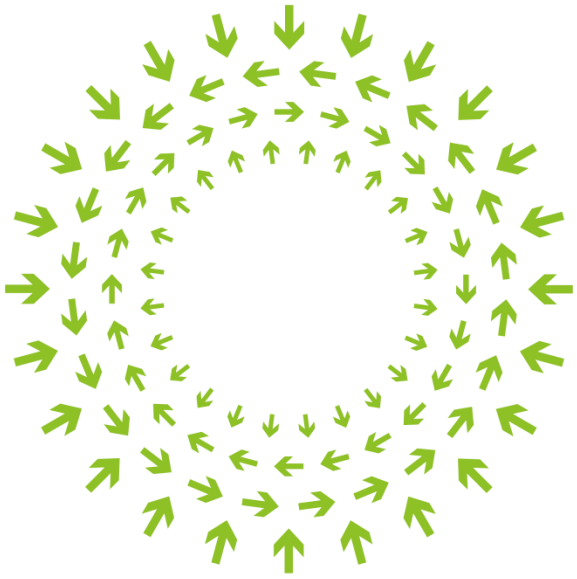

OPEN
Compute Project

Open CloudServer OCS Blade Specification Version 2.0

Author:
Mark Shaw, Director of Hardware Engineering, Microsoft

Revision History

Date	Description
10/30/2014	Version 2.0

Contents

Table of Figures

Table of Tables

1 Overview of V2.0 Open CloudServer Specifications

Table 1 lists the Open CloudServer (OCS) system specifications.

Table 1: List of specifications

Specification title	Description
Open CloudServer OCS Chassis Specification Version 2.0	Describes the hardware used in the Version 2.0 (V2.0) OCS system, including the chassis, tray, and systems management.
Open CloudServer OCS Blade Specification Version 2.0	Describes the blade used in the V2.0 OCS system, including interconnect and blade hardware and blade management.
Open CloudServer OCS Tray Mezzanine Specification Version 2.0	Describes the tray mezzanine card used in the V2.0 OCS system, including interconnect, hardware, and management.
Open CloudServer OCS NIC Mezzanine Specification Version 2.0	Describes the Network Interface Controller (NIC) mezzanine card used in the V2.0 OCS system.
Open CloudServer OCS Chassis Management Specification Version 2.0	Describes the chassis manager command-line interface (CLI).

This document is intended for designers and engineers who will be building blades for an OCS system.

2 Introduction to the Open CloudServer System

The V2.0 OCS system is a fully integrated rack of servers and IT equipment that is highly optimized and streamlined for large, web-scale deployments.

OCS is an off-the-shelf (OTS) commodity rack that is loaded with up to four modular chassis, each with trays, power supplies, power distribution, rack management, system fans, and two side-walls, as shown in Figure 1 and Figure 2.

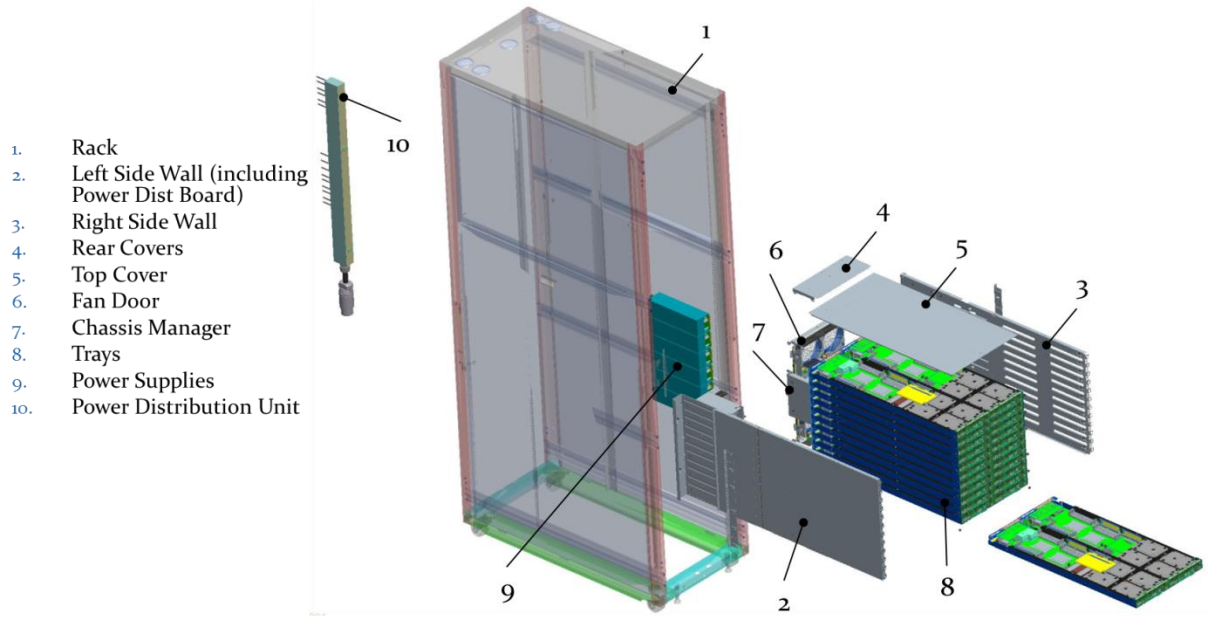

Figure 1: View of OCS with rack

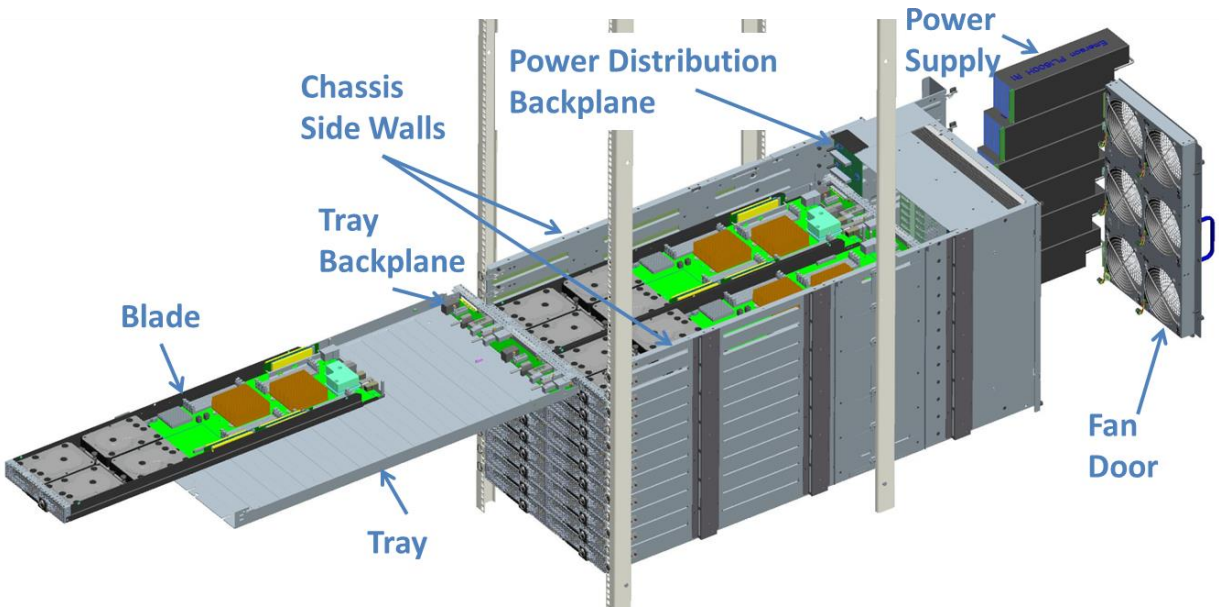

Figure 2: Second view of OCS

OCS blades are highly configurable, and are usually compute blades or storage "just a bunch of disks" (JBOD) blades.

Figure 3 shows an example of a V2.0 OCS blade.

Figure 3: V2.0 OCS blade

Each chassis supports 12 rack unit (EIA 310-E standard U or 1U, each 17.7" wide and 1.75" tall) trays that house up to 24 individual OCS blades (two blades per tray). Blades can be designed to use the full width of the tray. It is also possible to use multiple rack units to house a single tall blade, with certain restrictions.

Power, management, and networking are delivered through the tray backplane (TB) and the power distribution backplane (PDB). The tray backplane is located at the back of each tray. The power distribution backplane attaches vertically to the individual trays on one side and to the power supply unit (PSU) on the other side. This arrangement reduces the current carrying requirements of the distribution board, eliminates cabling, and reduces costs.

Power and management signals are received from the PDB and distributed to the blades by Ethernet networking cables that pass through a blind-mate connector and are routed to attachments at the rear of the chassis. Note that running the cables through the rear of the blade eliminates the need to connect directly to the servers. Once provisioned, the network cabling should only be touched when a cable or switch fails or the typology is changed. The type and number of networking switches depends on the specific deployment.

Following are the significant changes from the previous generation (V1) of the blade:

- Support for the Intel® Xeon® Processor E5-2600 v3 product family
- Support for Intel® C610 series chipset
- Increased number of Dual Inline Memory Module (DIMM) slots to 16
- Support for 8 M.2 Solid-State Drive (SSD) modules
- Support for a tray mezzanine on the tray backplane

3 Blade Block Diagram

Figure 4 shows the baseline configuration for the blade. Note that this diagram is provided for illustration and is not to scale. Note also that all requirements are specified in the text of this document.

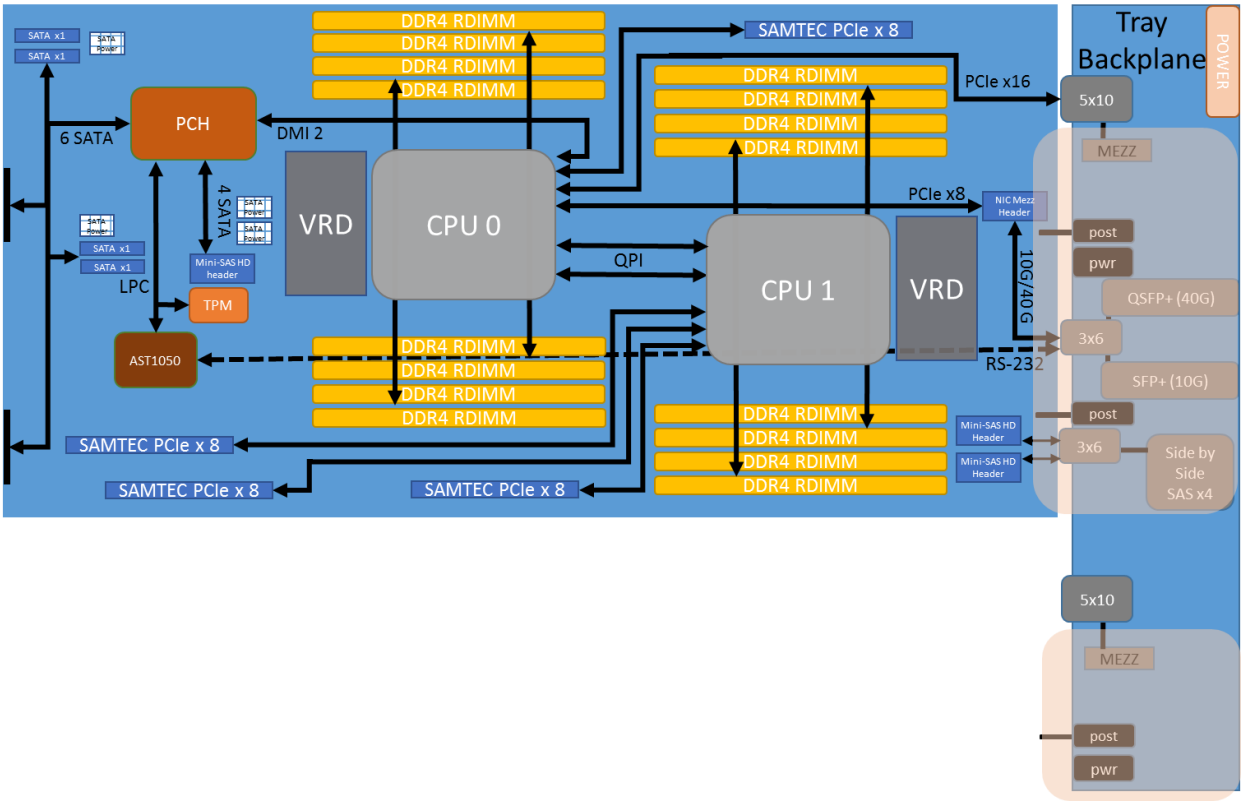

Figure 4: Baseline configuration of blade

Key features of the blade include:

- Support for up to four HDDs: two via motherboard connector, two via Serial AT Attachment (SATA) cables to the Platform Control Hub (PCH)
- Support for up to four SATA Small Form Factor (SFF) SSDs via SATA cable to the PCH
- Support for up to four Samtec Peripheral Component Interconnect Express (PCIe) x8 slots, with each slot capable of supporting two M.2 modules through an interposer board
- Support for a standard PCIe x8 riser card via a Samtec PCIe x8 edge connector
- Support for a Network Interface Controller (NIC) mezzanine card
- Support for a tray backplane mezzanine card

4 Blade Features

Table 2 lists features supported by the new blade design.

Table 2: Blade features

Processor	
Platform	Intel® Xeon® Processor E5-2600 v3 product family platform
CPU	Intel® Xeon® Processor E5-2600 v3 product family
Sockets	Dual and single socket operation
TDP Wattage	Up to 120W
Memory	
DIMM Slots	16 total DIMM slots 8 DIMMs per CPU 2 DIMMs per channel
DIMM Type	Double data rate fourth generation (DDR4) Registered DIMM (RDIMM) with Error-Correcting Code (ECC)
DIMM Speed	DDR4-1600, DDR4-1866, DDR4-2133
DIMM Size	16GB, 32GB
Capacities Supported	128GB, 192GB, 256GB, 512GB
On-board devices	
PCH	Intel® C610 series chipset
SATA	10 ports @ 6.0 Gb/s
Server management	
Chipset	Baseboard Management Controller (BMC)-lite BMC Aspeed AST1050, serial through the Chassis Manager
Interface	Representational State Transfer (REST) API Windows Management Interface (WMI) Open Management Interface (OMI) Command-Line Interface (CLI)
System firmware	
Version, Vendor	Unified Extensible Firmware Interface (UEFI) 2.3.1
Security	Trusted Platform Module (TPM 2.0) Secure Boot
PCI-Express expansion	
4 Samtec PCIe x8 slot	Slots 1-3 supports M.2 interposer modules Slot 4 supports 1 PCI-Express (PCIe) x8 riser or M.2 interposer module
Tray Mezzanine Card	Supports PCIe x16 mezzanine with bifurcation down to PCIe 4x4

NIC Mezzanine Card	Supports PCIe x8 mezzanine
Networking	
NIC mezzanine card	Single 10GbE Single 40GbE
Tray mezzanine card	Expansion slot funded by PCIe Gen3 x16

5 Labelling and Loading Guidelines

Figure 5 shows the approximate locations of major components on the blade motherboard, along with labelling information for referencing the CPU, DIMMs, PCIe riser slots, Hard-Disk Drives (HDDs), and SSDs. The sections that follow provide more detail about labelling and loading guidelines.

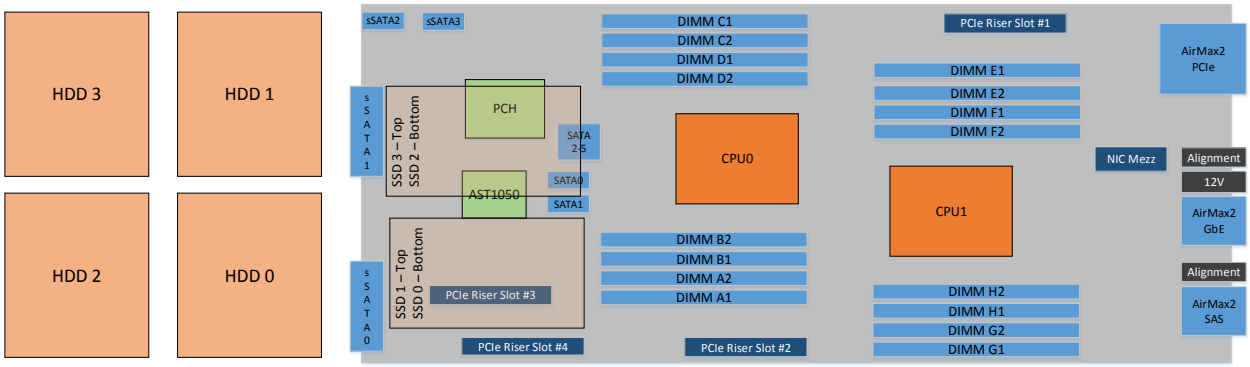

Figure 5: Major component labelling

5.1 DIMMS

Figure 5 shows how the DIMMs will be labelled. The DIMMs will be color coded to ease loading. Two colors will be used: one color for DIMMs A1,B1,C1,D1,E1,F1,G1,H1, and a second color for DIMMs A2,B2,C2,D2,E2,F2,G2,H2. Color coding can use either the DIMM connector body or the latch. Note that color selection is at the discretion of the manufacturer.

5.2 PCIe Riser Slots

The blade contains four PCIe riser slots, as shown in Figure 5. The slots use Samtec connectors to provide PCIe connectivity and to meet the height restrictions of the PCIe modules. All slots can support M.2 modules. Slot #4 can support either an M.2 module or a standard PCIe card (via PCIe riser). The blade can support both low profile and full height PCIe cards.

5.3 HDDs and SSDs

Table 3 shows how the HDD and SSD will be labelled. A label detailing the HDD numbers will be available on the frame for reference by service personnel. Table 3 shows the SATA port assignments, making it possible to readily locate and service failed drives. Loading of the drives will typically be governed by the configuration, but if a configuration supports partial loading, the drives should be loaded in numerical order.

Note that Table 3 details the assignment for two-SSD and four-SSD configurations: for a two-SSD configuration, it is anticipated that the optimal solution cables the SSDs from the two single port connectors, while for a four-SSD solution, it is anticipated that the optimal solution cables the four SSDs from the single four-port connector.

Table 3: Disk drive SATA port assignments

Signal name	I/O
SAS_B1_TX[8:1]	I
SAS_B1_RX[8:1]	O
BLADE_B1_EN1	O
SERIAL_B1_TX1	O
SERIAL_B1_RX1	I
JBOD_MATED_N	O

6 PCB Stackup

Figure 6 shows the recommended 10-layer dual stripline PCB stackup. The stackup uses standard FR4 PCB material. The PCB thickness requirement will be ~93mils.

Layer Name	Layer Type			Thickness	Copper Weight (oz)
		Soldermask		0.5	
Signal 1		SIGNAL		1.9	1.5
		Prepreg		2.7	
Plane 2	GND	VDD	GND	1.3	1.0
		Core		4.0	
Signal 3	SIGNAL	GND	SIGNAL	1.3	1.0
		Prepreg		25.0	
Signal 4	GND	SIGNAL	GND	1.3	1.0
		Core		4.0	
Plane 5	Power	GND	Power	2.6	2.0
		Prepreg		4.0	
Plane 6	Power	GND	Power	2.6	2.0
		Core		4.0	
Signal 7	GND	SIGNAL	GND	1.3	1.0
		Prepreg		25.0	
Signal 8	SIGNAL	GND	SIGNAL	1.3	1.0
		Core		4.0	
Plane 9	GND	VDD	GND	1.3	1.0
		Prepreg		2.7	
Signal 10		SIGNAL		1.9	1.5
		Soldermask		0.5	
		Total		93.2	+/-9

Figure 6: PCB Stackup

Table 4 lists the recommended PCIe mapping for the design. This mapping is used to determine feasibility of stackup support for the PCIe routing. Note that this is informational only; actual implementation may vary.

Table 4: PCIe port mapping

CPU	PCIe bus	Destination	Layer
0	PE1A	PCIe riser slot #1	1 and 10
0	PE1B	PCIe riser slot #1	1 and 10
0	PE2A	NIC	3 and 8
0	PE2B	NIC	3 and 8
0	PE2C	Not used	
0	PE2D	Not used	
0	PE3A	Tray backplane mezzanine	1 and 10
0	PE3B	Tray backplane mezzanine	1 and 10
0	PE3C	Tray backplane mezzanine	3 and 8
0	PE3D	Tray backplane mezzanine	3 and 8

CPU	PCIe bus	Destination	Layer
1	PE1A	PCIe riser slot #4	3 and 8
1	PE1B	PCIe riser slot #4	3 and 8
1	PE2A	PCIe riser slot #2	1 and 10
1	PE2B	PCIe riser slot #2	1 and 10
1	PE2C	PCIe riser slot #3	1 and 10
1	PE2D	PCIe riser slot #3	1 and 10
1	PE3A	Not used	----
1	PE3B	Not used	----
1	PE3C	Not used	----
1	PE3D	Not used	----

7 High-Speed Interface Topologies

The sections that follow detail the electrical topologies for the high-speed interfaces connecting the external boards/assemblies to the blade.

7.1 PCIe-to-Tray Backplane Mezzanine Topology

The blade contains a PCIe x16 Gen3 interface to a mezzanine card on the tray backplane. This link connects the CPU to a PCIe port on the mezzanine card. The block diagram in Figure 7 shows the connector interfaces and net length estimates for the tray backplane (TBP) and tray mezzanine card (TMC). The connector interface between the TBP and the TMC is defined in the *V2.0 Open CloudServer OCS Chassis Specification*. It is estimated that the blade motherboard can support <12” of trace between the CPU and the AirMax VS2 connector on the blade. (Note that this is an estimate derived for architectural planning and is not intended to replace signal integrity analysis.)

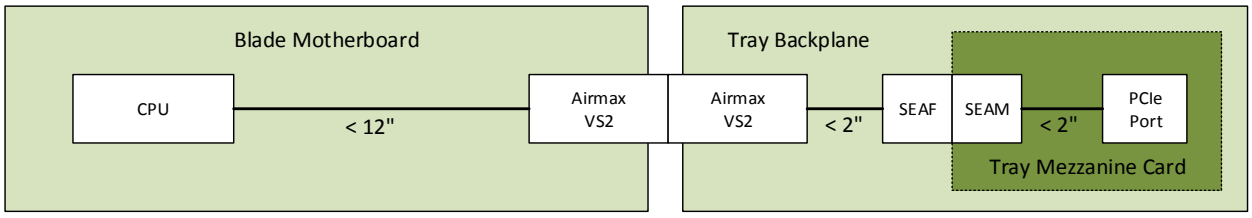

Figure 7: CPU-to-tray backplane mezzanine PCIe link topology

7.2 Blade-to-NIC Mezzanine Topology

The blade contains a NIC mezzanine card to provide small form-factor pluggable (SFP+) and Quad SFP (QSFP+) cable connectivity to the network switch, as shown in Figure 8. It is expected that this topology will require a retimer on the blade motherboard close to the AirMax VS2 connector to support the SFP+.

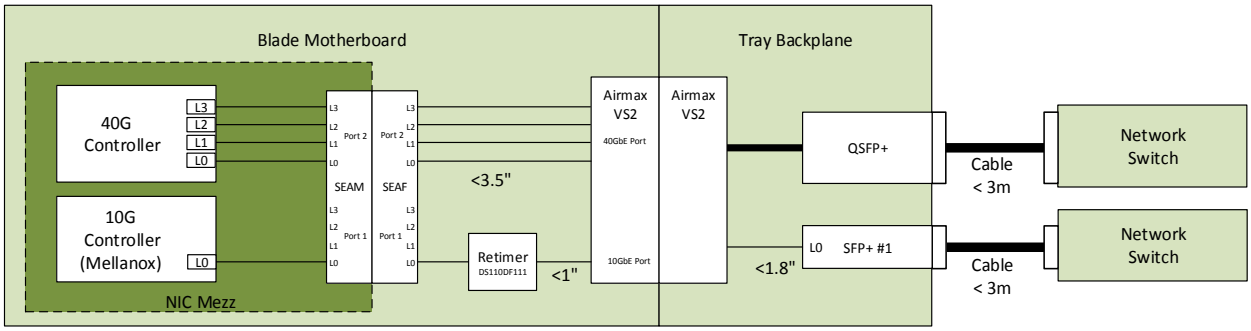

Figure 8: NIC 10GbE/40GbE topology

7.3 JBOD Storage Topology

The blade contains several topologies for delivering a high speed interface to SATA HDDs. The topologies are intended to support SATA at up to 6Gb/s or SAS at up to 12Gb/s. The block diagram in Figure 9 shows the connector and cable interfaces for the JBOD storage topology.

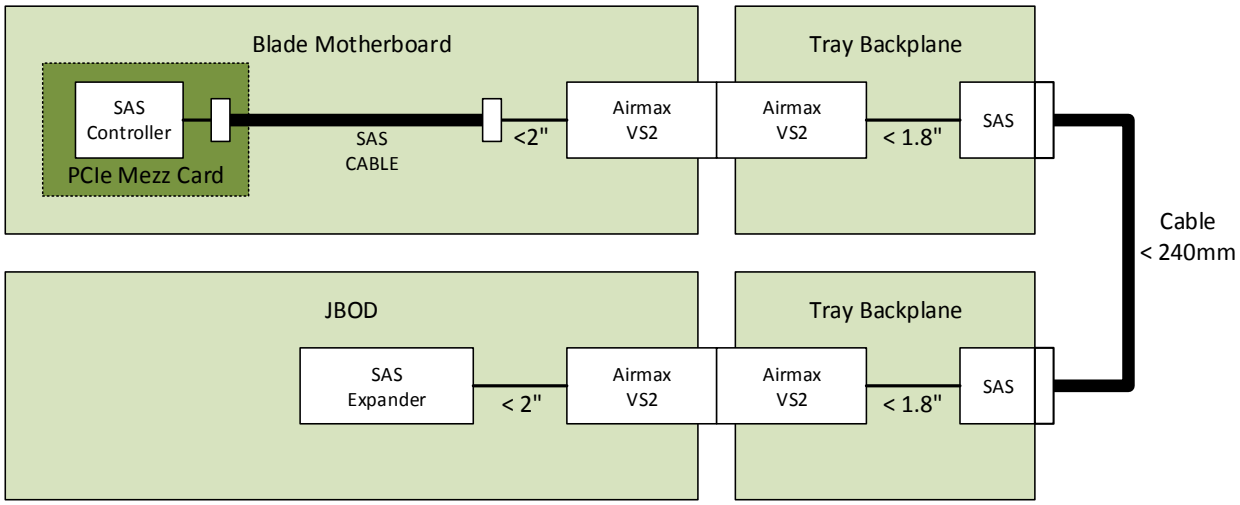

Figure 9: JBOD storage topology

Note that SAS re-timers will not be required on the blade motherboard and will be installed on the tray backplane if required.

7.4 PCIe M.2 Topology

The blade contains a PCIe Gen3 interface to an M.2 SSD module. The block diagram in Figure 10 shows the connector interface for this topology. It is estimated that the blade can support <13.5” of trace between the CPU and the HSEC8 high-speed edge card connector on the blade. (Note that this is an estimate derived for architectural planning and is not intended to replace signal integrity analysis.)

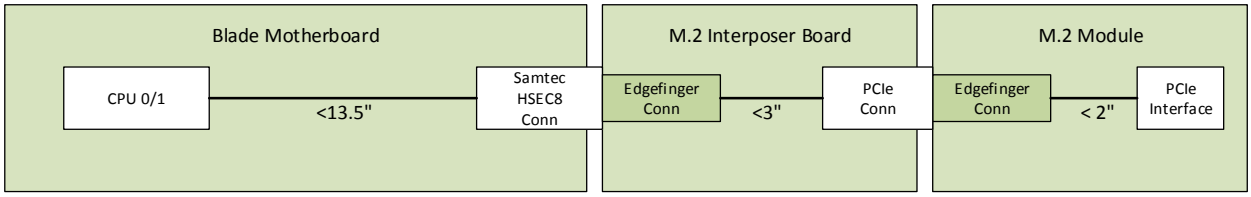

Figure 10: PCIe M.2 topology

7.5 Blade PCIe Riser Topology

The blade contains a PCIe Gen3 interface to standard PCIe card through a riser bridge board. The block diagram in Figure 11 shows the connector interface for this topology. It is estimated that the blade can support <12” of trace between the CPU and the standard PCIe connector on the blade. (Note that this is an estimate derived for architectural planning and is not intended to replace signal integrity analysis.)

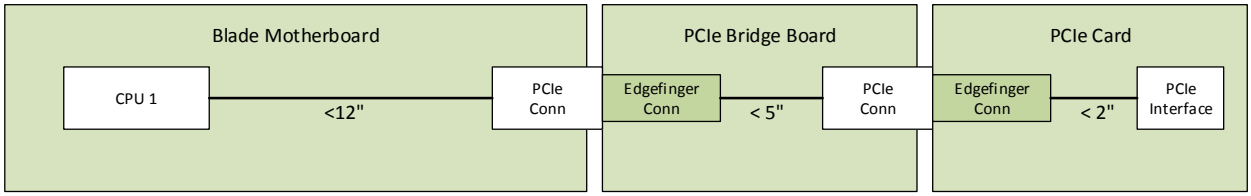

Figure 11: Blade PCIe riser topology

8 Datasafe Storage Support

The design will support datasafe storage solutions. A datasafe storage device provide nonvolatile storage backup for volatile memory, with backup power supplied by a device local energy source such as a lithium battery (Redundant Array of Independent Disks [RAID]) or Pfail circuit (M.2):

- **RAID controller**
  Uses device local battery solution. Save initiated by PERST#.
- **M.2**
  Uses device local PFAIL circuit. Save initiated by PERST#.

8.1 RAID

The design will support a RAID controller with battery backup in PCIe SLOT #4. The battery (if present) will reside within the volume space designated for the RAID PCIe solution.

8.2 M.2

The design will support M.2 storage in PCIe SLOTS 1-4. If the M.2 contains a local PFAIL solution, the solution will reside within the volume space designated for the M.2 module.

9 Blade Interconnects

The tray (or other supporting infrastructure) provides the electrical interface to the blade using the connectors listed in Table 5 (or their functional equivalents). Note that the choice of these connectors is based on a coplanar PCB for power and network distribution.

Table 5: Blade-to-tray backplane connector list

Qty	Connector description	Blade connector Manufacturer Part Number (MPN)	TBP mating connector MPN
1	AirMax VS power header, 1x2	(FCI) 10028918-001LF (or equivalent)	(FCI) 10052620-4555P00LF (or equivalent)
2	AirMax VS2 R/A header —3 pair x 6 column, 54 contact, 2mm spacing, 17mm pitch,	(FCI) 10123543-101LF (or equivalent)	(FCI) 10122643-101LF (or equivalent)
1	AirMax VS2 R/A header —5 pair x 10 column, 150 contact, 2mm spacing, 17mm pitch	(FCI) 10123529-101LF (or equivalent)	(FCI) 10122665-101LF (or equivalent)
2	Guide pin receptacle— 10.8mm right angle, 0° key	(FCI) 10037912-101LF (or equivalent)	(FCI) 10044366-101LF (or equivalent)

The chassis will provide electrical power and signaling to the tray. The tray will provide power to the blade. The blade interface to the tray backplane will provide power and high speed signaling for the tray mezzanine card on the tray backplane.

The interface to the tray backplane from a motherboard shall consist of four connectors:

1. An AirMax power connector for sourcing 12V power from PDB to blade motherboard.
2. An AirMax VS2 5x10 primarily to interface PCIe x16 Gen 3 to the tray backplane.
3. An AirMax VS2 3x6 to interface 10GbE and management signals to the tray backplane.

4. An AirMax VS2 3x6 primarily to interface SAS channels to the tray backplane.

The total amount of force required to mate the blade to the tray backplane will not exceed 18.6 pounds throughout the expected service life of the connector set. With the leverage provided by the latch at the face of the blade, the force required will not exceed 3.15 pounds. The retention force of the connectors is a minimum of 5.66 pounds, which equates to 0.94 pounds minimum force at the latch.

9.1 Tray Backplane Interface

The power connector used is in the FCI AirMax VS connector family. Figure 12 shows the power connector layout.

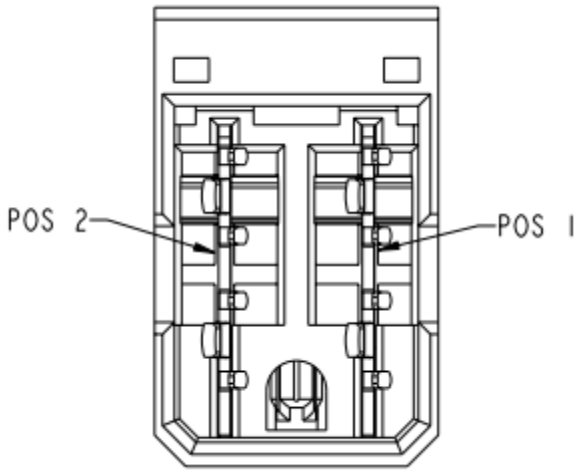

Figure 12: AirMax power receptacle pinout arrangement

Table 6 shows the power receptacle pinout

Table 6: Power connector pinout

Pin	Signal name	Capacity (in A)
Position 1	GND	40A
Position 2	12V supply	40A

The maximum power that can be delivered to a blade through this connector is 480W, assuming the connector supports 40A with 30^{o}C rise. Above this current, the Hot Swap Controller (HSC) should disable power to protect the hardware.

The signal connectors used are from the FCI AirMax VS2. Figure 13 shows an example of an AirMax coplanar connector pair.

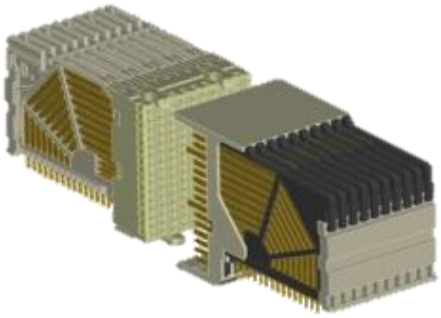

Figure 13: Example of a coplanar blade signal connector

There are three AirMax connectors interfacing the blade to the tray backplane:

- **AirMax VS2 5x10 connector**
  This connector is primarily responsible for interfacing PCIe x16 Gen 3 from the blade CPU to the tray mezzanine card on the tray backplane.
- **AirMax VS2 3x6 connector**
  This connector is responsible for interfacing 10GbE, 40GbE, and management signals from the Blade to the tray backplane. This connector is mate compatible with the V1.0 Open CloudServer blade.
- **AirMax VS2 3x6 connector**
  This connector is primarily responsible for interfacing eight 12G SAS channels to the tray backplane for storage expansion. This connector is mate compatible with the V1.0 Open CloudServer blade.

The previous-generation blade used the AirMax VS 3x6 connector. Note that the VS and VS2 family connectors are plug-in compatible.

The AirMax connector is organized as a grid, with rows A through L and columns 1 through 10. Note that the columns flip from header to receptacle so that the mated pairs match. Figure 14 shows the signal connector layout, with the blade header on the left and the tray backplane receptacle on the right.

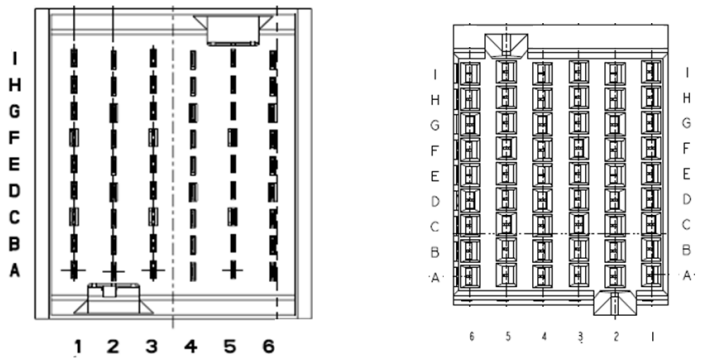

Figure 14: AirMax VS2 pinout arrangement

9.1.1 AirMax VS2 PCIe Connector

Table 7 shows the signal breakout for the AirMax PCIe header.

Table 7: AirMax VS2 PCIe connector pinout

GND	CLK_100M_P<0>	GND	PCIE_T2B_RX_DP<1>	GND	PCIE_T2B_RX_DP<3>	GND	PCIE_T2B_RX_DP<5>	GND	PCIE_T2B_RX_DP<7>	A
USB+	CLK_100M_N<0>	PCIE_T2B_RX_DP<0>	PCIE_T2B_RX_DN<1>	PCIE_T2B_RX_DP<2>	PCIE_T2B_RX_DN<3>	PCIE_T2B_RX_DP<4>	PCIE_T2B_RX_DN<5>	PCIE_T2B_RX_DP<6>	PCIE_T2B_RX_DN<7>	B
USB-	GND	PCIE_T2B_RX_DN<0>	GND	PCIE_T2B_RX_DN<2>	GND	PCIE_T2B_RX_DN<4>	GND	PCIE_T2B_RX_DN<6>	GND	C
GND	CLK_100M_P<1>	GND	PCIE_T2B_RX_DP<9>	GND	PCIE_T2B_RX_DP<11>	GND	PCIE_T2B_RX_DP<13>	GND	PCIE_T2B_RX_DP<15>	D
P5V_USB	CLK_100M_N<1>	PCIE_T2B_RX_DP<8>	PCIE_T2B_RX_DN<9>	PCIE_T2B_RX_DP<10>	PCIE_T2B_RX_DN<11>	PCIE_T2B_RX_DP<12>	PCIE_T2B_RX_DN<13>	PCIE_T2B_RX_DP<14>	PCIE_T2B_RX_DN<15>	E
GND	GND	PCIE_T2B_RX_DN<8>	GND	PCIE_T2B_RX_DN<10>	GND	PCIE_T2B_RX_DN<12>	GND	PCIE_T2B_RX_DN<14>	GND	F
GND	CLK_100M_P<2>	GND	RSVD	GND	MEZZ_SCL	GND	PCIE_CFG_ID0	GND	PCIE_RESET_N<3>	G
GND	CLK_100M_N<2>	RSVD	RSVD	MEZZ_PRESENT_N	MEZZ_EN	PCIE_WAKE_N	PCIE_CFG_ID1	PCIE_RESET_N<1>	PCIE_RESET_N<2>	H
P12V_MEZZ	GND	RSVD	GND	BLADE_MATED_N	GND	MEZZ_SDA	GND	PCIE_RESET_N<0>	GND	I
P12V_MEZZ	CLK_100M_P<3>	GND	PCIE_B2T_TX_DP<1>	GND	PCIE_B2T_TX_DP<3>	GND	PCIE_B2T_TX_DP<5>	GND	PCIE_B2T_TX_DP<7>	J
P12V_MEZZ	CLK_100M_N<3>	PCIE_B2T_TX_DP<0>	PCIE_B2T_TX_DN<1>	PCIE_B2T_TX_DP<2>	PCIE_B2T_TX_DN<3>	PCIE_B2T_TX_DP<4>	PCIE_B2T_TX_DN<5>	PCIE_B2T_TX_DP<6>	PCIE_B2T_TX_DN<7>	K
P12V_MEZZ	GND	PCIE_B2T_TX_DN<0>	GND	PCIE_B2T_TX_DN<2>	GND	PCIE_B2T_TX_DN<4>	GND	PCIE_B2T_TX_DN<6>	GND	L
GND	GND	GND	PCIE_B2T_TX_DP<9>	GND	PCIE_B2T_TX_DP<11>	GND	PCIE_B2T_TX_DP<13>	GND	PCIE_B2T_TX_DP<15>	M
P12V_MEZZ	GND	PCIE_B2T_TX_DP<8>	PCIE_B2T_TX_DN<9>	PCIE_B2T_TX_DP<10>	PCIE_B2T_TX_DN<11>	PCIE_B2T_TX_DP<12>	PCIE_B2T_TX_DN<13>	PCIE_B2T_TX_DP<14>	PCIE_B2T_TX_DN<15>	N
P12V_MEZZ	GND	PCIE_B2T_TX_DN<8>	GND	PCIE_B2T_TX_DN<10>	GND	PCIE_B2T_TX_DN<12>	GND	PCIE_B2T_TX_DN<14>	GND	O
GND	N/A	GND	N/A	GND	N/A	GND	N/A	GND	N/A	P
10	9	8	7	6	5	4	3	2	1	

Table 8 describes the signals used in this interface.

Table 8: AirMax VS2 PCIe signal definitions

Bus type	I/O	Logic	Definition for three pair, eight column
PCIE_B2T_TX_DP/N[15:0]	O	Current Mode Logic (CML)	PCIe Gen 3 data from blade to tray mezzanine
PCIE_T2B_RX_DP/N[15:0]	I	CML	PCIe Gen3 data from tray mezzanine to blade
CLK_100M_P/N[3:0]	O	CML	100MHz PCIe Gen 3 clocks
PCIE_RESET_N[3:0]	O	3.3V	PCIe reset signals
WAKE_PCIE_N	I	3.3V	PCIe wake signal
PCIE_CFG_ID[1:0]	I	3.3V	PCIe configuration ID bits Should be connected to General Purpose Input/Output (GPIO) on the PCH Should be pulled up with minimum 10K ohm resistor

Bus type	I/O	Logic	Definition for three pair, eight column
			Has 1K pulldown on tray mezzanine 00 = 1 x16 bifurcation 01 = 2 x8 bifurcation 10 = 4 x4 bifurcation 11 = N/A
PCIE_B2T_TX_DP/N[15:0]	O	CML	PCIe Gen 3 data from blade to tray mezzanine
MEZZ_SDA/SCL	I/O	3.3V	I2C from blade BMC to tray mezzanine
MEZZ_PRESENT_N	I	3.3V	Indicates tray mezzanine card is installed This signal should be pulled up on the blade and grounded on the tray mezzanine card
MEZZ_EN	O	3.3V	Enable for tray mezzanine card on-board power
P12V_MEZZ	O	12V	Tray mezzanine card 12V power from blade 12V HSC (3A)
BLADE_MATED_N	I	3.3V	Indicates blade is mated to tray backplane through AirMax 5x10 Grounded on tray backplane Should be pulled up on blade and used to qualify 12V in-rush turn on
USBP/N	I/O	CML	USB 2.0 from blade
P5V_USB	O	5V	5V USB power (500mA)
RSVD	--	--	Reserved—do not connect
GND	--	--	Ground

P12V_MEZZ and P5V_USB pins are assumed to support a derated 500mA per pin; therefore, P12V_MEZZ supports a derated maximum of 3A, and P5V_USB supports 500mA.

Table 9 shows how the reset and clock signals are mapped to support each of the bifurcation cases for PCIe to the tray mezzanine.

Table 9: PCIe bifurcation mapping

	1x16	2x8		4x4			
SIGNAL NAME	**15:0**	**15:8**	**7:0**	**15:12**	**11:8**	**7:4**	**3:0**
PCIE_RESET_N[0]	X	X		X			
PCIE_RESET_N[1]			X		X		
PCIE_RESET_N[2]						X	

	1x16	2x8		4x4			
SIGNAL NAME	**15:0**	**15:8**	**7:0**	**15:12**	**11:8**	**7:4**	**3:0**
PCIE_RESET_N[3]							X
CLK_100M_P/N[0]	X	X		X			
CLK_100M_P/N[1]			X		X		
CLK_100M_P/N[2]						X	
CLK_100M_P/N[3]							X

9.1.2 AirMax VS2 10/40GbE Connector

Table 10 shows the pinout for the AirMax 10Gb/40Gb Ethernet header.

Table 10: AirMax VS2 10/40GbE connector pinout

GND	ETH40G_RX1P	GND	ETH40G_RX2P	GND	ETH40G_TX3P	A
ETH40G_TX1P	ETH40G_RX1N	ETH40G_TX2P	ETH40G_RX2N	ETH40G_RX3P	ETH40G_TX3N	B
ETH40G_TX1N	GND	ETH40G_TX2N	GND	ETH40G_RX3N	GND	C
GND	P3V3_QSFP	GND	ETH10G_SDA	GND	ETH10G_RXP	D
ETH40G_RX0P	P3V3_QSFP	ETH40G_SDA	ETH10G_SCL	RSVD	ETH10G_RXN	E
ETH40G_RX0N	GND	ETH40G_SCL	SKU_ID0	SKU_ID2	GND	F
GND	PSU_ALERT_N	SERIAL_RX2	SERIAL_RX1	GND	ETH10G_TXP	G
ETH40G_TX0P	ETH40G_PRES_N	SERIAL_TX2	SERIAL_TX1	BLADE_EN1	ETH10G_TXN	H
ETH40G_TX0N	BLADE_MATED_N	SKU_ID1	P3V3_QSFP	ETH10G_PRES_N	GND	I
6	5	4	3	2	1	

Table 11 describes the signals used in this interface.

Table 11: AirMax VS2 10/40GbE signal definitions

Bus type	I/O	Logic	Definition for three pair, eight column
ETH10G_TXP/N	O	CML	10Gb Ethernet from blade NIC to SFP+ connector
ETH10G_RXP/N	I	CML	10Gb Ethernet from SFP+ connector to blade NIC
ETH40G_TX[3:0]P/N	O	CML	40Gb Ethernet from blade LOM to QSFP+ connector
ETH40G_RX[3:0]P/N	I	CML	40Gb Ethernet from QSFP+ connector to blade NIC mezz
ENET_40G_SDA/SCL	I/O	3.3V	I2C from Ethernet controller on blade to QSFP+ connector/cable
ENET40G_PRES_N	I	3.3V	Indicates presence of 40GbE cable from QSFP+ connector
ENET10G_SDA/SCL	I/O	3.3V	I2C from Ethernet controller on blade to SFP+ connector/cable

Bus type	I/O	Logic	Definition for three pair, eight column
ENET10G_PRES_N	I	3.3V	Indicates presence of 10GbE cable through DAC PRESENT_N signal from SFP+ connector
SKU_ID<2:0>	I	3.3V	3.3V Pu/Pd on TBP indicating the SKU of the tray backplane
P3V3_QSFP	O	3.3V	QSFP+ and SFP+ cable power
BLADE_MATED_N	I	3.3V	Indicates blade is mated to tray backplane through AirMax 3x6 Grounded on tray backplane Should be pulled up on blade and used to qualify 12V in-rush turn on
SERIAL_TX[2:1]	I	3.3V	Serial communication ports from the Chassis Manager (CM) to blade Two nodes per blade from CM
SERIAL_RX[2:1]	O	3.3V	Serial communication port from blade to CM Two nodes per blade from CM
BLADE_EN1	I	3.3V	Blade enable signal from CM Used to enable the 12V in-rush controller
PSU_ALERT_N	I	3.3V	Power supply failure alert This signal is active low (open drain) and should be pulled up on the motherboard
RSVD	I/O		Reserved—do not connect
Ground	--		Ground pins

9.1.3 AirMax VS2 SAS Connector

Table 12 shows the pinout for the AirMax SAS connector. Connector #1 supports SAS lanes 1-8.

Table 12: AirMax VS2 SAS connector pinout

GND	SAS_TX6P	GND	RSVD	GND	SAS_TX1P	A
SAS_RX8P	SAS_TX6N	RSVD	RSVD	SAS_RX3P	SAS_TX1N	B
SAS_RX8N	GND	RSVD	GND	SAS_RX3N	GND	C
GND	SAS_RX6P	GND	SAS_RX4P	GND	SAS_RX1P	D
SAS_TX8P	SAS_RX6N	SAS_TX5P	SAS_RX4N	SAS_TX3P	SAS_RX1N	E
SAS_TX8N	GND	SAS_TX5N	GND	SAS_TX3N	GND	F
GND	SAS_TX7P	GND	SAS_TX4P	GND	SAS_TX2P	G
SAS_RX7P	SAS_TX7N	SAS_RX5P	SAS_TX4N	SAS_RX2P	SAS_TX2N	H
SAS_RX7N	GND	SAS_RX5N	GND	SAS_RX2N	GND	I
6	5	4	3	2	1	

Table 13 describes the signals used in this interface.

Table 13: AirMax VS2 SAS signal definitions

Bus type	I/O	Logic	Definition for three pair, eight column
SAS_TX[8:1]	O	CML	SAS 8-channels from Blade to SAS cable
SAS_RX[8:1]	I	CML	SAS 8-channels from SAS cable to blade
RSVD	I/O	--	Reserved—do not connect
Ground	--	--	Ground pins

9.1.4 Connector Skew Compensation

Right-angle AirMax VS2 connector interfaces between the tray backplane and the blade motherboard will require signal delay compensation in the PCB to account for differences in pin length of a P/N differential pair in the mated connectors. The pin-pair skew can be obtained from the connector manufacturer.

Following are the skew compensation divisions between the blade motherboard and the tray backplane:

- **AirMax VS2 PCIe connector**
  All skew compensation is contained on the blade motherboard.
- **AirMax VS2 GbE connector**
  All skew compensation is contained on the Blade Motherboard.
- **AirMax VS2 SAS connector**
  Half (50%) of skew compensation is contained on the blade motherboard. Skew compensation for the VS2 connector pair is divided equally between the blade motherboard and the tray backplane.

9.2 PCIe M.2 Interposer Modules

The blade will support the M.2 interposer module, a custom edge card that supports two M.2 SSD modules (Next Generation Form Factor [NGFF] cards) in the connectorized SSD socket 3 format per the PCIe M.2 specification.

The interposer edge card interfaces to the blade motherboard through a Samtec HSEC8-150-01-S-DV-A connector (or equivalent). The interposer module supports the 60mm, 80mm, and 110mm form factors (Type 2260, 2280, and 22110). To support two M.2 modules, the connector interface is designed to support two PCIe Gen3x4 interfaces as well as the SSD specific signals, per the PCIE M.2 specification.

The interface is also designed to support a standard PCIe x8 interface through a separate riser card. The bifurcation is communicated through the LINK_WIDTH signal (Pin B3), which should be connected to the PCH. The PCIe card will not require I2C or Joint Test Action Group (JTAG) connections to the motherboard. Table 14 shows the pinout for supporting only the M.2 interposer module.

Table 14: PCIe M.2 interposer connector pinout

	Side B connector			Side B connector	
Pin	**Name**	**Description**	**Pin**	**Name**	**Description**
1	12V	12V power	2	PRSNT1#	Hot plug presence detect
3	12V	12V power	4	12V	12V power
5	LINK_WIDTH	0=1 x8, 1= 2x4 bifurcation	6	12V	12V power
7	GND	Ground	8	GND	Ground
9	REFCLK2+	Module 2 reference clock differential pair	10	CLKREQ1	Ref clock request (OD)
11	REFCLK2-		12	CLKREQ2	Ref clock request (OD)
13	GND	Ground	14	SUSCLK	Suspend cock (32.768Khz)
15	3.3V	3.3V power	16	DAS/DSS#	Drive active indicator
17	DEVSLP	Device sleep	18	3.3V	3.3V power
19	3.3V	3.3V power	20	3.3V	3.3V power
21	WAKE#	Ground	22	PERST#	PCIe reset
23	3.3V STBY	3.3V standby power	24	GND	Ground
25	GND	Ground	26	REFCLK1+	Module 1 reference clock differential pair
27	PETp(0)	Transmitter module 1 lane 0 differential pair	28	REFCLK1-	
29	PETn(0)		30	GND	Ground
31	GND	Ground	32	PERp(0)	Receiver module 1 lane 0 differential pair
33	PRSNT2#	Hotplug detect	34	PERn(0)	
35	GND	Ground	36	GND	Ground
37	PETp(1)	Transmitter module 1 lane 1 differential pair	38	RSVD	Reserved
39	PETn(1)		40	GND	Ground
41	GND	Ground	42	PERp(1)	Receiver module 1 lane 1 differential pair
43	GND	Ground	44	PERn(1)	
45	PETp(2)	Transmitter module 1	46	GND	Ground

	Side B connector			Side B connector	
Pin	**Name**	**Description**	**Pin**	**Name**	**Description**
47	PETn(2)	lane 2 differential pair	48	GND	Ground
49	GND	Ground	50	PERp(2)	Receiver module 1 lane 2 differential pair
51	GND	Ground	52	PERn(2)	
53	GND	Ground	54	GND	Ground
Mechanical key					
55	PETp(3)	Transmitter module 1 lane 3 differential pair	56	GND	Ground
57	PETn(3)		58	GND	Ground
59	GND	Ground	60	PERp(3)	Receiver module 1 lane 3 differential pair
61	RSVD	Reserved	62	PERn(3)	
63	PRSNT3#	Hot plug detect	64	GND	Ground
65	GND	Ground	66	RSVD	Reserved
67	PETp(4)	Transmitter module 2 lane 0 differential pair	68	RSVD	Reserved
69	PETn(4)		70	GND	Ground
71	GND	Ground	72	PERp(4)	Receiver module2 lane 0 differential pair
73	GND	Ground	74	PERn(4)	
75	PETp(5)	Transmitter module 2 lane 1 differential pair	76	GND	Ground
77	PETn(5)		78	GND	Ground
79	GND	Ground	80	PERp(5)	Receiver module2 lane 1 differential pair
81	GND	Ground	82	PERn(5)	
83	PETp(6)	Transmitter module 2 lane 2 differential pair	84	GND	Ground
85	PETn(6)		86	GND	Ground
87	GND	Ground	88	PERp(6)	Receiver module2 lane 2 differential pair
89	GND	Ground	90	PERn(6)	
91	PETp(7)	Transmitter module 2 lane 3 differential pair	92	GND	Ground
93	PETn(7)		94	GND	Ground
95	GND	Ground	96	PERp(7)	Receiver module2 lane 3 differential pair
97	PRSNT4#	Hot plug detect	98	PERn(7)	
99	GND	Ground	100	GND	Ground

Signals will satisfy the electrical requirements of the PCIe M.2 Specification and the PCIe Card Electromechanical Specification. Note that table includes columns to indicate whether a signal is required for use by the M.2 interposer module and/or the PCIe riser. Only slot 4 is required to support both the M.2 Interposer and PCIe Riser.

Table 15 provides a brief functional description of each signal.

Table 15: PCIe M.2 carrier signal definitions

Bus type	I/O	Voltage	M.2 interposer	PCIe riser	Definition for three pair, eight column
P12V	O	12V		X	12V Input
P3V3	O	3.3V	X	X	3.3V Input
WAKE#	I	3.3V	X	X	PCIe Wake for M.2 modules 1 and 2
PERST#	O	3.3V	X	X	PCIe Reset for M.2 modules 1 and 2
PETP/N(3:0)	O		X	X	PCIe Transmit to M.2 module 1
PERP/N[3:0]	O		X	X	PCIe Receiver from M.2 module 1
REFCLK1LP/N	O		X	X	Reference Clock for module 1
PETP/N[7:4]	O		X	X	PCIe Transmit to M.2 module 2
PERP/N[7:4]	O		X	X	PCIe Receive from M.2 module 2
REFCLK2P/N	O		X		Reference clock for module 2
SUSCLK	I		X		Suspend Clock for modules 1 and 2 (32.768Khz)
CLKREQ1	I		X		Reference Clock Request for module 1
CLKREQ2	I		X		Reference Clock Request for module 2
DEVSLP	O		X		Device Sleep to M.2 modules 1 and 2 (should be pulled low on MB)
RSVD	N/A		X	X	Reserved (do not connect)
PRSNT[3:1]	I/O		X	X	Present Should be connected on MB per PCIe specification
GND	I/O		X	X	Ground

9.3 NIC Mezzanine Connector

The blade will support a single NIC mezzanine card. The mezzanine will interface to the blade through a Samtec SEAF-20-06.5-L-08-2-A-K-TR connector or equivalent.

Table 16 lists the compatible connectors.

Table 16: NIC mezzanine connector part numbers

Manufacturer	Card connector MPN	Motherboard connector MPN
Samtec	SEAM-20-03.5-L-08-2-A-K-TR	SEAF-20-06.5-L-08-2-A-K-TR
Molex	45970-2385	45971-2385

Table 17 shows the pinout for the NIC mezzanine connector.

Table 17: NIC mezzanine connector pinout

8	CLK_100M_NIC_PE_DP	GND	SMB_ALERT_N	NIC_MEZZ_ID0	PCIE_WAKE_N	GND	PCIE_RESET_N	GND	1
16	CLK_100M_NIC_PE_DN	GND	NWK2_PRESENT_N	SMB_SCL	GND	P3E_CPU1_LAN_RX_DP<7>	GND	P3E_CPU1_LAN_TX_DP<0>	9
24	GND	MEZZ_PRESENT_N	GND	SMB_SDA	GND	P3E_CPU1_LAN_RX_DN<7>	GND	P3E_CPU1_LAN_TX_DN<0>	17
32	NWK_2_TX0P	GND	NWK_1_TX1P	GND	P3E_CPU1_LAN_RX_DP<6>	GND	P3E_CPU1_LAN_TX_DP<1>	GND	25
40	NWK_2_TX0N	GND	NWK_1_TX1N	GND	P3E_CPU1_LAN_RX_DN<6>	GND	P3E_CPU1_LAN_TX_DN<1>	GND	33
48	GND	NWK_2_TX1P	GND	NWK_1_TX2P	GND	P3E_CPU1_LAN_RX_DP<5>	GND	P3E_CPU1_LAN_TX_DP<2>	41
56	GND	NWK_2_TX1N	GND	NWK_1_TX2N	GND	P3E_CPU1_LAN_RX_DN<5>	GND	P3E_CPU1_LAN_TX_DN<2>	49
64	NWK_2_TX2P	GND	NWK_1_TX3P	GND	P3E_CPU1_LAN_RX_DP<4>	GND	P3E_CPU1_LAN_TX_DP<3>	GND	57
72	NWK_2_TX2N	GND	NWK_1_TX3N	GND	P3E_CPU1_LAN_RX_DN<4>	GND	P3E_CPU1_LAN_TX_DN<3>	GND	65
80	GND	NWK_2_TX3P	GND	NWK_1_RX1N	GND	P3E_CPU1_LAN_RX_DP<0>	GND	P3E_CPU1_LAN_TX_DP<4>	73
88	GND	NWK_2_TX3N	GND	NWK_1_RX1P	GND	P3E_CPU1_LAN_RX_DN<0>	GND	P3E_CPU1_LAN_TX_DN<4>	81
96	NWK_2_RX0P	GND	NWK_1_RX2N	GND	P3E_CPU1_LAN_RX_DP<1>	GND	P3E_CPU1_LAN_TX_DP<5>	GND	89
104	NWK_2_RX0N	GND	NWK_1_RX2P	GND	P3E_CPU1_LAN_RX_DN<1>	GND	P3E_CPU1_LAN_TX_DN<5>	GND	97
112	GND	NWK_2_RX1N	GND	NWK_1_RX3N	GND	P3E_CPU1_LAN_RX_DP<2>	GND	P3E_CPU1_LAN_TX_DP<6>	105
120	GND	NWK_2_RX1P	GND	NWK_1_RX3P	GND	P3E_CPU1_LAN_RX_DN<2>	GND	P3E_CPU1_LAN_TX_DN<6>	113
128	NWK_2_RX2N	GND	NWK_1_TX0P	GND	P3E_CPU1_LAN_RX_DP<3>	GND	P3E_CPU1_LAN_TX_DP<7>	GND	121
136	NWK_2_RX2P	GND	NWK_1_TX0N	GND	P3E_CPU1_LAN_RX_DN<3>	GND	P3E_CPU1_LAN_TX_DN<7>	NWK2_I2C_SDA	129
144	GND	NWK_2_RX3N	GND	NWK_1_RX0P	GND	NIC_MEZZ_ID1	GND	NWK2_I2C_SCL	137
152	NWK1_I2C_SDA	NWK_2_RX3P	GND	NWK_1_RX0N	GND	P3V3	P3V3_AUX	P12V_AUX	145
160	NWK1_I2C_SCL	GND	NWK1_PRESENT_N	GND	P3V3	P3V3	P3V3_AUX	P12V_AUX	153

Table 18 describes the signals used in this interface.

Table 18: NIC mezzanine signal descriptions

Bus type	I/O	Logic	Definition
P3E_CPU1_LAN_RX_DP/N[7:0]	I	CML	PCIe Gen3 from the NIC mezzanine to the CPU
P3E_CPU1_LAN_TX_DP/N[7:0]	O	CML	PCIe Gen3 from the CPU to the NIC mezzanine
CLK_100M_NIC_PE_DP/N	O	CML	100MHz PCIe clock
PCIE_RESET_N	O	3.3V	PCIe reset
MEZZ_PRESENT_N	I	3.3V	Mezzanine present Should be GND on mezzanine
NWK_1_TX[3:0]P/N	I	CML	Port 1 10GbE transmit from mezzanine to tray backplane

Bus type	I/O	Logic	Definition
NWK_1_RX[3:0]P/N	O	CML	Port 1 10GbE receive from motherboard to tray backplane
NWK_2_TX[3:0]P/N	I	CML	Port 2 40GbE transmit from mezzanine to tray backplane
NWK_2_RX[3:0]P/N	O	CML	Port 2 40GbE receive from tray backplane to mezzanine
SMB_ALERT_N	I	3.3V	I2C Alert from NIC mezzanine to BMC
NWK1_PRESENT_N	O	3.3V	Port 1 cable present indicator from tray backplane
NWK1_I2C_SDA	I/O	3.3V	Port1 I2C data to SFP+ cable
NWK1_I2C_SCL	I	3.3V	Port 1 I2C clock to SFP+ cable
NWK2_PRESENT_N	O	3.3V	Port 2 cable present Indicator from tray backplane
NWK2_I2C_SDA	I/O	3.3V	Port 2 I2C data to QSFP+ cable
NWK2_I2C_SCL	I	3.3V	Port 2 I2C clock to QSFP+ cable
PCIE_WAKE_N	I	3.3V	PCIe wake
SMB_SCL	O	3.3V	I2C to BMC
SMB_SDA	I/O	3.3V	I2C to BMC
NIC_MEZZ_ID[1:0]	I	3.3V	NIC mezzanine ID Should connect to BMC
P3V3	O	3.3V	3.3V input power
P3V3_AUX	O	3.3V	3.3V auxiliary input power
P12V_AUX	O	12V	12V input power
Ground	--	--	Ground pins

9.4 SATA Cable Ports

The board will include support for one x4 Mini-SAS HD connector, supporting the SFF-8643 (or equivalent) cable and four x1 SATA connectors for cabling SATA ports from the PCH to the HDDs and SSDs.

9.5 SATA Power Connector

The board will include four 4-pin Mini-Fit® Jr 5566 series power connectors, P/N: Molex 5566-04A (or equivalent). Two connectors are intended to provide power to the Large Form Factor (LFF) HDDs and SSDs. Each connector is intended to provide cabled power to two drives. Each connector pin has a maximum 13A current capacity.

The 12V and 5V power rails are supplied directly from the main 12V and 5V power supply rails. These rails need to drive at most six 3.5" LFF HDDs or Small Form Factor (SFF) SSDs. Table 19 lists the signal names and current capacities.

Table 19: SATA power connector signal names and current capacities

Pin	Signal name	Capacity (in A)
Pin 1 and 2	GND	26A (Black)
Pin 3	12V	13A (Yellow)
Pin 4	5V	13A (Red)

9.6 SATA HDD Connector

The board will include two connectors for connecting up to two SATA HDDs to the motherboard. The connectors will be FCI P/N 10029364001LF (or equivalent). Table 20 shows the industry standard pinout. The remaining two HDDs should be mounted to fixed connectors on the blade pan.

Table 20: SATA HDD connector pinout

Data		
Pin	**Signal Name**	**Signal Description**
1	GND	Ground
2	A+	Transmit +
3	A-	Transmit -
4	GND	Ground
5	B-	Receive -
6	B+	Receive +
7	GND	Ground
Power		
Pin	**Signal Name**	**Signal Description**
1	V33	3.3v Power
2	V33	3.3v Power
3	V33	3.3v Power, Pre-charge, 2nd mate
4	Ground	1st Mate
5	Ground	2nd Mate
6	Ground	3rd Mate
7	V5	5v Power, pre-charge, 2nd mate
8	V5	5v Power
9	V5	5v Power
10	Ground	2nd Mate
11	Optional GND	-
12	Ground	1st Mate
13	V12	12v Power, Pre-charge, 2nd mate
14	V12	12v Power
15	V12	12v Power

9.7 LED Cable Connector

The design will provide connector support for an LED cable for driving the two LEDs on the front panel.

10 Management Subsystem

Management circuitry for the blade uses the Intel® Manageability Engine (ME) combined with the Baseboard Management Controller (BMC). This section describes the requirements for management of the blade. Primary features include:

- Intel® Manageability Engine (ME)
- BMC chip (ASPEED AST1050)
- BMC serial link (universal asynchronous receiver/transmitter [UART]) for communication with Chassis Manager

- Low Pin Count (LPC) connection to the chipset to support in-band management
- Platform Environmental Control Interface (PECI) for CPU out-of-band environmental control
- FRUID Electrically Erasable Programmable Read-Only Memory (EEPROM) for storage of manufacturing data and event
- Thermal sensors for inlet and exhaust temperature monitoring
- Power monitoring through the 12V hot swap controller circuitry
- Service LEDs

Figure 15 shows the management block diagram.

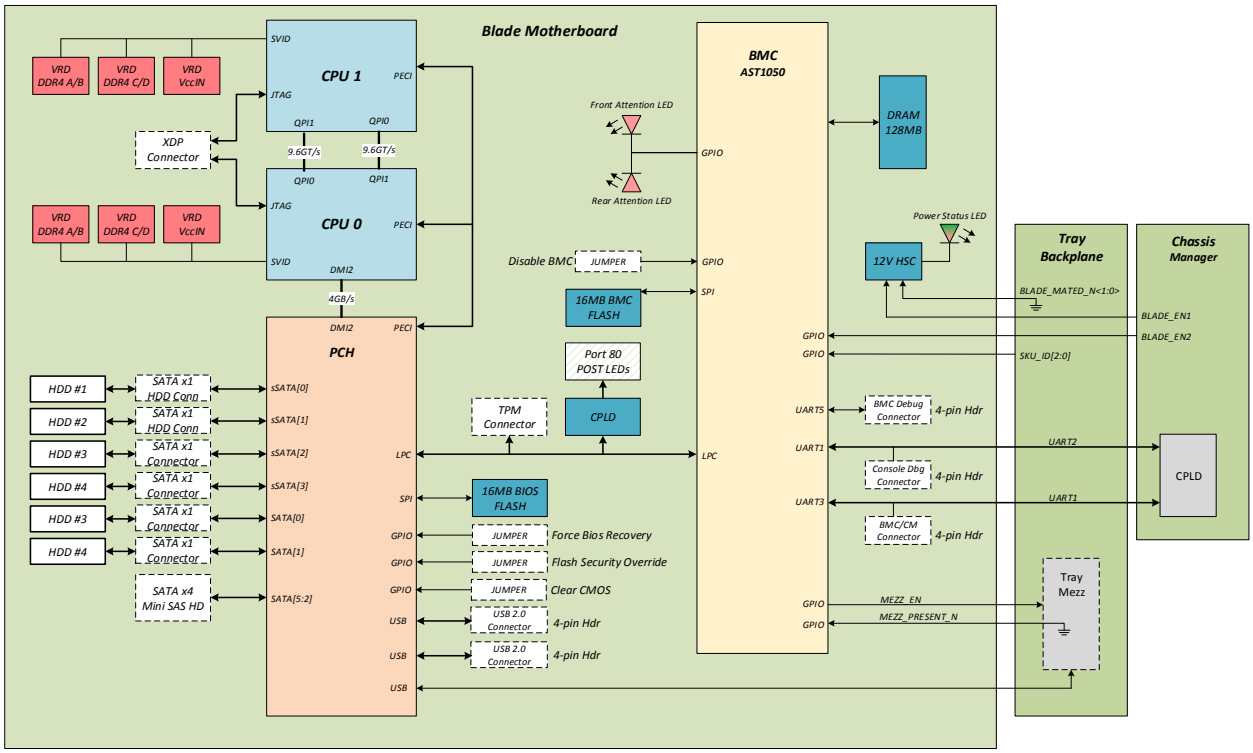

Figure 15: Blade management block diagram

10.1 Baseboard Management Controller

The design for the BMC is based on the ASPEED AST1050 processor. Primary features of interest for the AST1050 include:

- Embedded ARM926EJ
- Embedded 16KB/16KB cache
- Synchronous Dynamic Random Access Memory (SDRAM) memory up to 512MB
- NOR/NAND/Serial Peripheral Interface (SPI) flash memory
- I2C System Management Bus (SMBus) controller
- 5 UART 16550 controllers
- LPC bus interface

- Up to 152 GPIO pins

10.2 DRAM

The BMC requires 128 MB of Double Data Rate 3 (DDR3) memory.

10.3 BMC Boot Flash

The BMC boots from a flash memory device located on the SPI bus. The device size will be 128Mb (16MB) minimum. Winbond W25Q128 serial flash memory (or equivalent) is recommended. The design will support a socket for this device during pre-production.

10.4 BIOS Flash

The blade will support a 128Mb (16MB) flash BIOS located on the PCH SPI bus: Winbond MPN W25Q128BVGFIG (or equivalent). It is required to operate at 50 MHz and include Quad I/O Fast Read Support. The design will support a socket for this device during pre-production. The design shall also include a BIOS recover jumper connected to a PCH GPIO.

10.5 Serial Links

The blade is managed from Chassis Manager through a serial port connection to the BMC. There are two dedicated 3.3V serial links (UARTs) routed to the blade. The BMC supports a third UART for debug use only. The following is a brief description of the BMC UART port assignments and their uses:

- **UART3**
  Primary communication with CM for control of the blade.
- **UART1**
  Secondary communication with CM; primarily used for console debug.
- **UART5**
  BMC debug port; connects only to a debug connector on the blade.

10.6 PECI

The blade shall use the PECI host controller inside the CPU to support integrated thermal monitoring of the two CPU sockets.

10.7 TPM Module

The blade will include a connector to support a tamper-proof Trusted Platform Module (TPM) 2.0 solution.

10.8 PCH/BMC I2C

The design will include a number of I2C devices/functions available to the BMC and PCH. The block diagram is shown below in Figure 16. A brief description of the components follows. Some of the block components are specific with respect to requirements of the device and its slave address, while others (such as voltage regulators) are more general, requiring that the functionality be placed on a specific I2C bus but not requiring a specific solution (vendor part number) be used.

Care should be taken to electrically isolate components that are powered from separate power domains, but are located on the same I2C bus. Note that addresses shown are 8-bit address with the Read/Write (R/W) bit as the Least Significant Bit (LSB) set to 0 (0xA8=1010100x). Figure 16 shows a block diagram.

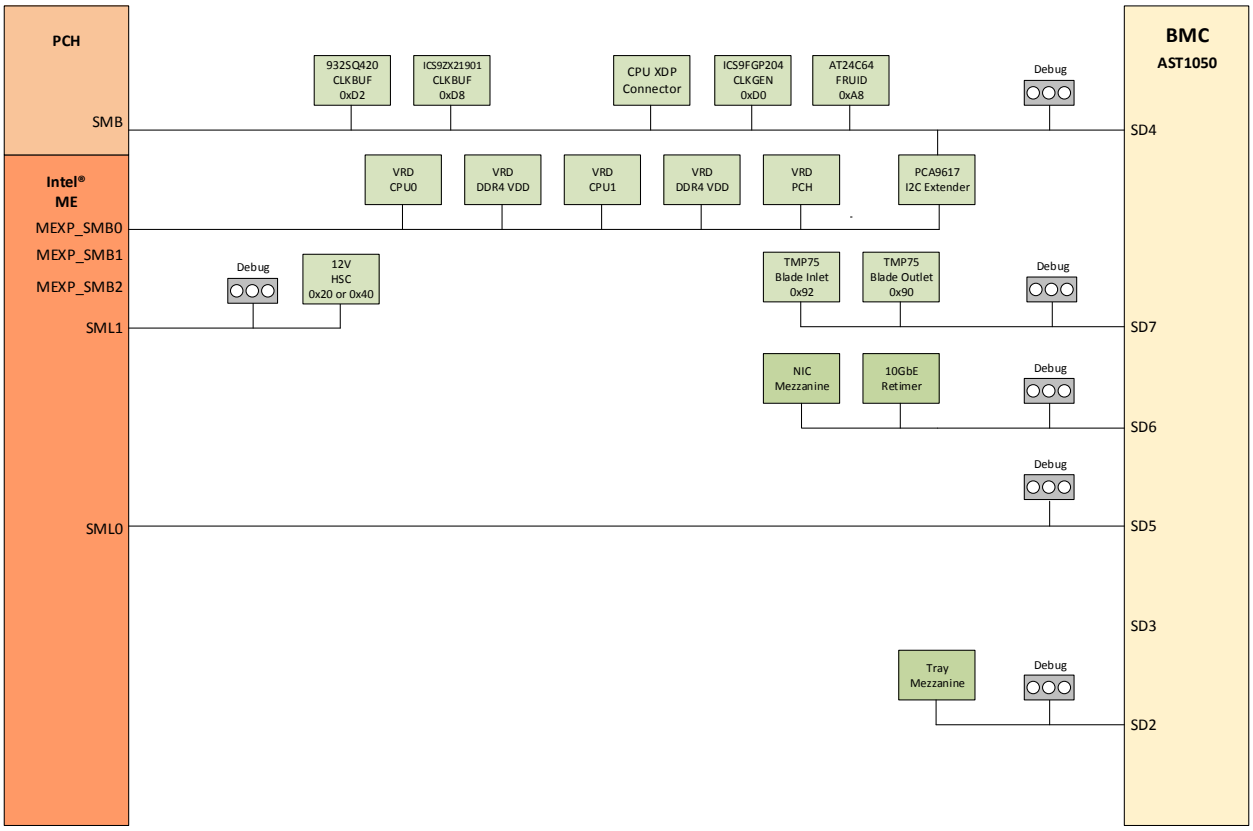

Figure 16: PCH / BMC I2C block diagram

10.8.1 Voltage Regulators

Voltage regulators that support I2C or PMBus should be available to PCH MEXP_SMB0, PCH SMB, and BMC I2C Port 3. This includes CPU and memory subsystem regulators, at a minimum. The Intel® ME is responsible for enabling power (through a programmable logic device) to any voltage regulators that are not on AUX power (i.e., enabled automatically by the presence of 12V).

10.8.2 Clock Generators and Buffers

Clock circuitry that supports I2C should be available to PCH MEXP_SMB0, PCH SMBus, and BMC I2C Port 3. The block diagram shows three clock devices that are common to this architecture (actual implementation may vary).

10.8.3 FRUID PROM

The blade will include a 64Kb serial EEPROM MPN AT24C64 (or equivalent) to store manufacturing data. The device should be available to PCH MEXP_SMB0, PCH SMB, and BMC I2C Port 3.

10.8.4 Temperature Sensors

The blade will include support for a minimum of two temperature sensors, MPN TMP75 (or equivalent), for monitoring the inlet and outlet temperatures of the blade. The sensors will be available on BMC I2C port 6. Figure 17 shows temperature sensor locations on the blade.

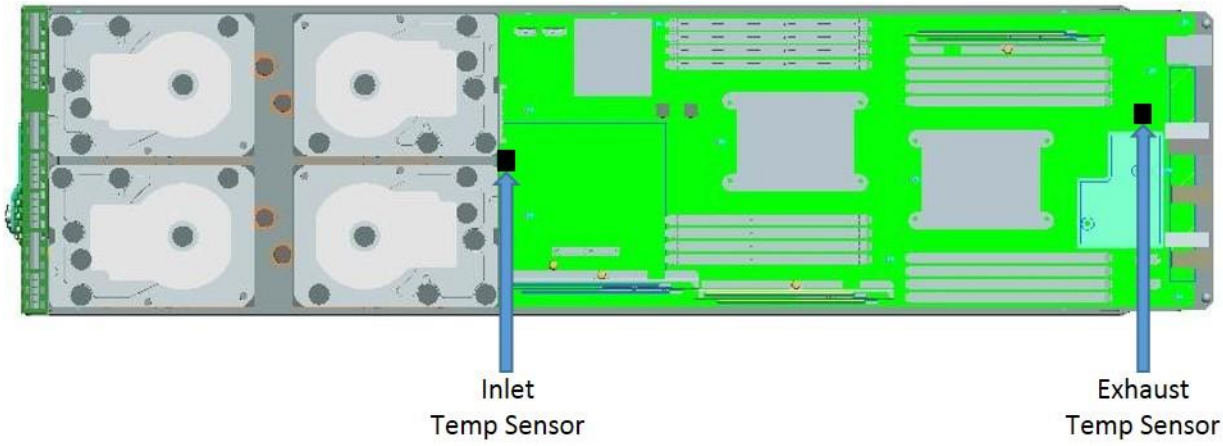

Figure 17: Temperature sensor locations

10.8.5 Hot Swap Controller

For in-rush current protection on the blade, an HSC that includes support for the PMBUS 1.2 interface will be used on the motherboard. The device will support an I2C polling rate of 10ms. The HSC will be available to the Intel® ME SML1 and will provide its ALERT# signal to a Multipath General-Purpose Input/Output (MGPIO) on the PCH. The Intel ME SML1 will be dedicated to the HSC to provide fast response time to power excursions.

Following is a list of devices known to be supported by the platform requirements (equivalent or newer devices may exist):

- Analog Devices ADM1276
- Texas Instruments LM25066i
- Texas Instruments LM25062
- Texas Instruments LM25063

Intel® Node Manager requires the ability to read blade power levels from the HSC to detect conditions where total blade power exceeds the desired level for power capping. Operation of the HSC is controlled by the BLADE_EN1 and BLADE_MATED signals from the tray backplane, as described in the sections that follow.

10.8.5.1 Blade Enable Signal

The BLADE_EN1 signal is sourced from the Chassis Management Card (CMC) through the tray backplane. The signal is active high, and is intended to be connected to the EN input of the HSC. The signal driver from the CMC is open-drain. In normal operation, it will be pulled high through a resistive termination when a 3.3V regulator on the blade is enabled.

To disable the power, the CMC pulls the Blade_EN1 signal to reference ground. The Blade_EN1 signal should not be pulled low on the blade or driven back to the Chassis Manager.

If BLADE_EN1 is used to disable power to the blade, the BMC will detect this event and disable logging to prevent spurious messages from propagating to the log files.

10.8.5.2 Blade Mated Signal

The active low Blade[1:2]_Mated_N signals are used to qualify Blade_EN to indicate that the connectors have properly mated during blade insertion before the 12V inrush power is turned on. These signals should be connected to the UV (undervoltage) input pin of the HSC. The mated signals are pulled to ground on the tray backplane through a 100 ohm (Ω) resistor. Both BLADE_MATED_N signals from the AirMax VS2 PCIe connector and the AirMax VS2 10/40GbE connector must be asserted before 12V inrush power is turned on.

To support potential blade or tray backplane SKUs in which connectors could be depopulated, the BLADE_MATED_N circuitry should allow for depopulation of components to meet the BLADE_MATED_N requirement from either connector.

Figure 18 shows the signal connections and contact sequence. The 3.3V pull-ups shown should be derived from the 12V raw input. Note that the block diagram is intended to be functional only and does not describe the actual circuitry required for the intended logic.

To support possible SKUs of the tray backplane in which connectors can be removed, the design should include a method for grounding the BLADE_EN1 pin from the PCIe AirMax connector in case this connector is depopulated. This could be accomplished through a BOM load or other methods.

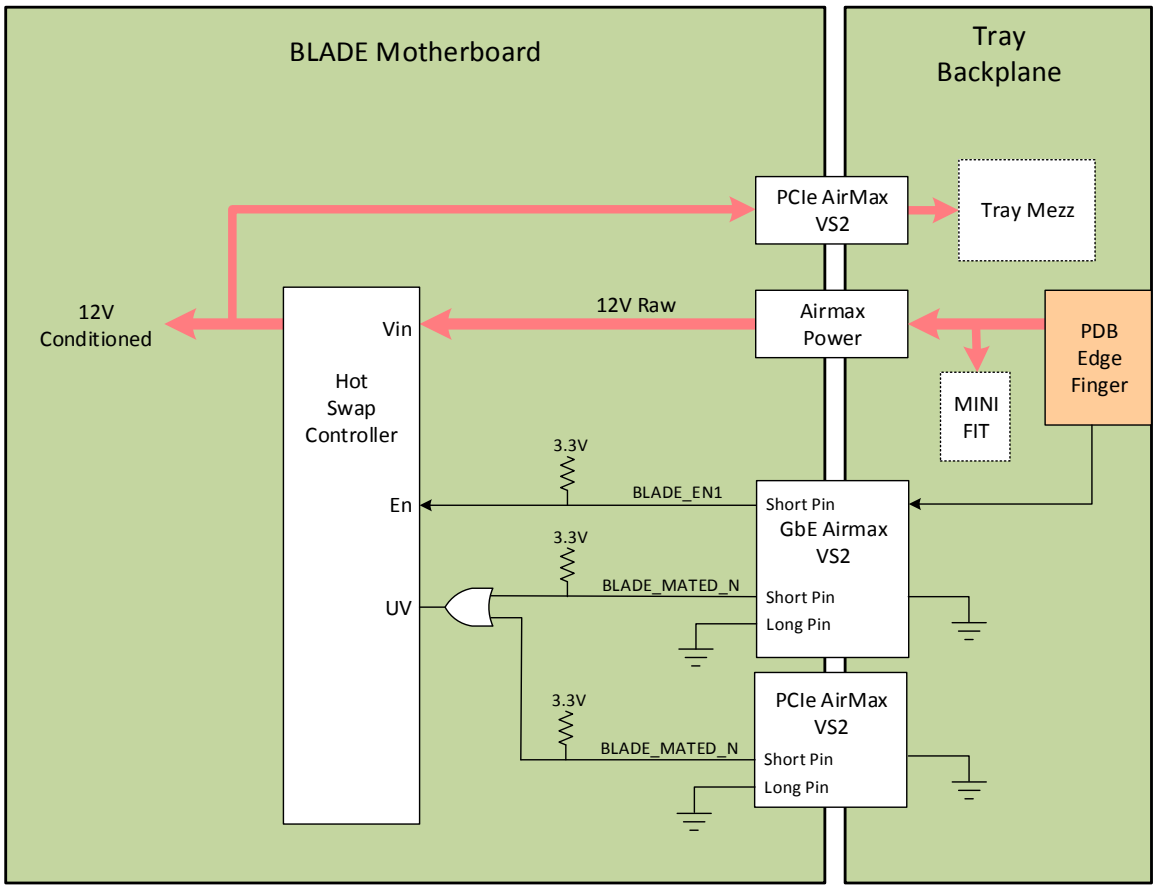

Figure 18: HSC functional block diagram

Following is the connector contact sequence:

1. Guide pin (GND) engages.
2. AirMax power connector (12V and GND) and AirMax VS2 long pin (GND) engage.
    - Blade_EN1 and both Blade_Mated_N are high on blade.
    - HSC is disabled.
3. AirMax VS2 short pins engage.

- Blade_EN, Blade_Mated_N and all other signal pins engage.
- Blade_EN1 is high and both Blade_Mated_N are low.
- HSC is enabled.

10.8.5.3 Signal Interpretation

Table 21 describes the HSC controller with regards to the Blade_EN1 and Blade_Mated_N signals.

Table 21: Signal interpretation

BLADE_EN1 on blade	BLADE_MATED_N on blade	Blade/hot swap controller status
1	1	• No 12V power, no signal connection to backplane • Blade was either not inserted or recently removed • Hot swap controller is DISABLED
1	0	• Blade is installed, 12V is present • Signal connector is mated • CM is either not present or is driving high • Hot swap controller is ENABLED
0	1	• Invalid or faulty state • The only way to drive BLADE_EN1 low is from the CM through the signal connector • If either BLADE_MATED_N is high, the connection was not made • Hot swap controller is DISABLED
0	0	• Blade is installed, 12V is present • Both Signal connector are mated • CM drives BLADE_EN1 low • Hot swap controller is DISABLED

10.8.6 PCIe I2C

The blade will support I2C for the tray mezzanine card and the NIC mezzanine card. It is expected that standard PCIe modules will support Address Resolution Protocol (ARP) per the I2C specification. This should be verified with the specification for specific mezzanine cards. The modules shall be accessible to the I2C ports as shown in the block diagram.

10.9 Intel® Node Manager Hardware Requirements

Blade management will meet the hardware requirements for support of Intel® Node Manager. The requirements include, but are not limited, to the following:

- The Intel ME shall be enabled (powered) in all states (i.e., must be power from AUX power).

- The following signals will be driven or received MGPIO signals of the PCH (additional SmaRT/CLST logic will be supported on the PCB):
    - 12V HSC ALERT#
    - SMBAlert#, SMBAlert_EN#
    - PROCHOT# and MEMHOT# for each CPU

10.10 Blade LEDs

The following sections describe the light-emitting diodes (LEDs) used as indicators on the blades and chassis.

Each blade has three LEDs: a front power/status LED that is green/amber, and two attention LEDs (front and back) that are red. Figure 19 shows the location of the front LEDs. The location of the rear LED is shown in Figure 20. The visible diameter and brightness requirements of the LEDs are TBD.

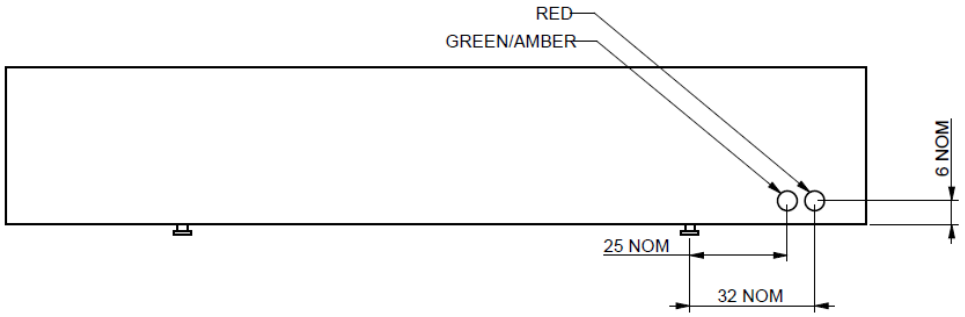

Figure 19: Front blade LED locations

The rear attention LED is located near the AirMax 3x6 SAS connector, as shown in Figure 20. The LED is placed so that it is visible externally through the blade and tray backplane connectors.

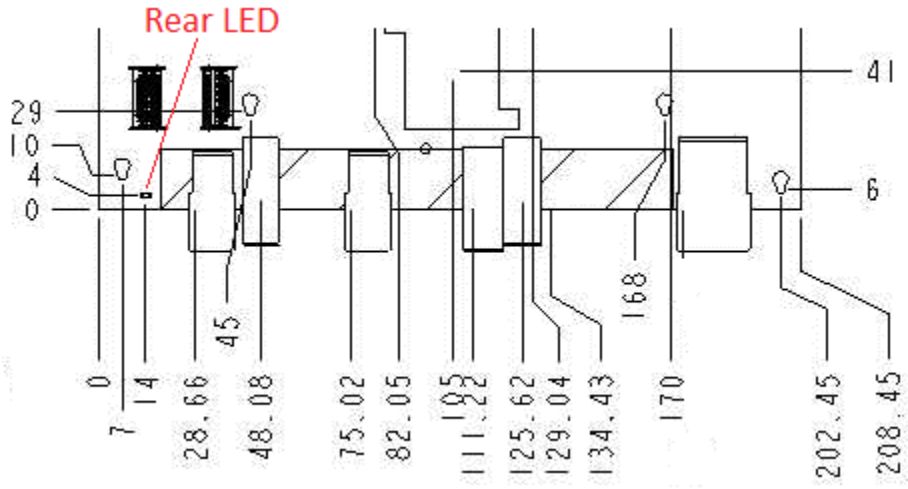

Figure 20: Rear attention LED location

10.10.1 Blade Power Status LED

When a blade is first inserted, the LED will turn amber if 12V is present at the output of the HSC. This ensures that the backplane has 12V power, the tray backplane connectors are mated, and the Blade_EN1 signal is asserted.

When the blade's management software turns on the system power (CPU/Memory/PCIe), the power status LED turns green. Note that the power status LED may be driven by an analog resistor network tied directly to a power rail, and is not an indication of the health of the blade. Table 22 describes the operation of the blade power status LED.

Table 22: Blade power status LED description

LED status	Condition
Off	Blade is not fully inserted, 12V power is absent, or Blade_EN1 is de-asserted.
Solid Amber ON	Blade is inserted, Blade_EN1 is asserted, 12V power output from the Hot Swap Controller is present.
Solid Green ON	Indicates that the BMC is booted and system power is enabled (CPU/Memory/PCIe).

10.10.2 Blade Attention LED

The blade attention LED directs the service technicians to the blade that requires repair. The technician can remove the blade from the rack and replace it with an operational blade.

The attention LED is driven by a single BMC GPIO. Table 23 describes the operation of the blade attention LED.

Table 23: Blade attention LED description

LED status	Condition
Off	No attention indicated.
Solid RED	System needs attention.

Note that a second red Attention LED is located at the rear of the blade so that it is visible through the tray when the fan door is opened. This LED is set at the same time as the front red attention LED, and indicates from the rear of the chassis which blade needs attention.

10.11 Debug Features

The following minimum debug features will be included on the motherboard during pre-production:

- I2C debug headers on the CPU SMBus, PCH SMBus, Intel ME SML I2C, and BMC I2C.
  Header will be a 3-pin compatible with standard I2C protocol analyzers (such as Beagle Protocol Analyzer or Aardvark Host Adapter).
  Note that the CPU SMBus can be eventually accessed through CPU XDP connector.
- Debug connector on all three BMC UARTS as shown in the block diagram (Figure 15).
  Connector TBD.
- Port 80 POST LED support on LPC bus as shown in the block diagram (Figure 15).
- BIOS debug support including:
  - Socketed BIOS Flash (to be removed for production)
  - BIOS recovery jumper connected to PCH GPIO
  - Flash security override driven from BMC GPIO (to PCH)
- Power button support.
- CPU XDP.
  Recommend support for ITP60c, or otherwise ensure that it will be mechanically possible to solder on the XDP connector post-production for debug.
- Dual USB 2.0 headers connected to PCH USB2 Port 1 and Port 9.
  Dongles will be readily available to get standard USB connectors (Port 1 on EHCI#1 and Port 9 on EHCI#2 [assuming 0-based numbering] of PCH).
- USB 2.0 debug port connected to PCH USB2 Port 0 for BIOS and OS debug.
  These can be implemented using 4-pin header connectors to save space.
- BMC disable jumper attached to GPIO on BMC.
- Intel ME recovery mode jumper.
- HW jumper to enable BIOS serial debug output.

Placement of the debug connectors will not obstruct the installation of any optional assemblies such as the PCIe RAID Card or SSDs. It is acceptable that a debug feature is not accessible if an optional assembly is installed unless the feature supports debug of that optional assembly.

10.12 Tray Mezzanine Power Control

The 12V power to the tray mezzanine is sourced directly from the output of the Hot Swap Controller, and is therefore always present when the HSC is enabled. MEZZ_EN is provided as a means to disable power to the tray mezzanine without removing power from the blade.

10.13 Connector Quality

The V2.0 Open CloudServer is designed for use in datacenters with a wide range of humidity (up to 90%). The connectors for these deployments must be capable of withstanding high humidity during shipping and installation. The baseline for plating DIMMs and PCIe connectors will be 30µ"-thick gold. DIMM connectors must also include lubricant/sealant applied by the connector manufacturer that can remain intact after soldering and other manufacturing processes. The sealant is required to displace any voids in the connector gold plating.

10.14 Overcurrent Protection

The hot swap controller (HSC) and associated circuitry is responsible for monitoring the 12V input current to the blade. There are two levels of overcurrent protection associated with this monitoring.

- **Over-current protection** is responsible for detecting a current level that indicates a catastrophic failure of the blade. In this event, the HSC should disable 12V to the blade typically by disabling the HSC's input FETS. The recommended current threshold for over-current protection is 50A.
- **Over-current monitor** is responsible for detecting a current level that which is higher than achievable limits by worst case applications. In this event, the blade typically throttles the CPU performance using the PROCHOT# input. It is expected that the current threshold be set high enough or with slow enough filter response to prevent throttling during performance or stress testing with worst case configuration. The recommended current threshold for over-current monitoring is 45A.

11 NIC Mezzanine

The V2.0 OCS motherboard will support a 40GbE or 10GbE NIC mezzanine card. The network controller on the NIC mezzanine will interface to CPU0 through a PCIe x8 channel. 10GbE is supported though Port 1 of the NIC to the blade connector interface, and 40GbE is supported through Port 2 of the NIC to the blade connector interface. The 10G port will connect to an SFP+ connector on the tray backplane. The 40G port will connect to a QSFP+ connector on the tray backplane. Figure 21 shows the NIC block diagram. The pinout for the NIC mezzanine is described in Section 9.3. Mechanical outlines can be found in the *Open CloudServer OCS NIC Mezzanine Specification Version 2.0*.

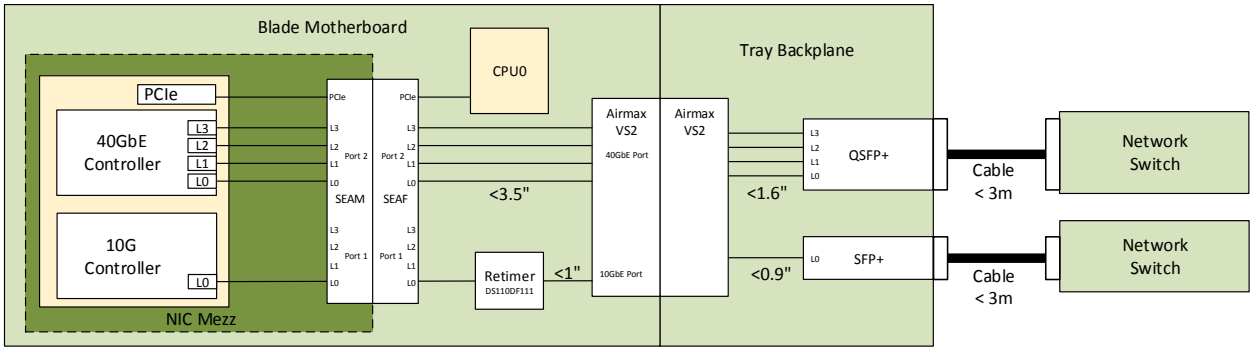

Figure 21: NIC block diagram

12 Blade Electrical Specifications

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

12.1 Input Voltage, Power, and Current

Table 24 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 24: Input voltage range

Nominal voltage	Maximum voltage	Minimum voltage
12.3V DC	14V DC	11V DC

The maximum amount of power allowed per blade is defined during system power allocation. The number of blades in a chassis might be limited by the capacity of the AC power cord or by the cooling capacity of the deployment. Table 25 lists the input power allocation for a low-power blade.

Table 25: Input power and current allocation

	Nominal voltage	System power allocation to blade assembly (in W)	Allocated current (in A)
Low-power blade (1x power connector)	12.3VDC	300W	30A

The blade 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 5 millisecond (ms) period.

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

The blade 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 will include power consumption measurements at the inrush controller. Accuracy of the measurement will be +/- 2%.

The blade also provides a way to interrupt current flow within 1 microsecond (µs) of exceeding the maximum current load.

12.3 Filtering and Capacitance

The blade provides sufficient filtering to operate with a 0.25A/ µs DC source.

Maximum capacitance allowed on the 12V input to the blade is 4,000µFarads (µF). Minimum capacitance allowed per blade is 300µF.

12.4 Grounding and Return

The blade chassis grounding/return is provided to the blade from the tray backplane through the tray backplane/blade interface connectors, including the 12V power connector, guide pins, and signal connectors. Chassis ground and logic ground are tied together on the tray backplane and should not be separated on the blade.

13 Blade Physical Specifications

The physical specifications for the blades include the dimensions of the space into which the blade is installed, blade weight limits, a description of the guiding and latching features, and information about electromagnetic interference shielding.

13.1 Blade Mechanical Control Outline

Figure 22 shows the outline for the blade mechanical controls; note that the locations shown for the connectors interfacing to the tray backplane are mandatory, but locations for other connectors, components, and mounting holes are suggestions and can be modified to meet the requirements of the PCB implementation.

- **Mounting holes**
  Figure 22 provides guidelines for mounting holes; actual mounting hole locations may vary to meet PCB implementation requirements.
- **NIC mezzanine keepout**
  Figure 22 does not include component height restrictions for the NIC mezzanine. For information concerning the component height restrictions, see *Open CloudServer NIC Mezzanine Specification Version 2.0*.
- **RAID/ROC card**
  Figure 22 shows the component height restrictions for supporting PCIe cards, such as the RAID/RAID on Chip (ROC). Note that the RAID card can include a local Supercap or battery backup solution; mounting for this should be accounted for by the mechanical design.

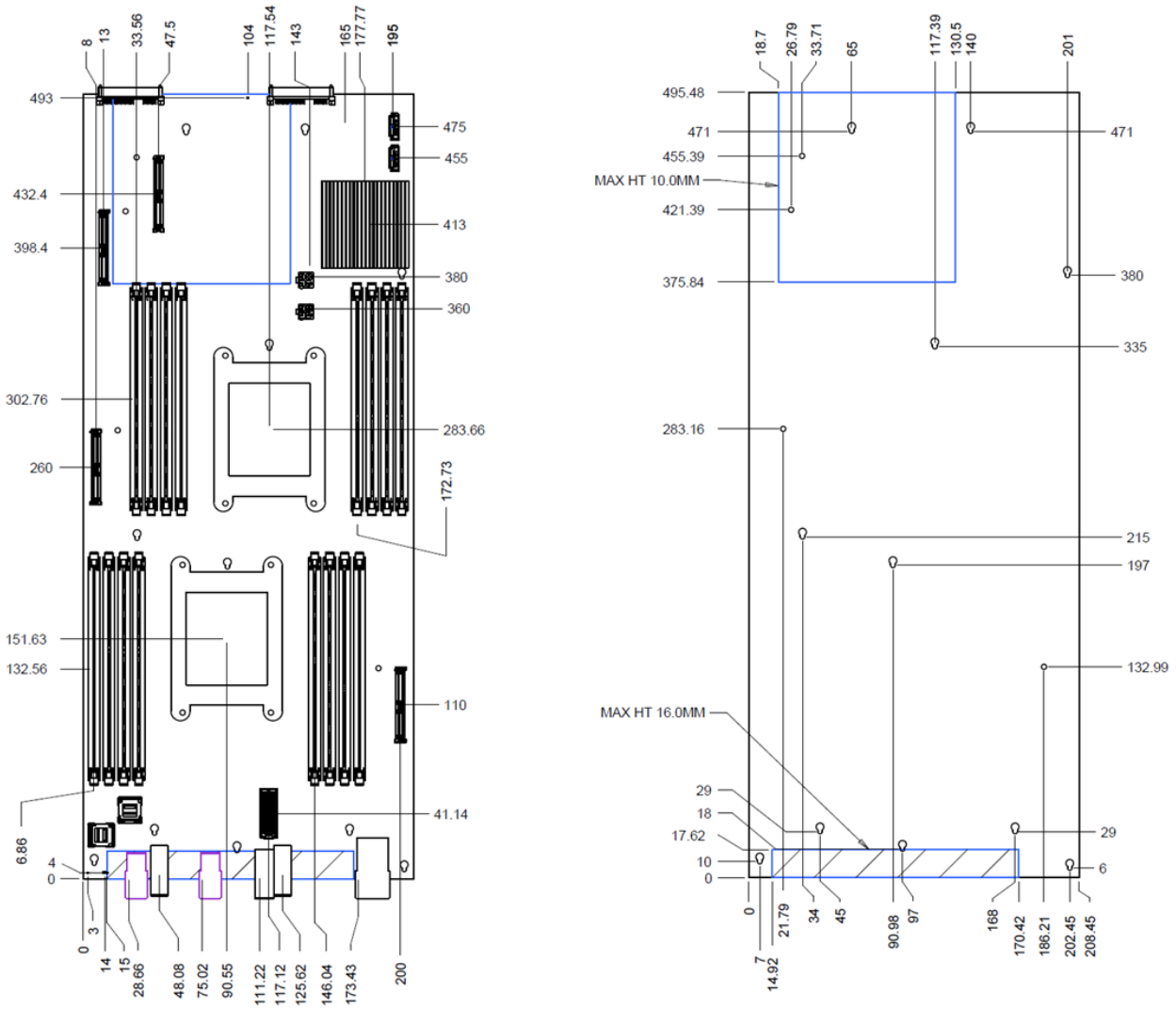

Figure 22: Mechanical control outline

13.2 Volumetric Specifications

The outer surfaces of the blade must fit into the volume defined in Figure 23, with all tolerances and manufacturing methods accounted for; no features of the blade (including the latch and the handle) can extend beyond the specified distance from the front column of the rack so that the front door of the rack can be closed without interference.

This specification assumes a coplanar power/network distribution PCB; if another structure is used, the connectors need to be form, fit, and functionally compatible with this specification. Note that if necessary, an electromagnetic interference (EMI) enclosure can be built to fit within the volumetric restrictions.

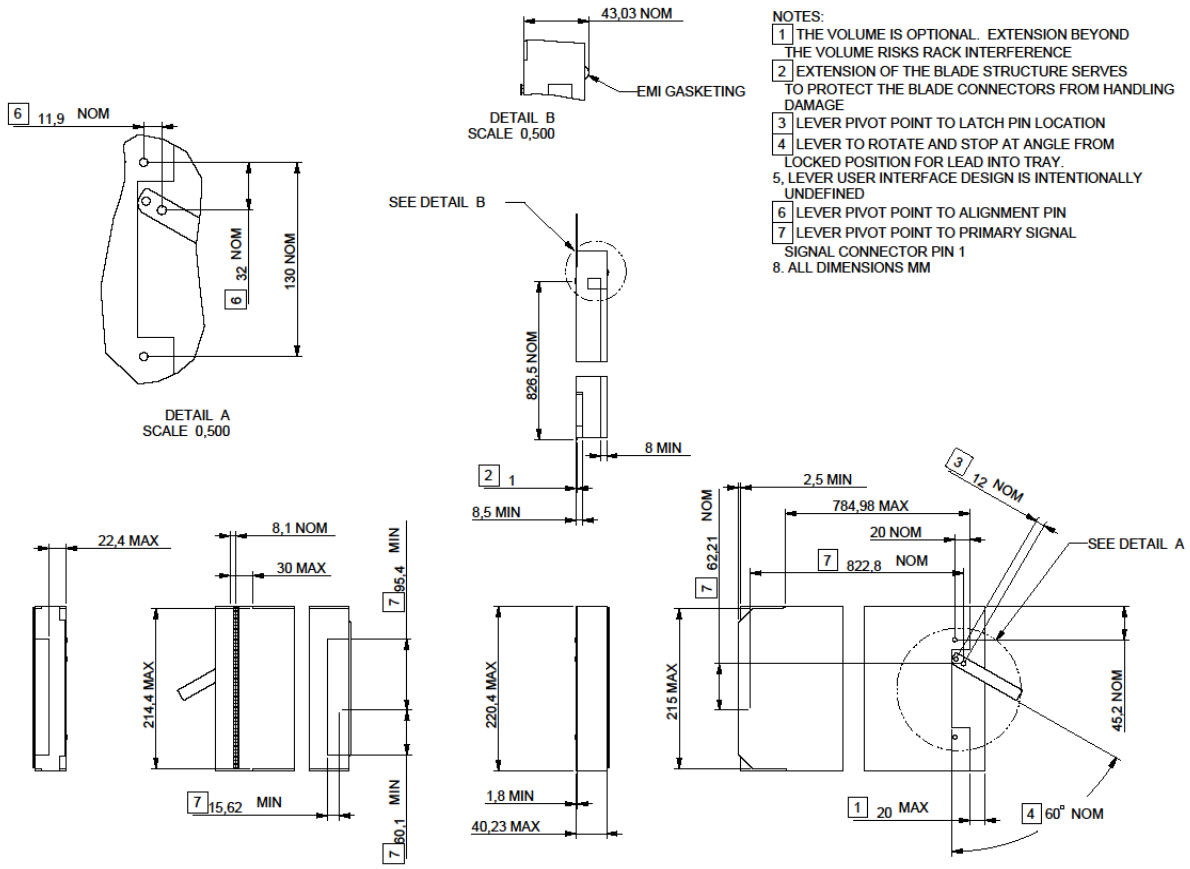

Figure 23: Dimensions of the volume that holds the blade

Each rack unit (or level) within the chassis enclosure can accommodate either two half-width blades inserted next to each other or a single full-width blade. Figure 24 shows two blades on a tray.

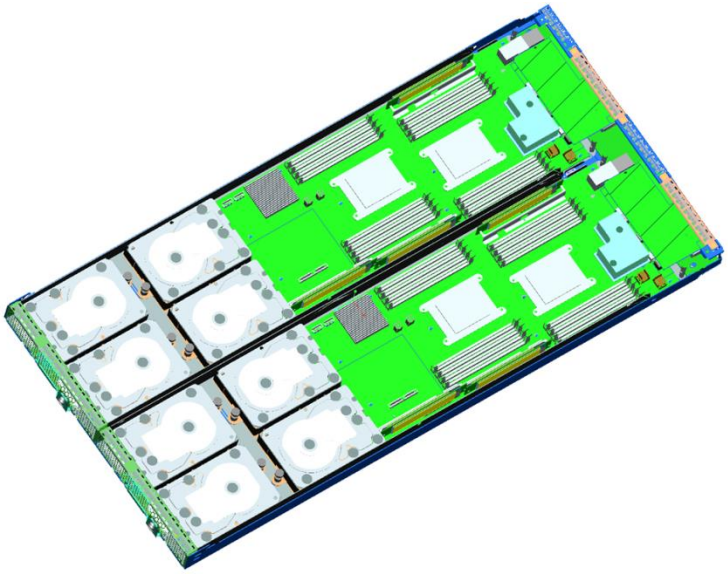

Figure 24: Two blades on a single tray

Figure 25 shows the dimensions of the blade-mounting envelope, which holds the blade on the tray.

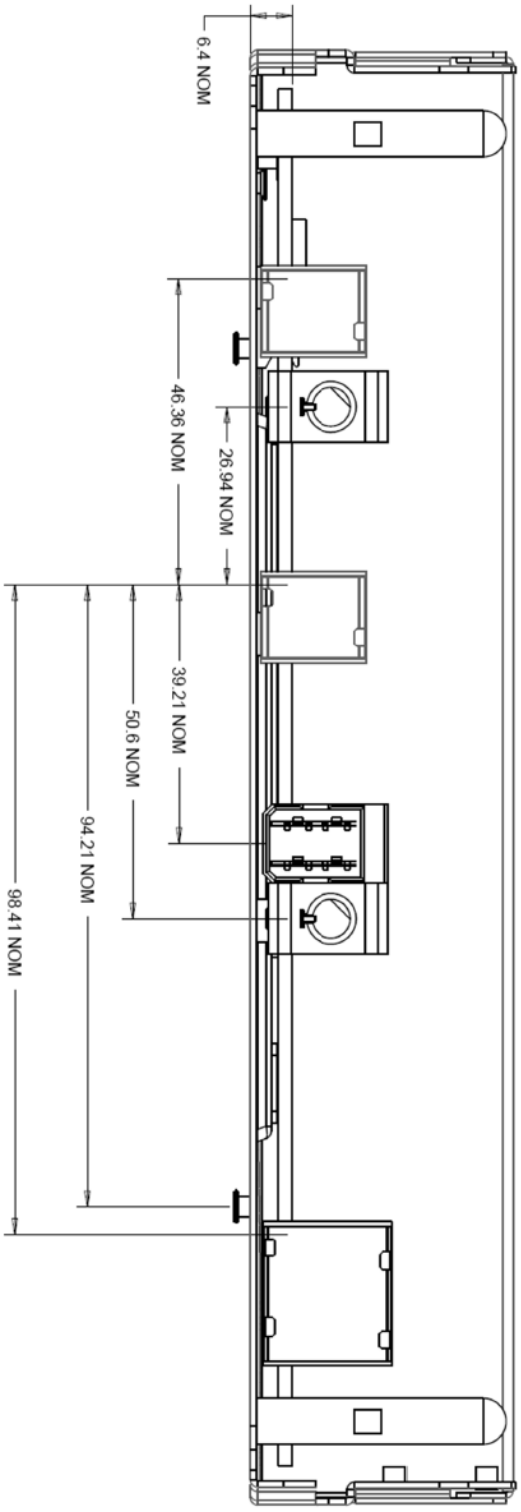

Figure 25: Blade-mounting envelope, rear view

Note that the only serviceable components in the blade are the hard drives, which can be replaced without tools. All other repairs to the blade are made in a bench-top environment. Components

such as heat sinks, hard drives, and mother boards can be attached so that removal requires a tool, if this option is less expensive.

13.3 Weight and Structure

Because the tray can support a total weight of 50 lbs, a standard blade should not exceed a static weight of 25 lbs when loaded with all options.

The blade must be mechanically stiff enough to support its own weight during shipping and handling. The side walls of the blade should be made as tall as possible within the given envelope to maximize the stiffness of the blade structure.

The blade shall contain a dimple on the side of the front surface to force the blades to fit tightly in the tray and prevent excessive bowing in the trays during a shock. Figure 26 shows an example.

Figure 26: Front surface dimple

13.4 Guiding and Latching Features

The sections that follow describe guiding and latching features for the blade.

13.4.1 Latching Feature Identification

Latches, as well as thumb screws and other components used to lock, unlock, or remove a subassembly from the chassis, are colored blue (Pantone code 285 C blue) to make them easy to identify. The locking feature used to secure the blade latch to the front of the blade must provide locking with minimal hand motion. A screw type lock with more than a half-turn rotation is not permitted.

If the identity and function of a latching feature for a particular field replaceable unit (FRU) is clear, coloring might not be required. For example, a lock/unlock indicator is sufficient for a blade lever.

13.4.2 Guiding and Latching

Because a single tray design without a center rail is used for both half-width and full-width blades, guides and alignment features are located in both the front and back of the enclosure. These features are especially important when inserting a half-width blade into an empty tray.

The blade enclosure interfaces with the side walls of the tray and the guide pin in the tray to make sure the blade is aligned within the ±3.50mm (±0.138”) tolerance necessary to engage the connector. The guide pin in the tray aligns the half-width blades with the correct side of the tray. The blade connectors are protected from damage by the guide pin by the blade sheet metal.

Figure 27 shows the blade guiding and latching features associated with the tray. Blade guide pins slide into slots on the front of the tray for alignment. A latch attached to the blade fits into a notch in the tray to secure the blade in place.

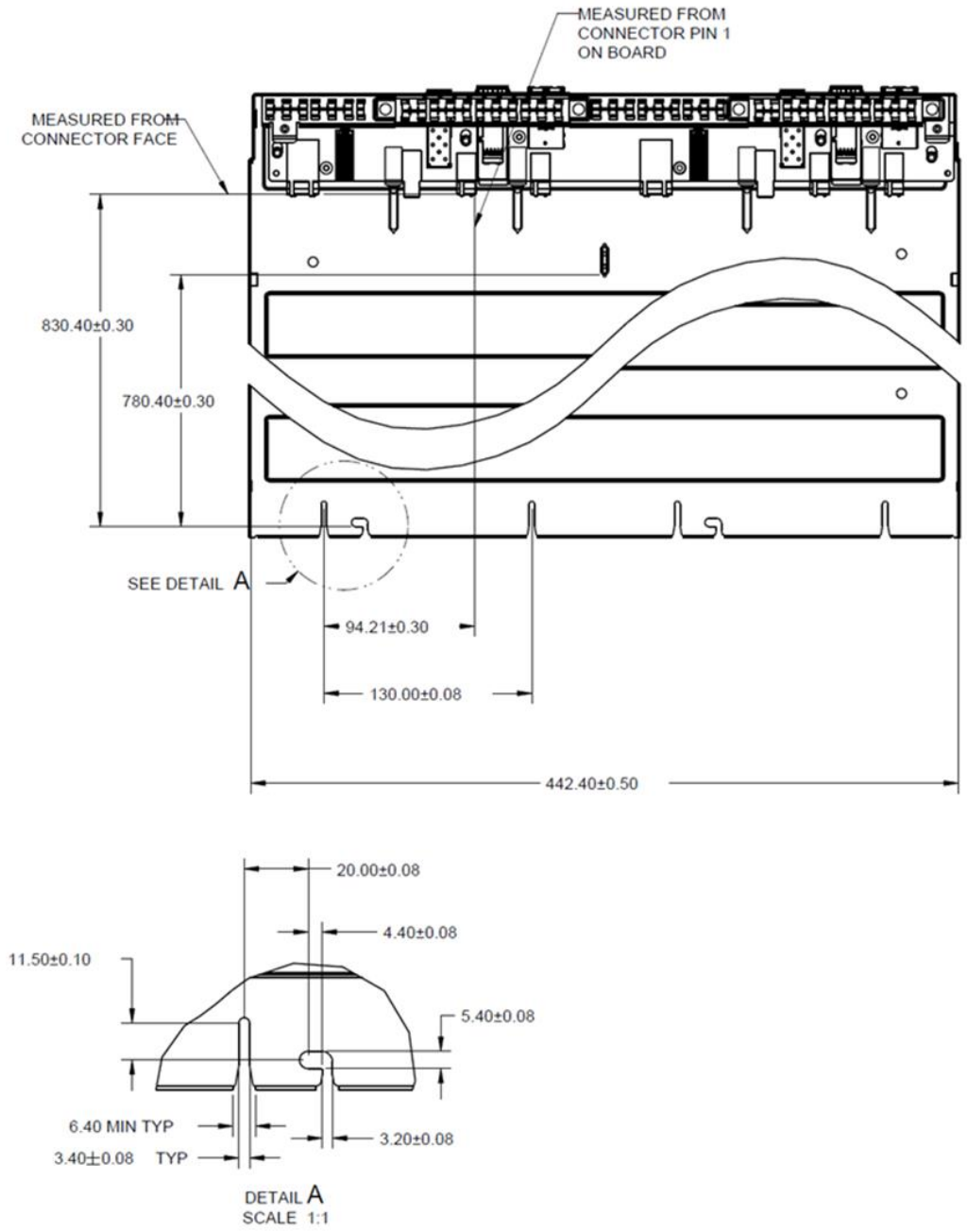

Figure 27: Guide and latch details, top view

Figure 28 and Figure 29 show examples of a front-blade guide and latch and of a rear-blade guide.

Note that a lever is included at the front of the blade to provide additional guidance when the blade is locking into the tray, and to help with the force required to install and remove the blade from the tray.

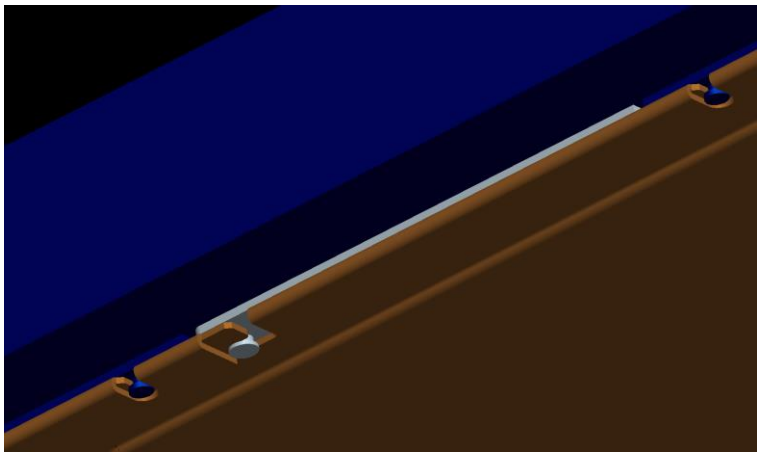

Figure 28. Example of a front-blade guide and latch

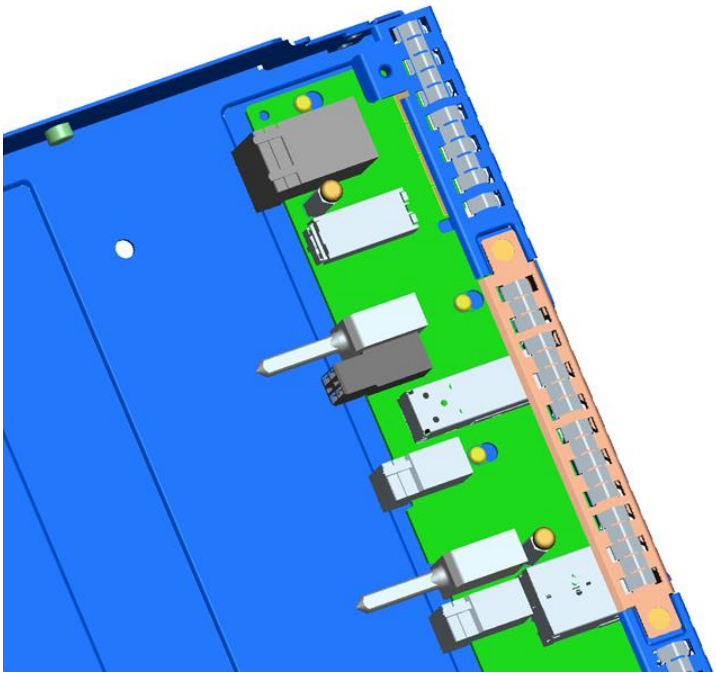

Figure 29: Example of a rear-blade guide pin

13.4.3 Electromagnetic Interference Mitigation

For EMI containment, an EMI shield is added to the top rear edge of the tray, as shown in Figure 30. EMI containment will be executed at the blade-assembly level, and EMI certification will be executed at the chassis level.

Figure 30: Tray with EMI enclosure, blade volume shown

When all trays are in place, they are electrically stitched together to prevent leakage of electromagnetic fields. The blade also has a gasket on the front of the top surface that provides electrical sealing, as shown in Figure 31. Dimensions for this gasket are shown in Figure 23.

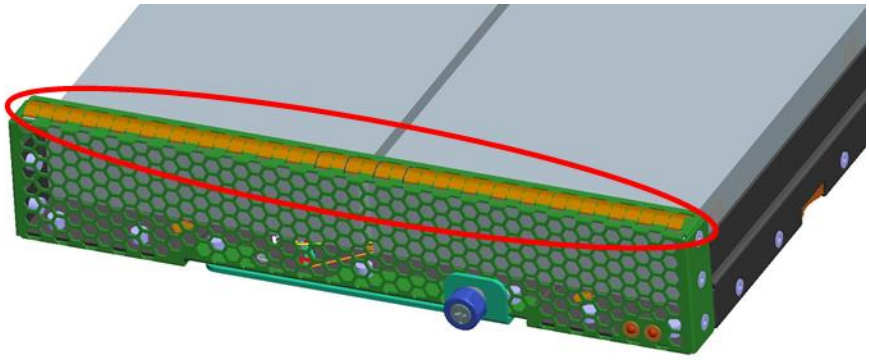

Figure 31: Blade EMI seal

14 Blade Thermal Design

To achieve high reliability required for use at scale, the blade must be designed to meet or exceed the thermal specifications of all components. Since the fans used to cool the blades are shared with other blades within the chassis, the blades are to be designed to operate according to the impedance and airflow curves provided in the Cooling System Specifications chapter of the Open CloudServer OCS Chassis Specification Version 2.0 document. Fan operation and communication with chassis management software to set the fan speed are fully described in that chapter.