.025
Top	Signal	0.5 oz+plating	2.1					
	Prepreg		3			3.9		0.009
L2	GND	1 oz	1.3					
	Core		4			3.9		0.009
L3	Signal	1 oz	1.3					
	Prepreg		16			4.2		0.009
L4	POWER	1 oz	1.3					
	Core		4			3.8	0	0.009
L5	POWER	1 oz	1.3	0.0	0			
	Prepreg		16	0	0	4.2	0	0.009
L6	Signal	1 oz	1.3	0	0			
	Core		4	0	0	3.9	0	0.009
L7	GND	1 oz	1.3	0	0			
	Prepreg		3	0	0	3.9	0	0.009
Bottom	Signal	0.5 oz+plating	2.1	0	0			
	Mask		0.5	0		3.4	0	0.025
Thickness requirement: 1.6 ± 10% mm		mil	63					
		mm	1.60					

Note:
1. Unit is mil
2. The total thickness includes trace and solder mask
3. Min hole copper thickness is 1.0 mil
4. Min surface copper thickness 1.5 mil
5. Impedance Cpk requirement: >=1.33 (Report should be provided)
6. PCB supplier should design Delta-L coupon and provide measurement results meeting following criteria:
    - **-0.65dB/inch** at 4GHz for stripline routing
    - **-0.69dB/inch** at 4GHz for microstrip routing

Figure 18 PEB Stack-up

	Single Ended Type(mil)			
Imp Variation	50 Inner/Outer+/-5ohm			
Bus	MISC./I2C			
Type Layers	SE 1,3,6,8			
	Wiwynn			
	Width	Imp	Width	Imp
S1	4.75(L2)	50.46		
P2	reference layer			
S3	5(L2/L4)	49.19		
P4	reference layer			
P5	reference layer			
S6	5(L5/L7)	49.19		
P7	reference layer			
S8	4.75(L7)	50.46		

	Differential Type(mil)					
Imp Variation	85 Inner/Outer+/-10%					
Bus	PCIe/Clock/USB					
Type Layers	DP 1,3,6,8					
	Wiwynn					
	Width	Space	Imp	Width	Space	Imp
S1	5.38(L2)	6.62	85.36			
P2	reference layer					
S3	5.5(L2/L4)	7.5	84.85			
P4	reference layer					
P5	reference layer					
S6	5.5(L5/L7)	7.5	84.85			
P7	reference layer					
S8	5.38(L7)	6.62	85.36			

	100 Inner/Outer+/-10%					
Bus	Clock/BMC DDR3					
Type Layers	DP 1,3,6,8					
	Wiwynn					
	Width	Space	Imp	Width	Space	Imp
S1	4.13(L2)	9.37	101.06			
P2	reference layer					
S3	4(L2/L4)	8	97.55			
P4	reference layer					
P5	reference layer					
S6	4(L5/L7)	8	97.55			
P7	reference layer					
S8	4.13(L7)	9.37	101.06			

Figure 19 PEB Impedance Control Information for Each Layer

6 Lightning PCIe Drive Plane Board

6.1 Block Diagram of PDPB

Figure 20 illustrates the functional block diagram of the PCIe Drive Plane Board (PDPB) for Lightning.

Notes:

- The PCIe Drive Plane Board (PDPB) is connected to the PEB through three 200 pin connectors (Conn1, Conn2 and Conn3) as described in section 5.5.1 and as shown in figure 19.
- There are 15 soft-start control circuits on the PDPB so that power control going into each x4 slot is independent to all of the other 14 slots, and inrush current for each slot is controllable.
- There are 15 Attention LEDs and 15 Activity LEDs so that each slot has an Attention LED and an Activity LED mapped to it.
- Voltage and temperature sensors are needed to monitor the status of the PDPB using Lightning BMC on the PEB.
- An EEPROM is used to store the FRUID.

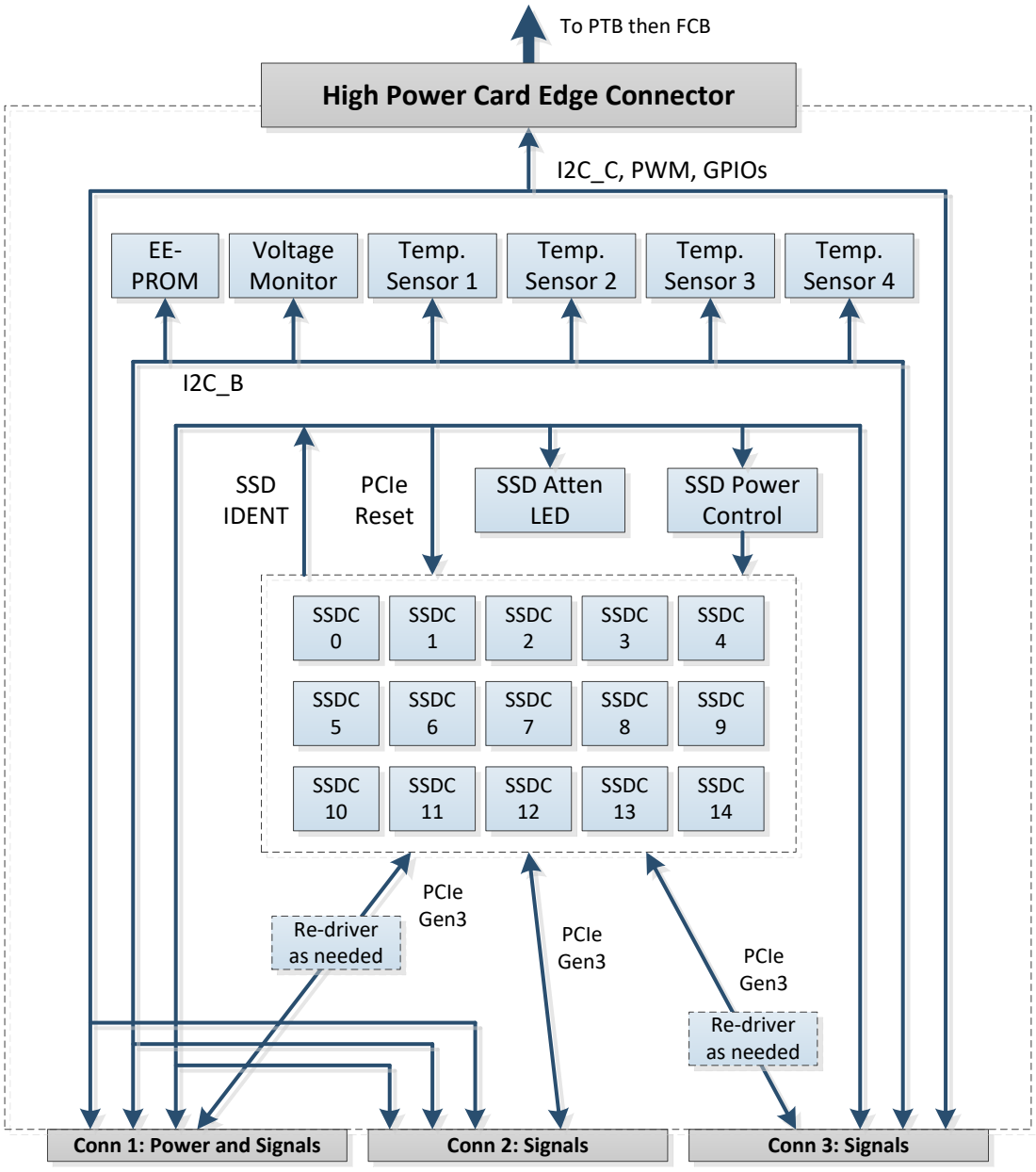


Figure 20 Lightning PCIe Drive Plane Board Block Diagram

6.2 PDPB Form Factor and Placement

The PCIe Drive Plane Board form factor is a 263.5mmX506mm board. Figure 21 illustrates the form factor and its rough placement.

The PCIe Drive Plane Board (PDPB) will be sized almost identically to the Knox and Honey Badger DPBs with the exception of the card edge connectors. See Section 6.2 of the OCP Open Vault Storage Specification v0.8 for the board placement drawing, and Section 5.3 of the OCP Honey Badger Specification v0.8 for modifications made for that

design. The two sets of PCIe card edge connectors (for the two SAS Expansion boards or the Honey Badger Baseboard) are replaced by higher density card edge connectors for the Lightning PDPB. The locations of all other connectors are exactly the same.

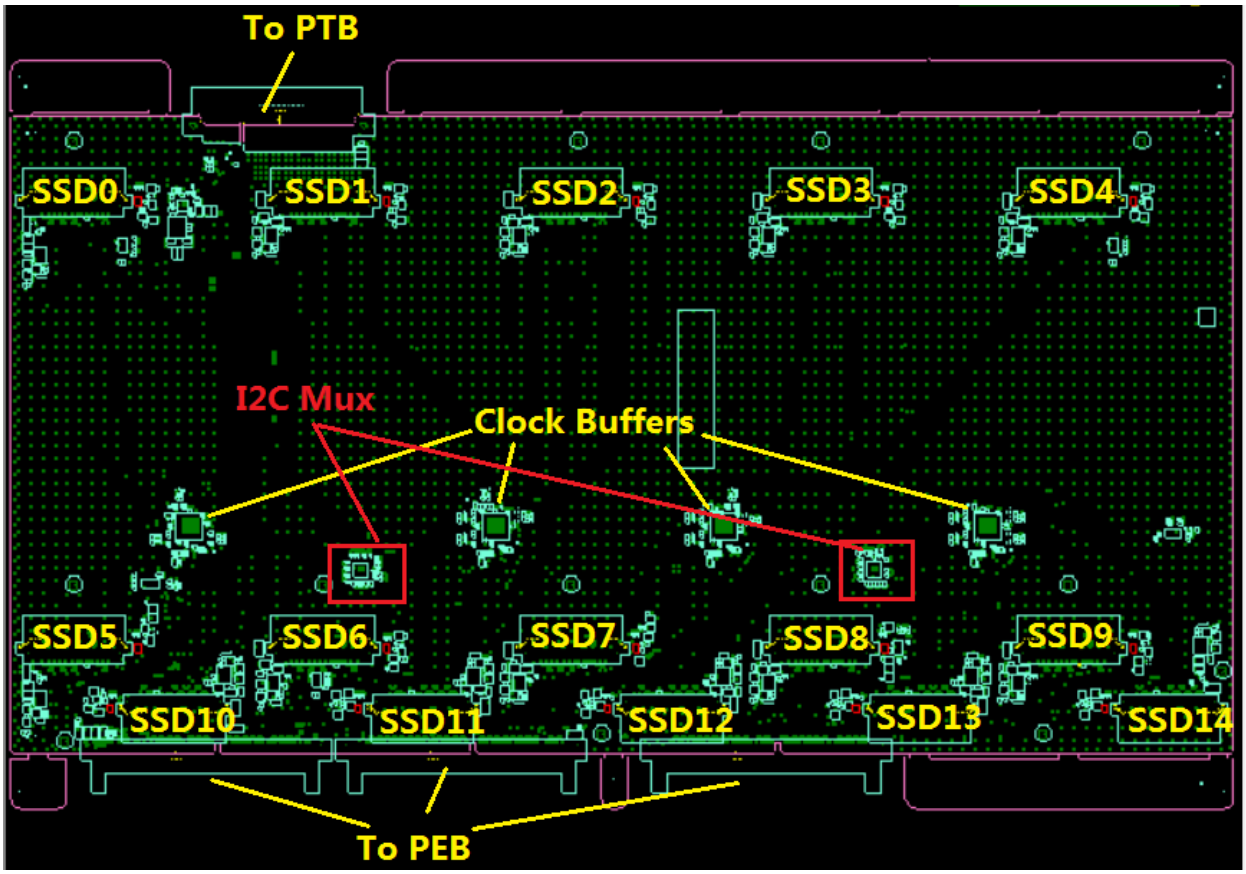


Figure 21 Lightning PCIe Drive Plane Board Placement

6.3 Thermal Sensors

LM75 or equivalent components (same as required in Honey Badger / Knox) are placed in locations as shown in Figure 22. For further details, see the thermal requirement section 11.

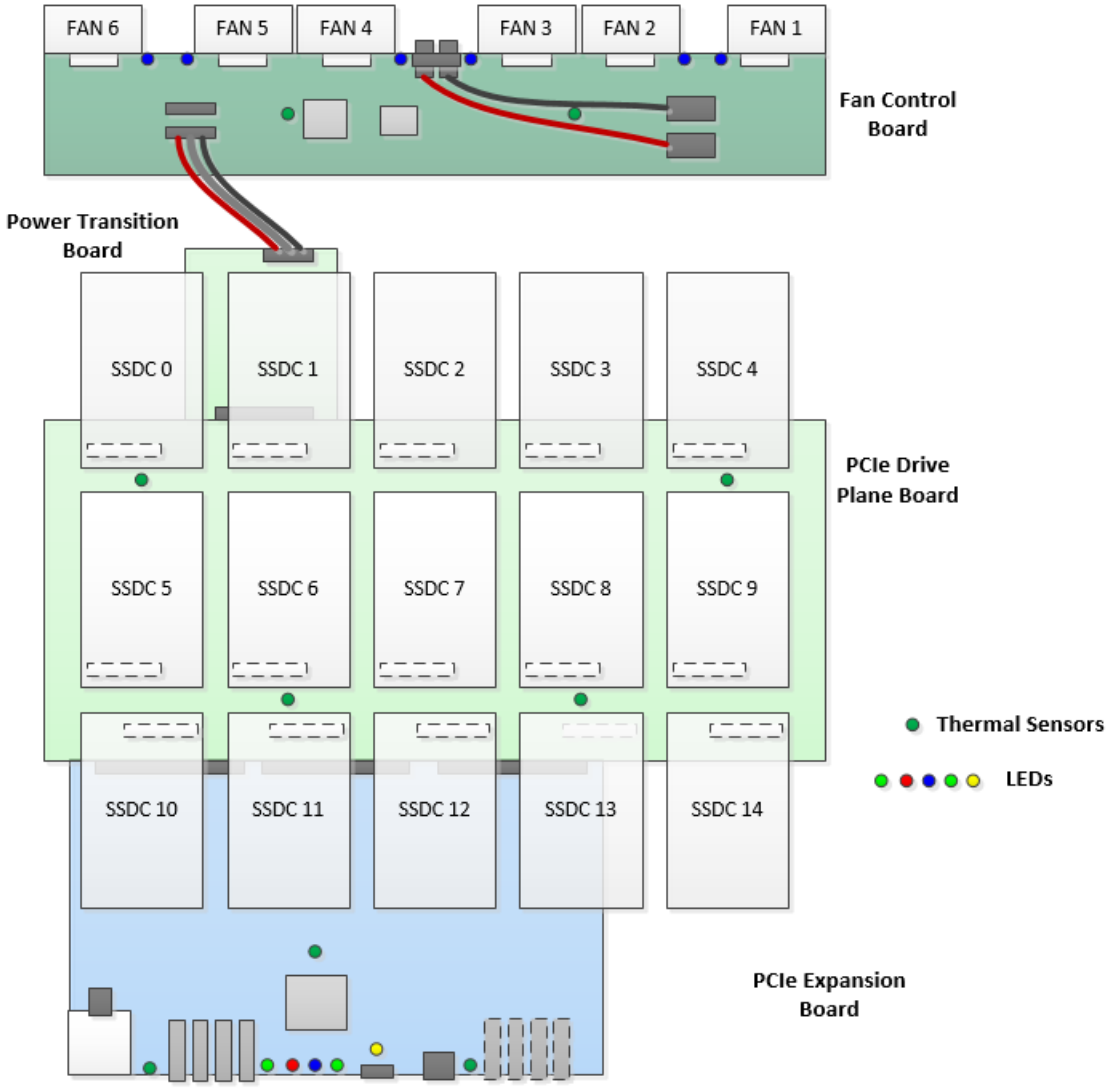


Figure 22 Drive Plane Board Thermal Sensor Locations

6.4 Voltage Monitor

A voltage monitor is required for the Lightning PCIe Drive Plane Board in order to ensure proper operation of all critical power rails at all times. The voltages will be reported as part of the enclosure status as described in Chapter 9 (Baseboard Management Controller). The voltage rails to be monitored are shown in table 9.

Power Rail	Voltage
P12V_PCIe	12.0V
P3V3	3.3V

Table 9 Monitored Voltage Rails on PDPB

6.5 Connectors

The sections below describe the connectors that reside on the Lightning PCIe Drive Plane Board.

6.5.1 Signal Connector to PCIe Expansion Board

As shown in Section 5.5.1, the PEB to PDPB connectors are Amphenol 200-pin 0.8mm straddle-mount connectors (G639f200221431HR). The design uses 3 of these connectors for both signal and power connections. The straddle-mount connectors as shown in Figure 23 reside on the PDPB.

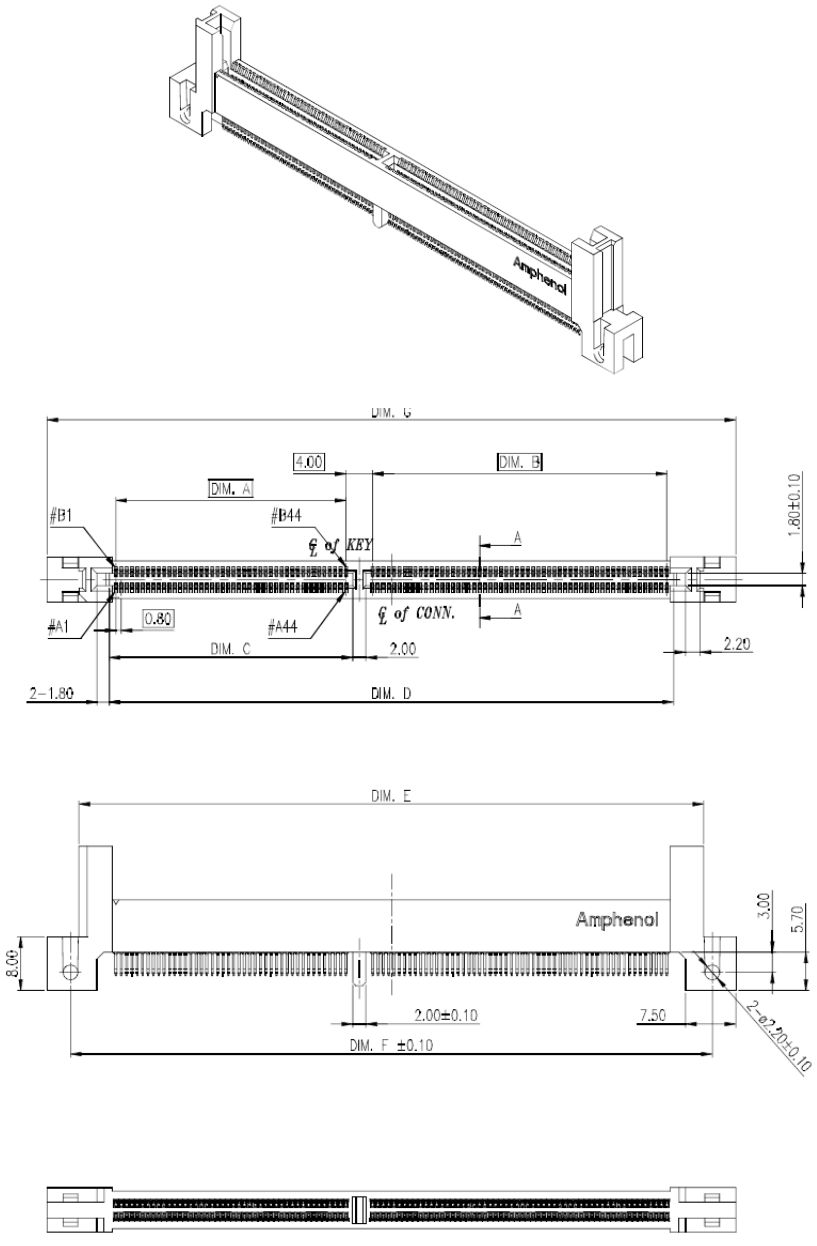


Figure 23 llustratioin of the 200 pin Straddle Connector

Due to the large number of signals, the full connector pin definition is provided in Section 15.1, part I, part II and part III.

6.5.2 SSD Connectors (SFF-8639)

The Lightning PCIe Drive Plane Board will use standard SFF-8639 connectors to connect with U.2 PCIe SSDs or Facebook's M.2 carriers. SFF-8639 is also known as the U.2 form factor. The connector illustration is shown in Figure 24, and pin out assignment is shown in Table 10.

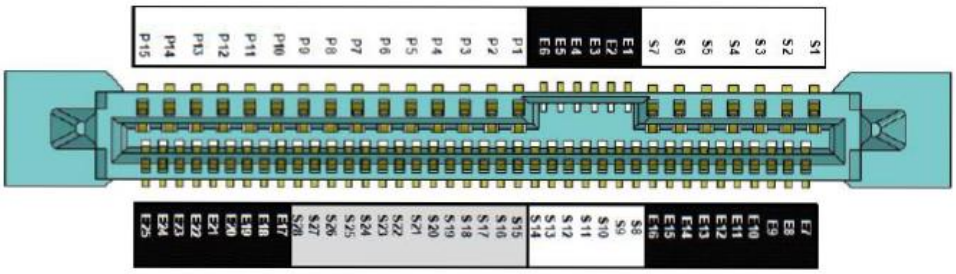
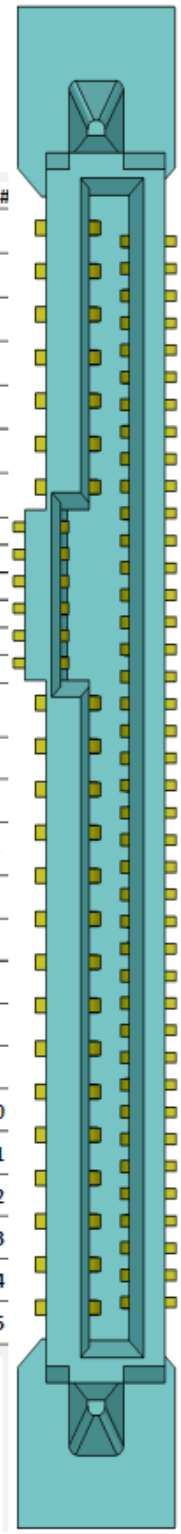


Figure 24 Right-Angle SFF-8639 SSD Connector

Pin Out Drawing (Receptacle pin naming)



Signal Description	Name	Mating	Pin #
Ground	GND	2nd	S1
SAS/SATA/SATAe 0 Tx+	S0T+ (A+)	3rd	S2
SAS/SATA/SATAe 0 Tx -	S0T- (A-)	3rd	S3
Ground	GND	2nd	S4
SAS/SATA/SATAe 0 Rcv -	S0R- (B-)	3rd	S5
SAS/SATA/SATAe 0 Rcv +	S0R+ (B+)	3rd	S6
Ground	GND	2nd	S7
ePCIe RefClk + (port B)	RefClk1+	3rd	E1
ePCIe RefClk – (port B)	RefClk1-	3rd	E2
3.3V for SM bus	3.3Vaux	3rd	E3
ePCIe Reset (port B)	ePERst1#	3rd	E4
ePCIe Reset (port A)	ePERst0#	3rd	E5
Reserved	RSVD	3rd	E6
Reserved(WAKE#/OBFF), SASAct2	RSVD(Wake#) /SASAct2	3rd	P1
SATAe Client /SAS reset	sPCIeRst/SAS	3rd	P2
Reserved (DevSLP#)	RSVD(DevSLP#)	2nd	P3
Interface Detect (Was GND-precharge)	IfDet#	1st	P4
Ground	GND	2nd	P5
		2nd	P6
Precharge	5 V	2nd	P7
SATA, SATAe, SAS only		3rd	P8
		3rd	P9
Presence (Drive type)	PRSNT#	2nd	P10
Activity(output)/Spinup	Activity	3rd	P11
Hot Plug Ground	GND	1st	P12
Precharge	12 V	2nd	P13
All – 12V Only power for ePCIe SSD		3rd	P14
		3rd	P15

Pin #	Mating	Name	Signal Description
E7	3rd	RefClk0+	ePCIe Primary RefClk +
E8	3rd	RefClk0-	ePCIe Primary RefClk -
E9	2nd	GND	Ground
E10	3rd	PETp0	ePCIe 0 Transmit +
E11	3rd	PETn0	ePCIe 0 Transmit -
E12	2nd	GND	Ground
E13	3rd	PERn0	ePCIe 0 Receive -
E14	3rd	PERp0	ePCIe 0 Receive +
E15	2nd	GND	Ground
E16	3rd	RSVD	Reserved
S8	2nd	GND	Ground
S9	3rd	S1T+	SAS/SATAe 1 Transmit +
S10	3rd	S1T-	SAS/SATAe 1 Transmit -
S11	2nd	GND	Ground
S12	3rd	S1R-	SAS/SATAe 1 Receive -
S13	3rd	S1R+	SAS/SATAe 1 Receive +
S14	2nd	GND	Ground
E17	3rd	RSVD	Reserved
E18	2nd	GND	Ground
E19	3rd	PETp1/S2T+	ePCIe 1 /SAS 2 Transmit +
E20	3rd	PETn1/S2T-	ePCIe 1 /SAS 2 Transmit -
E21	2nd	GND	Ground
E22	3rd	PERn1/S2R-	ePCIe 1 /SAS 2 Receive -
E23	3rd	PERp1/S2R+	ePCIe 1 /SAS 2 Receive +
E24	2nd	GND	Ground
E25	3rd	PETp2/S3T+	ePCIe2 / SAS 3 Transmit +
E26	3rd	PETn2/S3T-	ePCIe2 / SAS 3 Transmit -
E27	2nd	GND	Ground
E28	3rd	PERn2/S3R-	ePCIe 2 / SAS 3 Receive -
E29	3rd	PERp2/S3R+	ePCIe 2 / SAS 3 Receive +
E30	2nd	GND	Ground
E31	3rd	PETp3	ePCIe 3 Transmit +
E32	3rd	PETn3	ePCIe 3 Transmit -
E33	2nd	GND	Ground
E34	3rd	PERn3	ePCIe 3 Receive -
E35	3rd	PERp3	ePCIe 3 Receive +
E36	2nd	GND	Ground
E37	3rd	SMClk	SM-Bus Clock
E38	3rd	SMDat	SM-Bus Data
E39	3rd	DualPortEn#	ePCIe 2x2 Select

Table 10 SSD Connector Pin-Out

6.6 LEDs

On the PCIe Drive Plane Board, each SSD will have one bi-color(blue and red) LED to indicate its status. Please refer to o in section 4.8. The LED states are summarized below, and also shown in Table 11:

- The Blue LED is driven by the SSD ACTLED signal and the desired behaviors are shown below:
  - When the SSD is online and healthy, turn on the Blue LED;
  - When the SSD is active, the behavior follows the vendor's definition.
- The Red LED is driven by the PCIe Switch ATNLED# signal:
  - When there's any fault on the SSD, turn on the Red LED;
  - If the Red LED is on, the Blue LED is also forced off to avoid color contamination;
  - A blinking Red LED is used to identify a specific SSD slot.
- When there's no SSD inserted, turn off both LEDs.

Solid State Drive Status	Blue LED	Red LED
No Drive Inserted	OFF	OFF
Drive Online and Healthy	ON	OFF
Drive is Active	Follow Vendor's definition	OFF
Drive Failure	OFF	ON
Drive Identify	OFF	Blinking

Table 11 Drive plane board LED for SSD Status

6.7 PCB Stack-up

Figure 25 shows an example of PDPB stack-up, where medium-loss PCB material is chosen to ensure that we can drive the PCIe signals at 8GT/s without requiring re-drivers. Also, the detailed information of impedance control is listed for each layer.

		Version
		02
Board Number:	15669-1	
Project name:	Lightning_PMC	
Model Name:	PDPB	
Layer Count:	8 Layer	
Date:	10/19/2016	
Material:	TU863+VLP / NPG171+VLP	
Gold Finger(Y/N)	N	
Customer:	XXXXX	
EE engineer	Leon Du	
SI Engineer	Eason YS Chen	

	Layer	Cu oz	Thickness Wiwynn		Glass/Copper Style	Er Wiwynn		Df(1G)
	Mask		0.5			3.4		0.025
Top	Signal	0.5 oz+plating	2.1					
	Prepreg		4			3.8		0.009
L2	GND	1 oz	1.3					
	Core		5			3.8		0.009
L3	Signal	1 oz	1.3					
	Prepreg		20			4.5		0.009
L4	POWER	2 oz	2.6					
	Core		11.4			4.2	0	0.009
L5	POWER	2 oz	2.6	0.0	0			
	Prepreg		20	0	0	4.5	0	0.009
L6	Signal	1 oz	1.3	0	0			
	Core		5	0	0	3.8	0	0.009
L7	GND	1 oz	1.3	0	0			
	Prepreg		4	0	0	3.8	0	0.009
Bottom	Signal	0.5 oz+plating	2.1	0	0			
	Mask		0.5	0		3.4	0	0.025
Thickness requirement: 2.16 ± 0.1 mm		mil	85					
		mm	2.16					

	Single Ended Type(mil)			
Imp Variation	50 Inner/Outer+/-5ohm			
Bus	MISC./I2C			
Type Layers	SE 1,3,6,8			
	Wiwynn			
	Width	Imp	Width	Imp
S1	6.75(L2)	50.7		
P2	reference layer			
S3	6(L2/L4)	50.47		
P4	reference layer			
P5	reference layer			
S6	6(L5/L7)	50.47		
P7	reference layer			
S8	6.75(L7)	50.7		

Note:
1. Unit is mil
2. The total thickness includes trace and solder mask
3. Min hole copper thickness is 1.0 mil
4. Min surface copper thickness 1.5 mil
5. Impedance Cpk requirement: >=1.33 (Report should be provided)
6. PCB supplier should design Delta-L coupon and provide measurement results meeting following criteria:
   - **-0.65dB/inch** at 4GHz for stripline routing
   - **-0.69dB/inch** at 4GHz for microstrip routing

	Differential Type(mil)											
Imp Variation	85 Inner/Outer+/-10%						100 Inner/Outer+/-10%					
Bus	PCIe/Clock						Clock					
Type Layers	DP 1,3,6,8						DP 1,3,6,8					
	Wiwynn						Wiwynn					
	Width	Space	Imp	Width	Space	Imp	Width	Space	Imp	Width	Space	Imp
S1	6.88(L2)	5.13	84.07				5(L2)	6	99.3			
P2	reference layer						reference layer					
S3	6.5(L2/L4)	7.5	84.49				4.5(L2/L4)	9	100.81			
P4	reference layer						reference layer					
P5	reference layer						reference layer					
S6	6.5(L5/L7)	7.5	84.49				4.5(L5/L7)	9	100.81			
P7	reference layer						reference layer					
S8	6.88(L7)	5.13	84.07				5(L7)	6	99.3			

Figure 25 PCB Stack-up and Impedance Control for PDPB

7 Lightning Power System

7.1 PEB Power Budget

The total power consumption of the Lightning PEB is estimated as shown below in Table 12.

Major Device	TDP max (W)	Qty	Utilization(%)	Power(W)
PM8536	32.7	1	90%	29.43
AST2400	2.5	1	80%	2
Other Logics	5	1	80%	4
VR Losses	3.84	1	100%	3.84
Power Budget for PEB				44.04

Table 12 PEB Power Budget

7.2 System Level Power Budget

The total power consumption of the Lightning hardware system is shown below in table 13.

7.2.1 System Power Budge

Total maximum power consumption of the Lightning system can be as high as over 600W, and will vary based on workloads and SSD models. Noted that the data in column "Measured Power Consumption" in table 13 is based on the measurement of Facebook's Lightning system except that it is using Facebook's SSD maximum power budget 14W instead of the measured SSD power consumption:

Item	Qty in a 2U System	Power Consumption of Each Module	Total Power Consumption	Measured Power Consumption
SSD	30	14	420	420
Drive Plane Board	2	9	18	16.2
PCIe Expansion Board	2	44.04	88.08	54.98
Fan Control Board	1	2	2	1.6
Fan Module	6	20	120	116.6
Total				609.38

Table 13 System Power Budget for the whole Lightning Chassis

7.3 PEB Buck Converter Solutions

7.3.1 Key Power Rail Buck Converters

All key power rails' buck converters residing on the PCIe Expansion Board are designed with a 93% efficiency target under normal load. Selection of components for this solution is to be determined by the vendor. If a higher efficiency is available even at additional cost, the vendor shall present those options to Facebook.

7.4 Power Sequencing

The design shall follow key IC's power-up and power-down sequence requirements without any violations, and ensure power cycling with adequate reliability. The power sequence also needs to be considered between the Lightning PEB and the PDPB, where the PCIe switch and all end devices sit respectively. The sequence defined by PCIe CEM specification shall be strictly followed.

7.5 Power Button

No power button is implemented on the PEB or on the PDPB. The system shall always be in operation once power is supplied, or resume operation upon restoration of power in a power failure event.

There is a power button on the PTB of Lightning / Honey Badger, and it's always kept on. The use of that power button shall not be required to power on the system. It's only intended to enable an easy way to power cycle an entire tray during tray level service.

8 Mechanical of Tray and SSD Carrier

8.1 Changes for Tray Latches

The mechanical design of the chassis is mostly unchanged, this includes the six fan modules and fan control board (FCB), the power transition board (PTB), the bus bar clip, and the tray cable arms, etc. The tray latches remain unchanged, and the PEB will use the same card latches and same form factor as the Honey Badger Baseboard.

The card latches will be changed to accommodate the increased insertion force caused by the new 200 pin connectors between PEB and PDPB. Jack Screw will be used, and the card latch will be modified accordingly. Figure 26 shows the concept in development.

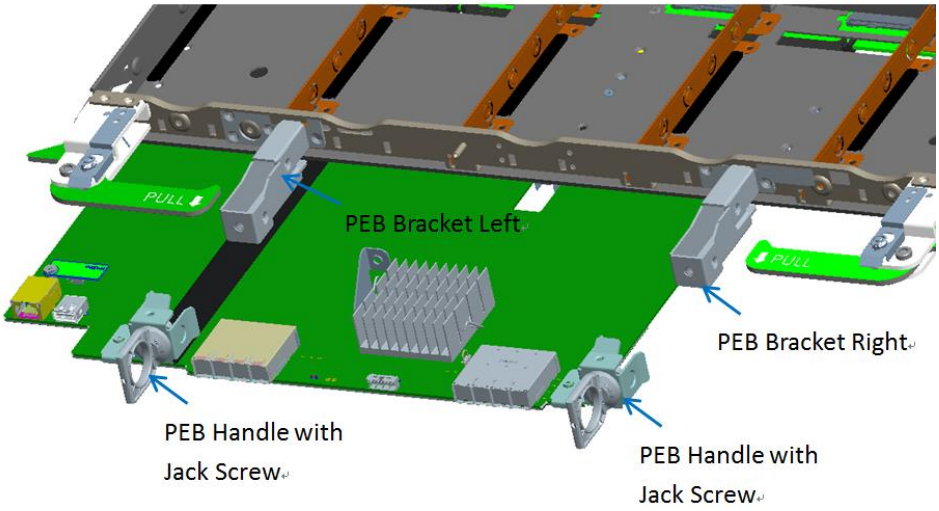


Figure 26 Card Latches and Jack Screw

8.2 SSD Connectors and Carriers

The SSD connectors will be SFF-8639 connectors (aka U.2). There are three SSD carriers supported by the Lightning PCIe Drive Plane Board. Two of them are for 2.5-inch form factor SSDs and one is a customized carrier design that carries two M.2 form factor drives. Both carriers support placing the 2.5-inch PCIe SSDs in the existing 3.5-inch hard disk drive bay supported by the Knox / Honey Badger mechanical design. The female connectors on the PDPB for the SSD carrier is the SFF-8639 connector which supports 4 lanes of PCIe.

8.3 Single SSD Carrier (15 mm)

The single SSD carrier provides the mechanical mounting required to fit a single 15mm SSD into the 3.5-inch HDD bay. The SSD docks directly into the SFF-8639 on the PDPB and

no electrical interposer is required. The drawings below show the design for the single SSD carrier. This 15mm carrier is an optional design for 2.5-inch SSD, but it is not Facebook's POR. Refer to Figure 27 for the 15mm carrier design.

8.4 Single SSD Carrier (7 mm)

The single SSD carrier provides the mechanical mounting required to fit a single 7mm SSD into the 3.5-inch HDD bay. The SSD docks directly into the SFF-8639 on the PDPB and no electrical interposer is required. The drawings below show the current design for the 7mm carrier design.

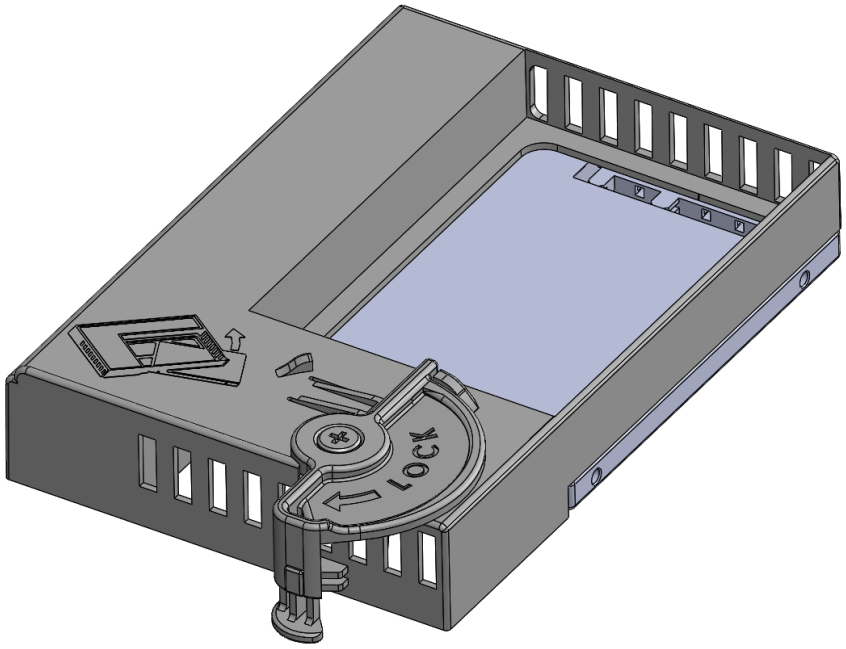


Figure 27 Lightning Single SSD carrier, 15mm and 7 mm

8.5 Dual M.2 Module Carrier

The M.2 carrier will contain an interposer board that allows the installation of 2x M.2 22110 D5+ SSDs. The interposer will also include a temperature sensor that is connected to SMBUS/I2C.

Figure 28 and Figure 29 show the mechanical design of the Dual M.2 carrier

The M.2 carrier is comprised of an adapter body, upon which is installed the interposer PCB and M.2 modules. 2 die cast heat sinks, with plastic, spring loaded latches, are installed on top of the M.2s. A tilting, removable cover is pressed down into place. The cover includes a rotating latch which, after the cartridge is installed into the Lightning drive tray, is rotated to fix the carrier into the SSD slot, preventing movement during transportation and service.

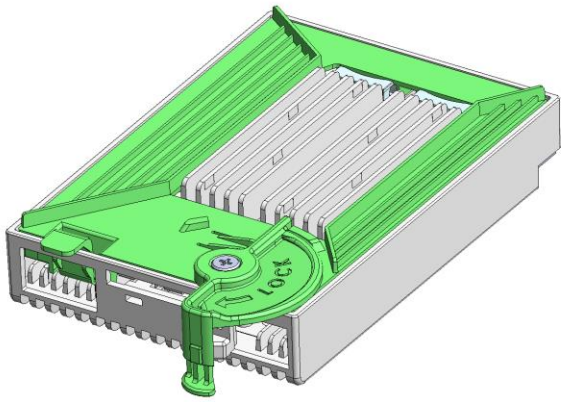


Figure 28 Lightning Dual M.2 Carrier External View

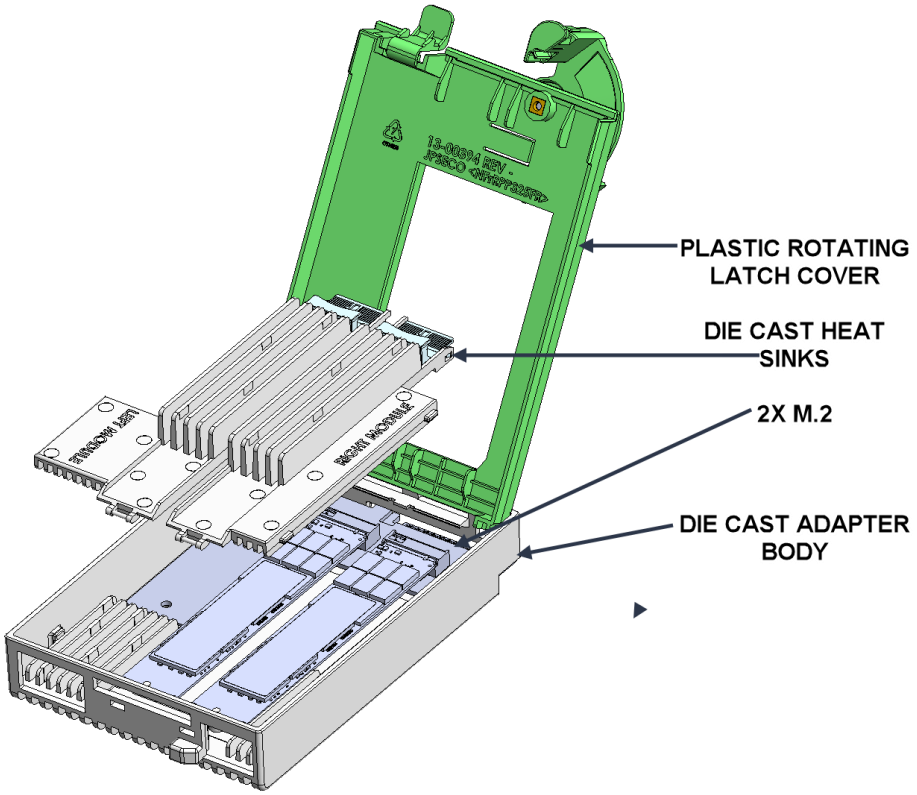


Figure 29 Lightning Dual M.2 Carrier Internal View

The sideband signals and the SMBus / I2C connection to the M.2 modules are shown in Figure 30.

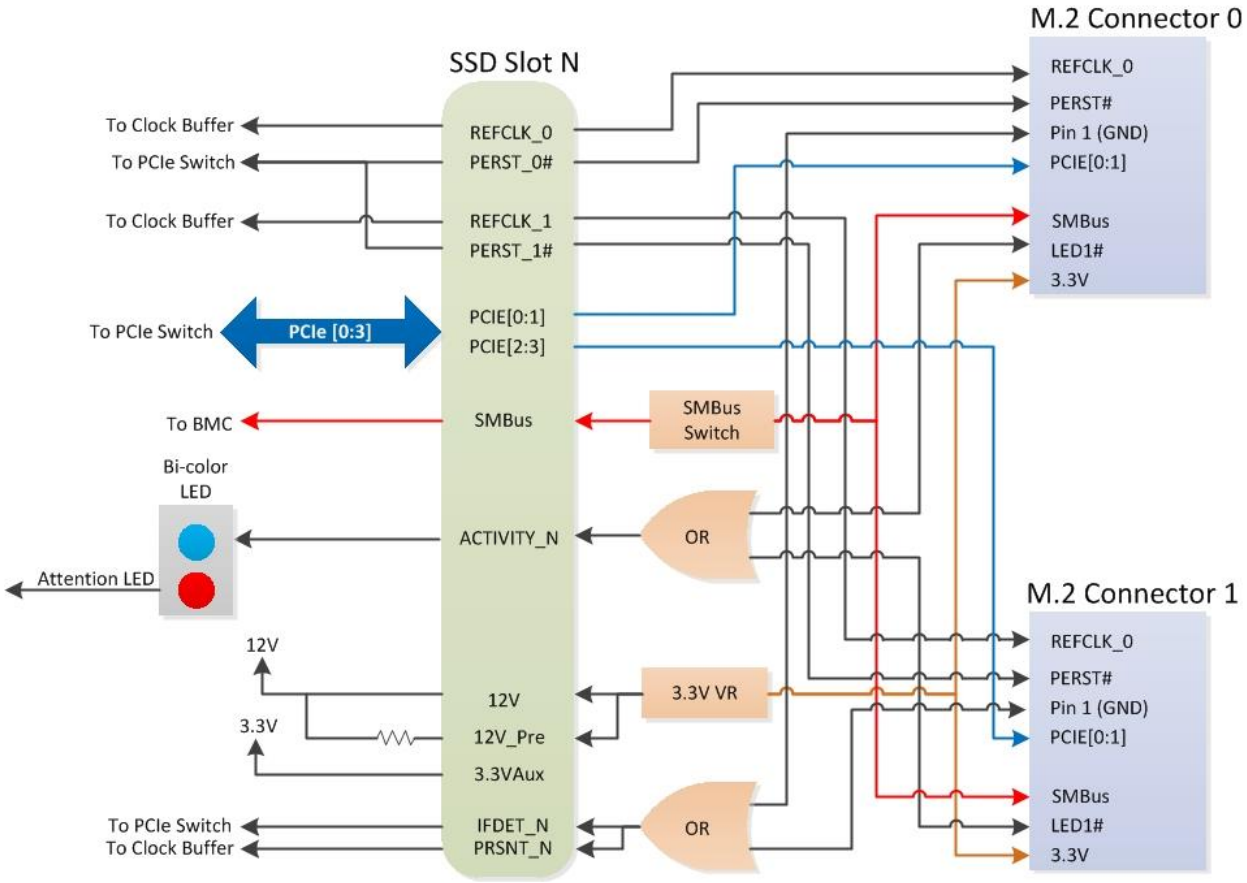


Figure 30 Lightning Dual SSD Carrier Interposer Connections

8.6 PCIe Expansion Board PCB Thickness

To ensure proper fit of the PCIe expansion board in the PCIe Drive Plane Board's connector, the PEB must be 63 mils thick.

8.7 Silk Screen

The silk screen shall be white in color and include labels for the components listed below. Additional items required on the silk screen will be available during product development phase.

- All SSDs
- Fan headers
- Debug headers
- PCIe Switches (numbered if more than one per board)
- All LEDs
- All buttons (push button or slide switch)

8.8 PCB Color

Different PCB colors shall be used to help identify the board revisions. Table 14 indicates the PCB color to be used for each development revision.

Revision	PCB Color
EVT	Red
DVT	Yellow
PVT	Green

Table 14 PCB Color for Different Revision

9 Baseboard Management Controller (BMC)

9.1 Overview

Since the current PCIe switch does not support any form of in-band management functions aside from switch specific functions, a BMC will be required. The BMC will be an ASPEED AST2400 and will leverage the Honey Badger design. The BMC will be accessible from the host server via USB bus from the host CPU subsystem and also through a x1 PCIe connection from the PCIe switch. There is no direct connection to Ethernet. An Ethernet connection is only provided as a BOM option for debug and bring-up purpose if needed. Additionally, the BMC FW will be based on OpenBMC and any changes will also be open sourced.

Here is the list of differences from the Honey Badger BMC:

Item	Add:	Remove:
1	I2C to PCIe switches	SMBUS to NIC
2	I2C to SSDs	All LAN and NIC support
3	PCIe switch configuration and reset	All micro-server connections and support
4	PCIe cable detection	All SAS HBA, re-driver, and switch connections
5	USB connection from Mini-SAS HD connector to BMC	
6	PCIe x1 connection from PCIe switch to BMC	
7	Support for 1 host (POR), or 2 or 4 hosts (Stretch Goal)	
8	PCIe switch configuration to enable multiple hosts (Stretch Goal)	

Table 15 Lightning BMC Compared to Honey Badger

9.2 Host to Lightning BMC (Prioritized)

Below lists the three ways for connections between host and Lightning BMC. Both "1" and "2" have been verified in Lightning:

1. Connection via USB
2. Connection via PCIe

3. Connection via I2C

9.2.1 Connection via USB

This connection will be the primary mode of communication between the Host and Lightning BMC. A USB connection runs from the Leopard to the USB hub on the Re-timer card which will have USB signals running over the mini SAS HD cable to the Lightning BMC. The connection between the Host OS and Lightning BMC will be using the IPv6 over USB interface. It is bi-directional. This will enable the Host OS to connect to OpenBMC via SSH and RESTful interface.

9.2.2 Connection via PCIe

The Lightning BMC will be connected to the PCIe switch via 1x PCIe lane, then to the host. In this case BMC will always be end device only. The Host OS will communicate with the Lightning BMC via MCTP (or any other PCIe protocol). This PCIe interface is mainly used for in-band programming of the BMC FW.

9.2.3 Connection via I2C

This is provided only as a backup option and is not POR. There are two portions included:

The Lightning BMC will be connected to one or more host BMCs via I2C. The I2C connection is routed through one or more PCIe cables to the host re-timer cards and then connected to the host BMC via the PCIe riser cards in Leopard servers. Up to 4x I2C connections (or 4x hosts) may be connected to a single Lightning BMC.

9.3 Lightning BMC Support

Lightning BMC features:

- Provide access to various FRUID devices on PEB, PDPB, retimer card, M.2 carrier and FCB via fru utility
- Provide the current reading from the requested current sensor via sensor utility
- Provide current firmware version by cat /etc/issue command
- Control LEDs via front panel control utility
- Support BMC reset and BMC FW update via firmware update tool
- Reset PCIe Switch via front panel control utility
- Monitor all the sensors and log errors in the System Event Log (SEL).
- Provide thermal management of the system by controlling the fans by making use of user provided Fan Speed Control configuration.

9.4 BMC Heartbeat

A BMC hearbeat LED circuit is used to provide a visual confirmation that the BMC is up and running.

9.5 BMC Watchdog Timer

Firmware shall enable the WDT capability on BMC hardware to make sure critical daemons on OpenBMC firmware are running properly and handle the errors by either restarting the offending application or rebooting BMC.

9.6 BMC UART

The BMC UART (pin header) will be connected to a debug header to enable BMC firmware debug or development. A dedicated header for BMC UART access also provided in case

debug header is configured to PCIe switch UART port.

9.7 Debug Support

An OCP debug header will be provided to provide the BMC UART, reset, and LED codes which can be used as diagnostic codes.

9.8 BMC Firmware Update Approach

BMC firmware can be updated through below approaches (also prioritized):

- Approach 1: Host CPU –> PCIe –> BMC –> SPI Flash. In-band; Doesn't need BMC to be functional.
- Approach 2: Host CPU –> USB –> BMC –> SPI Flash. In-band; Need BMC to be fully functional.

9.9 BMC Reset

If the BMC firmware is hung, or needs to be reset to reload firmware (for example, after it has been programmed), an I2C to GPIO expander is provided to allow the server BMC to reset the Lightning BMC.

9.10 Multiple Hosts (Stretch Goal)

The hardware design should make a "best effort" to support the future use case of multiple hosts at minimal changes or no change (preferably the same PCB with different BOM options).

Below is the stretch goal:

If multiple hosts are connected to the same Lightning BMC, it will respond differently to some commands (shared infrastructure vs. SSDs). The Lightning BMC will maintain the PCIe switch configuration and SSD to host mapping. For read only commands related to shared infrastructure, the BMC will respond with the same data to all hosts. If multiple hosts send a command that impacts the shared infrastructure (e.g. update firmware or fan tables), the BMC will only process the first command received and will ignore subsequent commands until TBD seconds after the first command has completed.

Since the SSD to host mapping changes depending on the number of hosts connected, the BMC must only allow the host BMC to see or control SSDs that apply to the current map. This mapping is defined by the switch configuration.

The current map can only be changed in TBD manner and must be saved in a non-volatile memory to ensure that a BMC reset will not lose the current configuration.

9.11 BMC Sensor List

Table 16 provides a list of all of the sensors that the BMC tracks, a description of each sensor and the corresponding thresholds that the BMC monitors.

BMC Sensor Name	Description	Lower Threshold	Upper Threshold
PEB_P12V	Current voltage of 12V rail on the PEB	11.25	13.63
PEB_P5V	Current voltage of 5V rail on the PEB	4.5	5.5

PEB_P3V3	Current voltage of 3.3V rail on the PEB	3.02	3.63
PEB_P1V8	Current voltage of 1.8V rail on the PEB	1.62	1.98
PEB_P1V53	Current voltage of 1.53V rail on the PEB	1.4	1.68
PEB_P0V9	Current voltage of 0.9V rail on the PEB	0.81	0.99
PEB_P0V9_E	Current voltage of 0.9V rail on the PEB	0.81	0.99
PEB_P1V26	Current voltage of 1.26V rail on the PEB	1.13	1.39
PEB_HSC_IN_VOLT	Current voltage of the 12V rail coming into the PEB hot-swap controller	11.25	13.63
PEB_HSC_OUT_CURR	Current current of the 12V rail going out of the PEB hot-swap controller	NA	8
PEB_HSC_IN_POWER	Current power consumption of the PEB	NA	96
PCIE_SW_TEMP	Current temperature of the PCIe switch	NA	95
SYS_INLET_TEMP	Current temperature from a temperature sensor on the PEB near the left PCIe connector. Used for fan control	NA	50
PDPB_P12V	Current voltage of the 12.V rail on the PDPB	11.25	13.63
PDPB_P3V3	Current voltage of the 3.3V rail on the PDPB	2.98	3.63
LEFT_REAR_TEMP	Current temperature from a temperature sensor on the PDPB	NA	55

LEFT_FRONT_TEMP	Current temperature from a temperature sensor on the PDPB	NA	55
RIGHT_REAR_TEMP	Current temperature from a temperature sensor on the PDPB	NA	55
RIGHT_FRONT_TEMP	Current temperature from a temperature sensor on the PDPB	NA	55
SSD_0_TEMP	Current temperature of the SSD in slot 0	NA	75
SSD_1_TEMP	Current temperature of the SSD in slot 1	NA	75
SSD_2_TEMP	Current temperature of the SSD in slot 2	NA	75
SSD_3_TEMP	Current temperature of the SSD in slot 3	NA	75
SSD_4_TEMP	Current temperature of the SSD in slot 4	NA	75
SSD_5_TEMP	Current temperature of the SSD in slot 5	NA	75
SSD_6_TEMP	Current temperature of the SSD in slot 6	NA	75
SSD_7_TEMP	Current temperature of the SSD in slot 7	NA	75
SSD_8_TEMP	Current temperature of the SSD in slot 8	NA	75
SSD_9_TEMP	Current temperature of the SSD in slot 9	NA	75
SSD_10_TEMP	Current temperature of the SSD in slot 10	NA	75

SSD_11_TEMP	Current temperature of the SSD in slot 11	NA	75
SSD_12_TEMP	Current temperature of the SSD in slot 12	NA	75
SSD_13_TEMP	Current temperature of the SSD in slot 13	NA	75
SSD_14_TEMP	Current temperature of the SSD in slot 14	NA	75
M.2_Amb_TEMP_0	Current temperature of the M.2 adapter card in slot 0	NA	65
M.2_Amb_TEMP_1	Current temperature of the M.2 adapter card in slot 1	NA	65
M.2_Amb_TEMP_2	Current temperature of the M.2 adapter card in slot 2	NA	65
M.2_Amb_TEMP_3	Current temperature of the M.2 adapter card in slot 3	NA	65
M.2_Amb_TEMP_4	Current temperature of the M.2 adapter card in slot 4	NA	65
M.2_Amb_TEMP_5	Current temperature of the M.2 adapter card in slot 5	NA	65
M.2_Amb_TEMP_6	Current temperature of the M.2 adapter card in slot 6	NA	65
M.2_Amb_TEMP_7	Current temperature of the M.2 adapter card in slot 7	NA	65
M.2_Amb_TEMP_8	Current temperature of the	NA	65

	M.2 adapter card in slot 8		
M.2_Amb_TEMP_9	Current temperature of the M.2 adapter card in slot 9	NA	65
M.2_Amb_TEMP_10	Current temperature of the M.2 adapter card in slot 10	NA	65
M.2_Amb_TEMP_11	Current temperature of the M.2 adapter card in slot 11	NA	65
M.2_Amb_TEMP_12	Current temperature of the M.2 adapter card in slot 12	NA	65
M.2_Amb_TEMP_13	Current temperature of the M.2 adapter card in slot 13	NA	65
M.2_Amb_TEMP_14	Current temperature of the M.2 adapter card in slot 14	NA	65
FCB_P12V_AUX	Current voltage of the 12.V rail on the FCB	11.25	13.63
FCB_P12VL	Current voltage of the 12.V rail on the FCB	11.25	13.63
FCB_P12VU	Current voltage of the 12.V rail on the FCB	11.25	13.63
FCB_P3V3	Current voltage of the 3.3V rail on the FCB	3.1	3.63
FCB_HSC_IN_VOLT	Current voltage of the 12V rail coming into the system from Open Rack	11.25	13.63
FCB_HSC_OUT_CURRENT	Current current of the 12V rail going out of the FCB hot-swap controller	NA	60

FCB_HSC_IN_POWER	Current power consumption of the entire system (both trays)	NA	750
FCB_BJT_TEMP_1	Current temperature from a temperature sensor on the FCB	NA	60
FCB_BJT_TEMP_2	Current temperature from a temperature sensor on the FCB	NA	60
FAN1_FRONT_SPEED	Current fan speed of the front rotor in fan 1	400	NA
FAN1_REAR_SPEED	Current fan speed of the rear rotor in fan 1	400	NA
FAN2_FRONT_SPEED	Current fan speed of the front rotor in fan 2	400	NA
FAN2_REAR_SPEED	Current fan speed of the rear rotor in fan 2	400	NA
FAN3_FRONT_SPEED	Current fan speed of the front rotor in fan 3	400	NA
FAN3_REAR_SPEED	Current fan speed of the rear rotor in fan 3	400	NA
FAN4_FRONT_SPEED	Current fan speed of the front rotor in fan 4	400	NA
FAN4_REAR_SPEED	Current fan speed of the rear rotor in fan 4	400	NA
FAN5_FRONT_SPEED	Current fan speed of the front rotor in fan 5	400	NA
FAN5_REAR_SPEED	Current fan speed of the rear rotor in fan 5	400	NA
FAN6_FRONT_SPEED	Current fan speed of the front rotor in fan 6	400	NA

FAN6_REAR_SPEED	Current fan speed of the rear rotor in fan 6	400	NA
AIRFLOW	Current airflow in Lightning	NA	NA

Table 16 BMC Sensor List

9.12 Error Code display on Debug Card

Similar to Knox or Honey Badger, the BMC will display an error code on the debug card when it's plugged in. For the specific error codes and their definition, please refer to Chapter 16, Appendix B.

10 High Level System Consideration

This chapter describes the high level system consideration when connecting Lightning to host(s).

10.1 Supported Servers

The first server to support the Lightning system will be Leopard populated with Broadwell-EP CPUs. Additional server support may be required in the future.

10.2 PCIe Re-timer Card

The re-timer card is installed in the host server and may be connected to one, two, or four Lightning trays. Figure 31 shows the block diagram. The upstream port of the re-timer is connected to the host through a standard PCIe card edge connector with x16 PCIe lanes and sideband signals. The downstream port of the re-timer is connected to four MiniSAS HD connectors, each of which has x4 PCIe lanes. The card can be configured to 3 modes: 1x16, 2x8 or 4x4. For detail information of the re-timer card, please refer to the re-timer card specification: Facebook PCIe Re-timer card_1.1.

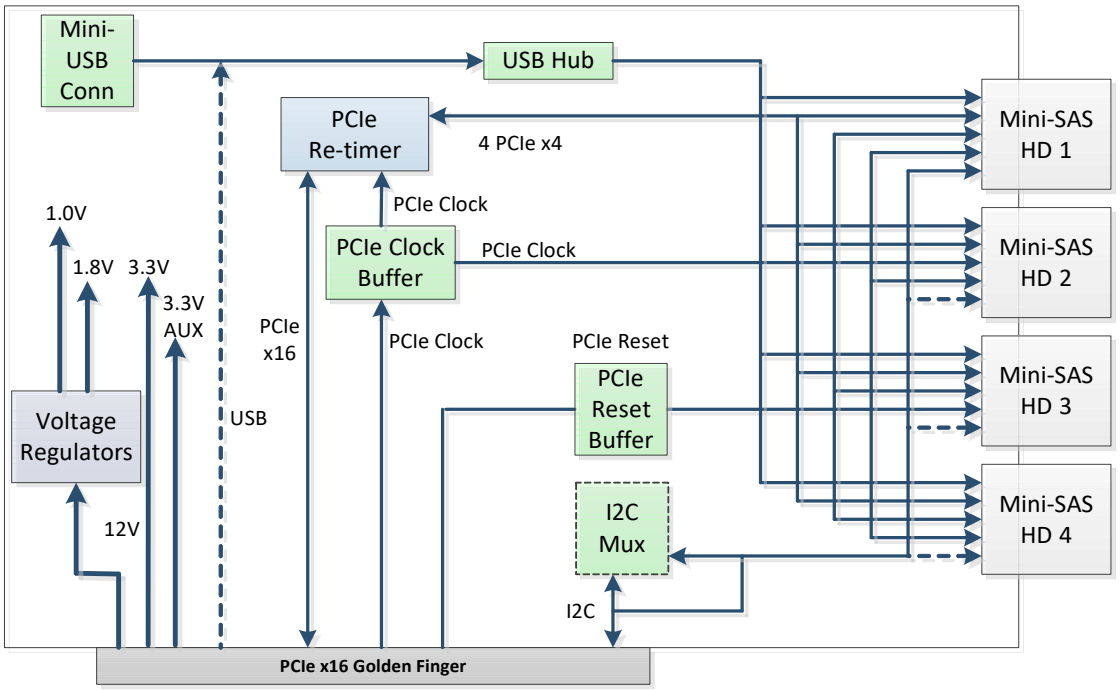


Figure 31 PCIe Re-timer Card Block Diagram

10.3 Supported SSDs

The following table provides the list of SSDs currently being investigated.

Supplier	Product	Capacity	Form Factor
Intel	P4501	4TB	2.5”
Samsung	PM963	4TB	2.5”
Seagate	XM1441	2TB	M.2

Table 17 Lightning Supported SSDs

10.4 PCIe Cables

10.4.1 Cabling Requirements

The host to PEB cables will be custom mini-SAS HD cables. The cables can be either x4 or x8 cables. Since the connectors are all x4, the x8 cables will actually bundle two x4 connectors together. The cables have a customized pinout to enable routing a full complement of sideband signals. The cables will be available in 1.5M length.

The following table shows the part numbers:

Supplier	Length	Type	Part Number
Molex	1.5M	X8	1110752201
Amphenol	1.5M	X8	NEETCT-F402

Table 18 Lightning PCIe Cable Solution

10.4.2 Cable Pinout and Connection

All cables have the same pinout and connections regardless whether they are x4 or x8 cables. The cables will have the following pinout and connections:

For Mini-SAS cable connector pin-out, please refer to Section 5.5.2.

11 Thermal Design Requirements

To meet thermal reliability requirement, the thermal and cooling solution should dissipate heat from the components when system operating at its maximum thermal power. The thermal solution should be found by setting a high-power target for initial design in order to avoid redesign of cooling solution; however, the final thermal solution of the system should be most optimized and energy efficient under data center environmental conditions with the lowest capital and operating costs. Thermal solution should not allow any overheating issue for any components in system. Switches or SSDs should not throttle due to any thermal issue under following environment.

- Inlet temperature lower than or equal to 35°C, and 0 inch H_2O datacenter pressure with all FANs in each thermal zone running properly
- Inlet temperature lower than or equal to 35°C, and 0.01 inch H_2O datacenter pressure with one FAN (or one rotor) in each thermal zone failed

11.1 Data Center Environmental Conditions

The thermal design for Lightning needs to satisfy the data center operational conditions as described below.

11.1.1 Location of Data Center/Altitude

Data centers may be located up to 2000 meters above sea level or higher. Any variation of air properties or environmental difference due to the high altitude needs to be considered when creating the thermal design.

11.1.2 Cold-Aisle temperature

Data centers will generally maintain cold aisle temperatures between 18°C and 30°C (65°F to 85°F). The mean temperature in the cold aisle is 24°C with 6°C standard deviation. The cold aisle temperature in a data center may fluctuate minutely depending to the outside air temperature of data center. Every component in system must be cooled and maintained below its maximum spec temperature in any of cold aisle temperature in a data center.

11.1.3 Cold-Aisle Pressurization

Data centers will maintain the cold aisle pressure to be between 0" H_2O and 0.05" H_2O. The thermal solution of the system should be considered the worst operational pressurization in a data center, which it is 0" H2O, and 0.01 "H2O with a single fan (or one rotor) failure.

11.1.4 R.H

Most data centers will maintain the relative humidity to be between 20% and 90%. In the thermal design, the environmental condition changes due to the high altitude may not be considered when the thermal design can meet the requirement with maximum relative humidity, 90%.

11.2 Lightning operational conditions

11.2.1 System loading

The power consumptions of individual components in the system will vary by application or by Lightning SKU. The total power consumption of system also may vary with use or with the number and type of SSDs in the system. Please see summary below.

- System loading: 100% reads to all SSDs to 100% writes to all SSDs
- Number of 2.5" SSDs: 30
- Number of M.2 SSDs: 60

Plan of record worst case configuration for thermal and power delivery design is with 60 x M.2 SSDs connected.

11.2.2 Inlet temperature

The inlet air temperature will vary. The cooling system should cover inlet temperatures as 20C, 25C, 30C, and 35C. Cooling above 30C is beyond operating specification, but used during validation to demonstrate design margin. SSD throttling is not allowed to activate over the validation range 0C – 35C.

11.2.3 Pressurization

Except for the condition when one fan in Lightning fails, the thermal solution should not be found with considering extra airflow from a data center cooling system. If and only if one fan fails, the negative or positive DC pressurization can be considered in the thermal solution in the hot aisle or in cold aisle respectively.

11.2.4 Fan Redundancy

The fans at N+1 rotor redundancy should be sufficient for cooling Lightning components to temperatures below their maximum spec to prevent system shut down or to prevent SSD throttling.

11.2.5 Delta T

The delta T is the air temperature difference across the system or the temperature difference between outlet air temperature of system and inlet air temperature of system. The delta T must be greater than 5.5°C (9.9°F). The desired delta T is 8.3°C (15°F) when the inlet air temperature to the system is lower than 30°C.

11.2.6 Thermal margin

The thermal margin is the difference between the maximum theoretical safe temperature and the actual temperature. Thermal design should maintain margins outlined below during steady-state operation up to 35°C ambient, 6000 feet above sea-level and under single-rotor failure.

For SSDs, the minimum thermal margin required is 5°C;

For PM8536, the minimum thermal margin required is 10°C.

11.3 Thermal kit requirements

Thermal testing must be performed up to 35°C (95°F) inlet temperature to guarantee high temperature reliability.

11.3.1 Heat Sink

The heat sink design should choose to be most optimized design with lowest cost. The heat sink design should be reliable and the most energy efficient design that satisfies all the conditions described above. The ODM can always propose for different heat sink type if there is alternative way to provide cost benefits. The heat sink should be without complex installation guidance, such as air-flow direction.

11.3.2 System Fan

The system fan must be highly power-efficient with dual bearing. The propagation of vibration caused by fan rotation should be minimized and limited. The minimum frame size of fan is 60x60mm and the maximum frame size is 80x80mm. ODM can propose larger frame size of fan than 80x80mm if and only if there is alternative way to provide cost benefits. Except for the condition when one fan (or one rotor) fails, the fan power consumption in system should not be exceeding 5% of total system power excluding the fan power for inlet temperatures up to 30°C.

11.3.3 Thermal sensor

The maximum allowable tolerance of thermal sensors in Lightning boards is ±2°C.

12 Environmental Requirements and Reliability

12.1 Environmental Requirements

All the boards in Lightning should support the Lightning system to meet the following environmental requirements:

- Gaseous contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Ambient operating temperature range for system without SSD: -5°C to +45°C
- Ambient operating temperature range for system with SSD: +5°C to +35°C
- Storage temperature range: -40°C to +70°C (long-term storage)
- Transportation temperature range: -55°C to +85°C (short-term storage)
- Operating and storage relative humidity: 10% to 90% (non-condensing)
- Operating altitude with no de-rating to 2,000m (6,600 feet)

12.2 Vibration and Shock

The Lightning system should meet shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) & IEC721-3-(*) Standard & Levels. The testing requirements are listed in table 19.

	Operating	Non-Operating
Vibration	0.4g acceleration, 5 to 500 Hz, 10 sweeps at 1 octave/minute per each of the three axes (one sweep is 5 to 500 to 5 Hz)	1g acceleration, 5 to 500 Hz, 10 sweeps at 1 octave/minute per each of the three axes (one sweep is 5 to 500 to 5 Hz)
Shock	6g, half-sine 11mS, 5 shocks per each of the three axes	12g, half-sine 11mS, 10 shocks per each of the three axes

Table 19 Vibration and Shock Requirements

Shock and vibration tests need to take place while Lightning is installed in an Open Rack for different types of shipping conditions.

12.3 Mean Time Between Failures (MTBF) Requirements

The system shall have a minimum calculated MTBF of 300,000 hours at 95% confidence level at 25°C ambient temperature while running at full load.

The system shall have a minimum service life of 3 years (24 hours/day, full load, at 35°C ambient temperature).

12.4 Regulations

The vendor needs to provide CB reports of Lightning at the component level. Facebook will need these documents to have rack level CE.

13 Labels and Markings

13.1 PCBA Labels and Markings

All Lightning PCBAs shall include the following labels on the component side of the boards. The labels shall not be placed in such a way that may cause them to disrupt the functionality or the airflow path of the system.

Description	Type	Barcode Required?
Safety markings	Silkscreen	No
Vendor P/N, S/N, REV (revision would increment for any approved changes)	Adhesive label	Yes
Vendor logo, name & country of origin	Silkscreen	No
PCB vendor logo, name	Silkscreen	No
Facebook P/N	Adhesive label	Yes
Date code (industry standard: WEEK/YEAR)	Adhesive label	Yes
DC input ratings	Silkscreen	No
RoHS compliance	Silkscreen	No

WEEE symbol:  The motherboard will have the crossed out wheeled bin symbol to indicate that it will be taken back by the manufacturer for recycling at the end of its useful life. This is defined in the European Union Directive 2002/96/EC of January 27, 2003 on Waste Electrical and Electronic Equipment (WEEE) and any subsequent amendments.	Silkscreen	No

Table 20 PCBA Label Requirements

13.2 Chassis Labels and Markings

The new Lightning chassis and trays shall carry the following adhesive barcoded labels in visible locations where they can be easily scanned during integration.

Vendor and Facebook will have an agreement for the label locations.

Description
Vendor P/N, S/N, REV (revision would increment for any approved changes)
Facebook P/N (or OCP customer P/N)
Date code (industry standard: WEEK/YEAR)
The assembly shall be marked “THIS SIDE UP”, “TOP SIDE”, “UP ^” or other approved marking in bright, large characters in a color to be defined by ODM and Facebook (or OCP customer). This printing may be on the PCB itself, or on an installed component such as an air baffle. The label should be clear and easy to read in low light conditions, when viewed from above or below from 2 feet away and at an angle of approximately 60 degrees off horizontal.

Table 21 Chassis Label Requirements

14 Prescribed Materials

14.1 Sustainable Materials

Materials and finishes that reduce the life cycle impact of Lightning system should be used where cost and performance are not compromised. This includes the use of non-hexavalent metal finishes, recycled and recyclable base materials and materials made from renewable resources, with associated material certifications.

Facebook identified plastic alternatives including polypropylene plus natural fiber (PP+NF) compounds that meet functional requirements while reducing environmental impact when compared to PC/ABS. GreenGranF023T is one acceptable alternate material. JPSECO also offers a PP+NF material that is acceptable; the model number will be available at a later date. It is strongly preferred that such alternatives are identified and used. If vendor is unable to use this, or a similar alternate material, vendor will provide a list of materials that are considered and why they are not successfully incorporated.

14.2 Disallowed Components

The following components shall not be used in the design of the motherboard:

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

14.3 Capacitors and Inductors

The following limitations shall be applied to the use of capacitors:

- Only aluminum organic polymer capacitors from high-quality manufacturers are used; they must be rated 105°C
- All capacitors have a predicted life of at least 50,000 hours at 45°C inlet air temperature, under worst conditions
- Tantalum capacitors are forbidden
- SMT ceramic capacitors with case size > 1206 are forbidden (size 1206 still allowed when installed far from PCB edge, and with a correct orientation that minimizes risks of cracks)
- Ceramics material for SMT capacitors must be X7R or better material (COG or NP0 type should be used in critical portions of the design)

Only SMT inductors may be used. The use of through-hole inductors is disallowed.

14.4 Component De-Rating

For inductors, capacitors, and FETs, de-rating analysis should be based on at least 20% de-rating.

15 Appendix A: Interconnect Pin Definitions

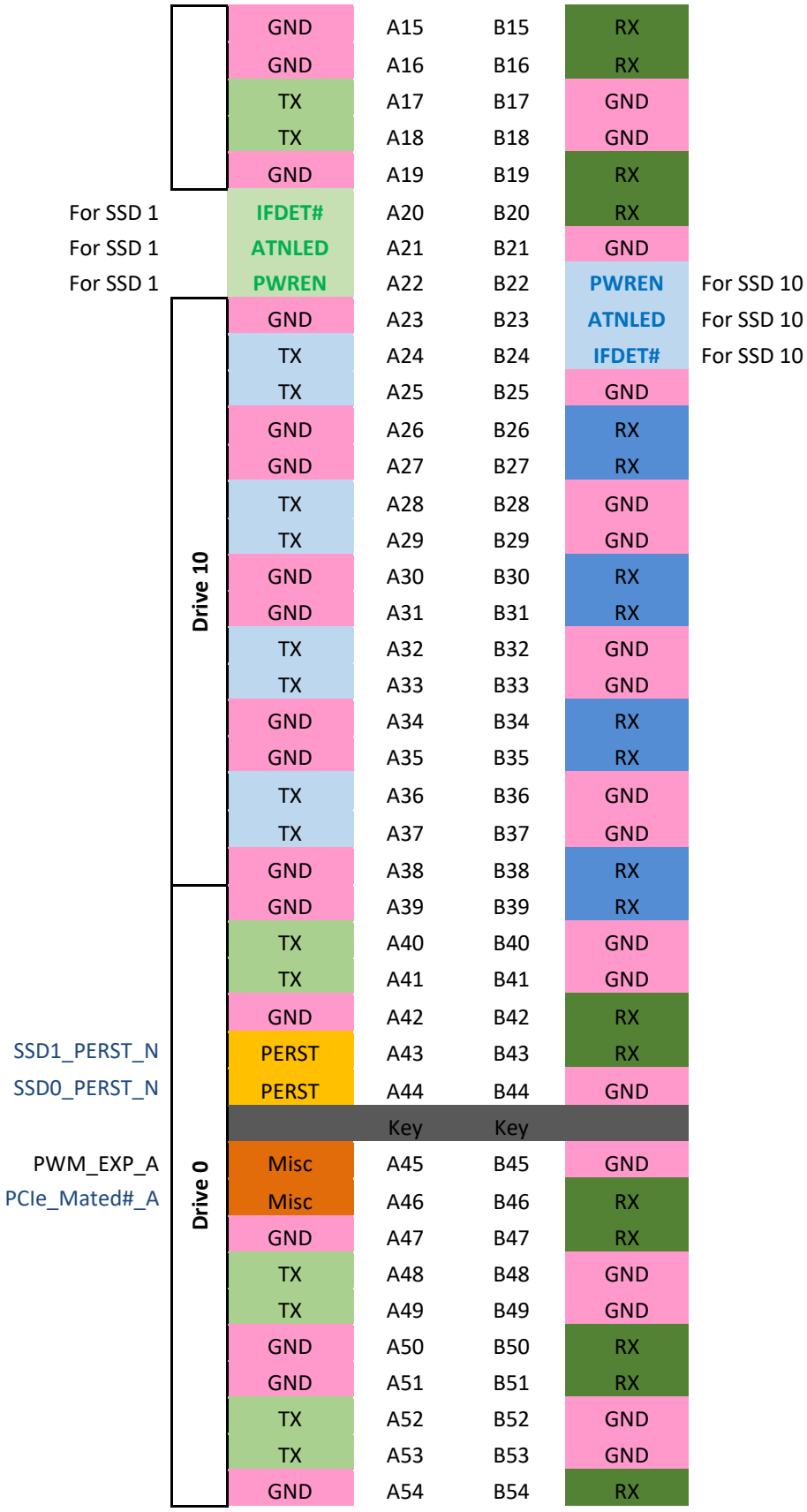
Below are the full interconnected pin definitions between PCIe Expansion Board and PCIe Drive Plane Board. The connectors are placed such that Pin1 is towards the right side of the chassis.

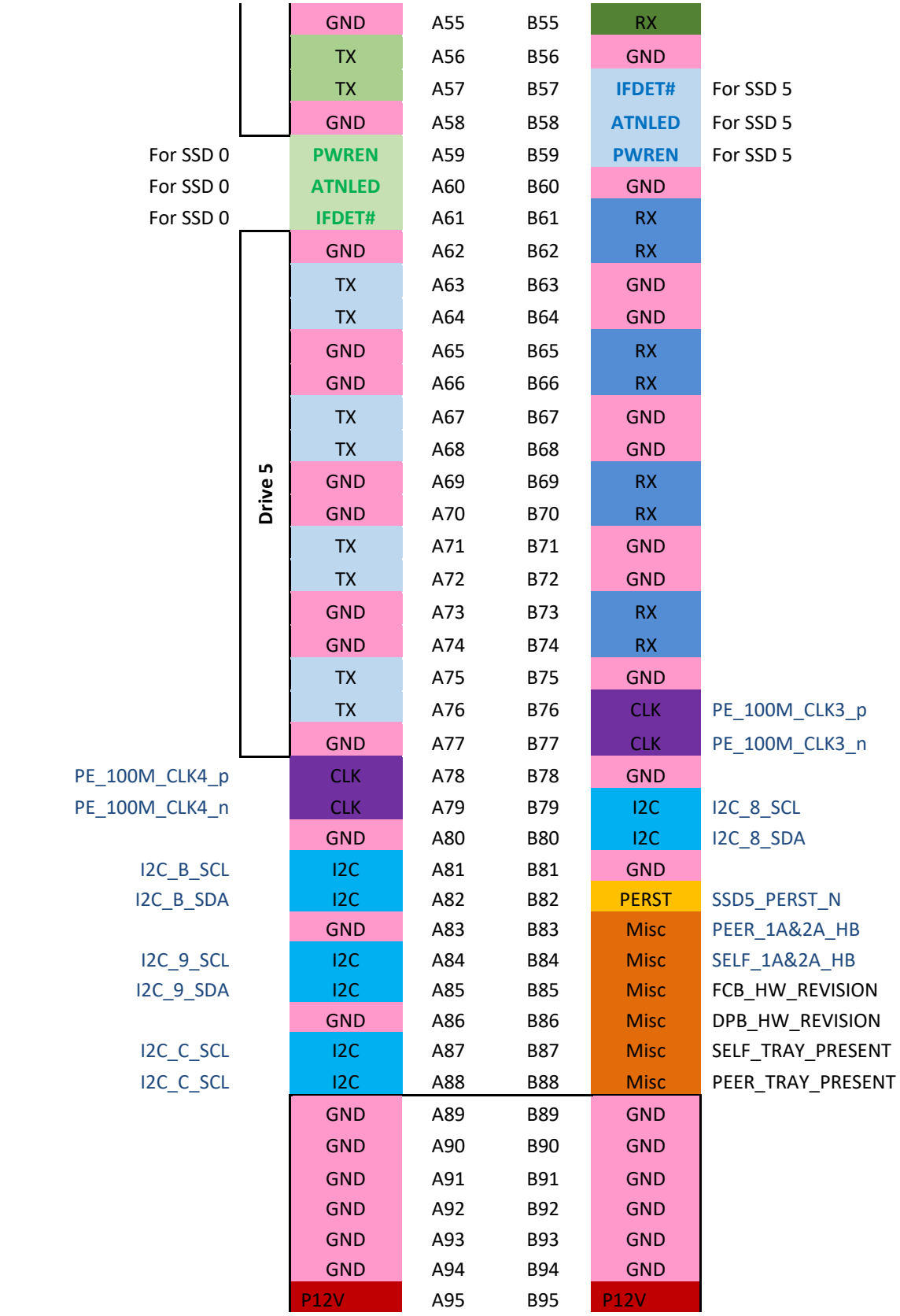
15.1 Pin definitions on PCIe Expansion Board

Below listed are the pin definitions on PCIe Expansion Board, to PCIe Drive Plane Board.

15.1.1 Pin Definition from PEB to PDPB (Connector 1)

		GND	A1	B1	GND	
PE_100M_CLK1_p		CLK	A2	B2	CLK	PE_100M_CLK2_p
PE_100M_CLK1_n		CLK	A3	B3	CLK	PE_100M_CLK2_n
	Drive 1	GND	A4	B4	GND	
		TX	A5	B5	PERST	SSD10_PERST_N
		TX	A6	B6	GND	
		GND	A7	B7	RX	
		GND	A8	B8	RX	
		TX	A9	B9	GND	
		TX	A10	B10	GND	
		GND	A11	B11	RX	
		GND	A12	B12	RX	
		TX	A13	B13	GND	
		TX	A14	B14	GND	


Drive 10
Drive 0
GND A15 B15 RX
GND A16 B16 RX
TX A17 B17 GND
TX A18 B18 GND
GND A19 B19 RX
For SSD 1 IFDET# A20 B20 RX
For SSD 1 ATNLED A21 B21 GND
For SSD 1 PWREN A22 B22 PWREN For SSD 10
GND A23 B23 ATNLED For SSD 10
TX A24 B24 IFDET# For SSD 10
TX A25 B25 GND
GND A26 B26 RX
GND A27 B27 RX
TX A28 B28 GND
TX A29 B29 GND
GND A30 B30 RX
GND A31 B31 RX
TX A32 B32 GND
TX A33 B33 GND
GND A34 B34 RX
GND A35 B35 RX
TX A36 B36 GND
TX A37 B37 GND
GND A38 B38 RX
GND A39 B39 RX
TX A40 B40 GND
TX A41 B41 GND
GND A42 B42 RX
SSD1_PERST_N PERST A43 B43 RX
SSD0_PERST_N PERST A44 B44 GND
Key Key
PWM_EXP_A Misc A45 B45 GND
PCIe_Mated#_A Misc A46 B46 RX
GND A47 B47 RX
TX A48 B48 GND
TX A49 B49 GND
GND A50 B50 RX
GND A51 B51 RX
TX A52 B52 GND
TX A53 B53 GND
GND A54 B54 RX

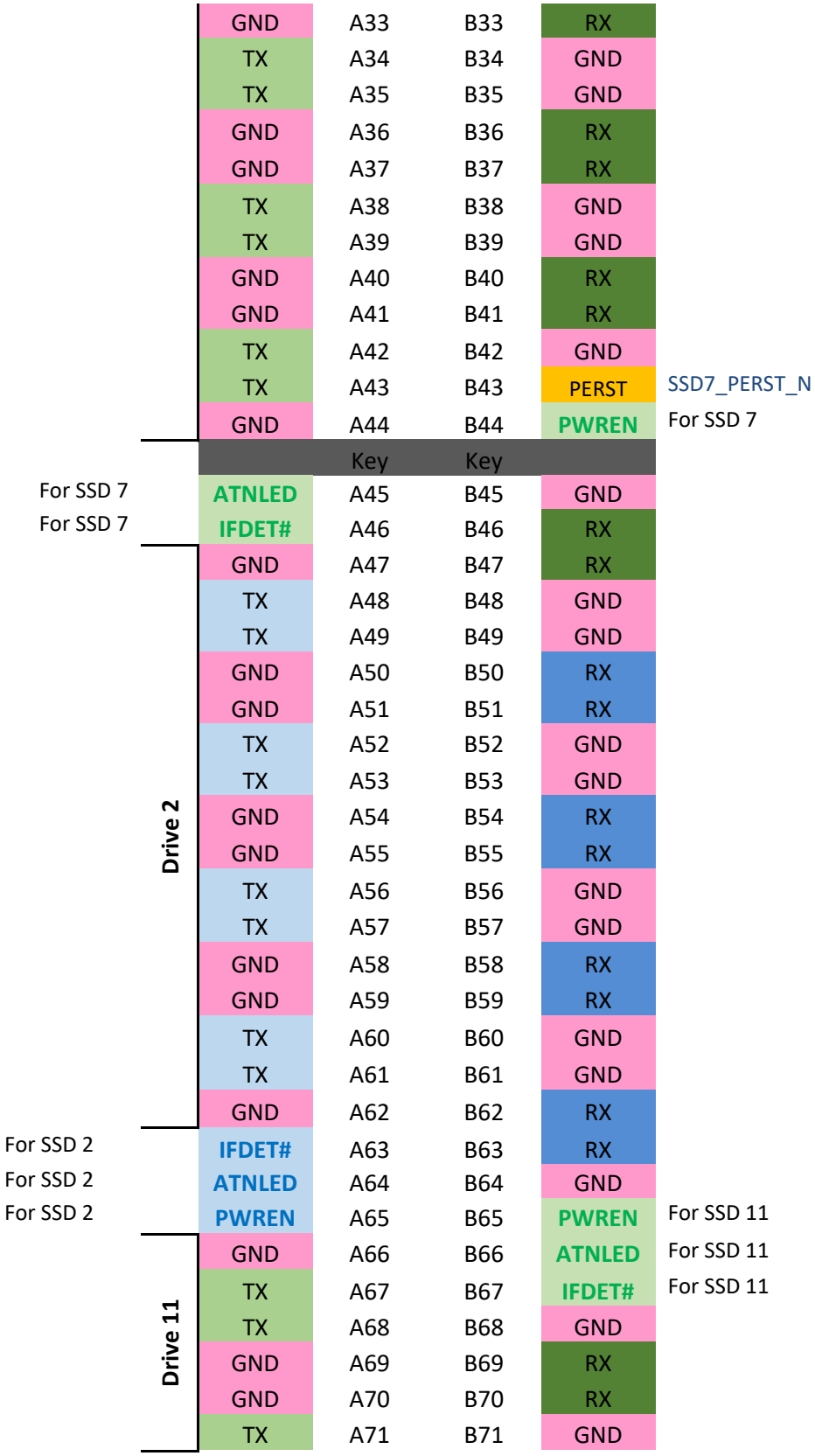

GND A55 B55 RX
TX A56 B56 GND
TX A57 B57 IFDET# For SSD 5
GND A58 B58 ATNLED For SSD 5
For SSD 0 PWREN A59 B59 PWREN For SSD 5
For SSD 0 ATNLED A60 B60 GND
For SSD 0 IFDET# A61 B61 RX
Drive 5
GND A62 B62 RX
TX A63 B63 GND
TX A64 B64 GND
GND A65 B65 RX
GND A66 B66 RX
TX A67 B67 GND
TX A68 B68 GND
GND A69 B69 RX
GND A70 B70 RX
TX A71 B71 GND
TX A72 B72 GND
GND A73 B73 RX
GND A74 B74 RX
TX A75 B75 GND
TX A76 B76 CLK PE_100M_CLK3_p
GND A77 B77 CLK PE_100M_CLK3_n
PE_100M_CLK4_p CLK A78 B78 GND
PE_100M_CLK4_n CLK A79 B79 I2C I2C_8_SCL
GND A80 B80 I2C I2C_8_SDA
I2C_B_SCL I2C A81 B81 GND
I2C_B_SDA I2C A82 B82 PERST SSD5_PERST_N
GND A83 B83 Misc PEER_1A&2A_HB
I2C_9_SCL I2C A84 B84 Misc SELF_1A&2A_HB
I2C_9_SDA I2C A85 B85 Misc FCB_HW_REVISION
GND A86 B86 Misc DPB_HW_REVISION
I2C_C_SCL I2C A87 B87 Misc SELF_TRAY_PRESENT
I2C_C_SCL I2C A88 B88 Misc PEER_TRAY_PRESENT
GND A89 B89 GND
GND A90 B90 GND
GND A91 B91 GND
GND A92 B92 GND
GND A93 B93 GND
GND A94 B94 GND
P12V A95 B95 P12V

P12V	A96	B96	P12V
P12V	A97	B97	P12V
P12V	A98	B98	P12V
P12V	A99	B99	P12V
P12V	A100	B100	P12V

Table 22 Pin Definition from PEB to PDPB (Connector 1)

15.1.2 Pin Definition from PEB to PDPB (Connector 2)

	Drive 4	GND	A1	B1	PERST	SSD11_PERST_N	
		TX	A2	B2	PERST	SSD2_PERST_N	
		TX	A3	B3	GND		
		GND	A4	B4	RX		
		GND	A5	B5	RX		
		TX	A6	B6	GND		
		TX	A7	B7	GND		
		GND	A8	B8	RX		
SSD3_PERST_N		PERST	A9	B9	RX		
For SSD 3		PWREN	A10	B10	GND		
For SSD 3		ATNLED	A11	B11	PESRT	SSD12_PERST_N	
For SSD 3		IFDET#	A12	B12	PWREN	For SSD 12	
	Drive 3	GND	A13	B13	ATNLED	For SSD 12	
		TX	A14	B14	IFDET#	For SSD 12	
		TX	A15	B15	GND		
		GND	A16	B16	RX		
		GND	A17	B17	RX		
		TX	A18	B18	GND		
		TX	A19	B19	GND		
		GND	A20	B20	RX		
		GND	A21	B21	RX		
		TX	A22	B22	GND		
		TX	A23	B23	GND		
		GND	A24	B24	RX		
		GND	A25	B25	RX		
		TX	A26	B26	GND		
		TX	A27	B27	GND		
		GND	A28	B28	RX		
	Drive 7	GND	A29	B29	RX		
		TX	A30	B30	GND		
		TX	A31	B31	GND		
		GND	A32	B32	RX		

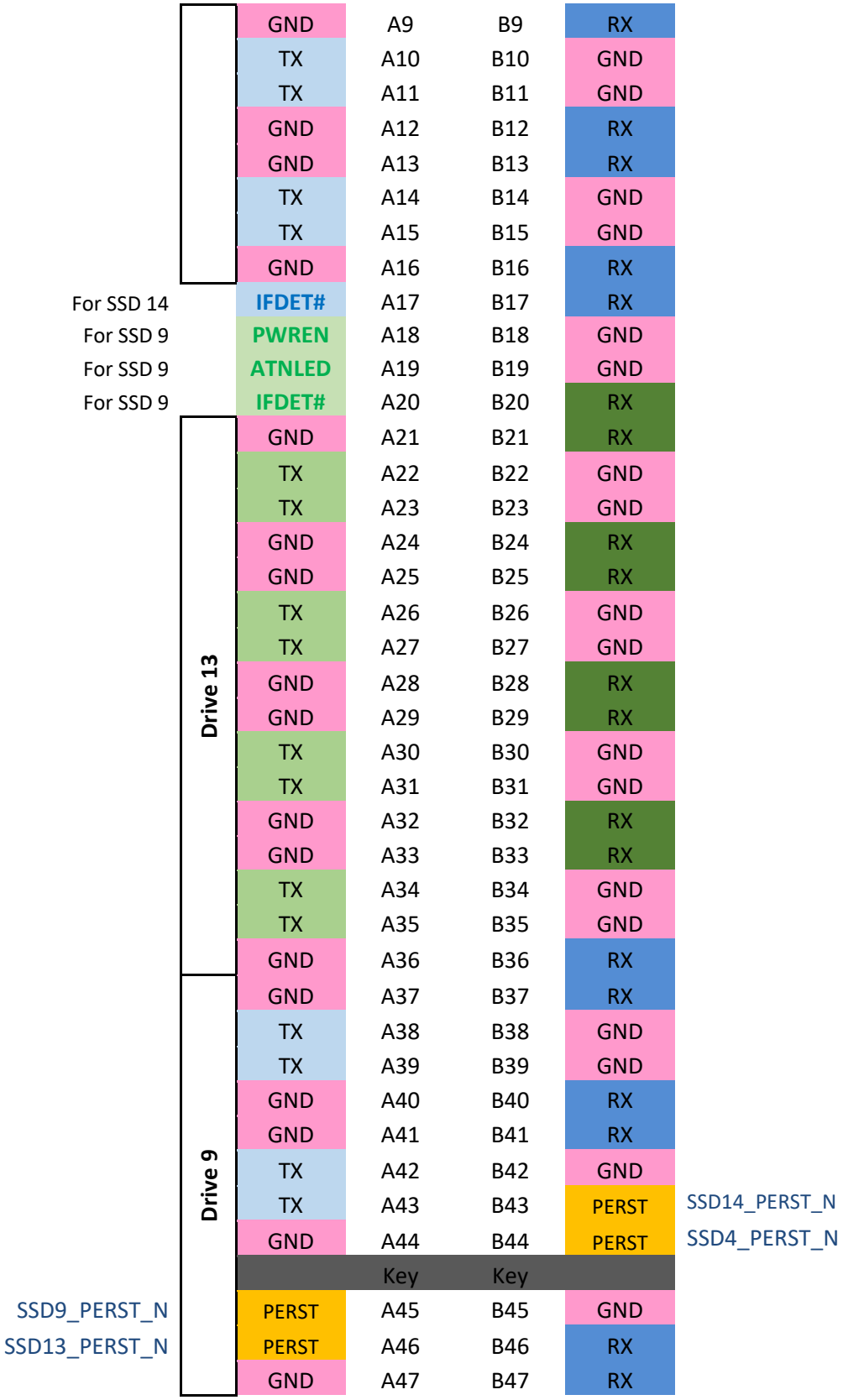

GND A33 B33 RX
TX A34 B34 GND
TX A35 B35 GND
GND A36 B36 RX
GND A37 B37 RX
TX A38 B38 GND
TX A39 B39 GND
GND A40 B40 RX
GND A41 B41 RX
TX A42 B42 GND
TX A43 B43 PERST
SSD7_PERST_N
GND A44 B44 PWREN
For SSD 7
Key Key
For SSD 7 ATNLED A45 B45 GND
For SSD 7 IFDET# A46 B46 RX
GND A47 B47 RX
TX A48 B48 GND
TX A49 B49 GND
GND A50 B50 RX
GND A51 B51 RX
TX A52 B52 GND
TX A53 B53 GND
GND A54 B54 RX
GND A55 B55 RX
TX A56 B56 GND
TX A57 B57 GND
GND A58 B58 RX
GND A59 B59 RX
TX A60 B60 GND
TX A61 B61 GND
GND A62 B62 RX
Drive 2
For SSD 2 IFDET# A63 B63 RX
For SSD 2 ATNLED A64 B64 GND
For SSD 2 PWREN A65 B65 PWREN
For SSD 11
GND A66 B66 ATNLED
For SSD 11
TX A67 B67 IFDET#
For SSD 11
TX A68 B68 GND
GND A69 B69 RX
GND A70 B70 RX
TX A71 B71 GND
Drive 11

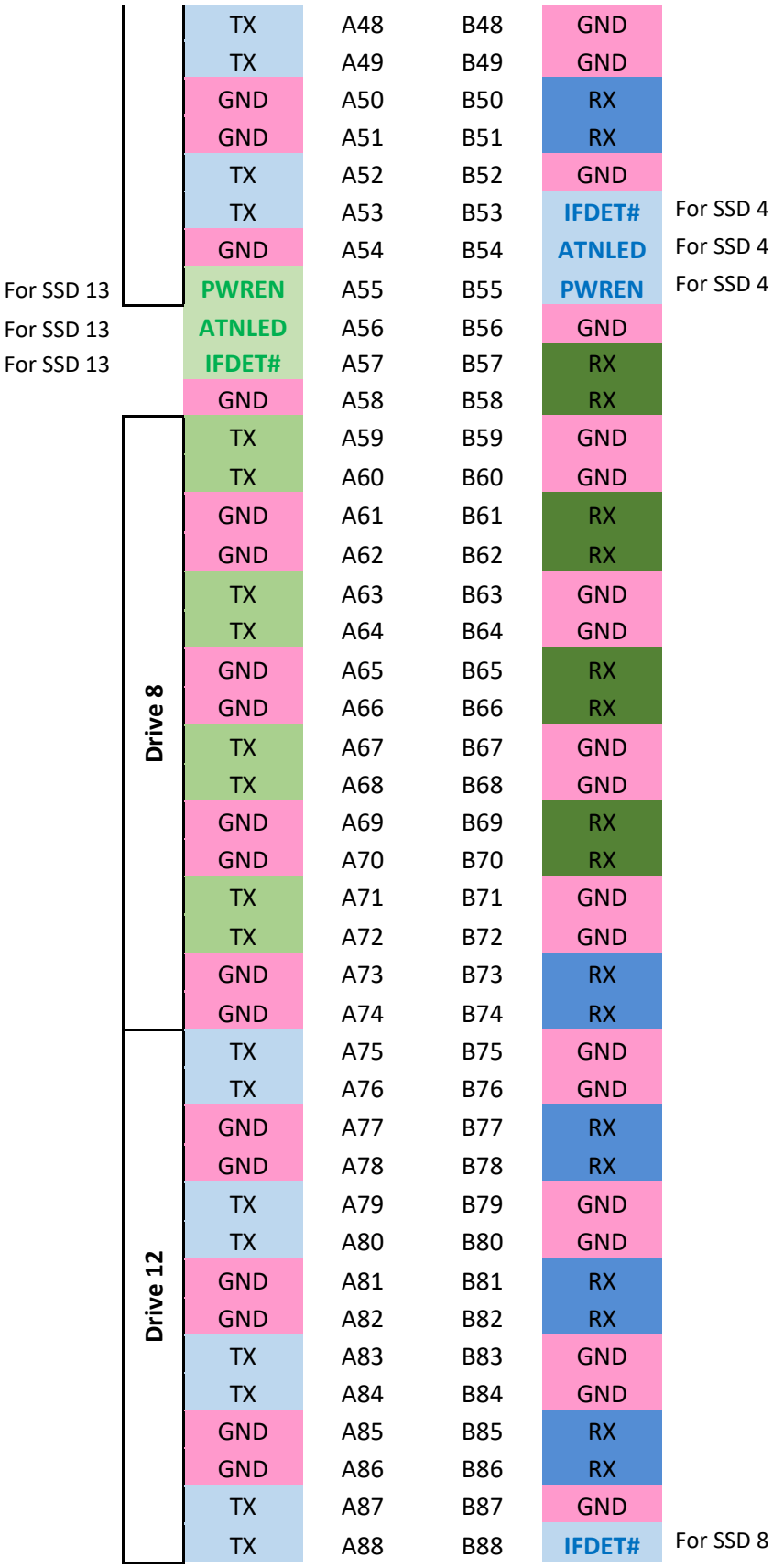
Drive 6	TX	A72	B72	GND		
	GND	A73	B73	RX		
	GND	A74	B74	RX		
	TX	A75	B75	GND		
	TX	A76	B76	GND		
	GND	A77	B77	RX		
	GND	A78	B78	RX		
	TX	A79	B79	GND		
	TX	A80	B80	GND		
	GND	A81	B81	RX		
	GND	A82	B82	RX		
	TX	A83	B83	GND		
	TX	A84	B84	GND		
	GND	A85	B85	RX		
	GND	A86	B86	RX		
	TX	A87	B87	GND		
	TX	A88	B88	GND		
	GND	A89	B89	RX		
	GND	A90	B90	RX		
	TX	A91	B91	GND		
	TX	A92	B92	GND		
	GND	A93	B93	RX		
	GND	A94	B94	RX		
	TX	A95	B95	GND		
	TX	A96	B96	GND		
	GND	A97	B97	RX		
For SSD 6	IFDET#	A98	B98	RX		
For SSD 6	ATNLED	A99	B99	GND		
For SSD 6	PWREN	A100	B100	PERST	SSD6_PERST_N	

Table 23 Pin Definition from PEB to PDPB (Connector 2)

15.1.3 Pin Definition from PEB to PDPB (Connector 3)

Drive 14	GND	A1	B1	ATNLED	For SSD 14	
	TX	A2	B2	PWREN	For SSD 14	
	TX	A3	B3	GND		
	GND	A4	B4	RX		
	GND	A5	B5	RX		
	TX	A6	B6	GND		
	TX	A7	B7	GND		
	GND	A8	B8	RX		


Drive 13
Drive 9
GND A9 B9 RX
TX A10 B10 GND
TX A11 B11 GND
GND A12 B12 RX
GND A13 B13 RX
TX A14 B14 GND
TX A15 B15 GND
GND A16 B16 RX
For SSD 14 IFDET# A17 B17 RX
For SSD 9 PWREN A18 B18 GND
For SSD 9 ATNLED A19 B19 GND
For SSD 9 IFDET# A20 B20 RX
GND A21 B21 RX
TX A22 B22 GND
TX A23 B23 GND
GND A24 B24 RX
GND A25 B25 RX
TX A26 B26 GND
TX A27 B27 GND
GND A28 B28 RX
GND A29 B29 RX
TX A30 B30 GND
TX A31 B31 GND
GND A32 B32 RX
GND A33 B33 RX
TX A34 B34 GND
TX A35 B35 GND
GND A36 B36 RX
GND A37 B37 RX
TX A38 B38 GND
TX A39 B39 GND
GND A40 B40 RX
GND A41 B41 RX
TX A42 B42 GND
TX A43 B43 PERST SSD14_PERST_N
GND A44 B44 PERST SSD4_PERST_N
Key Key
SSD9_PERST_N PERST A45 B45 GND
SSD13_PERST_N PERST A46 B46 RX
GND A47 B47 RX


Drive 8
Drive 12
TX A48 B48 GND
TX A49 B49 GND
GND A50 B50 RX
GND A51 B51 RX
TX A52 B52 GND
TX A53 B53 IFDET# For SSD 4
GND A54 B54 ATNLED For SSD 4
For SSD 13 PWREN A55 B55 PWREN For SSD 4
For SSD 13 ATNLED A56 B56 GND
For SSD 13 IFDET# A57 B57 RX
GND A58 B58 RX
TX A59 B59 GND
TX A60 B60 GND
GND A61 B61 RX
GND A62 B62 RX
TX A63 B63 GND
TX A64 B64 GND
GND A65 B65 RX
GND A66 B66 RX
TX A67 B67 GND
TX A68 B68 GND
GND A69 B69 RX
GND A70 B70 RX
TX A71 B71 GND
TX A72 B72 GND
GND A73 B73 RX
GND A74 B74 RX
TX A75 B75 GND
TX A76 B76 GND
GND A77 B77 RX
GND A78 B78 RX
TX A79 B79 GND
TX A80 B80 GND
GND A81 B81 RX
GND A82 B82 RX
TX A83 B83 GND
TX A84 B84 GND
GND A85 B85 RX
GND A86 B86 RX
TX A87 B87 GND
TX A88 B88 IFDET# For SSD 8

		GND	A89	B89	ATNLED	For SSD 8
		GND	A90	B90	PWREN	For SSD 8
	Drive 4	TX	A91	B91	PERST	SSD8_PERST_N
		TX	A92	B92	GND	
		GND	A93	B93	RX	
		GND	A94	B94	RX	
		TX	A95	B95	GND	
		TX	A96	B96	GND	
		GND	A97	B97	RX	
Tray_ID		Misc	A98	B98	RX	
SD_LATCH_RELEASE		Misc	A99	B99	GND	
PCIe_Mated#_B		Misc	A100	B100	Misc	I2C_MUX_RESET_PEB

Table 24 Pin Definition from PEB to PDPB (Connector 3)

16 Appendix B: Error Code Definition

Below listed is the full definition of Lightning Error Code. It will be displayed on the Debug Card as well as stored in system event log.

Error Code	Description
00	No error
01	PCIe internal switch fault
02	I2C bus 1 crash
03	I2C bus 2 crash
04	I2C bus 3 crash
05	I2C bus 4 crash
06	I2C bus 5 crash
07	I2C bus 6 crash
08	I2C bus 7 crash
09	I2C bus 8 crash
10	I2C bus 9 crash
11	I2C bus 10 crash
12	I2C bus 11 crash
13	I2C bus 12 crash
14	I2C bus 13 crash
15	I2C bus 14 crash
16	Reserved
17	Reserved
18	Reserved
19	Fan 1 front fault

20	Fan 1 rear fault
21	Fan 2 front fault
22	Fan 2 rear fault
23	Fan 3 front fault
24	Fan 3 rear fault
25	Fan 4 front fault
26	Fan 4 rear fault
27	Fan 5 front fault
28	Fan 5 rear fault
29	Fan 6 front fault
30	Fan 6 rear fault
31	Drive board Temp Sensor 1 (left front) critical
32	Drive board Temp Sensor 2 (left rear) critical
33	Drive board Temp Sensor 3 (right front) critical
34	Drive board Temp Sensor 4 (right rear) critical
35	PEB HSC Temp Sensor critical
36	BMC Temp Sensor critical
37	Left ambient temp sensor critical
38	Right ambient temp sensor critical
39	PCIe switch front temp tensor critical
40	PCIe switch temp sensor critical
41	Reserved
42	Reserved
43	Reserved
44	Any PEB voltage sensor critical
45	PEB current sensor critical
46	Any Drive plane board voltage sensor critical
47	Any Fan Control Board voltage sensor critical
48	Fan Control Board current sensor critical
49	Reserved
50	SSDC0 SMART Temp critical
51	SSDC1 SMART Temp critical
52	SSDC2 SMART Temp critical
53	SSDC3 SMART Temp critical
54	SSDC4 SMART Temp critical
55	SSDC5 SMART Temp critical
56	SSDC6 SMART Temp critical
57	SSDC7 SMART Temp critical
58	SSDC8 SMART Temp critical

59	SSDC9 SMART Temp critical
60	SSDC10 SMART Temp critical
61	SSDC11 SMART Temp critical
62	SSDC12 SMART Temp critical
63	SSDC13 SMART Temp critical
64	SSDC14 SMART Temp critical
65	Reserved
66	Reserved
67	Reserved
68	Reserved
69	Reserved
70	SSDC0 fault
71	SSDC1 fault
72	SSDC2 fault
73	SSDC3 fault
74	SSDC4 fault
75	SSDC5 fault
76	SSDC6 fault
77	SSDC7 fault
78	SSDC8 fault
79	SSDC9 fault
80	SSDC10 fault
81	SSDC11 fault
82	SSDC12 fault
83	SSDC13 fault
84	SSDC14 fault
85	Reserved
86	Reserved
87	Reserved
88	Reserved
89	Reserved
90	PCIe Uplink Error (left side)
91	PCIe Uplink Error (right side)
92	Reserved
93	Self Tray Pulled-out
94	Peer Tray Pulled-out
95	BMC CPU Utilization too high
96	Memory Utilization too high

97	ECC Recoverable Error
98	ECC Unrecoverable Error
99	Firmware and hardware mismatch

Table 25 Error Code Definition

17 Revision History

Author	Description	Revision Number	Date
Mike Yan	Initial draft version came from PRD for first discussion meeting.	0.01	6/25/15
Mike Yan	Updated version for EVT design implementation details.	0.1	9/11/15
Mike Yan	Updated according to EVT design finalizing: Single host as POR, Multi hosts as stretch goal; Keep GPIO expanders on PEB, route PCIe Reset to each SSD through connectors (the TBD and N/C pins); Change BMC to AST2400, connect PCIe x1 from BMC to PCIe switch.	0.2	9/25/15
Mike Yan	Updated according to further discussions and target to fix EVT design: Add USB connection through PCIe re-timer card, Mini-SAS HD cable and PEB (adding USB hub); Change I2C Mux as optional on Re-timer Card; Add single SSD Carrier for 7mm thickness; Take out 1m cable; Change location of USB connector from left to right; Details on BMC firmware update approach; A few more connectors' pin-outs.	0.3	10/09/15
Mike Yan	Updated according to further implementation before tape out EVT design: Separate PEX9716 into three I2C buses due to limited address option; Clock of PEX8780 change to clock buffer from clock gen; Sideband signal connection topology change for PLX solution; Two versions of pinout definition for Mini-SAS HD connector and cable; Reset button and slide switch on PEB; PEB to PDPB connector pin definition aligned with design; Voltage rails to be monitored on PDPB; Chapter 9 aligned with BMC PRD v0.3 with newly updates.	0.4	10/30/15
Chris Petersen Clark Shao	Removed all PLX references; Added Error code tables (Appendix B); Removed all EVT/TBD references; Removed re-drivers;	0.6	September 2016

	Table 10: updated PEB stackup to the latest; Table 6: deleted “Power off” scenarios; Table 4-2 and 4-3: added new tables for MiniSAS connection change and add comments; Figure 12: updated PEB placement, updated related description; Figure 17: updated PDPB placement; Chapter 4.1 “Lightning Key Features”: added BMC related feature updates in DVT design; Chapter 4.2 “Key Accessible Items”: modified and added some items on LED and header description; Chapter 5.5.3: Changed description RS-232 and reset items; Chapter 5.6: Updated switch and button description; Chapter 9.6: Updated description on BMC UART; Chapter 9.8: Delete USB to SPI support and renamed previous approach 4 to approach 3; Updated thermal section to include 20-90% RH and increased the altitude. Lots of clean-up!		
Clark Shao	Update all sections to reflect the latest changes in MP design, final clean up	1.0	December, 2017