OPEN

Compute Project

Facebook 2S Server

Tioga Pass

Rev 1.0

Author:

Whitney Zhao, Hardware Engineer, Facebook

Jia Ning, Hardware Engineer, Facebook

1 Revision History

Table 1

Date	Name	Description
6/3/2015	**Jia Ning**	**Version 0.1**
1/31/2017	Whitney Zhao	Version 0.30 Release version 0.30 for OCP summit 2017.
6/23/2017	Whitney Zhao	V0.6 - add 9.24 Intel® At Scale Debug
6/27/2017	Whitney Zhao	V0.7 - Typo updates - 18.2 update MLCC predicted life Inlet to 35C - 20.6 update manufacturing yield target
7/10/2017	Whitney Zhao	V0.8 - Typo updates - 15.2 update Over current trip point to 53.9A
9/15/2017	Whitney Zhao	V1.0 - Update block diagram, layout placement to latest version - Update BMC verified boot reference diagram to latest version
01/03/2018	Whitney Zhao	V1.0 - Update Fan numbering sequence - Update session 9.24 - 01/10/2018: Update Intel® code name based on feedback - 1/18/2018: update license section on page3

2 Scope

This specification describes Facebook dual sockets server Intel® Motherboard v4.0 (Project name: Tioga Pass) design and design requirement to integrate Tioga Pass into Open Rack V2[1].

3 Contents

1 http://files.opencompute.org/oc/public.php?service=files&t=348f3df2cc4ce573397fcc4424f68ca6&download

4 Overview

4.1 Overview

Tioga Pass (also referred to "motherboard" in this document, unless noted otherwise) is a derivative design from Crescent City Cloud Reference board based on Intel® Xeon® Scalable processor family (aka Skylake-SP) CPU architecture. The motherboard supports up to 24 DIMMs with double sided board SKU and up to 12 DIMMs with single sided sled design in ORv2.

Open Rack V2 is OCP Rack design in 2013. Tioga Pass-ORv2 sled is the compute sled compatible with Open rack V2.

4.2 Open Rack Introduction

This chapter gives background of Open Rack V2. The details of the motherboard's electrical and mechanical interfaces to Open Rack V2 are described in Chapter 14 and Chapter 15.

4.2.1 Open Rack V2 Introduction

Open Rack V2 has 2x power zones. Each power zone has 16x OU for IT equipment (server, storage, etc.), and 3x OU for power shelf. Each ORv2 power shelf has 2+1x 3.3KW PSUs, and 3x Battery Backup Units (BBU) and provides 6.3KW[3] continuous max loading through 1x bus bar to the power zone it is attached to.

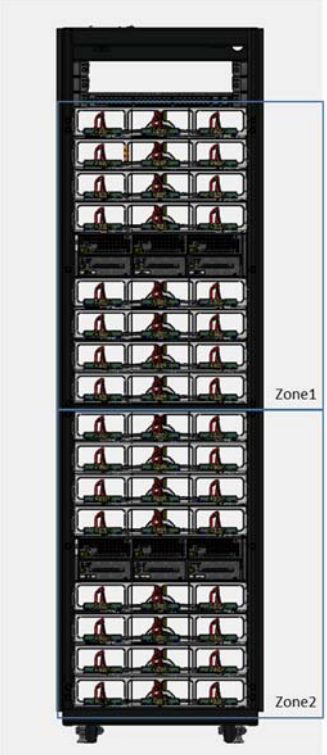


Figure 4-1 Open Rack V2 with 2x power zones

[3] Not 6.6KW due to current balancing between supplies are not perfect

5 Physical Specifications

5.1 Block Diagram

Figure 5-1 illustrates the functional block diagram of the Tioga Pass. The dashed lines are for reserved connection, dual layout, and high-speed mid-plane option.

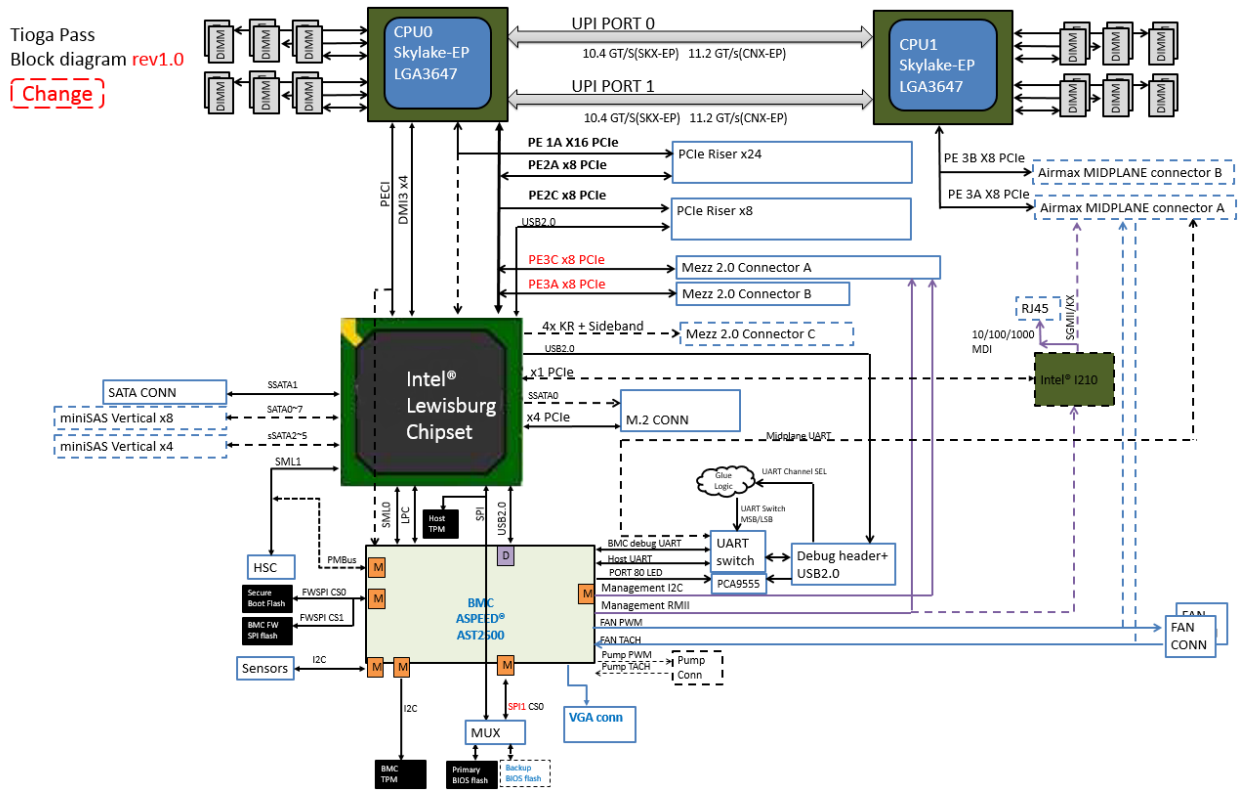


Figure 5-1 Tioga Pass Block Diagram

5.2 Placement and Form Factor

Board form factor is 6.5 inch by 20 inch (6.5”x20”). And Figure 5-2 illustrates board placement. The placement is meant to show key components’ relative positions, while exact dimension and position information would be exchanged by DXF format for layout and 3D model for mechanical. The vendor is responsible to do complete component placement. Form factor, CPU and DIMM socket location, PCIe slot position, front I/O ports’ position, Power and Reset buttons, PCIe Mezzanine Card connector position, mid-plane connectors, battery connector, and mounting holes should be followed strictly, while other components can be shifted based on layout routing as long as relative position is maintained.

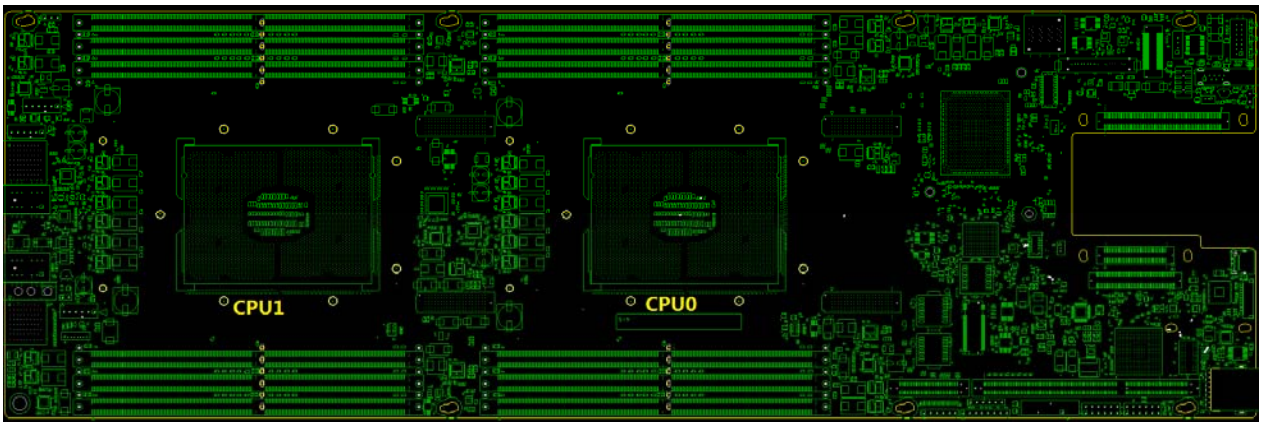
Following internal connectors should be placed as close as possible to front of the board in order to have easy front access:

- 1x vertical combo SATA signal and power connector
- 1x 14-pin Debug card header
- 1X right angle USB3 Type A connector
- 1X SMD switch to enable/disable Intel® Intel® At Scale Debug
- 1x M.2 connector with 2280 and 22110 support
- 1x RJ45
- 1x USB type C
- 1X customized VGA connector

Following mid-plane connector footprints should be placed at backside of board to provide mid-plane connection co-layout:

- 2x AirMax VS2® x8
- 1x AirMax® Guide
- 2x AirMax® VS Power 2x2

Figure 5-2 Tioga Pass Placement



Refer to DXF for critical component placement.

5.3 CPU and Memory

5.3.1 CPU

The motherboard supports all Intel® Xeon® Scalable processor family (aka Skylake-SP) processors with TDP up to 165W. The motherboard shall provision the support of all future CPUs in Intel® Xeon® Scalable processor Family Platform and the Next gen Intel® Xeon® Scalable processor Family Platform unless noted otherwise.

The features listed below must be supported by the motherboard:

- Support two Intel® Xeon® Scalable processor family (aka Skylake-SP) processors up to 165W TDP, and vendors should engage with Intel® to ensure the design ready for future processors

- Two full-width Intel® UPI links up to 10.4 GT/s/direction for Intel® Xeon® Scalable processor family (aka Skylake-SP) processor
- Up to 28 cores per CPU (up to 56 threads with Intel® Hyper-Threading Technology).
- Single Processor mode is supported
- Integrated Intel® Omni-Path Architecture Fabric is supported on CPU0 with BOM option with MCP VRM for CPU0 socket

5.3.2 DIMM

The motherboard has DIMM subsystem designed as below:

- DDR4 direct attach memory support on CPU0 and CPU1.
- 6x channels DDR4 registered memory interface on each CPU
- For Double side SKU, 2x DDR4 slots per channel (total 24x DIMM)
- The farthest DIMM slot of each channel is on the component side of PCB
- The nearest DIMM slot of each channel is on the solder side of PCB
- For Single side SKU, 1x DDR4 slots per channel (total 12x DIMM)
- Support DDR4 speeds up to 2666
- Support RDIMM, LRDIMM DDP, LRDIMM 3DS TSV-4H
- Support SR, DR, QR and 8R DIMM
- Up to maximum 3072 GB with 128 GB DRAM DIMM
- Follow updated JEDEC DDR4 specification with 288 pin DIMM socket

Board design shall allow Intel® Xeon® Processor to run at maximum POR memory speeds for Intel® Xeon® Scalable processor Family Platform and the Next gen Intel® Xeon® Scalable processor Family Platform with 2DIMM per channel configuration.

5.3.3 Non-Volatile DIMM

Besides traditional DDR4 DIMM, the motherboard shall support Non-Volatile DIMM (NVDIMM) on all DIMM slots as described:

- A power failure detection circuit needs to be implemented to initiate 3x actions related to data transferring:
  1. CPU cache flush
  2. Memory controller write pending queue flush and ADR mechanism
  3. Issue SAVE# signal to NVDIMM pin 230 to move DRAM data to NAND

Due to system energy storage and timing requirement constraints, the logic of item 1 *(CPU cache flush)* is disabled by default with the resistor option to be enabled. The logic of items 2 and 3 is enabled by default with resistor option set to disable. The Original Design Manufacturer (ODM) will work with NVDIMM vendor to implement the Basic Input/Output System (BIOS) design.

The under voltage based power failure detection circuit should also trigger separate CPU/DIMM/FAN throttling with separate resistor enable and disable options – the default is disable.

The NVDIMM shall cover data protection test cases described below:

- AC power disruption by removal of the node from the bus bar

- Issuing a raw, write-read command to the Baseboard Management Controller (BMC) to power cycle the hot-swap controller
- DC power off triggered by a four second override from either the front panel switch or BMC
- DC power cycle from the BMC via the chassis power cycle command
- Warm-reboot triggered by either the front panel switch or the BMC
- Power off and reboot triggered by the Operating System (OS)
- DC cycle from a host partition reset (write 0xE to 0xCF9)

5.4 Intel® PCH

The motherboard uses Intel® PCH chipset, which supports following features:

- 4x USB 3.0/2.0 ports: one type A for front connector; one type C for front connector; one for BMC in-band firmware update; one to X32 riser connector
- 1x M.2 connector; 1x individual SATA 6Gps port; 1x miniSAS HD x8 port or 1x miniSAS HD x4 port
- 1x PCIe x4 ports to M.2 connector ,colayout with SATA port to M.2 connector
- SPI interface, mux with BMC to enable BMC the capability to perform BIOS upgrade and recovery
- SPI interface for TPM header
- SMBUS interface (master & slave)
- Intel® Server Platform Services 4.0 Firmware with Intel® Node Manager
    - PECI access to CPU
    - SMLinko connect to BMC
    - Intel® Manageability Engine (ME) obtain HSC PMBus related information directly. Intel® Manageability Engine SMLink1 connects to Hotswap controller PMBus interface by default. BMC has connection to HSC PMBus to have flexibility of HSC PMBus related feature support, default is disconnected
    - Power capping capability
    - Temperature sensors reading from BMC
- PCH SKUs
    - Board design shall support all PCH SKUs in terms of power delivery and thermal design

5.5 PCIe Usage

5.5.1 PCIe lane mapping

Intel® Xeon® Scalable processor family (aka Skylake-SP) provides 48x PCIe Gen3 lanes and Intel® PCH provides up to x16 PCIe Gen3 uplinks and up to x20 PCIe. PCIe lanes are configured according to Table 2.

Table 2 Motherboard CPU and PCH PCIe Lane Usage

SS 2X16			SS 2X8 + 1X16			DS 2X16		
	PCIE down			PCIE down			PCIE down	
PCH	PCIe4	BMC PCIe Gen2 x1	PCH	PCIe4	BMC PCIe Gen2 x1	PCH	PCIe4	BMC PCIe Gen2 x1
		I210AT Gen1 x1(To RJ45)			I210AT Gen1 x1(To RJ45)			I210AT Gen1 x1(To RJ45)
PCH	PCIe5	I210AS Gen1 x1(to Airmax)	PCH	PCIe5	I210AS Gen1 x1(to Airmax)	PCH	PCIe5	I210AS Gen1 x1(to Airmax)
	PCIe Plug In			PCIe Plug In			PCIe Plug In	
PCH	PCIe[3..0]	M.2 PCIe x4	PCH	PCIe[3..0]	M.2 PCIe x4	PCH	PCIe[3..0]	M.2 PCIe x4
CPU0	PE3C x8	Mezz Conn A x8	CPU0	PE3C x8	Mezz Conn A x8	CPU0	PE3C x8	Mezz Conn A x8
CPU0	PE3A x16	Mezz Conn A+B x16	CPU0	PE3A x16	Mezz Conn A+B x16	CPU0	PE3A x16	Mezz Conn A+B x16
CPU0	PE1A x16	SS 2 slots Riser Lower x16	CPU0	PE1C x8	SS 3 slots Riser Lower x8	CPU0	PE1A x16	DS 2 slots Riser Lower x16
CPU0	PE2A x16	SS 2 slots Riser Upper x16	CPU0	PE1A x8	SS 3 slots Riser Middle x8	CPU0	PE2A x16	DS 2 slots Riser Upper x16
			CPU0	PE2A x16	SS 3 slots Riser Upper x16			
CPU1	PE 3A	Airmax Conn A x8	CPU1	PE 3A	Airmax Conn A x8	CPU1	PE 3A	Airmax Conn A x8
CPU1	PE 3B	Airmax Conn B x8	CPU1	PE 3B	Airmax Conn B x8	CPU1	PE 3B	Airmax Conn B x8

5.5.2 PCIe Hot Plug

The x16 PCIe in Airmax connectors from CPU1, and the x32 PCIe in riser slot shall support Standard PCIe signal hot swap defined in Intel® Xeon® Scalable processor family (aka Skylake-SP) CPU design specification.

Motherboard design shall connect PE_HP_SCL/SDA of CPU0 and CPU1 to Airmax connector or BMC in order for CPU to get access to the expender logic directly through Airmax connector, or through BMC. Refer to SMBus block diagram for connection.

PCIe power hotswap, and the expender logic is not in the scope of the motherboard design.

5.6 PCB Stack Up

Following PCB stack up should be followed for motherboard design. The vendor needs to check with PCB fab vendors to fine tune the impedance based on the impedance control table below before starting PCB design.

Table 3 Motherboard PCB Stack Up

	Material : IT170GRA1			
Layer	Layer type	Material requirement	Thickness	Thickness tolerance
			(mil)	(mil)
	solder mask		0.50	
1	Top	0.5oz+plating	1.80	
	prepreg	1080 62%	2.70	+/- 0.4
2	PLANE	1.0 oz RTF	1.30	
	core (10 mils)	2116 50.2% x 2	10.00	
3	SIGNAL	1.0 oz RTF	1.30	
	prepreg	106 71.5% x 2	3.00	
4	PLANE	1.0 oz RTF	1.30	
	core (10 mils)	2116 50.2% x 2	10.00	
5	SIGNAL	1.0 oz RTF	1.30	
	prepreg	106 71.5% x 2	3.00	
6	Plane	1.0 oz RTF	1.30	
	core	1086 57.9% x 1	3.00	
7	PLANE	2.0 oz	2.60	
	prepreg	1080 RC 68% x2	5.80	> 5 mil min.
8	PLANE/Signal	2.0 oz	2.60	
	core	1086 57.9% x 1	3.00	
9	Plane	1.0 oz RTF	1.30	
	prepreg	106 71.5% x 2	3.00	
10	SIGNAL	1.0 oz RTF	1.30	
	core (10 mils)	2116 50.2% x 2	10.00	
11	Plane	1.0 oz RTF	1.30	
	prepreg	106 71.5% x 2	3.00	
12	SIGNAL	1.0 oz RTF	1.30	
	core (10 mils)	2116 50.2% x 2	10.00	
13	Plane	1.0 oz RTF	1.30	
	prepreg	1080 62%	2.70	+/- 0.4
14	Bottom	0.5oz+plating	1.80	
	solder mask		0.50	
		Total Thickness	**92+/-9**	
			mils	

Important : Vendor must follow below items and discuss with FB if there is difficulty to meet

1. Total board thickness should be **92.1 mils +/- 9 mil**
2. Please keep the thickness between (L2 & L3) be target **3X** of the thickness between (L3 & L4) **or absolute ratio > 2.46 minimum.**
3. Please keep the thickness between (L4 & L5) be at least **3X** of the thickness between (L5 & L6) **or absolute ratio > 2.46 minimum.**
4. Please keep the thickness between (L10 & L11) be at least **3X** of the thickness between (L9 & L10) **or absolute ratio > 2.46 minimum.**
5. Please keep the thickness between (L12 & L13) be at least **3X** of the thickness between (L11 & L12) **or absolute ratio > 2.46 minimum.**
6. **Must meet width design +/-20% tolerance with +/-10 impedance control for signle-end and differential trace**
7. The prepreq between L1 & L2 and L13 & L14 should be controlled with **+/- 0.4 mil** tolerance
8. The prepreq between L7 & L8 should be controlled thickness **min. > 5 mil**

Table 4 Motherboard PCB Impedance Control

	Target Design Single Track Break Out Geometry				Target Design SE 40 ohms				Target Design SE 50 ohms				Target Design Diff Break Out Geometry			
Layer	Width	Ohms	ref. plane	Sim. Zo	Width	Ohms	ref. plane	Sim. Zo	Width	Ohms	ref. plane	Sim. Zo	Width/ SPACE	Ohms	ref. plane	Sim. Zo
1	3.5	NA	L2		6.5	40±10%	L2		4.25	50±10%	L2		3.5/4.0	NA	L2	
2																
3	3.5	NA	L2/4		5.6	40±10%	L2/4		3.52	50±10%	L2/4		3.5/4.0	NA	L2/4	
4																
5	3.5	NA	L4/6		5.6	40±10%	L4/6		3.52	50±10%	L4/6		3.5/4.0	NA	L4/6	
6																
7																
8																
9																
10	3.5	NA	L9/11		5.6	40±10%	L9/11		3.52	50±10%	L9/11		3.5/4.0	NA	L9/11	
11																
12	3.5	NA	L11/13		5.6	40±10%	L11/13		3.52	50±10%	L11/13		3.5/4.0	NA	L11/13	
13																
14	3.5	NA	L13		6.5	40±10%	L13		4.25	50±10%	L13		3.5/4.0	NA	L13	

	Target Design DIFF 85 ohms				Insertion Loss Spec (db/inch)		Target Design DIFF 100 ohms				Insertion Loss Spec (db/inch)	
Layer	Width/ SPACE	Ohms	ref. plane	Sim. Zo	@4GHz	@8GHz	Width/ SPACE	Ohms	ref. plane	Sim. Zo	@4GHz	@8GHz
1	5.0/7.0 & 4.5/5 (UPI)	85±10%	L2		-0.69	-1.38	3.95/14.05	100±10%	L2		-0.69	-1.38
2												
3	4.5/7	85±10%	L2/4		-0.65	-1.25	3.99/12.52	100±10%	L2/4		-0.65	-1.25
4												
5	4.5/7	85±10%	L4/6		-0.65	-1.25	3.99/12.52	100±10%	L4/6		-0.65	-1.25
6												
7												
8												
9												
10	4.5/7	85±10%	L9/11		-0.65	-1.25	3.99/12.52	100±10%	L9/11		-0.65	-1.25
11												
12	4.5/7	85±10%	L11/13		-0.65	-1.25	3.99/12.52	100±10%	L11/13		-0.65	-1.25
13												
14	5.0/7.0 & 4.5/5 (UPI)	85±10%	L13		-0.69	-1.38	3.95/14.05	100±10%	L13		-0.69	-1.38

6 BIOS

Vendors shall be responsible for supplying and customizing the BIOS for the motherboard. The specific BIOS requirements are outlined in this section. Vendors must make changes to the BIOS at any point in the motherboard's life cycle (development, production, and sustaining) upon request.

6.1 BIOS Chip

The BIOS chip should use PCH's SPI interface through BMC controlled MUX for BMC to perform offline BIOS update or recovery. The vendor is responsible for selecting a specific BIOS chip, which should fit the required functionality in addition to potential additional features that may be required in the future. 32MByte size is recommended considering space needed for both BIOS and Intel® Manageability Engine firmware. Vendor should provide flash region plan for different code and current used size for each region to justify the sizing of the SPI flash.

A socket on the motherboard should be used to hold BIOS chip, so BIOS chip can be manually replaced. The BIOS socket must be easily accessible; other components on the motherboard must not interfere with the insertion or removal of the BIOS chip; The BIOS socket needs to fit JEDEC specification package considering tolerance, and fit major available SPI flash vendors' package drawing.

A secondary identical BIOS chip is designed in sharing the same SPI bus with multiplexed CS pin; MUX is controlled by BMC. More detail of this scheme is described in section 9.15.

6.2 BIOS Source Code

BIOS should be a UEFI system firmware. The vendor shall be responsible to maintain BIOS source code to make sure it has latest code release from Intel® and UEFI system firmware code base vendors. Vendor shall provide an updated version tracker with each release.

6.3 BIOS Feature Requirements

6.3.1 Optimization

The BIOS should be tuned to minimize system power consumption and maximize performance. This includes:

- Disable any unused devices, such as unused PCI, PCIe ports, USB ports, SATA/SAS ports, clock generator and buffer ports.
- Tuning CPU/Chipset settings to reach minimized power consumption and best performance in a data center environment.
- Open Turbo Mode tuning option for PL1, PL2, PL1 clamp bit, PL2 clamp bit, short and long time duration.
- SPEC power should be used as guidance by ODM to validate BIOS tuning results.

6.3.2 Setup Menu

The vendor should provide a BIOS specification, which includes a complete BIOS, setup menu and default settings. Requirements include but are not limited to:

- Settings for adjusting memory speed, UPI speed, Speed-step/Turbo mode and CPU C-state power state: The default follows the CPU and chipset vendor's POR unless otherwise mentioned.
- Settings to enable different Turbo mode tuning settings based on CPU SKU and memory configuration: The default is Turbo enabled with the CPU vendor's POR, unless otherwise mentioned.
- Setting for the power feature after AC failure: The default is set to restore last power state.
- Setting for the local physical COM port (COM0) and SOL (COM1): The default is enable console redirection on both ports with baud rate 57600, no flow control, terminal type VT100, 8 data bits, No Parity, 1 Stop Bit.
- Setting for legacy console redirection to be local physical COM port (COM0) and SOL(COM1). The default is SOL(COM1)
- Setting for the altitude of the server deployment location: The default is 300M.
- Setting for the watchdog timer: The default setting for EVT/DVT/PVT is disabled. The default setting for MP is enabled. The timeout value is 15 minutes and reset the system after the timer expires. The watchdog timer is always disabled after POST.
- Setting for ECC error threshold: Available settings are 1, 4, 10 and 1000. The default setting is 1 for EVT, DVT, and PVT and 1000 for MP.
- Setting for ECC error event log threshold: Available settings are disabled, 10, 50, 100. The default setting is 10.
- If a CMOS CRC error happens, the BIOS should load the system default automatically and log the CMOS clear event in SEL.
- The default setting to disable all "wait for keyboard input to continue" types of features is "not to wait for keyboard input to continue".
- Calculate checksum of BIOS setting, display in BIOS setup menu, and output to SMBIOS table.

- Setting to save and load 10 different sets of user default.
- Setting of UEFI and Legacy boot options: The default is UEFI.
- Display SKU and hardware revision in main page based on BOARD ID and FAB ID.
- Setting of PPIN (Protected Processor Inventory Number) Control: The default setting is unlock/enable.
- Display RC version in main page.
- Display CPU information in main page including CPU signature, processor cores, and microcode patch version.
- Display memory information in main page including current memory speed, total memory capacity and type (DDR4).
- Display PCH information in main page including name and stepping.
- Setting of Setup Prompt timeout: The default is "7 seconds".

6.3.3 Boot Options

The BIOS must support PXE Boot capability in both IPv4 and IPv6 environment at the same time, and boot from SATA/SAS and USB interface. BIOS should provide boot option selection capability. The default boot device priority is:

- 1st: USB device if available
- 2nd: Mezzanine card NIC IPv6
- 3rd: Mezzanine card NIC IPv4
- 4rd: LOM Intel® I210 IPv6
- 5th: LOM Intel® I210 IPv4
- 6th: PCIe M.2 or SATA M.2
- 7th: SATA HDD
- 8th: SATA-CDROM
- 9th: Reserved

If there is no bootable device found, BIOS should keep loop searching for bootable device.

BIOS should support UEFI and legacy boot mode options and default is UEFI. UEFI and legacy boot mode have independent boot loop.

Boot mode and boot order can be displayed and changed from BMC with OEM command.

6.3.4 Board SKU ID

The motherboard should provide 5 strapping pins to be used as BOARD_SKU_ID[4:0], so BIOS can do correct board initialization based on different board configurations. Board ID is also routed to BMC to be accessed by BMC firmware.

Note:

Board SKU ID4 is to identify MB SKU, 0=single side, 1=double side

Board SKU ID3 is to identify VR power stage, 0=Fairchild, 1=Infineon

Board SKU ID2 is to identify QAT, 0= no support, 1=support

Board SKU ID1 is to identify M.2 Type, 0=SATA, 1=PCIe

Board SKU ID0 is to identify ODM. 0=Vendor1, 1=Vendor2

SKU_ID[4:0]					Description
4	3	2	1	0	
0	0	0	0	0	Single Side/Fairchild/no QAT/M.2=SATA
0	0	0	1	0	Single Side/Fairchild/no QAT/M.2=PCIe
0	0	1	0	0	Single Side/Fairchild/ QAT /M.2=SATA
0	0	1	1	0	Single Side/Fairchild/ QAT /M.2=PCIe
0	1	0	0	0	Single Side/Infineon/no QAT/M.2=SATA
0	1	0	1	0	Single Side/Infineon/no QAT/M.2=PCIe
0	1	1	0	0	Single Side/Infineon/ QAT /M.2=SATA
0	1	1	1	0	Single Side/Infineon/ QAT /M.2=PCIe
1	0	0	0	0	Double Side/Fairchild/no QAT/ M.2=SATA
1	0	0	1	0	Double Side/Fairchild/no QAT/ M.2=PCIe
1	0	1	0	0	Double Side/Fairchild/ QAT / M.2=SATA
1	0	1	1	0	Double Side/Fairchild/ QAT / M.2=PCIe
1	1	0	0	0	Double Side/Infineon/no QAT/ M.2=SATA
1	1	0	1	0	Double Side/Infineon/no QAT/ M.2=PCIe
1	1	1	0	0	Double Side/Infineon/ QAT / M.2=SATA
1	1	1	1	0	Double Side/Infineon/ QAT / M.2=PCIe

The motherboard has 1Kbit EEPROM(Address 0xA2 in 8 bit format) for soft-strap board ID to be accessed by BIOS on host SMBus. The definition of the soft-strap is TBD. Vendor shall keep this EEPROM blank until the definition is provided by Facebook. If no definition defined during DVT, vendor shall remove it from BOM.

6.3.5 FAB Revision ID

The motherboard should provide 3 strapping pins to be used as FAB_REVISION_ID [2:0], so BIOS can differentiate correct board FAB versions. FAB revision ID is also routed to BMC to be accessed by BMC firmware.

FAB_ID[2:0]			Description
0	0	0	FAB1
0	0	1	FAB2
0	1	0	FAB3
0	1	1	FAB4
1	0	0	Fab5

6.3.6 Remote BIOS Update Requirement

Vendors should provide tool(s) to implement remote BIOS update function. Vendor must validate update tools on each BIOS release during development and production. Tool(s) provided should support four update scenarios:

- Scenario 1: Sample/Audit BIOS settings
  - Return current BIOS settings, or
  - Save/Export BIOS settings in a human-readable form that can be restored/imported (i.e. Scenario 2). Output must include detailed value-meaning description for each setting. Setting must include pre-production setup menus/options too.
- Scenario 2: Update BIOS with pre-configured set of BIOS settings
  - Update/Change multiple BIOS settings. Setting must include pre-production setup menus/options. Tool(s) should provide detailed value-meaning description for each setting.
  - Reboot
- Scenario 3: BIOS update with a new revision
  - Load new BIOS on machine and Update, retaining current BIOS settings
  - Reboot
- Scenario 4: use BMC to update BIOS in PCH flash(also described in section 9.15)
  - Update BIOS flash from BMC.
  - Update need to be done with command line script in Linux environment from a remote server. Web GUI interface is not accepted.

Additionally, the update tools and scripts should have the following capabilities:

- Update from the Operating System through ssh; the current OS based is CentOS 7.x 64-bit with updated Kernel specified by customer
- Require no more than one reset cycle to the system to complete BIOS update or BIOS setup option change
- Require no physical access to system
- BIOS update or BIOS setup option change should not take more than 5 minutes to complete
- BIOS update procedure can be scripted and propagated to multiple machines

6.3.7 Event log requirement

BIOS should do event log through BMC SEL with Generator ID 0x0001 and the combination of BIOS and BMC should meet the SEL log requirements in section 9.11.

6.3.8 BIOS Error Code Requirement

BIOS fatal error codes listed in following table should be enabled for POST CODE output. Vendor should display major and minor code alternatively.

Table 5 BIOS Error Code

Fatal Errors	Major Code	Minor Code	Error Description
ERR_NO_MEMORY	0E8h		
ERR_NO_MEMORY_MINOR_NO_MEMORY		01h	1. No memory was detected via SPD read. No warning log entries available.

			2. Invalid configuration that causes no operable memory. Refer to warning log entries for details.
ERR_NO_MEMORY_MINOR_ALL_CH_DISABLED		02h	Memory on all channels of all sockets is disabled due to hardware memtest error.
ERR_NO_MEMORY_MINOR_ALL_CH_DISABLED_MIXED		03h	No memory installed. All channels are disabled.
ERR_LT_LOCK	0E9h		Memory is locked by LT, inaccessible.
ERR_DDR_INIT	0EAh		DDR training did complete successfully
ERR_RD_DQ_DQS		01h	Error on read DQ/DQS init
ERR_RC_EN		02h	Error on Receive Enable
ERR_WR_LEVEL		03h	Error on Write Leveling
ERR_WR_DQ_DQS		04h	Error on write DQ/DQS
ERR_MEM_TEST	0EBh		Memory test failure
ERR_MEM_TEST_MINOR_SOFTWARE		01h	Software memtest failure
ERR_MEM_TEST_MINOR_HARDWARE		02h	Hardware memtest failure
ERR_MEM_TEST_MINOR_LOCKSTEP_MODE		03h	Hardware memtest failure in Lockstep channel mode requiring a channel to be disabled. This is a fatal error which requires a reset and calling BIOS with a different RAS mode to retry
ERR_VENDOR_SPECIFIC	0ECh		
ERR_DIMM_COMPAT	0EDh		RDIMMs is present DIMM vendor-specific errors
ERR_MIXED_MEM_TYPE		01h	Different DIMM types are detected installed in the system
ERR_INVALID_POP		02h	Violation of population rules
ERR_INVALID_POP_MINOR_UNSUPPORTED_VOLTAGE		05h	Unsupported DIMM Voltage
Reserved	0EEh		Reserved
ERR_INVALID_BOOT_MODE		01h	Boot mode is unknown
ERR_INVALID_SUB_BOOT_MODE		02h	Sub boot mode is unknown

6.3.9 POST Code for Option ROM Entry and Exit

Special BIOS post codes are assigned to indicate the entry and exit of option ROM. Two Bytes sequence are assigned for the entry and the exit to avoid same sequence used on other BIOS Post code. For example, use AA-C0 indicates entry, and use BB-C1 indicates exit. These two sequences should be avoided to be used in other post codes process.

6.3.10 PPIN BIOS Implementation

BIOS shall support PPIN and set default to [Unlock and Enable].

There are two ways for user to access PPIN:

- BIOS shall map PPIN of CPU0 and CPU1 to SMBIOS OEM Strings (Type 11) String 5 and String 6. User can view PPIN value from Linux's "*dmidecode" command*.
- BIOS shall implement Send_PPIN and Get_PPIN OEM Command to communicate to BMC, per BMC's request. User can retrieve PPIN information from BMC through OEM command.

BIOS shall perform 2x actions to synchronize PPIN value to BMC:

- Serves SMI# signal from BMC and use Send_PPIN OEM command to communicate PPIN to BMC.

- Use Send_PPIN OEM command to communicate PPIN to BMC when BIOS POST COMPLETE.

7 PCH Intel® Server Platform Services Firmware Plan of Record

Tioga Pass uses Intel® PCH chipset. Its Management Engine (Intel® Manageability Engine) runs Intel® Server Platform Services 4.0 firmware. Intel® Server Platform Services firmware is required for system operation.

Intel® Server Platform Services 4.0 Firmware (FW) consists of two parts. Intel® Server Platform Services Silicon Enabling FW is required to boot the platform and have a fully functional platform. Intel® Server Platform Services Node Manager FW provides power, thermal and compute utilization statistics, P/T-state control, simple and multi-policy power limiting at platform, memory, processor and core level with assist of a BMC. Both parts are required for Tioga Pass. All Intel® Xeon® Scalable processor Family Platform Server Platform Services FW features, such as Fast NM, Predictive power capping, NM PTU, etc., shall be supported and validated.

Intel® Server Platform Services firmware is stored in PCH flash. Vendor should provide utility to update Intel® Server Platform Services firmware in CentOS 7.x 64-bit with updated Kernel specified by customer through ssh. Utility should support updating Intel® Server Platform Services firmware and BIOS together or separately, and also provide option to update only Intel® Server Platform Services Firmware's operation region or entire Intel® Server Platform Services firmware region. Vendor should also implement BMC to update PCH flash where Intel® Server Platform Services firmware is located as described in section 9.15.

8 Innovation Engine (IE)

Tioga Pass does not plan to enable Innovation Engine. IE will be fused off in MP manufacturing flow as IE is not used.

9 BMC

Tioga Pass uses an ASPEED® AST2500 BMC with one x16 4Gb DDR4 Dram for various platform management services and interfaces with hardware, BIOS, Intel® Server Platform Services firmware.

BMC should be a standalone system in parallel to host (dual processor x86). The health status of host system should not affect the normal operation and network connectivity of BMC. BMC cannot share memory with host system. BMC management connectivity should work independently from host, and has no NIC driver dependency for Out-of-Band communication if using a shared-NIC.

9.1 Management Network Interface

BMC should have both I2C port and RMII/NCSI port for Out-of-Band access.

Three options of Out-of-Band access methods should be supported. Option 2 and Option 3 share same device footprint as co-layout.

Option 1: Shared-NIC uses I2C or RMII/NCSI interfaces to pass management traffic on data network of Mezzanine 25GE NIC.

Option 2: Shared-NIC uses RMII/NCSI interfaces to pass management traffic on data network of Intel® I210-AS. Intel® I210-AS has SGMII/KX interface to mid-plane.

Option 3: Shared-NIC uses RMII/NCSI interfaces to pass management traffic on data network of Intel® I210-AT. Intel® I210-AT has 10/100/1000 MDI interface to RJ45.

BMC firmware needs to be flexible about which interface and device to activate by hardware strapping or a preset priority policy. BMC firmware needs to make sure the unused interfaces and devices are disabled and do not interfere with the activated management interface and device.

MAC address of Out-of-Band should use NIC's data network MAC plus an offset defined by NIC vendors.

BMC management network firmware and utility need to support all features defined in this specification in both IPv4 and IPv6 network environment.

9.2 Local Serial Console and SOL

BMC needs to support two paths to access the serial console:

A local serial console on debug header (described in section 11.8) and a remote console, also known as Serial-Over-LAN (SOL) through the management network (described in section 0). It is preferred that both interfaces are functional at all stages of system operation.

During system booting, POST codes will be sent to port 80 and decoded by the BMC to drive the LED display as described in section 9.5. POST codes should be displayed in SOL console during system POST. Before the system has the first screen, POST codes are dumped to and displayed in the SOL console in sequence. For example, display as "[00] [01] [02] [E0]…" etc. After the system has the first screen in the SOL console, the last POST code received on port 80 is displayed in the lower right corner of the SOL console screen.

A serial console buffer feature is required. A buffer needs to save at least the last 5x screens of local and 5x screens of remote console output; 80 column x24 row for each screen. The buffer has to be stored in volatile media, such as internal or external SDRAM of BMC. SOL buffer data is cleared within 5 seconds of the removal of standby power. The SOL buffer should NOT be stored in any non-volatile media for security and privacy. SOL buffer implementation shall allow the SOL buffer being dumped by script with OEM command to a file (for scaling of data collection).

9.3 Graphic and GUI

As the graphic user interface (GUI) is not scalable, all the BMC features need to be available in command line model or SOL. Tioga Pass adds support of GUI and KVM on hardware level to accommodate the OCP customers whose environment requires using of VGA and KVM.

9.4 Remote Power Control and Power policy

The vendor should implement BMC firmware to support remote system power on/off/cycle and warm reboot through BMC.

The vendor should implement BMC firmware to support power on policy to be last-state, always-on, and always-off. The default setting is last-state. The change of power policy should be supported through BMC and take effect without BMC a firmware cold reset or a system reboot.

It should take less than 3 seconds from AC on, for BMC to process power button signal and power up system for POST. A long waiting period from AC on, for BMC firmware to get ready before allowing system POST start is NOT allowed.

9.5 Port 80 POST

The vendor should implement BMC to support port 80 POST code display to drive 8 bit HEX GPIO to debug header. The BMC post function need to be ready before system BIOS starts to send 1st POST code to port 80. POST code should also be sent to SOL as mentioned in section 9.2.

BMC should have access to POST code and record up to 256x POST codes and user should be able to retrieve last 256x POST code from BMC.

9.6 Power and System Identification LED

The motherboard combines Power LED and System Identification LED to a single blue LED at front side.

Power LED on is defined by the readiness of major run time power rails (P12V, P5V, and P3V3) but NOT the readiness of all run time power rails; for example, CPU core power rail being ready is not required for Power LED on indication.

Power LED blinking is used as system identification. The on time is different during power on and power off.

There are 4 states of Power/system identification LED depending on system power state, and chassis identify status.

Power off, Chassis identify off:	LED consistently off
Power off, Chassis identify on:	LED on for 0.1sec, off for 0.9sec, and loop
Power on, Chassis identify off:	LED consistently on
Power on, Chassis identify on:	LED on for 0.9sec, off for 0.1sec, and loop

9.7 Platform Environment Control Interface (PECI)

BMC should access Platform Environment Control Interface (PECI) through PCH SMLink0 by default. PECI connection implementation should follow Intel® guidelines. BMC should be able to execute PECI raw command by using Intel® Manageability Engine as a proxy.

Vendor should implement board design to connect CPU PECI interface to PCH PECI or BMC PECI by adding an analog switch controlled by BMC GPIOAA4(ASPEED 2500 Pin U22). CPU PECI is accessed by Intel® Manageability Engine firmware by default when GPIO is low and it will switch to BMC access PECI when ME is not responsive.

9.8 Power and Thermal Monitoring and power limiting

The vendor should implement BMC firmware to support platform power monitoring. To enable power limiting for processor, memory, and platform, Intel® Server Platform Services-NM is required.

The vendor should implement BMC FW to support thermal monitoring, including processor, memory, chipset, VRs, PCIe card, Mezzanine cards, Inlet/outlet air temperature, and airflow sensor. To make sure of temperature report accuracy, TI TMP421 with external transistor is preferred to be used for detect Inlet and Outlet temperature. Caution could be taken for inlet air sensor implementation to avoid preheats of nearby components, and reduce heat conducted through PCB. Airflow sensor is not a physical sensor; Airflow is calculated based on system FAN PWM.

9.9 SMBUS Diagram

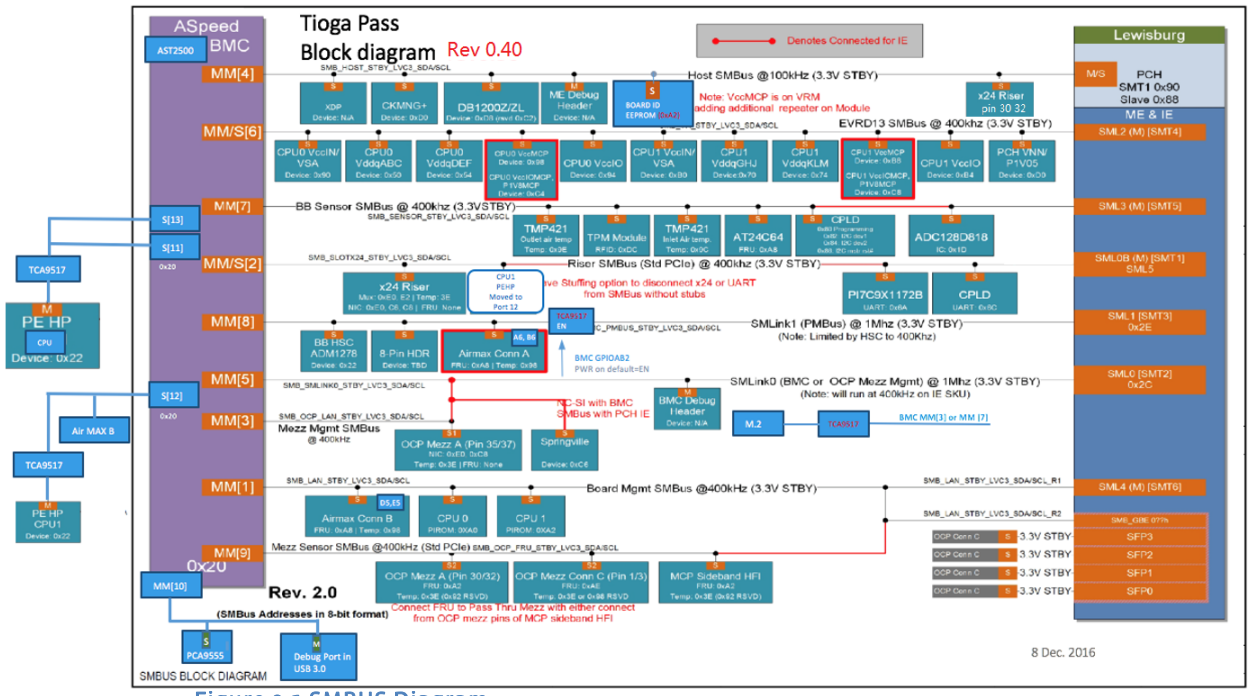


Figure 9-1 SMBUS Diagram

9.10 Sensors

This chapter describes all Analog sensors required. It does not include all the detail requirements. Please refer to sensor table below for more detail requirements.

BMC has access to all analog sensors on the motherboard directly or through PCH Management Engine. All analog sensors need to be displayed in sensor data record repository.

The analog sensors required are list as below. The lower and upper critical threshold is listed for system event logging purpose. Please refer to section 9.11 for logging requirements.

Table 9-1 Analog Sensor Table with Lower and Upper Critical

Sensor Description	Sensor Name	Sensor#	LCR	UCR	Reading Available	Access Info
HSC Input Power	MB_HSC_IN_POWER	0x29	NA	792	Standby	
HSC Input Vol	MB_HSC_IN_VOLT	0x2A	10.8	13.2	Standby	
HSC Output Current	MB_HSC_OUT_CURR	0xC1	NA	47.71	Standby	I2C:

	MB_HSC_TEMP	0xC2				
P0 Vccin VR Vol	MB_VR_CPU0_VCCIN_VOLT	0xB0	1.45	2.05	DC	I2C:
P0 Vccin VR Temp	MB_VR_CPU0_VCCIN_TEMP	0xB1	NA	100	DC	
P0 Vccin VR Curr	MB_VR_CPU0_VCCIN_CURR	0xB2	NA	235	DC	
P0 Vccin VR Pwr	MB_VR_CPU0_VCCIN_POWER	0xB3	NA	414	DC	
P0 Vsa VR Vol	MB_VR_CPU0_VSA_VOLT	0xB4	0.45	1.2	DC	I2C:
P0 Vsa VR Temp	MB_VR_CPU0_VSA_TEMP	0xB5	NA	100	DC	
P0 Vsa VR Curr	MB_VR_CPU0_VSA_CURR	0xB6	NA	20	DC	
P0 Vsa VR Pwr	MB_VR_CPU0_VSA_POWER	0xB7	NA	25	DC	
P0 Vccio VR Vol	MB_VR_CPU0_VCCIO_VOLT	0xB8	0.8	1.2	DC	I2C:
P0 Vccio VR Temp	MB_VR_CPU0_VCCIO_TEMP	0xB9	NA	100	DC	
P0 Vccio VR Curr	MB_VR_CPU0_VCCIO_CURR	0xBA	NA	24	DC	
P0 Vccio VR Pwr	MB_VR_CPU0_VCCIO_POWER	0xBB	NA	32	DC	
DIMM ABC VR Vol	MB_VR_CPU0_VDDQ_GRPA_VOLT	0xBC	1.08	1.32	DC	I2C:
DIMM ABC VR Temp	MB_VR_CPU0_VDDQ_GRPA_TEMP	0xBD	NA	90	DC	
DIMM ABC VR Curr	MB_VR_CPU0_VDDQ_GRPA_CURR	0xBE	NA	40 [SS] 95 [DS]	DC	
DIMM ABC VR Pwr	MB_VR_CPU0_VDDQ_GRPA_POWER	0xBF	NA	66 [SS] 115 [DS]	DC	
DIMM DEF VR Vol	MB_VR_CPU0_VDDQ_GRPB_VOLT	0xCC	1.08	1.32	DC	I2C:
DIMM DEF VR Temp	MB_VR_CPU0_VDDQ_GRPB_TEMP	0xCD	NA	90	DC	
DIMM DEF VR Curr	MB_VR_CPU0_VDDQ_GRPB_CURR	0xCE	NA	40 [SS] 95 [DS]	DC	
DIMM DEF VR Pwr	MB_VR_CPU0_VDDQ_GRPB_POWER	0xCF	NA	66 [SS] 115 [DS]	DC	
P0 DIMM ABC Temp	MB_CPU0_DIMM_GRPA_TEMP	0xAC	NA	81	DC	via ME
P0 DIMM DEF Temp	MB_CPU0_DIMM_GRPB_TEMP	0xAD	NA	81	DC	via ME
P0 Temp	MB_CPU0_TEMP	0xAA	NA	Tjmax - 4	DC	via ME
P0 Tjmax	MB_CPU0_TJMAX	0x30	NA	NA	DC	via ME
P1 Vccin VR Vol	MB_VR_CPU1_VCCIN_VOLT	0xF0	1.45	2.05	DC	I2C:
P1 Vccin VR Temp	MB_VR_CPU1_VCCIN_TEMP	0xF1	NA	100	DC	
P1 Vccin VR Curr	MB_VR_CPU1_VCCIN_CURR	0xF2	NA	235	DC	
P1 Vccin VR Pwr	MB_VR_CPU1_VCCIN_POWER	0xF3	NA	420	DC	
P1 Vsa VR Vol	MB_VR_CPU1_VSA_VOLT	0xF4	0.45	1.2	DC	I2C:
P1 Vsa VR Temp	MB_VR_CPU1_VSA_TEMP	0xF5	NA	100	DC	
P1 Vsa VR Curr	MB_VR_CPU1_VSA_CURR	0xF6	NA	20	DC	
P1 Vsa VR Pwr	MB_VR_CPU1_VSA_POWER	0xF7	NA	25	DC	

P1 Vccio VR Vol	MB_VR_CPU1_VCCIO_VOLT	0xF8	0.8	1.2	DC	I2C:
P1 Vccio VR Temp	MB_VR_CPU1_VCCIO_TEMP	0xF9	NA	100	DC	
P1 Vccio VR Curr	MB_VR_CPU1_VCCIO_CURR	0xFA	NA	24	DC	
P1 Vccio VR Pwr	MB_VR_CPU1_VCCIO_POWER	0xFB	NA	32	DC	
DIMM GHJ VR Vol	MB_VR_CPU1_VDDQ_GRPC_VOLT	0xFC	1.08	1.32	DC	I2C:
DIMM GHJ VR Temp	MB_VR_CPU1_VDDQ_GRPC_TEMP	0xFD	NA	90	DC	
DIMM GHJ VR Curr	MB_VR_CPU1_VDDQ_GRPC_CURR	0xFE	NA	40 [SS] 95 [DS]	DC	
DIMM GHJ VR Pwr	MB_VR_CPU1_VDDQ_GRPC_POWER	0xD8	NA	66 [SS] 115 [DS]	DC	
DIMM KLM VR Vol	MB_VR_CPU1_VDDQ_GRPD_VOLT	0xD9	1.08	1.32	DC	I2C:
DIMM KLM VR Temp	MB_VR_CPU1_VDDQ_GRPD_TEMP	0xDA	NA	90	DC	
DIMM KLM VR Curr	MB_VR_CPU1_VDDQ_GRPD_CURR	0xDB	NA	40 [SS] 95 [DS]	DC	
DIMM KLM VR Pwr	MB_VR_CPU1_VDDQ_GRPD_POWER	0xDC	NA	66 [SS] 115 [DS]	DC	
P1 DIMM GHJ Temp	MB_CPU1_DIMM_GRPC_TEMP	0xAE	NA	81	DC	via ME
P1 DIMM KLM Temp	MB_CPU1_DIMM_GRPD_TEMP	0xAF	NA	81	DC	via ME
P1 Temp	MB_CPU1_TEMP	0xAB	NA	Tjmax - 4	DC	via ME
P1 Tjmax	MB_CPU1_TJMAX	0x31	NA	NA	DC	via ME
PCH PVNN VR Vol	MB_VR_PCH_PVNN_VOLT	0xC8	0.76	1.1	DC	I2C:
PCH PVNN VR Temp	MB_VR_PCH_PVNN_TEMP	0xC9	NA	80	DC	
PCH PVNN VR Curr	MB_VR_PCH_PVNN_CURR	0xCA	NA	23	DC	
PCH PVNN VR Pwr	MB_VR_PCH_PVNN_POWER	0xCB	NA	28	DC	
PCH P1V05 VR Vol	MB_VR_PCH_P1V05_VOLT	0xC4	0.94	1.15	DC	I2C:
PCH P1V05 VR Temp	MB_VR_PCH_P1V05_TEMP	0xC5	NA	80	DC	
PCH P1V05 VR Curr	MB_VR_PCH_P1V05_CURR	0xC6	NA	19	DC	
PCH P1V05 VR Pwr	MB_VR_PCH_P1V05_POWER	0xC7	NA	26	DC	
P3V BAT	MB_P3V_BAT	0xD0	2.73	3.74	Standby	
P3V3	MB_P3V3	0xD1	2.97	3.62	DC	
P5V	MB_P5V	0xD2	4.52	5.49	DC	
P12V STBY	MB_P12V	0xD3	10.77	13.23	Standby	
P1V05 STBY PCH	MB_P1V05	0xD4	0.94	1.15	Standby	
PVNN PCH	MB_PVNN_PCH_STBY	0xD5	0.76	1.1	Standby	
P3V3 STBY	MB_P3V3_STBY	0xD6	2.97	3.62	Standby	
P5V STBY	MB_P5V_STBY	0xD7	4.52	5.49	Standby	
PCH Temp	MB_PCH_TEMP	0x8	NA	84	DC	via ME
Inlet Temp	MB_INLET_TEMP	0xA0	NA	40	Standby	

Outlet Temp	MB_OUTLET_TEMP	0xA1	NA	90	Standby	
Inlet Remote Temp	MB_INLET_REMOTE_TEMP	0xA3	NA	40	Standby	
Outlet Remote Temp	MB_OUTLET_REMOTE_TEMP	0xA4	NA	75	Standby	
C1 Remote Temp	MEZZ_SENSOR_TEMP	0xA2	NA	95	Standby	
SYS FAN0	MB_FAN0_TACH	0xC0	500	9000[SS] 13500[DS]	DC	
SYS FAN1	MB_FAN1_TACH	0xC3	500	9000[SS] 13500[DS]	DC	
SYS PUMP						
P0 Package Power	MB_CPU0_PKG_POWER	0x2C				
P1 Package Power	MB_CPU1_PKG_POWER	0x2D				
INA230 Sensors	MB_C2_P12V_INA230_VOL	0x90	11.04	12.96	DC	
	MB_C2_P12V_INA230_CURR	0x91	NA	5.5	DC	
	MB_C2_P12V_INA230_PWR	0x92	NA	75	DC	
	MB_C3_P12V_INA230_VOL	0x93	11.04	12.96	DC	
	MB_C3_P12V_INA230_CURR	0x94	NA	5.5	DC	
	MB_C3_P12V_INA230_PWR	0x95	NA	75	DC	
	MB_C4_P12V_INA230_VOL	0x96	11.04	12.96	DC	
	MB_C4_P12V_INA230_CURR	0x97	NA	5.5	DC	
	MB_C4_P12V_INA230_PWR	0x98	NA	75	DC	
	MB_CONN_P12V_INA230_VOL	0x99	11.04	12.96	DC	
	MB_CONN_P12V_INA230_CURR	0x9A	NA	20	DC	
	MB_CONN_P12V_INA230_PWR	0x9B	NA	250	DC	

9.11 SEL

Vendor should implement BMC to support System Event Log (SEL). Please refer to Tioga pass Feature list v1.7 for detail requirements.

9.11.1 Error to be logged

- **Analog Sensors**
  SEL is logged when Analog Sensor triggers upper or lower threshold.
- **Machine Check Error**
  MCE shall be logged on running time when MCE generates SMI#. BIOS shall check MCE banks after a warm reboot and log error before clearing MCE
- **Memory Error**
  Both correctable ECC and uncorrectable ECC errors should be logged into SEL. Each log entry should indicate location of DIMM by CPU socket#, Channel # and slot #. Memory error reporting need to be tested by both XDP injection and reworked ECC DIMM.
- **UPI error**
  All errors, which have status, register should be logged into Event Log. Fatal or non-fatal classification follows Chipset vendor's recommendation.
- **PCIe error**
  All errors, which have status register, should be logged into Event Log, including root complex, endpoint device and any switch upstream/downstream ports if available. Link

disable on errors should also be logged. Fatal, non-fatal, or correctable classification follows Chipset vendor's recommendation.

- **POST error**
  All POST errors, which are detected by BIOS during POST, should be logged into Event Log.
- **Power error**
  Two kinds of power errors should be logged:
  - SYS_PWROK FAIL: SYS_PWROK has a falling edge when SLP_S3_N is high; normal AC/DC cycle/or HSC cycle shall not trigger this event
  - PCH_PWROK FAIL = PCH_PWROK has a falling edge when SLP_S3_N is high; normal AC/DC cycle/or HSC cycle shall not trigger this event
- **MEMHOT# , PROCHOT# and VRHOT#**
  Memory hot error and Processor hot error should be logged. Error log should identify error source as internally coming from processor or memory, externally coming from voltage regulator, over-current triggered throttling, or under-voltage triggered throttling.
- **FAN Failure**
  FAN failure error should be logged if FAN speed reading is out of expected ranges between lower and upper critical threshold. Error log should also identify which FAN fails.
- **PMBus status error**
  PMBus status sensors checks PMBus controller health status and log error if abnormal value is detected. PMBus controller can be a DC Hot Swap Controller (HSC) or PMBus AC to DC power supply unit.
- **Intel® Server Platform Services FW related Error logging**

For all above error logging and reporting, you may select to enable or disable each one of them based on needs.

9.11.2 Error Threshold Setting

Enable error threshold setting for both correctable and uncorrectable errors, once programmed threshold is reached, an event should be trigged and logged.

- Memory Correctable ECC. Suggest setting threshold value to be [1000] in MP stage and [1] for evaluation, development, pilot run stage, with option of [1, 4, 10, 1000]. When threshold is reached, BIOS should log the event including DIMM location information and output DIMM location code through a debug card.
- ECC error event log threshold, defines the max number of correctable DIMM ECC is logged in the same boot. Default value is 10, with option Disable, 10, 50, and 100.
- UPI Error. Follow chipset vendor's suggestion.
- PCIE Error. Follow chipset vendor's suggestion.

9.11.3 Critical SEL Filter

OEM commands are required to set and display two different level of SEL filtering. Default is to log all error during EVT/DVT/PVT with option to log only critical SEL that needs service or indicates power cycle state change, and SEL clear and overflow.

9.12 FSC in BMC

The vendor should enable Fan Speed Control (FSC) on BMC. The BMC samples thermal related analog sensors in real time. The FSC algorithm processes these inputs and drives two PWM outputs to optimized speed.

9.12.1 Data gathering for FSC

BMC needs to gather data as possible input of FSC. The data to be gathered includes:

Type of data	Data used for FSC input
Temperature	CPU0/1 core temperature from PECI
Temperature	TSOD of all DIMMs from PECI
Temperature	PCH temperature through SMLINK0
Temperature	Inlet and outlet air
Temperature	VR of CPU and DIMM
Temperature	Hot Swap Controller
Temperature	Mezz card and PCIe card support thermal reporting interface
Power	CPU0/1 package power through PECI
Power	DIMM power through VR
Power	Platform power from HSC
Fan speed	2 FAN tachometer inputs
Airflow	Airflow sensor

The BMC sensor monitor interval is 2s.

9.12.2 FSC in BMC

Vendor shall follow and implement FSC and FSC update interface per *Facebook Server Fan Speed Control Interface*[5]. The BMC should support FSC configuration updates. Updates should take effect immediately without requiring a reboot.

9.12.3 Fan Connection

The motherboard has 2x FAN and 1XPump headers on motherboard. The motherboard and mid-plane interface both have optional FAN tachometer and PWM connections. Fan number should start from 0 and numbering starts from left to right by looking from inlet sensor.

9.13 OEM commands

Vendor shall implement OEM features with OEM command listed in the table below

Table 9-2 OEM command list

NetFn	Command	Comments

[5] http://files.opencompute.org/oc/public.php?service=files&t=d48482b8b87a596dd93ac5b80e9fa3a2&download

Chassis	Get Chassis Status (0x01)	BIOS to read the current power restore policy configuration
	Set Power Restore Policy (0x06)	BIOS to set power restore policy configuration
	Get System Restart Cause(0x07)	BIOS to know why the system got restarted e.g. user command vs. power button vs. power policy vs. WDT
	Get Boot Options(0x09)	Boot Order sequence;CMOS settings clearance
App	Get Device ID (0x01)	Generic info purpose
	Cold Reset (0x02)	Reset BMC from Host
	Get Selftest Results (0x04)	BIOS decides weather to install IPMI protocol or not
	Manufacturing Test On (0x05)	Allowed Sled-cycle via KCS
	Get Device GUID (0x08)	To read unique UID for device
	Reset WDT(0x22)	WDT commands are used by FRB2 timer
	Set WDT(0x24)	
	Get WDT(0x25)	
	Set BMC Global Enables (0x2e)	Used by ipmi_si driver
	Get Global Enables (0x2f)	Used by BIOS for deciding to send SEL events to BMC or not
	Clear Message Flags (0x30)	Used by ipmi_si driver
	Get System GUID (0x37)	To read unique UID for System
	Set System Info Params(0x58)	Inform BIOS version information to BMC
	Get System Info Params(0x59)	Read BIOS version info from Host
Storage	Get FRUID Info(0x10)	FRUID info is needed by BIOS to populate SMBIOS tables and used by dmideode command for provisioning

	Get FRUID Data (0x11)	
	Get SEL Info (0x40)	SEL commands are needed by BIOS to log critical events that are found during POST/runtime e.g. PCIe, DIMM ECC etc.
	Reserve SEL (0x42)	
	Get SEL Entry (0x43)	
	Add SEL Entry (0x44)	
	Clear SEL (0x47)	
	Get SEL UTC Offset(0x5C)	
Transport	Get LAN Config (0x02)	Needed by software agent on Host to read IP address of BMC; Supports only three parameters: "Lan Address Enables", "read IPv4", and "read IPv6 address"
	Get SOL Config (0x22)	Needed by BIOS to output console data on both ports during bootup
OEM (0x30)	Set DIMM Info (0x1C)	(OPTIONAL based on BIOS need) Inform BMC about DIMM info Request: Byte 1 – DIMM Index, 1 based Byte 2 – DIMM Type [7:6] Voltage type 00: Normal voltage(1.5V) 01: Ultra Low Voltage DIMM(1.25V) 10: Low voltage(1.35V) 11: DDR4 Normal voltage(1.35V) [5:0] DIMM type 0x00: SDRAM 0x01: DDR-1 RAM 0x02: Rambus 0x03: DDR-2 RAM 0x04: FBDIMM 0x05: DDR-3 RAM 0x06: DDR-4 RAM 0x3F – No DIMM present Byte 3..4 – DIMM speed in MHz, LS-byte first Byte 5..8 – DIMM size in Mbytes, LS-byte first

		Response: Byte 1 – Completion Code
	Get Board ID(0x37)	(OPTIONAL based on BIOS need) Get Board Information from BMC Request: None Response: Byte 1 – Completion Code Byte2 - Board SKU ID Byte3 - Board Revision ID Byte4 - MB Slot ID Byte5 - Slot Config ID
	Set POST Start (0x73)	Inform BMC when POST begins Request: None Response: Byte1 - Completion Code
	Set POST End (0x74)	Inform BMC when POST Ends Request: None Response: Byte1 - Completion Code
	Set PPIN(0x77)	Inform BMC about PPIN data of 8 bytes for each CPU Request: Byte 1:8 – CPU0 PPIN data Optional: Byte 9:16 – CPU1 PPIN data Response: Byte 1 – Completion Code
	ADR Trigger(0x7A)	BIOS to inform BMC to Trigger ADR low/high Request: Byte 1..3 - 0x00 0x9C 0x9C Byte 4 - Pull LOW/HIGH 0x00 - Pull Low 0x01 - Pull High

		Response: Byte 1 - Completion Code
	Set PPR Parameter (0x90)	DDR4 PPR/sPPR support;Need to repair DDR4 memory by using extra row as part of repair flow Request: Byte 1 - PPR Command Selector Byte 2..N - Configuration parameter data per table below Response: Byte 1 - Completion Code Command: PPR Action Selector: 0x01 Byte 1 - Enable or disable PPR function in next host system reboot Bit[7]: 0 for disable;1 for enable Bit[6..0] - 01h for Soft PPR; 02h for Hard PPR If there is no candidate row for repair: When get, response data is zero for PPR function When Set, response completion code: D5h for parameter not support in current state Command: PPR Candidate Row Count Selector: 0x2 Byte1 - PPR Candidate row count, range from 0 to 100 0h - No candidate row Command: PPR Candidate Row Address Selector: 0x3 Byte1 - Set Selector: PPR candidate row index, from 0 to ("PPR Candidate Row Count" - 1). Byte 2..7 - PPR candidate row address Byte 2: Bit[7:2] - Reserved Bit[1:0] - Logical Rank Byte 3: Bit[7:5] - Socket Number Bit[4:2] - Channel Number Bit[0:1] - DIMM Number Byte 4: Devicde ID Byte 5: Bit[7:4] - Bank Bit[3:0] - Bank Group Byte 6: Row (LSB) Byte 7: Row (MSB)

		Command: PPR History Data Selector: 0x04 Byte1: Set Selector: PPR history index, from 0 to ("PPR History Count" - 1) Byte 2..17: PPR history data Byte 2..5 - PPR Table Signature 0xbfeeeefb Byte 6 - Checksum Byte 7..10 - Seconds since Epoch for First PPR Entry Byte 11..14 - Seconds since Epoch for the Most Recent Entry Byte 15 - Number of Successful PPR (LSB) Byte 16 - Number of Successful PPR (MSB) Byte 17 - DIMM ID
	Get PPR Parameter (0x91)	DDR4 PPR/sPPR support;Need to repair DDR4 memory by using extra row as part of repair flow Request: Byte 1 - PPR Command Selector Byte 2 - Set Selector. Selects a given set of parameters under a parameter selector value. 00h if parameter does not require a set selector Response: Byte 1 - Completion Code Byte 2..N - Configuration parameter data per table below Command: PPR Action Selector: 0x01 Byte 1 - Enable or disable PPR function in next host system reboot Bit[7]: 0 for disable;1 for enable Bit[6..0] - 01h for Soft PPR; 02h for Hard PPR If there is no candidate row for repair: When get, response data is zero for PPR function When Set, response completion code: D5h for parameter not support in current state Command: PPR Candidate Row Count Selector: 0x2 Byte1 - PPR Candidate row count, range from 0 to 100

		0h - No candidate row Command: PPR Candidate Row Address Selector: 0x3 Byte1 - Set Selector: PPR candidate row index, from 0 to ("PPR Candidate Row Count" - 1). Byte 2..7 - PPR candidate row address Byte 2: Bit[7:2] - Reserved Bit[1:0] - Logical Rank Byte 3: Bit[7:5] - Socket Number Bit[4:2] - Channel Number Bit[0:1] - DIMM Number Byte 4: Devicde ID Byte 5: Bit[7:4] - Bank Bit[3:0] - Bank Group Byte 6: Row (LSB) Byte 7: Row (MSB) Command: PPR History Data Selector: 0x04 Byte1: Set Selector: PPR history index, from 0 to ("PPR History Count" - 1) Byte 2..17: PPR history data Byte 2..5 - PPR Table Signature 0xbfeeeefb Byte 6 - Checksum Byte 7..10 - Seconds since Epoch for First PPR Entry Byte 11..14 - Seconds since Epoch for the Most Recent Entry Byte 15 - Number of Successful PPR (LSB) Byte 16 - Number of Successful PPR (MSB) Byte 17 - DIMM ID
	Write Machine Configuration Information (0x6A)	BIOS to inform BMC about System configuration: #of CPUs, #of DIMMs etc..to help in calculation of Airflow Request: Byte 1 - Chassis Type 00h - ORv1 01h - ORv2 (Tioga pass) FFh - Unknown Byte 2 - Motherboard Type 00h - SS 01h - DS 02h - Type3

		03h:FEh - Reserved FFh - Unknown Byte 3 - Processor Count Byte 4 - Memory Count Byte 5 - 3.5" HDD Count (FFh - Unknown) (0 or 1 for Tioga Pass) Byte 6 - 2.5" HDD Count (FFh - Unknown) (0 in tioga pass) Byte 7 - Riser Type 00h - Not Installed 01h - 2-Slot Riser Card 02h - 3-Slot Riser Card FFh - Unknown Byte 8 - PCIe Card Location Bit 0 - SLOT1 (1h - Present; 0h - Absent) Bit 1 - SLOT2 (1h - Present; 0h - Absent) Bit 2 - SLOT3 (1h - Present; 0h - Absent) Bit 3 - SLOT4 (1h -Present; 0h - Absent) Byte 9 - SLOT 1 Byte 10 – SLOT 2 00h - Absent 01h - AVA (2 x m.2) 02h - AVA (3 x m.2) 03h - AVA (4 x m.2) 04h - Re-timer 05h - HBA 06h - Other flash cards (Intel, HGST) 80h - Unknown Byte 11 – SLOT 3 Same as Byte 10 Byte 12 – SLOT 4 Same as Byte 10 Byte 13 – NV Memory count Note: Bit 0 - presence of C1 [OCP Mezzanine Card] Bit 1 - presence of C2 [bottom card in 2(or 3)-slot riser] Bit 2 - presence of C3 [top card in 2-slot riser or middle slot in 3-slot riser] Bit 3 - presence of C4 [top card in 3-slot riser] Resonse: Byte 1 - Completion Code 00h - Normal C7h - Request data exceeds beyond data length C9h - Invalid data type of data length specified

		Bit 3 - Card 4 (1h -Present; 0h - Absent)
	Set BIOS Boot Order (0x52h)	BIOS to set boot order to BMC Request: Byte 1– Boot mode Bit 0 – 0 : Legacy, 1 : UEFI Bit 1 – CMOS clear (Optional, BIOS implementation dependent) Bit 6:2 – reserved Bit 7 – boot flags valid Byte 2-6– Boot sequence Bit 2:0 – boot device id 000b: USB device 001b: Network 010b: SATA HDD 011b: SATA-CDROM 100b: Other removable Device Bit 7:3 – reserve for boot device special request If Bit 2:0 is 001b (Network), Bit3 is IPv4/IPv6 order Bit3=0b: IPv4 first Bit3=1b: IPv6 first Response: Byte1 - Completion Code
	Get BIOS Boot Order (0x53h)	BIOS to get boot order from BMC Request: None Response: Byte 1 – Completion Code Byte 2– Boot mode Bit 0 – 0 : Legacy, 1 : UEFI Bit 1 – CMOS clear (Optional, BIOS implementation dependent) Bit 6:2 – reserved Bit 7 – boot flags valid Byte 3-7– Boot sequence Bit 2:0 – boot device id 000b: USB device 001b: Network 010b: SATA HDD 011b: SATA-CDROM 100b: Other removable Device Bit 7:3 – reserve for boot device special request If Bit 2:0 is 001b (Network), Bit3 is IPv4/IPv6

		order Bit3=0b: IPv4 first Bit3=1b: IPv6 first
	Get FLASH Info (0x55h)	Request: None Response: Byte 1 – Completion Code Byte 2:4 – Flash Device ID (*Note1) Byte 5:6 – Flash Status Register (*Note2) *Note1: MX25L12873FM2I-10G device id is 0x1820C2 W25Q128FVFIQ device id is 0x1840EF *Note2: MX25L12873FM2I-10G status register: BIT7 BIT6 BIT5 BIT4 BIT3 BIT2 BIT1 BIT0 SRWD QE BP3 BP2 BP1 BP0 WEL WIP W25Q128FVFIQ status register: BIT7 BIT6 BIT5 BIT4 BIT3 BIT2 BIT1 BIT0 SRP0 SEC TB BP2 BP1 BP0 WEL BUSY BIT15 BIT14 BIT13 BIT12 BIT11 BIT10 BIT9 BIT8 SUS CMP LB3 LB2 LB1 (R) QE SRP1
	Get PCIe Configuration (0xF4)	Request: Response: Byte 1 - Completion Code Byte 2 - Configuration Numbe 0x00: Empty/Unknown 0x01: Glacier Point 0x0F: Crane Flat (1) For the configuration of 4x Twin Lakes: BMC replies “0x00” (2) For the configuration of 2x GP + 2x Twin Lakes: BMC (slot2 and slot4) replies “0x01” (3) For the configuration of 2x CF + 2x Twin Lakes: BMC (slot2 and slot4) replies “0x0F”
OEM (0x36)	Set Processor Information (0x10)	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZ h, LSB first Byte 4 – Processor Index, 0 base

		Byte 5 – Parameter Selector Byte 6..N – Configuration parameter data (see below for Parameters of Processor Information) **Response:** Byte 1 – Completion code **Parameter#1: (Processor Product Name)** Byte 1..48 –Product name(ASCII code) Ex. Intel(R) Xeon(R) CPU E5-2685 v3 @ 2.60GHz **Param#2: Processor Basic Information** Byte 1 – Core Number Byte 2 – Thread Number (LSB) Byte 3 – Thread Number (MSB) Byte 4 – Processor frequency in MHz (LSB) Byte 5 – Processor frequency in MHz (MSB) Byte 6..7 – Revision
	Get Processor Information (0x11)	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZ h, LSB first Byte 4 – Processor Index, 0 base Byte 5 – Parameter Selector **Response:** Byte 1 – Completion code Byte 2..N – Configuration Parameter Data (see below for Parameters of Processor Information) **Parameter#1: (Processor Product Name)** Byte 1..48 –Product name(ASCII code) Ex. Intel(R) Xeon(R) CPU E5-2685 v3 @ 2.60GHz **Param#2: Processor Basic Information** Byte 1 – Core Number Byte 2 – Thread Number (LSB) Byte 3 – Thread Number (MSB) Byte 4 – Processor frequency in MHz (LSB) Byte 5 – Processor frequency in MHz (MSB) Byte 6..7 – Revision
	Set DIMM Information (0x12)	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZh, LSB first Byte 4 – DIMM Index, 0 base Byte 5 – Parameter Selector Byte 6..N – Configuration parameter data (see below for Parameters of DIMM Information) **Response:** Byte 1 – Completion code

		Param#1 (DIMM Location): Byte 1 – DIMM Present Byte 1 – DIMM Present 01h – Present FFh – Not Present Byte 2 – Node Number, 0 base Byte 3 – Channel Number , 0 base Byte 4 – DIMM Number , 0 base **Param#2 (DIMM Type):** Byte 1 – DIMM Type Bit [7:6] For DDR3 00 – Normal Voltage (1.5V) 01 – Ultra Low Voltage (1.25V) 10 – Low Voltage (1.35V) 11 – Reserved For DDR4 00 – Reserved 01 – Reserved 10 – Reserved 11 – Normal Voltage (1.2V) Bit [5:0] 0x00 – SDRAM 0x01 – DDR-1 RAM 0x02 – Rambus 0x03 – DDR-2 RAM 0x04 – FBDIMM 0x05 –DDR-3 RAM 0x06 –DDR-4 RAM **Param#3 (DIMM Speed):** Byte 1..2 – DIMM speed in MHz, LSB Byte 3..6 – DIMM size in Mbytes, LSB **Param#4 (Module Part Number):** Byte 1..20 –Module Part Number (JEDEC Standard No. 21-C) **Param#5 (Module Serial Number):** Byte 1..4 –Module Serial Number (JEDEC Standard No. 21-C) **Param#6 (Module Manufacturer ID):** Byte 1 - Module Manufacturer ID, LSB Byte 2 - Module Manufacturer ID, MSB

	Get DIMM Information (0x13)	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZh, LSB first Byte 4 –DIMM Index, 0 base Byte 5 – Parameter Selector **Response:** Byte 1 – Completion code Byte 2..N – Configuration Parameter Data (see Table_1213h Parameters of DIMM Information) **Param#1 (DIMM Location):** Byte 1 – DIMM Present Byte 1 – DIMM Present 01h – Present FFh – Not Present Byte 2 – Node Number, 0 base Byte 3 – Channel Number , 0 base Byte 4 – DIMM Number , 0 base **Param#2 (DIMM Type):** Byte 1 – DIMM Type Bit [7:6] For DDR3 00 – Normal Voltage (1.5V) 01 – Ultra Low Voltage (1.25V) 10 – Low Voltage (1.35V) 11 – Reserved For DDR4 00 – Reserved 01 – Reserved 10 – Reserved 11 – Normal Voltage (1.2V) Bit [5:0] 0x00 – SDRAM 0x01 – DDR-1 RAM 0x02 – Rambus 0x03 – DDR-2 RAM 0x04 – FBDIMM 0x05 –DDR-3 RAM 0x06 –DDR-4 RAM **Param#3 (DIMM Speed):** Byte 1..2 – DIMM speed in MHz, LSB Byte 3..6 – DIMM size in Mbytes, LSB **Param#4 (Module Part Number):** Byte 1..20 –Module Part Number (JEDEC Standard No. 21-C)

		Param#5 (Module Serial Number): Byte 1..4 –Module Serial Number (JEDEC Standard No. 21-C) **Param#6 (Module Manufacturer ID):** Byte 1 - Module Manufacturer ID, LSB Byte 2 - Module Manufacturer ID, MSB

9.14 BMC FW chip and Firmware Update

The BMC FW flash chip should use BMC's SPI interface. The vendor is responsible for selecting a specific flash chip, which should fit the required functionality in addition to potential additional features that may be required in the future. 32Mbyte size is recommended.

Vendors should provide tool(s) to implement remote BMC firmware update, which will not require any physical access to the system. Remote update means either through management network or through In-Band by logging into local OS (CentOS) with data network. Tool(s) shall support CentOS 7.x 64-bit with updated Kernel specified by customer.

A remote BMC firmware update may take a maximum of 5 minutes to complete. I2C sideband has bottle neck to achieve this requirement and NC-SI interface is needed. BMC firmware update process and BMC reset process require no reboot or power down of host system and should have no impact to normal operation of host system. BMC need to be fully functional with updated firmware after the update and reset without further configuration.

Default update should recovery BMC to factory default; option need to be provided to preserve SEL, configuration. MAC address is based on NIC MAC, so it would not be cleared with BMC firmware update.

9.15 BMC Update PCH flash

Vendor should implement BMC to be able to access host system PCH flash and recovery a host system PCH flash from corruption in remote. The PCH flash stores BIOS and Intel® Server Platform Services Firmware code; both BIOS and Intel® Server Platform Services Firmware region should be updated. Dual PCH flash is an optional implementation in Tioga Pass.

9.15.1 Hardware Scheme

The hardware scheme contains one or two multiplexers controlled by BMC GPO. GPO_A controls PCH or BMC has access to flash. GPO_B controls primary or backup flash to be accessed by CS# signal if dual PCH flash design in. Primary flash is the default flash PCH is using to access BIOS and Intel® Server Platform Services Firmware code. Backup flash is stored with a known good image and usually is not modified. Backup flash can still be modified if needed by manual command line operations.

The rules below should be followed to avoid unexpected system behavior:

- Default status of GPO_A and GPO_B shall be [0,0] during system AC on or during BMC booting and reset to make sure PCH has access to Primary Flash by default.
- BMC shall check system power status and not to change GPO_A and GPO_B status when system is in S0.
- BMC shall set Intel® Manageability Engine to recovery mode, before changing GPO_A and GPO_B status from [0,0] to another other value.
- BMC shall set Intel® Manageability Engine to normal mode, after changing GPO_A and GPO_B status from other value to [0,0].

GPO_A	Flash Selection
0	PCH SPI / BIOS Flash
1	BMC SPI / BIOS Flash

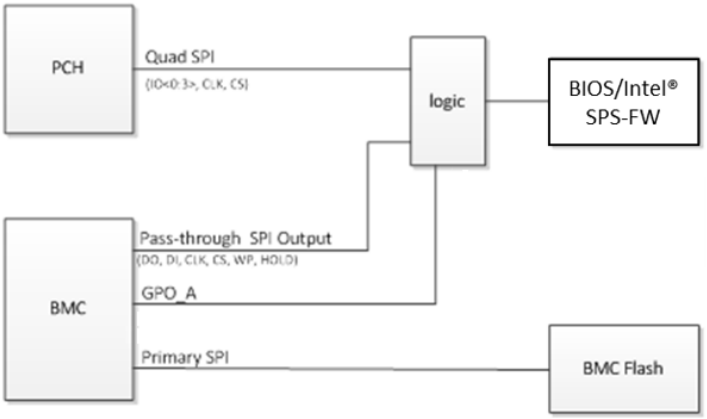


Figure 9-2 BMC and PCH flash diagram

GPO_A	GPO_B	Flash Selection
0	0	PCH SPI / BIOS Primary Flash
0	1	PCH SPI / BIOS Backup Flash
1	0	BMC SPI / BIOS Primary Flash
1	1	BMC SPI / BIOS Backup Flash

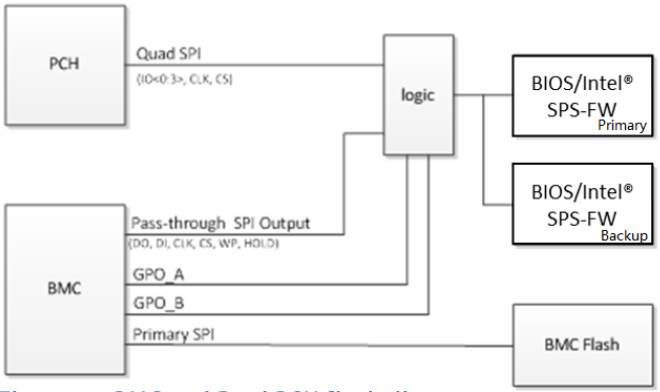


Figure 9-3 BMC and Dual PCH flash diagram

9.15.2 Software Scheme for Dual PCH flash

The following software features must be supported if dual bios hard scheme is designed:

1) Auto BIOS recovery from backup flash to primary flash:

When 1st boot up after BIOS update fails and the BMC watchdog timer (WDT) times out, the BMC determines that the BIOS needs to recover. It takes over control of flash access and duplicates backup flash content to primary flash. After that, BMC will return flash access to PCH and power cycle the system. This recovery attempt is only executed once after a failed BIOS update. BMC shall generate SEL for each recovery attempt.

2) Manual BIOS recovery from backup flash to primary flash :

User uses the IPMI OEM command to trigger BMC taking over control of flash access and duplicate backup flash content to primary flash.

3) Manual BIOS save from primary flash to backup flash:

User uses the IPMI OEM command to trigger BMC taking over control of flash access and duplicate primary flash to backup flash.

4) Manual Mux control:

User uses the IPMI OEM command to trigger the BMC to change the state of two multiplexers.

5) Manual BIOS recovery from network to flash

User uses the IPMI OEM command to trigger a PCH flash binary to be transferred to BMC and written to primary or backup flash, depending on the state of the multiplexer.

6) Enable/Disable Auto BIOS recovery

User to use IPMI OEM command to Enable and Disable feature of Auto BIOS recovery; default is Disable for EVT/DVT/PVT and Enable for MP.

7) Get dual BIOS feature status:

User uses the IPMI OEM command to get the current state of two multiplexers, and auto recovery Enable/Disable status.

8) Get dual BIOS version:

User uses the IPMI OEM command to access and return BIOS version of both primary and secondary flash.

Vendor shall provide an update utility that supports CentOS 7.x 64-bit with updated Kernel specified by customer.

9.16 BMC Update and Access CPLD

Vendor should implement BMC to access motherboard CPLD through JTAG and perform CPLD code upload from remote control server to BMC, update code from BMC to CPLD through CPLD JTAG interface, and verify CPLD code. All the steps above shall be done from BMC command line mode. Vendor shall provide an update utility that supports CentOS 7.x 64-bit with updated Kernel specified by customer.

BMC shall implement OEM command to read CPLD FW checksum, device ID, and FW version.

9.17 BMC Time Sync(change follow BMC feature list)

During BMC initialization:

- If there is NTP server, BMC should sync up BMC RTC with NTP server.
- Otherwise, BMC should sync up BMC RTC with system's RTC. BIOS should not overwrite System RTC from BMC(Caused SEV);
    - For BIOS, Do not use 'Get SEL Time' to update System RTC
    - BMC to get System RTC after POST completes
    - No need to have a separate process to sync BMC time with ME periodically
    - Do not respond to 'Get SEL Time' Command from Host to avoid corrupting System RTC

9.18 PCIe and Mezzanine card Thermal monitoring

BMC should implement thermal monitoring feature for PCIe card on riser, and Mezzanine card. BMC reads the temperature of key components of PCIe and Mezzanine cards through its SMBus ports in the format as TMP421 temperature sensor. BMC uses the temperature reading in FSC and sensor reporting.

Refer to "*Add-on-Card Thermal Interface Spec for Facebook Servers*" for detail requirement and implementation.

9.19 BMC PPIN Implementation

BMC shall send SMI# to PCH after BMC reset; BIOS shall serve the request by using Set_PPIN (OEM command) to write PPIN of CPU0 and CPU1 to BMC. BIOS also Set_PPIN when BIOS POST COMPLETE

BMC shall return the PPIN to user with Get_PPIN (OEM command).

9.20 BMC Average Power Reporting

Intel® Manageability Engine shall refresh READ_EIN_EXT(register) every 0.1 seconds. BMC shall get and record EIN_EXT value from Intel® Manageability Engine through SMBUS every 0.1 seconds in a ring buffer. The ring buffer size shall support any duration from 0.1 second to 60 seconds, in 0.1 second increments. BMC shall support the command such as Get_PIN to calculate and report the average power between current and a duration defined in the Get_PIN command. The return shall have a resolution of 0.1W.

9.21 BMC Access and Update VR Firmware

Vendor should implement BMC to access motherboard CPU VCCIN/VCCIO, PCH PVNN/P1V05 and memory VDDQ VR controllers' firmware.

Vendor shall implement a script or utility through BMC to perform a VR firmware code update and verify it from the BMC command line mode. The script or utility shall support CentOS 7.x 64-bit with updated Kernel specified by customer.

During the VR firmware update, BMC and/or the update script shall stop the sensor polling to related VR. The VR firmware upgrade can be performed in S5. The capability to have a VR firmware upgrade in S0 is optional, and a DC cycle is allowed after a VR firmware upgrade in S0 to activate the new VR firmware.

Vendor shall implement OEM command to read VR FW version.

9.22 BMC MSR Dump

Vendor shall implement BMC to dump MSR from both CPUs through Intel® Manageability Engine and PECI. Vendor shall provide utility that supports CentOS 7.x 64-bit with updated Kernel specified by customer.

This is a debug feature to allow the user to access critical debug information from faulty SUT on a server rack, without removing the system from a failure status and risking the loss of critical debug information.

BMC firmware shall apply MSR dump automatically when there is IERR or MCERR. The BMC shall store the MSR dump in BMC flash. During the dump, the BMC shall reject chassis power related commands to avoid interrupting the dump.

9.23 BMC Verified boot

Vendor shall implement hardware and Firmware to support BMC verified boot.

9.23.1 BMC Verified boot hardware

BMC verified boot hardware is shown in diagram below.

Tioga Pass has one TPM module with 2x independent TPM ICs on the module. SLB9645 TPM1.2 is for BMC and SLB9670 TPM2.0 is for system firmware.

Pin 8 has a special implementation to allow it being used as both Physical Presence pin and reset pin for SLB9645.

The design is based on a few key points:

- BMC verified boot SPI flash device(same size as BMC FW flash 32MB) can be set to read only by hardware strapping and SPI configure command in secure flash code if needed
- Design allows hardware strapping (WP#) being over written in EVT/DVT when needed
- The image in BMC Firmware flash shall not have privilege to perform any write/erase operation on verified boot SPI flash device
- The image in BMC Firmware flash shall not have privilege to reset SLB9645 hardware without reset BMC processor itself and trigger code loading from CS0 with verified Boot image

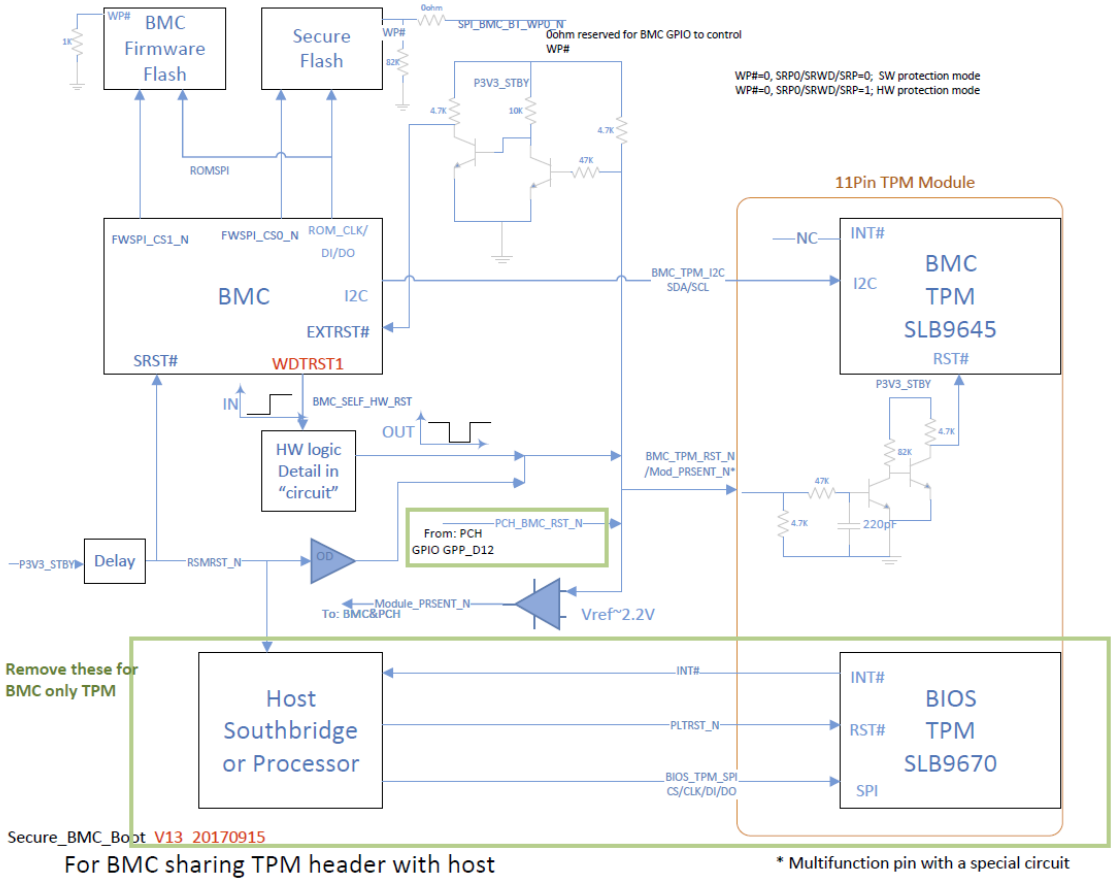


Figure 9-4 BMC verified boot hardware

9.23.2 BMC verified boot Firmware flow

BMC verified boot firmware flow is as diagram below.

BMC verified boot firmware shall utilize hardware hash and crypto engine in AST2500 for SHA-256 or RSA4096 hash generation to shorten the BMC booting time.

Vendor shall provide data to compare the decrypt time with software RSA-4064 algorithm and hardware accelerated RSA-4064. Facebook considers to reduce the requirement from RST-4096 to RSA-4064 based on the data saved by hardware accelerated decryption.

For below Firmware flow:

- SPI0.0 is the Secure flash, or ROM;
- SPI0.1 is the field-upgradable flash, containing U-boot, kernel, rootfs, etc.
- Both SPI chips contain the same layout and content. A BMC build that supports verified boot may be placed into either location.

```
P0.0 0x0000:0000 - 0x0001:5000: 64kB   ROM, KEK RSA public key
P0.1 0x0001:5000 - 0x0006:0000: 320kB  recovery U-Boot
P1   0x0006:0000 - 0x0008:0000: 128kB  environment (updated whenever)
P2.0 0x0008:0000 - 0x0008:4000: 16kB   signed firmware FIT for U-Boot and signed subordinate keys
P2.1 0x0008:4000 - 0x000E:0000: 368kB  U-Boot
P3   0x000E:0000 - 0x0148:0000: 19.6M  Signed FIT for kernel, rootfs, and device tree
P4   0x0148:0000 - 0x0200:0000: 11.5M  data
```

Figure 9-5 Flash Chips Symmetrical Layout

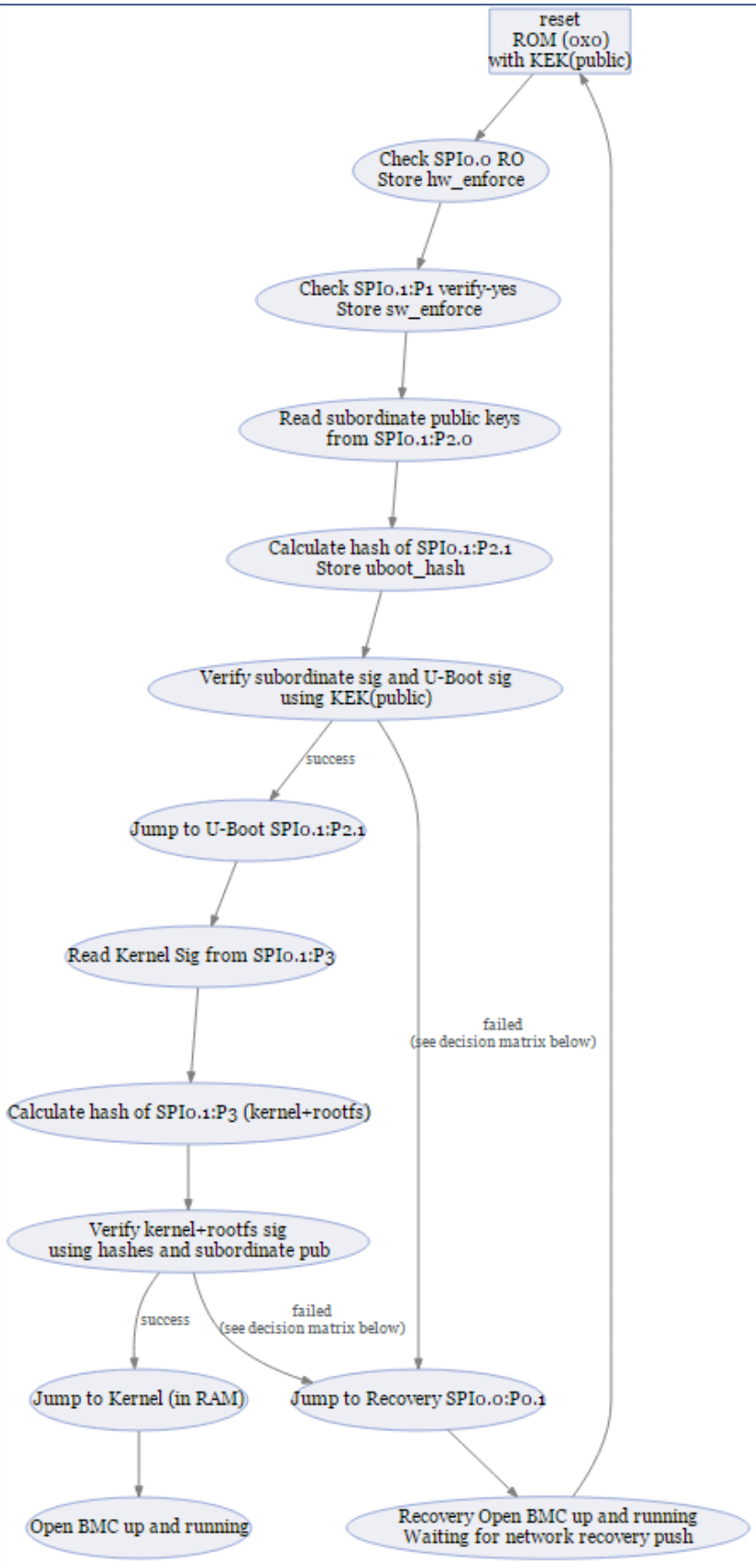

reset
ROM (0x0)
with KEK(public)
Check SPI0.0 RO
Store hw_enforce
Check SPI0.1:P1 verify-yes
Store sw_enforce
Read subordinate public keys
from SPI0.1:P2.0
Calculate hash of SPI0.1:P2.1
Store uboot_hash
Verify subordinate sig and U-Boot sig
using KEK(public)
success
failed
(see decision matrix below)
Jump to U-Boot SPI0.1:P2.1
Read Kernel Sig from SPI0.1:P3
Calculate hash of SPI0.1:P3 (kernel+rootfs)
Verify kernel+rootfs sig
using hashes and subordinate pub
success
failed
(see decision matrix below)
Jump to Kernel (in RAM)
Jump to Recovery SPI0.0:P0.1
Open BMC up and running
Recovery Open BMC up and running
Waiting for network recovery push

Figure 9-6 BMC Verified boot firmware flow

9.23.3 Key Management

Facebook generates FB.pub.key/FB.pri.key key pair-0 to sign all the subordinate key pairs generated for BMC FW signing.

Facebook generates FB.pub.key/FB.pri.key key pair-1 to sign BMC FW. FB.pri.key is used to sign the Hash of BMC FW Uboot/Kernel+rootfs. FB.pub.key is used to verify the signature.

Facebook will generate FB.pub.key/FB.pri.key key pair-2/3 to ODMs to build debug BMC FW during development stage.

9.24 Intel® At Scale Debug

Traditional Intel® In-Target Probe (ITP) and DCI requires local HW and close-proximity technician for processor or chipset related debug. The new feature Intel® At Scale Debug is to address the proximity problem with BMC to JTAG connection over a networked path. It is assumed BMC manageability network is secured.

The Below diagram shows the connectivity and flow.

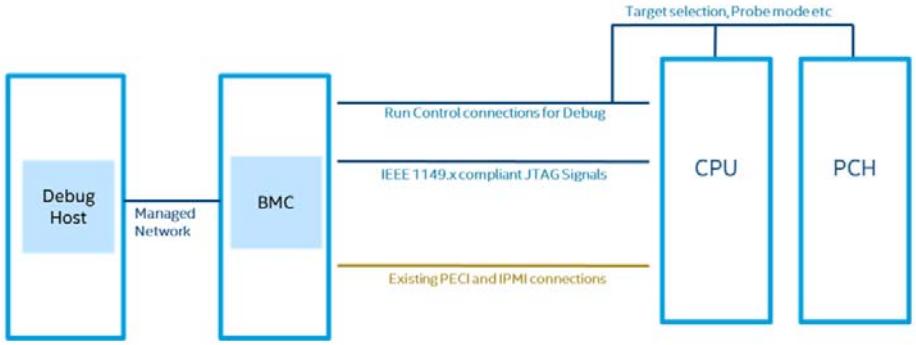


Figure 9-7 Intel® At Scale Debug connectivity and flow

Debug host based Intel® debug software utilizes data center management network transport to communicate with BMC.

Vendor needs to have NDA with Intel® to get the design guidance/whitepaper.

10 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. CPU or memory 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 0.001 inch H_2O datacenter pressure with one FAN (or one rotor) in each thermal zone failed

10.1 Data Center Environmental Conditions

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

10.1.1 Location of Data Center/Altitude

Data centers may be located 6000 feet 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.

10.1.2 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 3°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.

10.1.3 Cold-Aisle Pressurization

Data centers will maintain the cold aisle pressure to be between 0" H_2O and 0.005" H_2O. The thermal solution of the system accommodates the worst-case operational pressurization in the data centers, which is 0" H2O with no fan failures and 0.001" H2O with a single fan (or rotor) failure.

10.1.4 R.H

Most data centers will maintain relative humidity between 20% and 90%. The thermal solution must sustain uninterrupted operation of the system across the aforementioned RH range.

10.2 Server operational condition

10.2.1 System loading

The power consumptions of individual components on the system motherboard will vary by application or by motherboard SKU. The total power consumption of system also may vary with use or with the number of existence of PCIe cards on the system. Please see summary below.

- System loading: idle to 100%
- Number of PCIe full height or half height cards can be installed: 0 to 3
- Number of PCIe Mezz card can be installed: 0 to 1
- Number of 3.5" HDD: 0 to 1

Plan of record worst case configuration for thermal and power delivery design is 2 x 165W TDP CPU with 24 x 128G DDR4 DIMM, or 12x 64G DDR4 DIMM + 12x 512G NVDIMM.

A unified thermal solution that can cover up to 100% system loading is preferred. However, an ODM can propose non-unified thermal solution if there is alternative way to provide cost benefits. At least the air-duct design should be unified for all SKU.

10.2.2 DDR DIMM DRAM Operation

Thermal design should meet DIMM max operating temperature as 85°C with single refresh rate. Thermal test should be done based on a DIMM module's AVL (Approved Vendor List). The vendor should implement BIOS and memory subsystem to have optimized refresh rate and utilize optional DIMM Auto-Self-Refresh (ASR) based on DIMM temperature. The implementation should follow updated DDR4 memory controller and DIMM vendor's specification.

10.2.3 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. CPU throttling is not allowed over the validation range 0C – 35C.

10.2.4 Pressurization

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

10.2.5 Fan Redundancy

The server fans at N+1 redundancy should be sufficient for cooling server components to temperatures below their maximum spec to prevent server shut down or to prevent either CPU or memory throttling.

10.2.6 System airflow or volumetric flow

The unit of airflow (or volumetric flow) used for this spec is CFM (cubic feet per minute). The maximum allowable airflow per watt in the system must be 0.107. The desired airflow per watt is 0.1 or lower in the system at the mean temperature (plus or minus standard deviation). See section 10.1.2 for the temperature definitions.

10.2.7 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 desired rack-level delta T must be greater than 13.9°C (25°F). The desired server-level delta T is 17°C (31°F) when the inlet air temperature to the system is equal to or lower than 30°C.

10.2.8 Thermal margin

The thermal margin is the difference between the maximum theoretical safe temperature and the actual temperature. The board design operates at an inlet temperature of 35°C (95°F) outside of the system with a minimum 2% thermal margin for

every component in the system. Otherwise, the thermal margin for every component in the system is at least 7% for temperatures up to 30°C.

10.3 Thermal kit requirements

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

10.3.1 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 number of heat pipes in the heat sink should not be more than three. 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.

10.3.2 System Fan

The system fan must be highly power-efficient with dual bearing. The propagation of vibration cause 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. The maximum thickness of fan should not be greater than 38mm. Each rotor in the fan should have maximum 5 wires. 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.

System fan should not have back rush current in all condition. System fan should have an inrush current of less than 1A on 12V per fan. When there is a step change on fan PWM signal from low PWM to high PWM, there should be less than 10% of overshoot or no overshoot for fan input current. System should stay within its power envelope in all condition of fan operation.

10.3.3 Air-Duct

The air duct needs to be part of the motherboard tray cover, and must be most energy efficient design. The air-duct design should be simple and easily serviceable. The air-duct design should be unified for all SKUs. Using highly green material or reusable material for the air duct is preferred.

10.3.4 Thermal sensor

The maximum allowable tolerance of thermal sensors in the motherboard is ±3°C.

11 I/O System

This section describes the motherboard I/O requirements.

11.1 PCIe x32 Slot/Riser Card

11.1.1 Riser slot interface between riser card and motherboard

The motherboard has a single, 2x sockets to be used by PCIe riser cards:

- A x24 slot Samtec/HSEC8-1100-01-L-DV-A-K(200pin) is used for x24 PCIe and power delivery
- A x8 slot Samtec/HSEC8-130-01-L-DV-A-TR(60pin) is used for x8 PCIe, and USB signal

Slot location must follow mechanical requirement that will be delivered in DXF format.

All PCIe lanes to x32 PCIe is from CPU0. SMBus to PCH connects to PCH host SMBUS; SMBALT_N_PCH from riser to PCH connects to PCH GPP_C2_SMBALERT_N(pin BY27).

1	GND	rsvd	2	
3	PETp(16)	GND	4	
5	PETn(16)	PERp(16)	6	**PE2C lane 0**
7	GND	PERn(16)	8	
9	PETp(17)	GND	10	
11	PETn(17)	PERp(17)	12	
13	GND	PERn(17)	14	
15	PETp(18)	GND	16	
17	PETn(18)	PERp(18)	18	
19	GND	PERn(18)	20	
21	PETp(19)	GND	22	
23	PETn(19)	PERp(19)	24	
25	GND	PERn(19)	26	
27	PETp(20)	GND	28	
29	PETn(20)	PERp(20)	30	
31	GND	PERn(20)	32	
33	PETp(21)	GND	34	
35	PETn(21)	PERp(21)	36	
37	GND	PERn(21)	38	
39	PETp(22)	GND	40	
41	PETn(22)	PERp(22)	42	
43	GND	PERn(22)	44	
45	PETp(23)	GND	46	
47	PETn(23)	PERp(23)	48	
49	GND	PERn(23)	50	**PE2C lane 7**
51	JTAG_TDI	GND	52	
53	JTAG_TDO	JTAG_TMS	54	
55	GND	JTAG_TCK	56	
57	USB2.0_N	GND	58	
59	USB2.0_P	JTAG_TRST	60	

Figure 11-1 X8 PCIe slot pin definition

1	12V	12V	2
3	12V	12V	4
5	12V	12V	6
7	12V	12V	8
9	12V	12V	10
11	12V	12V	12
13	GND	GND	14
15	SMDAT_BMC	SMALT#	16
17	SMCLK_BMC	SLT_CFG1	18
19	SLT_CFG0	PWR_BREAK#	20
21	+3.3v	+3.3v	22
23	+3.3v	+3.3v	24
25	3.3Vaux	GND	26
27	SMBALT_M_PCH	WAKE#	28
29	PERST#	SMBDAT_PCH	30
31	GND	SMBCLK_PCH	32
33	REFCLK1+	GND	34
35	REFCLK1-	REFCLK2+	36
37	GND	REFCLK2-	38
39	REFCLK3+	GND	40
41	REFCLK3-	REFCLK4+	42
43	GND	REFCLK4-	44
45	REFCLK5+	GND	46
47	REFCLK5-	REFCLK6+	48
49	GND	REFCLK6-	50
51	PETp(0)	GND	52
53	PETn(0)	PERp(0)	54
55	GND	PERn(0)	56
57	PETp(1)	GND	58
59	PETn(1)	PERp(1)	60
61	GND	PERn(1)	62
63	RSVD	GND	64
	KEY	KEY	
65	GND	RSVD	66
67	PETp(2)	GND	68
69	PETn(2)	PERp(2)	70
71	GND	PERn(2)	72
73	PETp(3)	GND	74
75	PETn(3)	PERp(3)	76
77	GND	PERn(3)	78
79	PETp(4)	GND	80
81	PETn(4)	PERp(4)	82
83	GND	PERn(4)	84
85	PETp(5)	GND	86
87	PETn(5)	PERp(5)	88
89	GND	PERn(5)	90
91	PETp(6)	GND	92
93	PETn(6)	PERp(6)	94
95	GND	PERn(6)	96
97	PETp(7)	GND	98
99	PETn(7)	PERp(7)	100

PE1A lane 15

101	GND	PERn(7)	102
103	PETp(8)	GND	104
105	PETn(8)	PERp(8)	106
107	GND	PERn(8)	108
109	PETp(9)	GND	110
111	PETn(9)	PERp(9)	112
113	GND	PERn(9)	114
115	PETp(10)	GND	116
117	PETn(10)	PERp(10)	118
119	GND	PERn(10)	120
121	PETp(11)	GND	122
123	PETn(11)	PERp(11)	124
125	GND	PERn(11)	126
127	PETp(12)	GND	128
129	PETn(12)	PERp(12)	130
131	GND	PERn(12)	132
133	PETp(13)	GND	134
135	PETn(13)	PERp(13)	136
137	GND	PERn(13)	138
139	PETp(14)	GND	140
141	PETn(14)	PERp(14)	142
143	GND	PERn(14)	144
145	PETp(15)	GND	146
147	PETn(15)	PERp(15)	148
149	GND	PERn(15)	150
151	PETp(16)	GND	152
153	PETn(16)	PERp(16)	154
155	GND	PERn(16)	156
157	PETp(17)	GND	158
159	PETn(17)	PERp(17)	160
161	GND	PERn(17)	162
163	PETp(18)	GND	164
165	PETn(18)	PERp(18)	166
167	GND	PERn(18)	168
169	PETp(19)	GND	170
171	PETn(19)	PERp(19)	172
173	GND	PERn(19)	174
175	PETp(20)	GND	176
177	PETn(20)	PERp(20)	178
179	GND	PERn(20)	180
181	PETp(21)	GND	182
183	PETn(21)	PERp(21)	184
185	GND	PERn(21)	186
187	PETp(22)	GND	188
189	PETn(22)	PERp(22)	190
191	GND	PERn(22)	192
193	PETp(23)	GND	194
195	PETn(23)	PERp(23)	196
197	GND	PERn(23)	198
199	RISER_PRSNT	GND	200

PE1A lane 0

PE2A lane 0

PE2A lane 7

Figure 11-2 X16 PCIe slot pin definition

11.1.2 Riser card types

Vendor shall enable 3x riser cards below for single side and double side use cases.

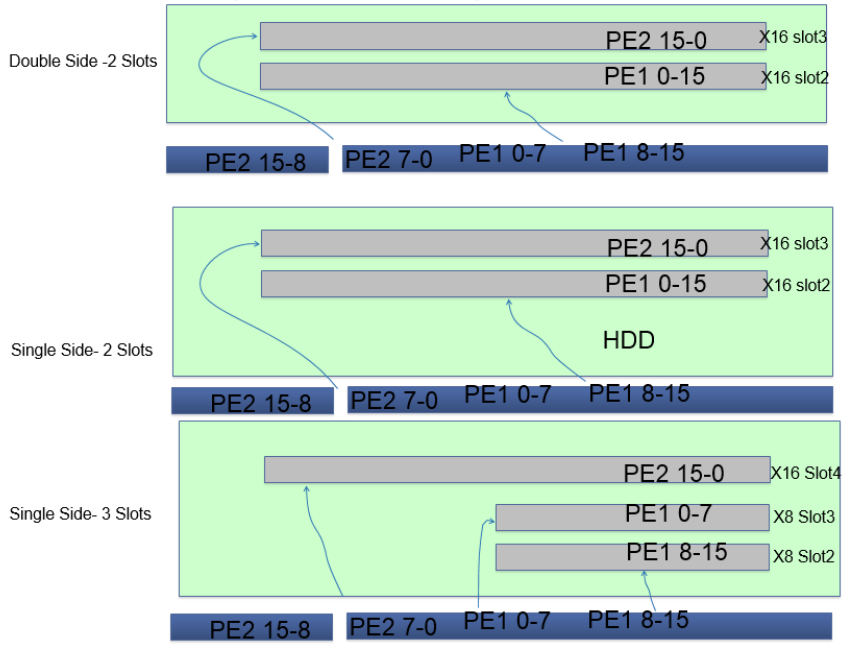


Figure 11-3 Riser Card PCIe Lanes Allocation

Please follow Table 13-1 to Table 13-5 to use reserved pins on PCIe x8 or PCIe x16 slot 2(low), slot3(middle) and slot 4(high) on 2 slot or 3 slot riser card, and follow 12-6 to set configuration pins on riser card, to indicate which riser on motherboard.

Table 11-1 2 slot PCIe x16 Slot 2 (low) Reserved Pin Usage on Single Side 2 slots Riser

Pin	Pin Defined	Description
A7	SMB_PCH_SCL	PCH SMBUS Clock
A8	SMB_PCH_SDA	PCH SMBUS Data
A19	RSVD	Reserve
A32	USB2.0_P1	USB2.0 Port from Hub port 1
A33	USB2.0_N1	USB2.0 Port from Hub port 1
B12	SMB_ALERT_N	SMBUS Alert from PCIe card to system, OD, low active
B17	SLOT2_PRSNT-1	SLOT2 PRESENT1
B30	PWR_BRKN_N	Power Brake from system to PCIe card, OD, low active
B31	SLOT2_PRSNT-2	SLOT2 PRESENT2
B48	SLOT2_PRSNT-3	SLOT2 PRESENT3
B81	SLOT2_PRSNT-4	SLOT2 PRESENT4

Table 11-2 2 slot PCIe x16 Slot 3 (high) Reserved Pin Usage on Single Side 2 slots Riser

Pin	Pin Defined	Description
A7	SMB_PCH_SCL	PCH SMBUS Clock
A8	SMB_PCH_SDA	PCH SMBUS Data
A19	RSVD	Reserve
A32	USB2.0_P2	USB2.0 Port from Hub port 2
A33	USB2.0_N2	USB2.0 Port from Hub port 2
B12	SMB_ALERT_N	SMBUS Alert from PCIe card to system, OD, low active
B17	SLOT3_PRSNT-1	SLOT3 PRESENT1
B30	PWR_BRKN_N	Power Brake from system to PCIe card, OD, low active
B31	SLOT3_PRSNT-2	SLOT3 PRESENT2

B48	SLOT3_PRSNT-3	SLOT3 PRESENT3
B81	SLOT3_PRSNT-4	SLOT3 PRESENT4

Table 11-3 3 slot PCIe x8 Slot 2 (low) Reserved Pin Usage on Single Side 3 slots Riser

Pin	Pin Defined	Description
A7	SMB_PCH_SCL	PCH SMBUS Clock
A8	SMB_PCH_SDA	PCH SMBUS Data
A19	RSVD	Reserve
A32	USB2.0_P1	USB2.0 Port from Hub port 1
A33	USB2.0_N1	USB2.0 Port from Hub port 1
B12	SMB_ALERT_N	SMBUS Alert from PCIe card to system, OD, low active
B17	SLOT2_PRSNT-1	SLOT2 PRESENT1
B30	PWR_BRKN_N	Power Brake from system to PCIe card, OD, low active
B31	SLOT2_PRSNT-2	SLOT2 PRESENT2
B48	SLOT2_PRSNT-3	SLOT2 PRESENT3

Table 11-4 3 slot PCIe x8 Slot 3 (middle) Reserved Pin Usage on Single Side 3 slots Riser

Pin	Pin Defined	Description
A7	SMB_PCH_SCL	PCH SMBUS Clock
A8	SMB_PCH_SDA	PCH SMBUS Data
A19	RSVD	Reserve
A32	USB2.0_P2	USB2.0 Port from Hub port 2
A33	USB2.0_N2	USB2.0 Port from Hub port 2
B12	SMB_ALERT_N	SMBUS Alert from PCIe card to system, OD, low active
B17	SLOT3_PRSNT-1	SLOT3 PRESENT1
B30	PWR_BRKN_N	Power Brake from system to PCIe card, OD, low active
B31	SLOT3_PRSNT-2	SLOT3 PRESENT2
B48	SLOT3_PRSNT-3	SLOT3 PRESENT3

Table 11-5 3 slot PCIe x16 Slot 4 (high) Reserved Pin Usage on Single Side 3 slots Riser

Pin	Pin Defined	Description
A7	SMB_PCH_SCL	PCH SMBUS Clock
A8	SMB_PCH_SDA	PCH SMBUS Data
A19	RSVD	Reserve
A32	USB2.0_P3	USB2.0 Port from Hub port 3
A33	USB2.0_N3	USB2.0 Port from Hub port 3
B12	SMB_ALERT_N	SMBUS Alert from PCIe card to system, OD, low active
B17	SLOT_PRSNT4-1	SLOT PRESENT1
B30	PWR_BRKN_N	Power Brake from system to PCIe card, OD, low active
B31	SLOT4_PRSNT-2	SLOT4 PRESENT2
B48	SLOT4_PRSNT-3	SLOT4 PRESENT3
B81	SLOT4_PRSNT-4	SLOT4 PRESENT4

Table 11-6 SLOT CONFIG DVT Definition

SLOT_CONFIG0	SLOT_CONFIG1	Riser Type
0	0	Single Side 2 slots riser 2x 16 slots
0	1	Single Side 3 slots riser

		3 x 8 slots
1	0	Double side 2 slots riser 2 x16 slots
1	1	RSVD

Riser card should implement SMBus Mux to avoid address confliction of PCIe cards. Follow Figure 13-4 , and for implementation and slave address assignment in 8 bit.

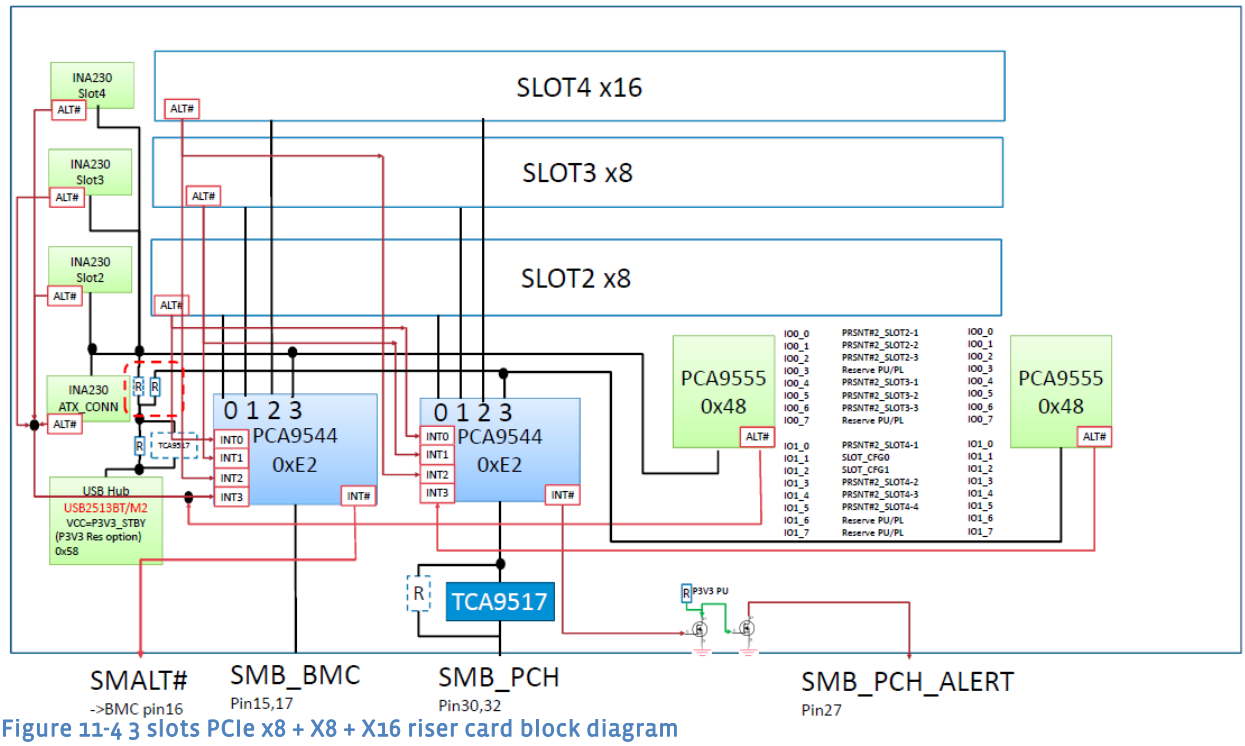


Figure 11-4 3 slots PCIe x8 + X8 + X16 riser card block diagram

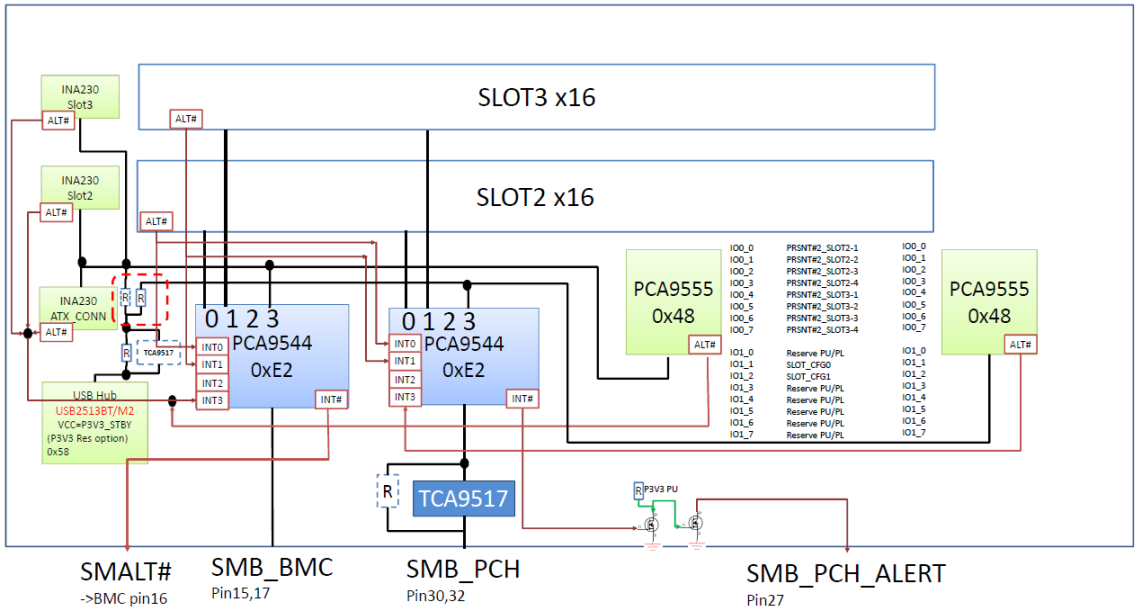


Figure 11-5 2 slots PCIe X16 riser card block diagram

Riser card should implement I2C to GPIO expander (PCA9555) to be accessed by both BIOS and BMC on motherboard, to tell what AICs are currently on each slot. BIOS shall follow below table 12-7 and table 12-8 to do PCIe bifurcation accordingly. Vendor should follow the SMBUS addresses defined in the diagram to avoid address conflict. The addresses are defined as 8-bit address.

Table 13-7 X16 PCIe card bios bifurcation table

	Normal PCIe Card			M.2 Carrier	Retimer			Empty
X16 PCIe slot PRSNT#	1X16	1X8	1X4	4X4 M.2	4X4	2X8	1X16	1X16
SLOT_PRSNT_N_1 (b17)	1	1	1	1	1	1	1	1
SLOT_PRSNT_N_2 (b31)	1	1	0	0	0	1	0	1
SLOT_PRSNT_N_3 (b48)	1	0	1	1	0	0	0	1
SLOT_PRSNT_N_4 (b81)	0	1	1	0	0	0	1	1

Table 13-8 X8 PCIe card bios bifurcation table

	Normal PCIe Card		M.2 Carrier	Empty
X8 PCIe slot PRSNT#	1X8	1X4	2X4 M.2	1X16
SLOT_PRSNT_N_1 (b17)	1	1	1	1
SLOT_PRSNT_N_2 (b31)	1	0	0	1
SLOT_PRSNT_N_3 (b48)	0	1	1	1

Riser card should implement one I2C power monitoring device (TI/INA230) for each slots' P12V rail and one for the 2x2 power connector to be accessed by BMC on motherboard.

Slave address is 0x80, 0x82 for 2 slots riser, and 0x80, 0x82, 0x88 for 3 slots riser, and 0x8A for 2x2 power connector.

Riser type	Slot2	Slot3	Slot4	2x2 Power Conn
2 slots riser	0x80	0x82	N/A	0x8A
3 slots riser	0x80	0x82	0x88	0x8A

11.1.3 Riser card Power outlet

Add 2x2 right angle power connector (Molex/46991-1004 or equivalent) for deliver power from riser to PCIe cards that need higher power than PCIe slot allows.

Table 69 Riser power connector pin define

Position	Type	Description
1,2	Ground	Ground return
3,4	P12V	P12V from riser to Add-on-Cards

11.1.4 USB on riser card

There is a USB2.0 port from motherboard PCH USB2.0 Port 2 (BW61/CA61) connected to motherboard's riser card interface. An USB hub is needed to fan out the USB2.0 to 2 or 3 PCIe slots on riser card. The Hub p/n is Microchip USB2513BT/M2 for both 2 slots riser and 3 slots riser.

11.1.5 PCIe super slots optional implementation

This section describes an optional implementation of PCIe super slots in Tioga Pass. This implementation is to add 2x optional footprint in the bottom of x24 and x8 HSEC8 connectors. The purpose of this implementation is to enable riser card slot at bottom side of motherboard.

Vendor shall perform careful evaluation of SI risk during implementation.

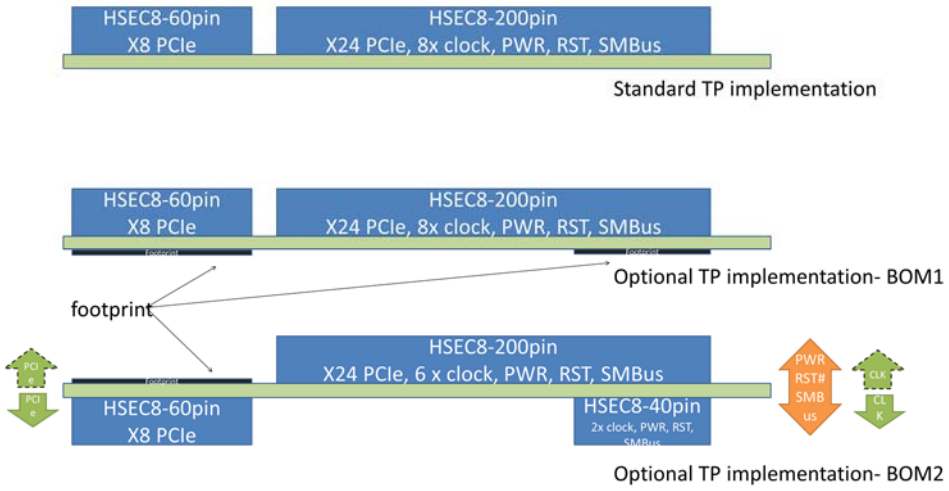


Figure 11-6 PCIe optional super slot on bottom side

11.2 DIMM Sockets

The motherboard requires 15u" gold contact for DDR4 SMT DIMM socket. Socket uses Yellow housing and white/nature latch for far end DIMM in a DDR channel, Black

housing and white/nature latch for near end DIMM in a DDR channel. Vendor shall announce if the color selection will increase the cost of the DIMM Socket.

11.3 Mezzanine Card

Tioga Pass is compatible with *Mezzanine Card for* Intel® *v2.0 Motherboard*[6] and *OCP Mezzanine Card 2.0 design specifications*[7].

The motherboard has Connector A and Connector B as OCP Mezz 2.0 specification to provide up to x16 PCIe Gen3 connection to Mezzanine card.

The motherboard also has Connector C as OCP Mezz 2.0 specification to provide up to x4 KR. Connector C can be used independently on Mezzanine card side.

Motherboard mezzanine connector A is named as Slot1 in system.

11.3.1 Connector A

Below table shows connector A pin definition:

Table 7 Mezzanine Connector A Pin Definition

Signal	Description	Pin	Pin	Signal	Description
P12V_AUX	Aux Power	61	1	MEZZ_PRSNT1_N	Present pin1, short to Pin120 on Mezz card
P12V_AUX	Aux Power	62	2	P5V_AUX	Aux Power
P12V_AUX	Aux Power	63	3	P5V_AUX	Aux Power
GND	Ground	64	4	P5V_AUX	Aux Power
GND	Ground	65	5	GND	Ground
P3V3_AUX	Aux Power	66	6	GND	Ground
GND	Ground	67	7	P3V3_AUX	Aux Power
GND	Ground	68	8	GND	Ground
P3V3	Power	69	9	GND	Ground
P3V3	Power	70	10	P3V3	Power
P3V3	Power	71	11	P3V3	Power
P3V3	Power	72	12	P3V3	Power
GND	Ground	73	13	P3V3	Power
LAN_3V3STB_ALERT_N	SMBus Alert for OOB	74	14	NCSI_RCSDV	BMC NCSI
SMB_LAN_3V3STB_CLK	SMBus Clock for OOB	75	15	NCSI_RCLK	BMC NCSI
SMB_LAN_3V3STB_DAT	SMBus Data for OOB	76	16	NCSI_TXEN	BMC NCSI
PCIE_WAKE_N	PCIE wake up	77	17	RST_PLT_MEZZ_N	PCIe reset signal
NCSI_RXER	BMC NCSI	78	18	RSVD (MEZZ_SMCLK)	Reserved(PCIe slot SMBus Clock)
GND	Ground	79	19	RSVD (MEZZ_SMDATA)	Reserved(PCIe slot SMBus Data)
NCSI_TXD0	BMC NCSI	80	20	GND	Ground
NCSI_TXD1	BMC NCSI	81	21	GND	Ground
GND	Ground	82	22	NCSI_RXD0	BMC NCSI
GND	Ground	83	23	NCSI_RXD1	BMC NCSI
CLK_100M_MEZZ1_DP	100MHz PCIe clock	84	24	GND	Ground
CLK_100M_MEZZ1_DN	100MHz PCIe clock	85	25	GND	Ground
GND	Ground	86	26	RSVD(CLK_100M_MEZZ2_DP)	Reserved(2nd set of 100MHz PCIe clock)
GND	Ground	87	27	RSVD(CLK_100M_MEZZ2_DN)	Reserved(2nd set of 100MHz PCIe clock)
MEZZ_TX_DP_C<0>	PCIE TX signal	88	28	GND	Ground
MEZZ_TX_DN_C<0>	PCIE TX signal	89	29	GND	Ground
GND	Ground	90	30	MEZZ_RX_DP<0>	PCIE RX signal
GND	Ground	91	31	MEZZ_RX_DN<0>	PCIE RX signal
MEZZ_TX_DP_C<1>	PCIE TX signal	92	32	GND	Ground

[6] http://files.opencompute.org/oc/public.php?service=files&t=2047e49112f6109c0f7e595cc93af8ae&download

[7] http://files.opencompute.org/oc/public.php?service=files&t=b9b9b1892b8584c52aeaf53bf8706ce0&download

MEZZ_TX_DN_C<1>	PCIE TX signal	93	33	GND	Ground
GND	Ground	94	34	MEZZ_RX_DP<1>	PCIE RX signal
GND	Ground	95	35	MEZZ_RX_DN<1>	PCIE RX signal
MEZZ_TX_DP_C<2>	PCIE TX signal	96	36	GND	Ground
MEZZ_TX_DN_C<2>	PCIE TX signal	97	37	GND	Ground
GND	Ground	98	38	MEZZ_RX_DP<2>	PCIE RX signal
GND	Ground	99	39	MEZZ_RX_DN<2>	PCIE RX signal
MEZZ_TX_DP_C<3>	PCIE TX signal	100	40	GND	Ground
MEZZ_TX_DN_C<3>	PCIE TX signal	101	41	GND	Ground
GND	Ground	102	42	MEZZ_RX_DP<3>	PCIE RX signal
GND	Ground	103	43	MEZZ_RX_DN<3>	PCIE RX signal
MEZZ_TX_DP_C<4>	PCIE TX signal	104	44	GND	Ground
MEZZ_TX_DN_C<4>	PCIE TX signal	105	45	GND	Ground
GND	Ground	106	46	MEZZ_RX_DP<4>	PCIE RX signal
GND	Ground	107	47	MEZZ_RX_DN<4>	PCIE RX signal
MEZZ_TX_DP_C<5>	PCIE TX signal	108	48	GND	Ground
MEZZ_TX_DN_C<5>	PCIE TX signal	109	49	GND	Ground
GND	Ground	110	50	MEZZ_RX_DP<5>	PCIE RX signal
GND	Ground	111	51	MEZZ_RX_DN<5>	PCIE RX signal
MEZZ_TX_DP_C<6>	PCIE TX signal	112	52	GND	Ground
MEZZ_TX_DN_C<6>	PCIE TX signal	113	53	GND	Ground
GND	Ground	114	54	MEZZ_RX_DP<6>	PCIE RX signal
GND	Ground	115	55	MEZZ_RX_DN<6>	PCIE RX signal
MEZZ_TX_DP_C<7>	PCIE TX signal	116	56	GND	Ground
MEZZ_TX_DN_C<7>	PCIE TX signal	117	57	GND	Ground
GND	Ground	118	58	MEZZ_RX_DP<7>	PCIE RX signal
GND	Ground	119	59	MEZZ_RX_DN<7>	PCIE RX signal
MEZZ_PRSNT2_N	Present pin2, short to Pin1 on Mezz card	120	60	GND	Ground

Note: For x16 PCIe, lane 0~7 is mapped to connector A and lane 8~15 is mapped to connector B.

11.3.2 Connector B

Table 8 Mezzanine Connector B Pin Definition

Signal	Description	Pin	Pin	Signal	Description
P12V_AUX/P12V	Aux Power	B41	B1	MEZZ_PRSNTB1_N /BASEBOARD_B_ID	Present pin1, short to Pin120 on Mezz card
P12V_AUX/P12V	Aux Power	B42	B2	GND	Ground
RSVD		B43	B3	MEZZ_RX_DP<8>	Aux Power
GND	Ground	B44	B4	MEZZ_RX_DN<8>	Aux Power
MEZZ_TX_DP<8>	Ground	B45	B5	GND	Ground
MEZZ_TX_DN<8>	Aux Power	B46	B6	GND	Ground
GND	Ground	B47	B7	MEZZ_RX_DP<9>	Aux Power
GND	Ground	B48	B8	MEZZ_RX_DN<9>	Ground
MEZZ_TX_DP<9>	Power	B49	B9	GND	Ground
MEZZ_TX_DN<9>	Power	B50	B10	GND	Power
GND	Power	B51	B11	MEZZ_RX_DP<10>	Power
GND	Power	B52	B12	MEZZ_RX_DN<10>	Power
MEZZ_TX_DP<10>	Ground	B53	B13	GND	Power
MEZZ_TX_DN<10>	SMBus Alert for OOB	B54	B14	GND	BMC NCSI
GND	SMBus Clock for OOB	B55	B15	MEZZ_RX_DP<11>	BMC NCSI
GND	SMBus Data for OOB	B56	B16	MEZZ_RX_DN<11>	BMC NCSI
MEZZ_TX_DP<11>	PCIE wake up	B57	B17	GND	PCIe reset signal
MEZZ_TX_DN<11>	BMC NCSI	B58	B18	GND	Reserved(PCIe slot SMBus Clock)
GND	Ground	B59	B19	MEZZ_RX_DP<12>	Reserved(PCIe slot SMBus Data)
GND	BMC NCSI	B60	B20	MEZZ_RX_DN<12>	Ground
MEZZ_TX_DP<12>	BMC NCSI	B61	B21	GND	Ground
MEZZ_TX_DN<12>	Ground	B62	B22	GND	BMC NCSI
GND	Ground	B63	B23	MEZZ_RX_DP<13>	BMC NCSI
GND	100MHz PCIe clock	B64	B24	MEZZ_RX_DN<13>	Ground
MEZZ_TX_DP<13>	100MHz PCIe clock	B65	B25	GND	Ground

MEZZ_TX_DN<13>	Ground	B66	B26	GND	Reserved(2nd set of 100MHz PCIe clock)
GND	Ground	B67	B27	MEZZ_RX_DP<14>	Reserved(2nd set of 100MHz PCIe clock)
GND	PCIE TX signal	B68	B28	MEZZ_RX_DN<14>	Ground
MEZZ_TX_DP<14>	PCIE TX signal	B69	B29	GND	Ground
MEZZ_TX_DN<14>	Ground	B70	B30	GND	PCIE RX signal
GND	Ground	B71	B31	MEZZ_RX_DP<15>	PCIE RX signal
GND	PCIE TX signal	B72	B32	MEZZ_RX_DN<15>	Ground
MEZZ_TX_DP<15>	PCIE TX signal	B73	B33	GND	Ground
MEZZ_TX_DN<15>	Ground	B74	B34	GND	PCIE RX signal
GND	Ground	B75	B35	CLK_100M_MEZZ2_DP	PCIE RX signal
GND	PCIE TX signal	B76	B36	CLK_100M_MEZZ2_DN	Ground
CLK_100M_MEZZ3_DP	PCIE TX signal	B77	B37	GND	Ground
CLK_100M_MEZZ3_DN	Ground	B78	B38	PERST_N1	PCIE RX signal
GND	Ground	B79	B39	PERST_N2	PCIE RX signal
MEZZ_PRSNTB2_N	PCIE TX signal	B80	B40	PERST_N3	Ground
P12V_AUX/P12V	PCIE TX signal	B41	B1	MEZZ_PRSNTB1_N /BASEBOARD_B_ID	Ground
P12V_AUX/P12V	Ground	B42	B2	GND	PCIE RX signal
RSVD	Ground	B43	B3	MEZZ_RX_DP<8>	PCIE RX signal
GND	PCIE TX signal	B44	B4	MEZZ_RX_DN<8>	Ground
MEZZ_TX_DP<8>	PCIE TX signal	B45	B5	GND	Ground
MEZZ_TX_DN<8>	Ground	B46	B6	GND	PCIE RX signal
GND	Ground	B47	B7	MEZZ_RX_DP<9>	PCIE RX signal
GND	PCIE TX signal	B48	B8	MEZZ_RX_DN<9>	Ground
MEZZ_TX_DP<9>	PCIE TX signal	B49	B9	GND	Ground
MEZZ_TX_DN<9>	Ground	B50	B10	GND	PCIE RX signal
GND	Ground	B51	B11	MEZZ_RX_DP<10>	PCIE RX signal
GND	PCIE TX signal	B52	B12	MEZZ_RX_DN<10>	Ground
MEZZ_TX_DP<10>	PCIE TX signal	B53	B13	GND	Ground
MEZZ_TX_DN<10>	Ground	B54	B14	GND	PCIE RX signal
GND	Ground	B55	B15	MEZZ_RX_DP<11>	PCIE RX signal
GND	PCIE TX signal	B56	B16	MEZZ_RX_DN<11>	Ground
MEZZ_TX_DP<11>	PCIE TX signal	B57	B17	GND	Ground
MEZZ_TX_DN<11>	Ground	B58	B18	GND	PCIE RX signal
GND	Ground	B59	B19	MEZZ_RX_DP<12>	PCIE RX signal
GND	Present pin2, short to Pin1 on Mezz card	B60	B20	MEZZ_RX_DN<12>	Ground
MEZZ_TX_DP<12>		B61	B21	GND	
MEZZ_TX_DN<12>		B62	B22	GND	
GND		B63	B23	MEZZ_RX_DP<13>	
GND		B64	B24	MEZZ_RX_DN<13>	
MEZZ_TX_DP<13>		B65	B25	GND	
MEZZ_TX_DN<13>		B66	B26	GND	
GND		B67	B27	MEZZ_RX_DP<14>	
GND		B68	B28	MEZZ_RX_DN<14>	
MEZZ_TX_DP<14>		B69	B29	GND	
MEZZ_TX_DN<14>		B70	B30	GND	
GND		B71	B31	MEZZ_RX_DP<15>	
GND		B72	B32	MEZZ_RX_DN<15>	
MEZZ_TX_DP<15>		B73	B33	GND	
MEZZ_TX_DN<15>		B74	B34	GND	
GND		B75	B35	CLK_100M_MEZZ2_DP	
GND		B76	B36	CLK_100M_MEZZ2_DN	
CLK_100M_MEZZ3_DP		B77	B37	GND	
CLK_100M_MEZZ3_DN		B78	B38	PERST_N1	
GND		B79	B39	PERST_N2	
MEZZ_PRSNTB2_N		B80	B40	PERST_N3	

11.3.3 Connector C

Below table shows NIC connector C pin definition:

Table 9 Mezzanine Connector C Pin Definition

Signal	Description	Pin	Pin	Signal	Description
P12V_AUX/P5V_AUX-P12V	Aux Power	C33	C1	MEZZ_SMCLK	
P12V_AUX/P5V_AUX-P12V	Aux Power	C34	C2	MEZZ_SMDATA	
P12V_AUX/P5V_AUX-P12V	Aux Power	C35	C3	EXT_MDIO_I2C_SEL	
RSVD	Ground	C36	C4	GND	
SDP0	Ground	C37	C5	KR_TX_DP<2>	
SDP1	Aux Power	C38	C6	KR_TX_DN<2>	
GND	Ground	C39	C7	GND	
KR_TX_DP<0>	Ground	C40	C8	LED_P1_0_N	
KR_TX_DN<0>	Power	C41	C9	LED_P1_1_N	
GND	Power	C42	C10	GND	
LED_P0_0_N	Power	C43	C11	KR_TX_DP<3>	
LED_P0_1_N	Power	C44	C12	KR_TX_DN<3>	
GND	Ground	C45	C13	GND	
KR_TX_DP<1>	SMBus Alert for OOB	C46	C14	LED_P2_0_N	
KR_TX_DN<1>	SMBus Clock for OOB	C47	C15	LED_P2_1_N	
GND	SMBus Data for OOB	C48	C16	GND	
SHARED_KR_MDC_0	PCIE wake up	C49	C17	KR_RX_DP<2>	
SHARED_KR_MDIO_0	BMC NCSI	C50	C18	KR_RX_DN<2>	
GND	Ground	C51	C19	GND	
KR_RX_DP<0>	BMC NCSI	C52	C20	Module_SCL0	
KR_RX_DN<0>	BMC NCSI	C53	C21	Module_SDA0	
GND	Ground	C54	C22	GND	
LED_P3_0_N	Ground	C55	C23	KR_RX_DP<3>	
LED_P3_1_N	100MHz PCIe clock	C56	C24	KR_RX_DN<3>	
GND	100MHz PCIe clock	C57	C25	GND	
KR_RX_DP<1>	Ground	C58	C26	Module_SCL1	
KR_RX_DN<1>	Ground	C59	C27	Module_SDA1	
GND	PCIE TX signal	C60	C28	GND	
Module_SCL2	PCIE TX signal	C61	C29	Module_SCL3	
Module_SDA2	Ground	C62	C30	Module_SDA3	
GND	Ground	C63	C31	SDP2	
MEZZ_PRSNTC2_N	PCIE TX signal	C64	C32	SDP3	

11.3.4 Base board ID

Baseboard ID is to let Mezzanine card has awareness of different base board. Baseboard ID only applies connector A and connector B.

The implementation of Baseboard ID circuit on Tioga Pass is shown below figure 13-7. R1 is 10K and R2 is 8870hm on both connect A and connector B.

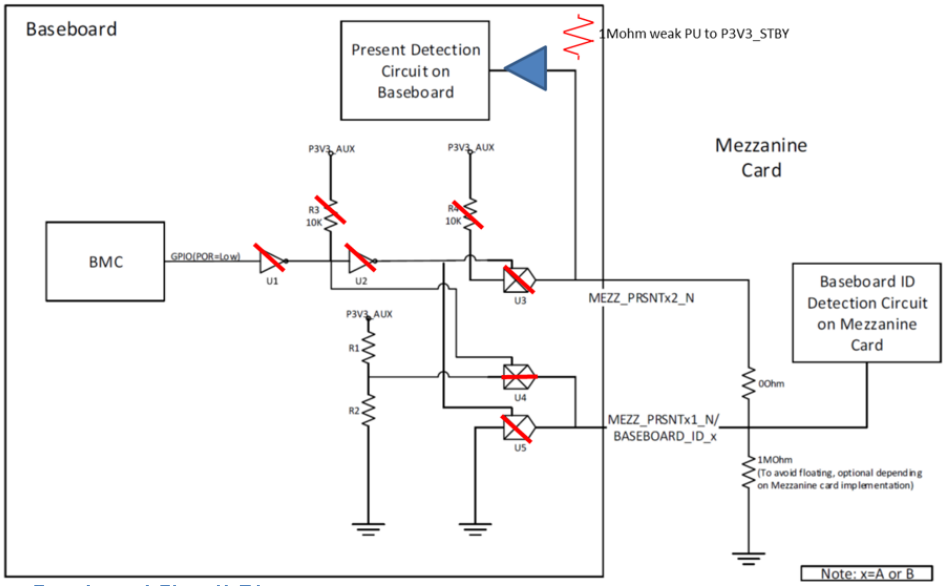


Figure 11-7: Baseboard Circuit Diagram

The Mezzanine card identifies different of Baseboard on the resistor pair R1/R2 shown in tables below:

Table: Connector A Baseboard Types

ConnA R1	ConnA R2	Baseboard type on Connector A
NC	0 Ω	One x8 PCIe Root Port on baseboard Connector A; No Connector B on Baseboard
10 KΩ	887 Ω	One x16 PCIe Root Ports on Baseboard Connector A and B
10 KΩ	2.10 KΩ	One x8 PCIe Root Port on baseboard Connector A; Connector B presents on Baseboard
10 KΩ	3.83 KΩ	Two x4 PCIe Root Ports on baseboard Connector A
10 KΩ	6.49 KΩ	Four x2 PCIe Root Ports on baseboard Connector A
10 KΩ	11 KΩ	Eight x1 PCIe Root Ports on baseboard Connector A
10 KΩ	20.5 KΩ	RFU
10 KΩ	48.7 KΩ	RFU
10 KΩ	NC	Up to 8x KR on baseboard Connector A

Table: Connector B Baseboard Types

ConnB R1	ConnB R2	Baseboard type on Connector B
NC	NC	No Connector B on baseboard; Mezzanine card samples Baseboard_ID_B as 0V with weak pull low on Mezzanine card side
10 KΩ	887 Ω	One x16 PCIe Root Ports on Baseboard Connector A and B
10 KΩ	2.10 KΩ	One x8 PCIe Root Port on baseboard Connector B
10 KΩ	3.83 KΩ	Two x4 PCIe Root Ports on baseboard Connector B
10 KΩ	6.49 KΩ	Four x2 PCIe Root Ports on baseboard Connector B
10 KΩ	11 KΩ	Eight x1 PCIe Root Ports on baseboard Connector B

10 KΩ	20.5 KΩ	RFU
10 KΩ	48.7 KΩ	RFU
10 KΩ	NC	Up to 8x KR on baseboard Connector B

11.3.5 Mezzanine present pin

Tioga pass Mezzanine present pins' connection to BMC and PCH is shown below:

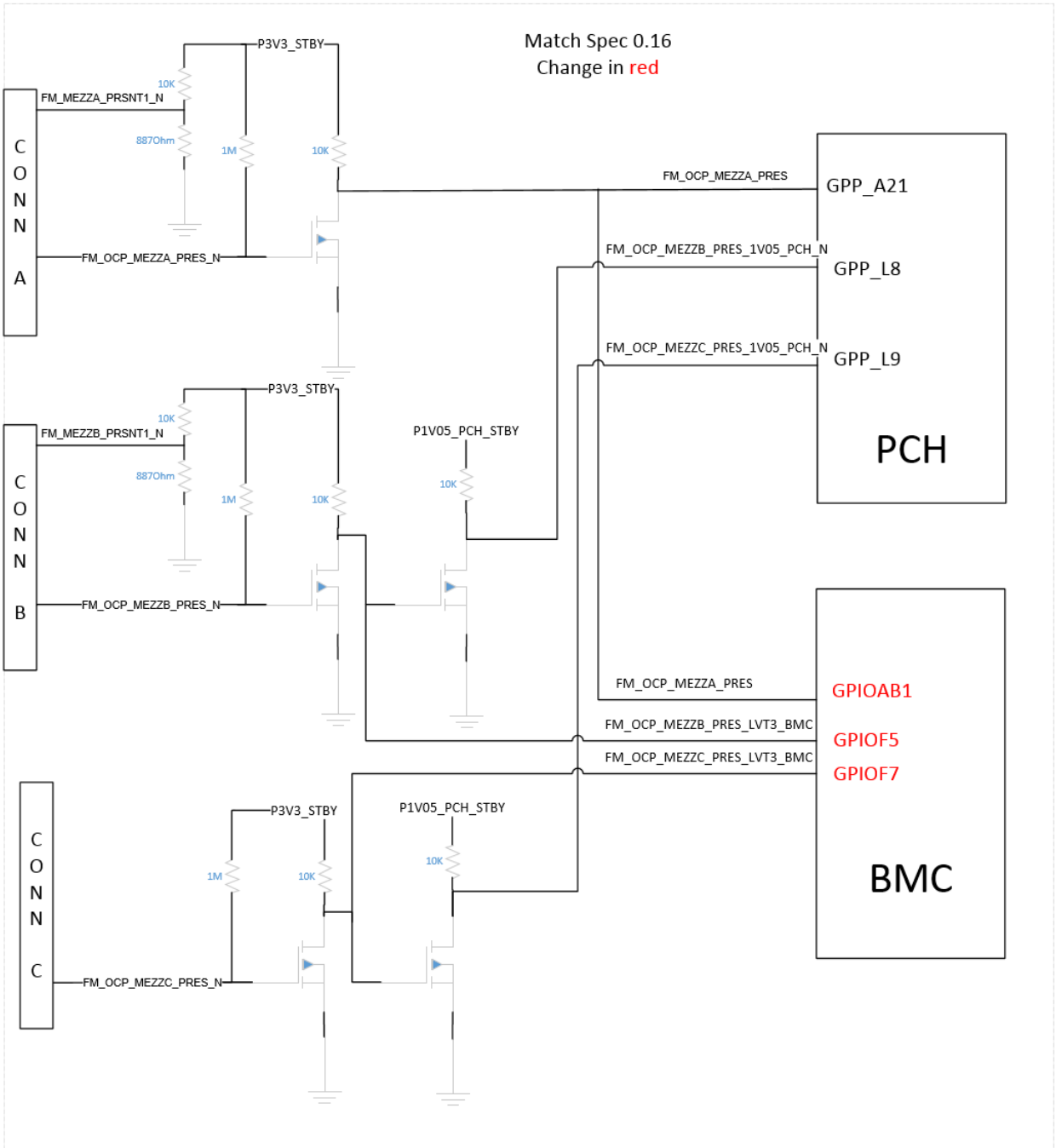


Figure 11-8: Mezzanine Pins Schematic Diagram

11.4 Network

11.4.1 Data network

The motherboard uses OCP Mezzanine 2.0 card as its primary data network interface at I/O side. There is option of single port or dual port.

One Intel® I210 1G NIC is placed on board to provide optional 10/100/1000 data connection.

11.4.2 Management Network

The motherboard has 3x options of management network interface for BMC's connection. Management network shares data network's physical interface. Management connection should be independent from data traffic, and OS/driver condition. The information below echoes the requirement in section 0, while emphasis on physical connection.

- SFP+ shared-NIC from Mezzanine 10G NIC or PCIe NIC, driven by BMC through RMII/NC-SI or I2C. I2C being default
- SGMII/KX shared-NIC connected to mid-plane interface from Intel® I210-AS, driven by BMC through RMII/NCSI
- 10/100/1000 MDI shared-NIC connected to RJ45 from Intel® I210-AT(co-layout with Intel® I210-AS), driven by BMC through RMII/NCSI

11.4.3 IPv4/IPv6 Support

The system need to have the capability to be deployed in both IPv4 and IPv6 network environment. All data network and management network should have this capability. This includes, but not limit to: DHCP and static IP setting, PXE booting capability, NIC and BMC firmware support, OS driver, and utility in both IPv4 and IPv6.

11.5 USB

The motherboard has one external type A right angle USB2.0/3.0 port and one usb3.0 type C port located in front of the motherboard. BIOS should support following devices on USB ports available on motherboard:

- USB Keyboard and mouse
- USB flash drive (bootable)
- USB hard drive (bootable)
- USB optical drive (bootable)

  Type C USB port is reserved for Intel® DCI(Direct Connect Interface) debug.

The external USB2.0/3.0 port shall have the 5x USB3.0 signals remapped with debug interface. Refer to section 13.8.2 for details.

USB power enabling is controlled by CPLD to allow debug port being powered in S5 to perform BMC related trouble shooting. In Tioga Pass USB power should stay on in S5 state to allow user able to power on system through debug card.

11.6 SATA

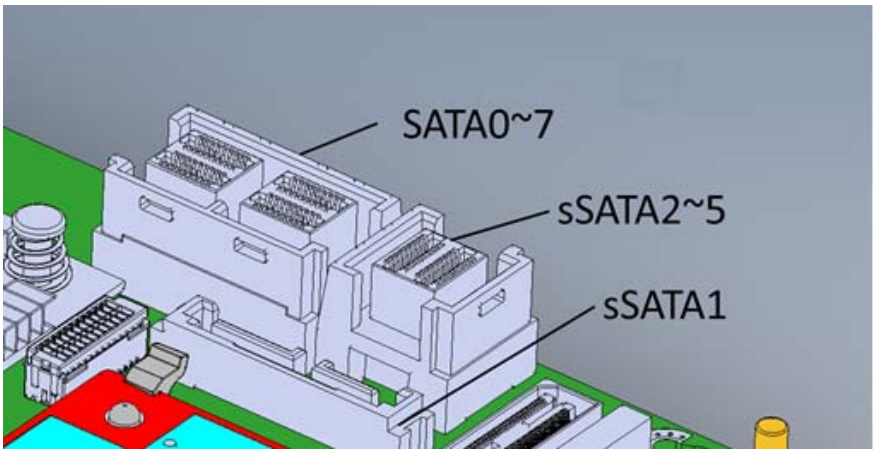
The motherboard has Intel® PCH on board. Intel® PCH has a SATA controller support 8x SATA3 ports, and an sSATA controller support 6x SATA3 ports.

SATA Port 0~7 can be connected to one vertical miniSAS HD 8 ports connector. sSATA Port 2~5 can be connected to one miniSAS HD 4 ports connector. This is optional implementation for Tioga Pass.

sSATA Port 0 is connected to 1x M.2 connector on motherboard with dual layout with PCIe x4 to M2. PCIe x4 to M.2 is default population.

sSATA Port 1 is connected to 1x vertical SATA connector on motherboard. This vertical SATA port is using a power-signal combined SATA connector with latch to allow 1-step and secure operation. Connector is Alltop/C18625-11331-L.

Vendor shall follow DXF for SATA ports placement



Both mSATA connector and vertical SATA connector need to be placed near IO side of the motherboard for easy access. HDDs attached to all SATA connectors need to follow spin-up delay requirement described in section 15.7.1.

11.7 M.2

The motherboard has 1x M.2 connector with Key ID=M and H6.7 Type. M.2 connector has option connection of PCIe x4 from PCH or sSATA Port 0 from PCH. On board M.2 connector supports 2280, 22110 card form factor with both single side and double side.

Sideband signals such as WAKE# and SATA_ACTIVITY should be connected when it applies.

PERST# shall go active before the power on M.2 connector is removed per PCI CEM specification.

Vendor shall add SMBUS and Alert connection to BMC base on the latest M.2 spec. Please be aware of the M.2 SMBUS is at 1.8v level. Vendor shall use shunt regulator to create 1.8V for SMBUS pull up and add level shift to connect to BMC SMB3 or SMB7.

11.8 Debug Header

The motherboard has 2 debug headers to work with 2 types of debug card, 14pin V1 debug card and OCP debug card with LCD.

11.8.1 14pin V1 debug card header

The 1st debug header is placed in front right of the motherboard. Debug card can be plugged into this header directly or through a cable. This debug header should support hot plug. Through this debug header, debug card should provide one UART serial port connector, two 7-segment LED displays, one reset button and one UART channel selection button. UART should provide console redirection function. Two 7-segment LED displays show BIOS POST code and DIMM error information. One reset button will trigger system reset when pressed. Debug card pin 14 power P5V_AUX on or off is controlled by CPLD. Default is enabled.

UART channel selection button sends negative pulse to motherboard CPLD to select and rotate UART console in a loop of host console->BMC debug console->mid-plane console. Default after debug card plugged in or system power up is host console. CPLD counts the negative pulse and turns the state machine states in loop of 00->01->10->11 to control the MSB and LSB. MSB and LSB will controls the output of FSA3357 to switch between host, BMC and Mid-plane console.

The connector for the debug header is a 14 pin, shrouded, vertical, 2mm pitch connector. Figure 11-9 is an illustration of the headers. Debug card should have a key to match with the notch to avoid pin shift when plugging in.

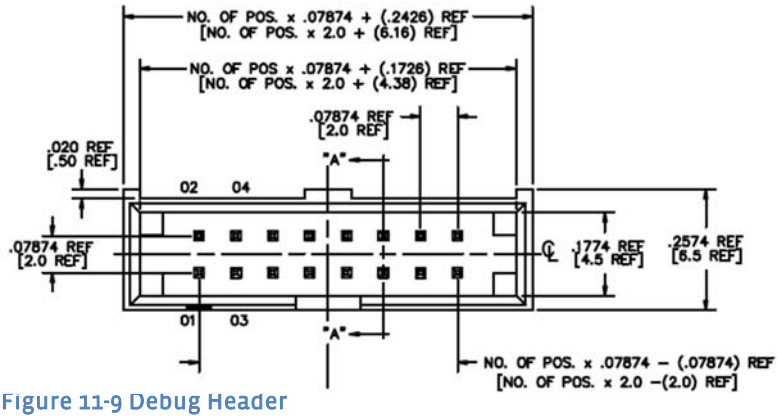


Figure 11-9 Debug Header

Table 10 Debug Header Pin Definition

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 (motherboard transmit, 3.3V signal level)
10	Serial Receive (motherboard receive, 3.3V/5V tolerant)

11	System Reset
12	UART channel selection
13	GND
14	P5V(default)/P5V_AUX

11.8.2 Debug Port Dual Layout to USB3.0 Connector

The 2nd debug header is to remap 5x USB3.0 signals to UART signals to pass console data, and I2C signals to access post code through an I2C GPIO expender on motherboard. The block diagram and pin definitions is shown below. Please refer to "debug_card_v2_fb_sch_v07_20170123.pdf" in the OCP spec package for the detailed schematic.

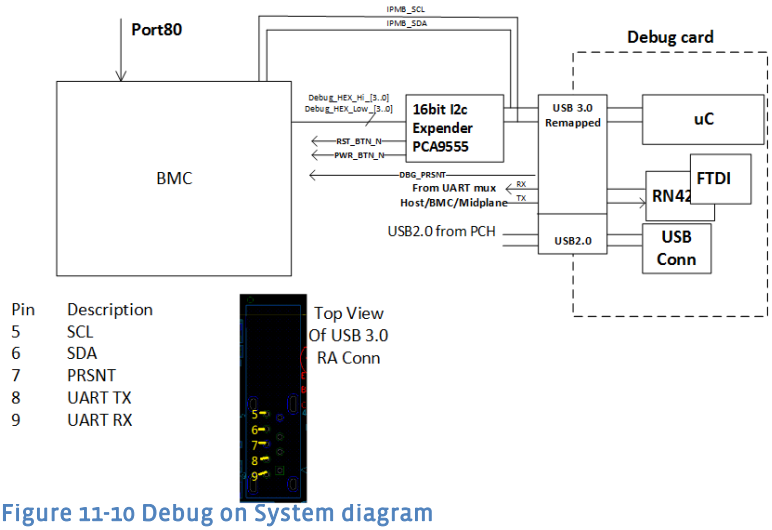


Figure 11-10 Debug on System diagram

The IPMB from remapped USB3.0 port is connected to port 10 of BMC.

Vendor shall follow the diagram above to implement the SMBus address: BMC uses 0x20h, PCA9555 uses 0x4Eh, and micro controller on debug card uses 0x60h. PCA9555 GPIO mapping as below:

Bit	Usage	Direction In the perspective of PCA9555
IO0_0-3	HEX Low 0-3	Input
IO0_4-7	Hex High 0-3	Input
IO1_0	RST_BTN_N	Output
IO1_1	PWR_BTN_N	Output
IO1_2	PWRGD_SYS_PWROK	Input
IO1_3	RST_PLTRST_N	Input
IO1_4	PWRGD_DSW_PWROK	Input
IO1_5	FM_CPU_CATERR_MSMI_LVT3_N	Input
IO1_6	FM_SLPS3_N	Input
IO1_7	FM_SOL_UART_CH_SEL	Output

11.8.3 Debug Port power policy

User should be able to power on system by pressing power button on debug card with LCD. The USB port power should be standby power rail to support debug card MCU to function at S5 state. Vendor should implement debug card FW to control the 7 segment LED to be off when system is in S5 state.

11.8.4 Post Codes

The post codes are brought to debug header in HEX format via two HEX codes. The HEX codes can be driven by either the legacy parallel port (port 80) on SIO, or 8 GPIO pins. A debug card with two seven segment displays, two HEX-to-seven segment converters, logic level to RS-232 shifter and a RS-232 connector shall interface the debug header.

During POST, BIOS should also output POST code to BMC SOL. So when SOL session is available during POST, remote console should show POST code as mentioned in section 9.2.

During the boot sequence the BIOS shall initialize and test each DIMM module. If a module fails initialization or does not pass the BIOS test the following post codes should flash on the debug card to indicate which DIMM has failed. The first hex character indicates which CPU interfaces the DIMM module; the second hex character indicates the number of the DIMM module. POST code will also display error major code and minor code from Intel® Manageability Enginemory reference code. The display sequence will be “00”, DIMM location, Major code and Minor code with 1 second delay for every code displayed. The BIOS shall repeat the display sequence indefinitely to allow time for a technician to service the system. DIMM location code table is as Table 11. DIMM number count starts from the furthest DIMM from CPU.

Table 11 DIMM Error Code Table

Code	Result
CPU0 (Channel 0 & 1)	A0
	C0
	A1
	C1
	A2
	C2
CPU0 (Channel 2 & 3)	A3
	C3
	A4
	C4
	A5
	C5
CPU1 (Channel 0 & 1)	B0
	D0
	B1
	D1
	B2
	D2
CPU1 (Channel 2 & 3)	B3

	D3
	B4
	D4
	B5
	D5

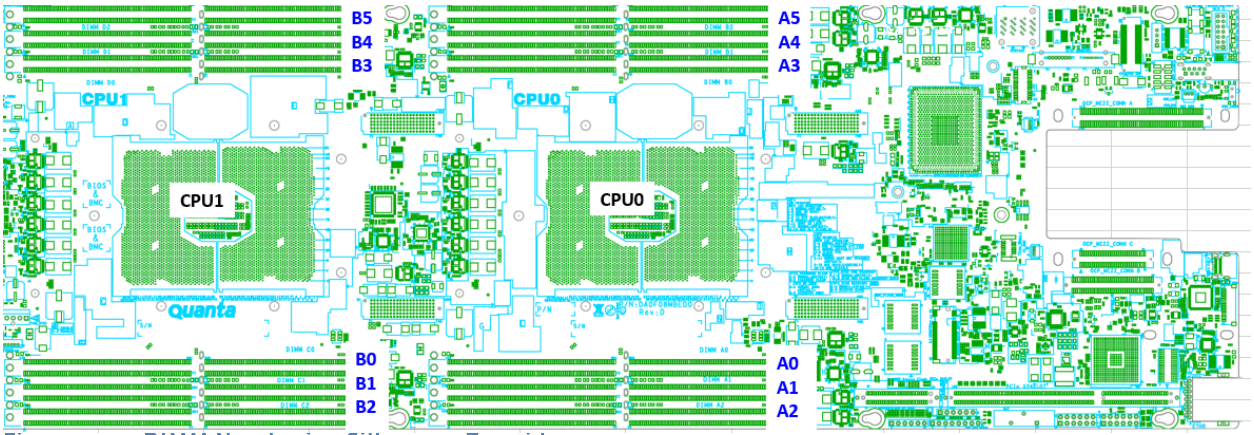


Figure 11-11 DIMM Numbering Silkscreen Top side

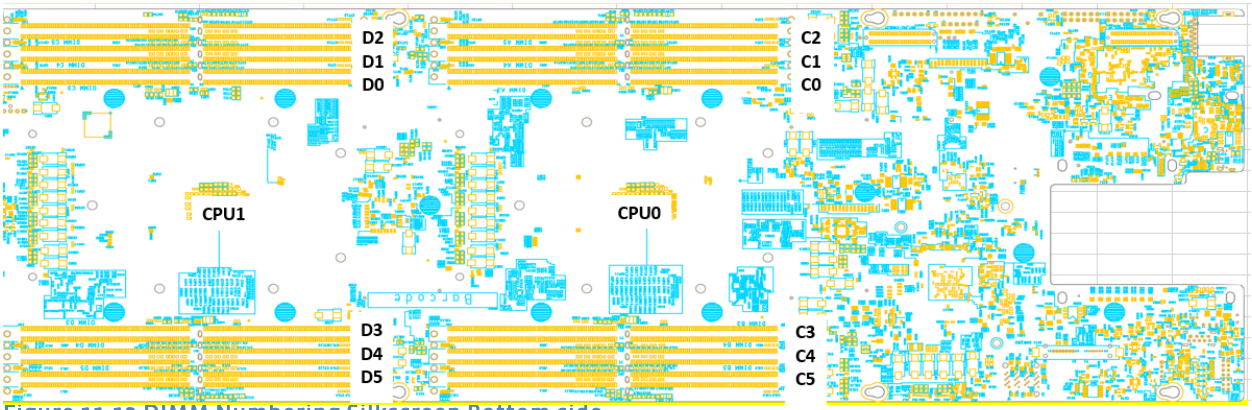


Figure 11-12 DIMM Numbering Silkscreen Bottom side

11.8.5 Serial Console (To be updated base on latest debug card)

The output stage of the systems serial console shall be contained on the debug card. The TX and RX signals from the system UART shall be brought to the debug header at the chips logic levels (+3.3V). The debug card will contain mini USB type connector with pin definition shown in Table 12. A separate convertor is needed to provide RS-232 transceiver and DB9 connector.

Table 12 Debug card mini-USB UART Pin Define

Pin	Function
1	VCC (+5VDC)
2	Serial Transmit (motherboard transmit)
3	Serial Receive (motherboard receive)
4	NC
5	GND

The Debug card will contain a vertical receptacle female 6x pin 0.1" pitch header (FCI/ 68685-306LF or equivalent) to provide connection to an optional Class 2 Bluetooth module (Roving Networks/RN42SM[8] or equivalent).

Table 13 Bluetooth header pin define

Pin	Function
1	Serial Transmit (motherboard transmit)
2	Serial Receive (motherboard receive)
3	NC
4	NC
5	VCC (+5VDC)
6	GND

11.8.6 UART channel selection

When debug card is plugged, debug card pin 12 shall be used to notice motherboard BMC GPIO for its presence. When UART Channel selection button on debug card is pressed, same pin is used to send pulses to motherboard to trigger UART connection change among host console (Default) -> BMC debug console->mid-plane console and loop. Falling edge triggers change.

Debug card has 10K pull down for pin 12, with a white UART Channel selection button between Pin 12 and ground.

Motherboard side should implement logic to detect debug card presence when 10K or stronger pull down presents on Pin 12.

Motherboard side should implement logic to trigger UART connection change when UART Channel selection button is pressed. UART channel selection has a power on reset value of 00, which means host console.

Motherboard should stop its original POST code display for 1 sec when a falling edge to ground is detected, and give a 1sec display of Channel number to debug port POST code as an indication of UART channel change. If system POST code is not changed within this 1 sec, motherboard outputs original display. If system POST code is changed within this 1 sec, the latest POST code should be displayed. Motherboard also has two LEDs for displaying of UART connection status as described in Table 14.

Table 14 UART channel and connection

Channel	UART Connection
00	Host console
01	BMC debug console
10/11	Reserved

[8] http://www.rovingnetworks.com/products/RN42SM

11.8.7 Other debug use design requirements on motherboard

XDP header is required for BIOS debug and should be populated in EVT/DVT/PVT samples. The access to the XDP header should not be mechanically blocked by CPU heat sink or other components.

SMBUS debug header should be inserted for SMBUS on motherboard based on SMBUS topology vendor designs. SMBUS debug headers for PCH host bus and CPU/DIMM VR PMBus are required.

If any other testing/debugging header is needed based on Intel® platform development requirement, it should be added and populated in EVT/DVT samples.

11.9 Switches and LEDs

The motherboard shall include Power switch, Reset switch, Power LED, HDD activity LED and Beep error LED.

11.9.1 Switches

Vertical tactile switches are placed behind debug header. The push button actuator has a minimum 2.5mm diameter and protrudes 9mm+/-1mm from top of actuator to PCB surface. System power button is red and on the left. Systems reset button is black and on the right.

If the power switch is depressed for durations less than four seconds a Power Management Event indicating that the power switch has been triggered shall be issued. If the power switch is depressed for durations longer than four seconds the motherboard shall perform a hard power off.

If the reset switch is depresses for any duration of time the motherboard shall perform a hard reset and begin executing BIOS initialization code.

Power switch and Reset switch function should not be gated by BMC FW or have dependency to BMC firmware readiness.

The functionality of each switch shall be indicated by a label on the motherboard's silk screen. The labels PWR and RST are acceptable.

11.9.2 LEDs

The table below indicates the color and function of each LED. The motherboard's silkscreen shall indicate the functionality of each of these LEDs. Silk screen labels are included in Table 15. Looking from I/O towards LED, from right to left, the sequence is Blue, Green, Yellow, Green, Green.

Table 15 LED Functionality

LED Color	Function	Silk Screen Label
Blue	Power LED. This LED shall illuminate if the motherboard is in the power on state. This LED is also used as chassis identify.	PWR
Green	Hard drive activity. This LED shall illuminate when there is activity on the motherboards SATA hard drive interfaces, or onboard mSATA and M.2 connector interface.	HDD

Yellow	BEEP/Error LED. This LED shall illuminate when PCH speaker has output, or, BIOS_ERR_TRIGGER_N asserts. BIOS_ERR_TRIGGER_N is for debug purpose to have a predefined error identified from LED. It can also be used as oscilloscope trigger. It is disabled in production BIOS.	BEEP/ERR
Green	UART Channel status LEDs. Two LEDs indicates the UART channel number's binary code. Both LEDs should stay off by default to indicate UART channel is on host console. Smaller package should be used for these two LEDs compare to the other three.	UART_CH[1..0]

11.10 FAN connector

The motherboard has 2x system FAN connectors. FAN connector signals should follow "4-Wire Pulse Width Modulation (PWM) Controlled Fans Specification" Revision 1.3 September 2005 published by Intel® Corporation. Each FAN has 6 pins, which is compatible with standard 4-Wire FAN connector, and also can be used to support dual rotor FAN that shares PWM control signal but has separate TACH signals. FAN connector pin definition is as in Table 16. LOTES APHD0019-P001A or equivalent shall be used as Fan connector. Its mating part is a LOTES GAP-ABA-WAF-038 or equivalent. The fan power should be connected at the downstream side of hot swap controller (HSC). The fan power needs to be turned off during S5.

The motherboard has 1Xpump header to support water cooling solution. Pump header has same form factor as the fan connectors. Pump header will have a cap to avoid to be connected to a system fan.

Table 16 FAN Connector Pin Definition

Pin	Description
1	GND
2	12.5VDC
3	Sense #1
4	Control
5	Sense #2 (optional)
6	No Connect

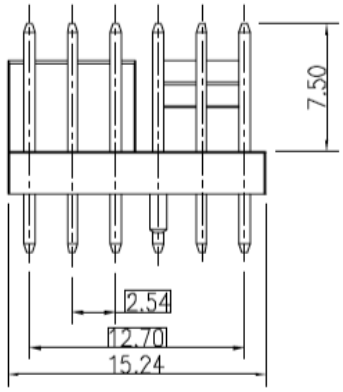


Figure 11-13 Fan Connector

11.11 TPM Connector and Module

A 11pin vertical receptacle connector is defined on MB for SPI and I2C TPM module. Connector pin definition on motherboard side is shown in Table 17. FCI/91931-31111LF Receptacle or equivalent should be used on motherboard. Connector is shown in Figure 11-14.

TPM module is a 32.3mm (L) x 13mm (W) x 0.8mm (T) PCB with FCI/91911-31511LF header or equivalent in the center of the module. Please refer to the 3D for detail.

SPI_TPM_CLK	1	7	P3V3_STBY
SPI_TPM_PLTRST_N	2	8	Module_PRSNT_N/I2C_TPM_RST_N
SPI_TPM_MOSI	3	9	SPI_TPM_IRQ_N
SPI_TPM_MISO	4	10	SMB CLK
SPI_TPM_CS_N	5	11	GND
I2C_TPM_DAT	6		
I2C_TPM for BMC			
SPI_TPM for BIOS			

Table 17 TPM Header Pin Definition

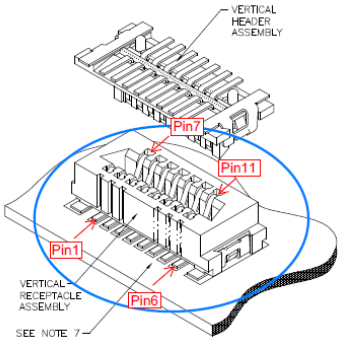


Figure 11-14 TPM Header

11.12 Sideband Connector

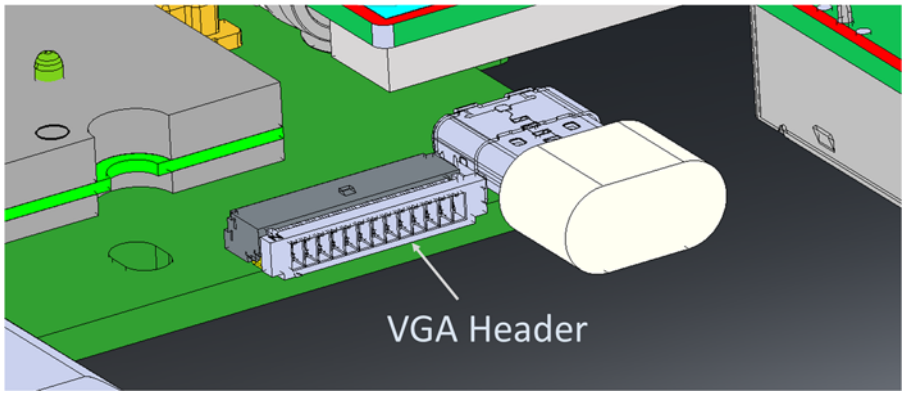
An 8pin connector is defined for side band signals. Tyco/2-1734598-8 or equivalent is used.

Pin	Signal Name	Description
1	P3V3_AUX	3.3V Aux Power
2	PS_REDUDENT_LOST_N	Power Shelf redundant lost status, low active; connect to BMC GPIO
3	PS_FAIL_N	Power Shelf fail, low active; connect to BMC GPIO. It should also trigger NVDIMM SAVE sequence
4	MATED_IN_N	Mate detection; low active; to enable HSC and triggers NVDIMM SAVE sequence if used.
5	PMBUS_ALERT	PMBus Alert signal

6	PMBUS_SDA	PMBus data signal
7	PMBUS_SCL	PMBus clock signal
8	GND	Ground

11.13 VGA header

OCP FB server motherboard v4.0 added a VGA header based on community feedback. Due to the board I/O space limitation, a standard DB15 VGA connector cannot be used. Alternatively, Samtec/T1M-13-GF-S-RA-TR right angle style header is used for VGA connector on motherboard side. Vendor shall enable the adaptor from this VGA header to standard DB15 VGA connector.



Vendor shall refer to the DXF for placement of VGA header and table below for signal assignment.

	Sig
1	RED
2	RED_RTN(GND)
3	Green
4	Green_RTN (GND)
5	Blue
6	Blue_RTN (GND)
7	V-Sync
8	GND(Vsync-DDC)
9	H-Sync
10	GND(H-sync)
11	SDA
12	SCL
13	PWR

12 Rear Side Power, I/O and mid-plane

12.1 Overview of Footprint and Population Options

There are 4x footprints at rear side of the motherboard to provide power to motherboard, and I/O to a mid-plane. The population of the footprints is flexible to fit the need of different use cases.

Population options and major differences are listed in Figure 12-1. The high speed mid-plane is not covered in this document. The ORv2 implementation is described in Chapter 13.

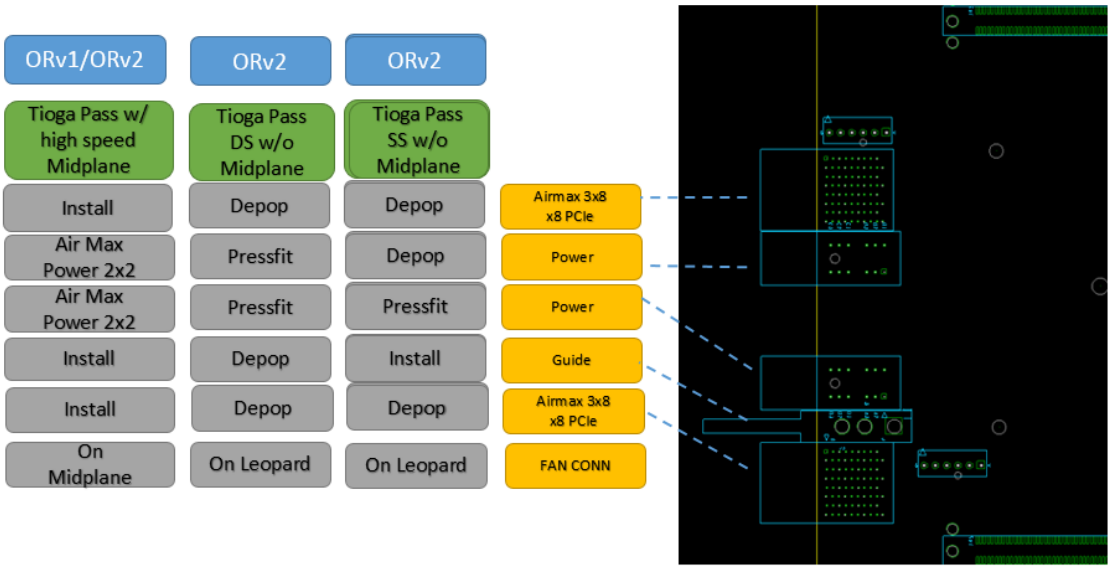


Figure 12-1 Motherboard midplane connector population options

12.2 Rear Side Connectors

12.2.1 Footprints and Connectors/Pressfit Cable

There are 3x types of connector footprints. A design can install a combination of 3x types of connectors and 1x type of Pressfit cable to the footprints. The AVL is listed in Table 18. The placement of connector footprints is shown in Figure 12-2.

Table 18

Connector Type	Sled side P/N	Mid-plane side P/N
AirMax® Guide	FCI/10045588-101LF	FCI/10045367-101LF
AirMax VS® Power 2x2	FCI/10124648-001LF	FCI/10124620-5545P00LF
AirMax VS2® 3x8 press-fit/E4 short	FCI/10124755-111LF	FCI/10124756-101LF
Pressfit Cable	TE/2159562-1	N/A

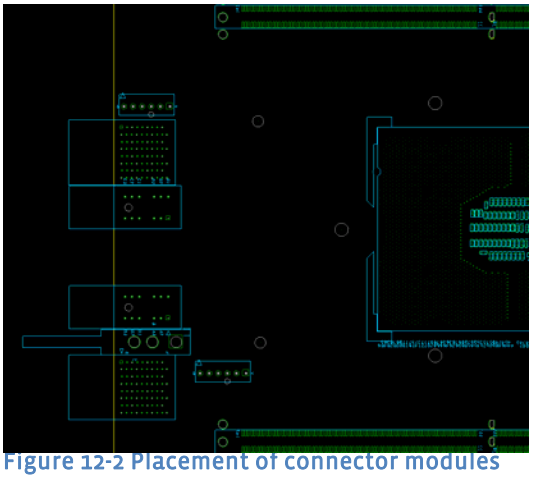


Figure 12-2 Placement of connector modules

12.2.2 AirMax® power 2x2

In use case with less or equal to 49 A[9] on 12VDC, 1x pair of AirMax® power 2x2 R/A low profile headers and receptacle are populated. 2x pairs of such connector supports up to 98A on 12VDC.

12.2.2.1 Sled side

Up to 2x AirMax® power 2x2 R/A low profile headers are used at sled side, and shown in Figure 12-3. Pin definition is as Table 19.

Figure 12-3 Power connector at sled side

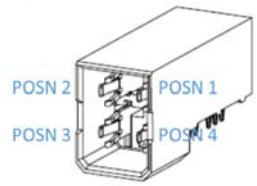


Table 19 Pin definition on both sled and midplane side

Position	Type	Description
1,2	Power	P12V_AUX power rail from mid-plane to motherboard or uServer sled. Hotswap controller is on motherboard or uServer sled. This rail is a standby rail and NOT controlled by MB_ON_N.
3,4	Ground	Ground return

12.2.2.2 Mid-plane side

Up to 2x AirMax® power 2x2 R/A receptacles for co-planar application are used at mid-plane side, and shown in Figure 12-4. Receptacle has long and short pin to control mating sequence. Part number with S-S-L-S pattern is used to ensure at least one ground pin mates before any power pin mates. Refer to Table 20 for detail.

[9] In 65C ambient with PCB stackup of the motherboard. Current rating need to be re-evaluated in different design.

Figure 12-4 Power connector at midplane side

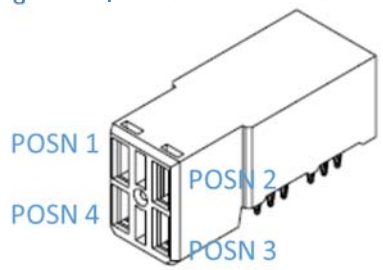


Table 20 Part number with short and long pattern on midplane side

PRODUCT NUMBER ("LF" DENOTES LEAD-FREE)	CONTACT PLATING NOTE	CONTACT DEPTH (SEE SECTION A-A)				ROHS COMPATIBILITY
		POSN 1	POSN 2	POSN 3	POSN 4	
10124620-4444P00LF	2a	LONG	LONG	LONG	LONG	SEE NOTE 8
10124620-5555P00LF	2a	SHORT	SHORT	SHORT	SHORT	SEE NOTE 8
10124620-4555P00LF	2a	LONG	SHORT	SHORT	SHORT	SEE NOTE 8
10124620-5455P00LF	2a	SHORT	LONG	SHORT	SHORT	SEE NOTE 8
10124620-5545P00LF	2a	SHORT	SHORT	LONG	SHORT	SEE NOTE 8
10124620-5554P00LF	2a	SHORT	SHORT	SHORT	LONG	SEE NOTE 8
10124620-4554P00LF	2a	LONG	SHORT	SHORT	LONG	SEE NOTE 8
10124620-5455P00	2b	SHORT	LONG	SHORT	SHORT	NA

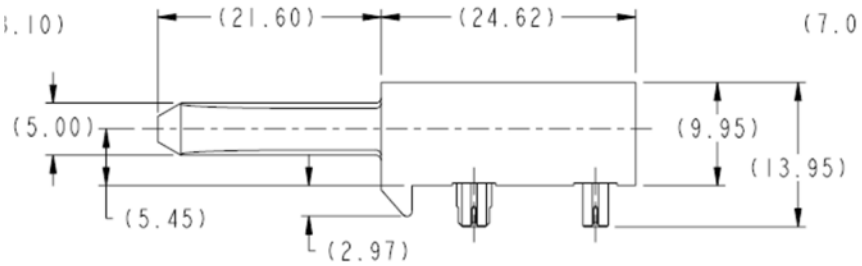
12.2.3 AirMax® Guide

1x pair of AirMax® 7.2mm R/A guide is used on the motherboard and mid-plane ONLY in the use cases that blind mate is needed.

12.2.3.1 Sled side

Sled side uses 1x AirMax® 7.2mm R/A guide blade as Figure 12-5.

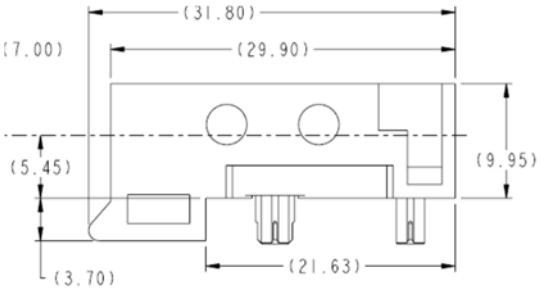
Figure 12-5 AirMax® 7.2mm R/A guide blade



12.2.3.2 Mid-plane side

Mid-plane side uses 1x AirMax® 7.2mm R/A guide socket as Figure 12-6

Figure 12-6 AirMax® 7.2mm R/A guide socket



12.2.4 AirMax® 3x8 signal

1x AirMax® VS/VS2 3x8 connectors follows the definition in Table 21 in the perspective of sled side. Signal description is shown in Table 22. It supports up to x16 PCIe plus side band signals and management interface. This connector is optional to support a high speed mid-plane with PCIe interface.

Table 21 Pin Definition for AirMax® 3x8 signal

10124755-111LF								
Signal 3x8 VS/VS2								
8		6		4		2		CONN_A2
GND	PCIE_TX_DP3	PMBUS_ALERT_N	MNG_TX_DP	FAN_TACH3	PCIE_RX_DP6	GND	PCIE_RX_DP0	A
PCIE_TX_DP0	PCIE_TX_DN3	PMBUS_DATA	MNG_TX_DN	FAN_TACH2	PCIE_RX_DN6	PCIE_RX_DP3	PCIE_RX_DN0	B
PCIE_TX_DN0	GND	PMBUS_CLK	GND	FAN_TACH1	GND	PCIE_RX_DN3	GND	C
GND	PCIE_TX_DP4	GND	MNG_RX_DP	FAN_TACH0	PCIE_RX_DP7	GND	PCIE_RX_DP1	D
PCIE_TX_DP1	PCIE_TX_DN4	PCIE_TX_DP6	MNG_RX_DN	MATED_IN_N	PCIE_RX_DN7	PCIE_RX_DP4	PCIE_RX_DN1	E
PCIE_TX_DN1	GND	PCIE_TX_DN6	MB_SLOT_ID0	FAN_PWM0	GND	PCIE_RX_DN4	GND	F
GND	PCIE_TX_DP5	GND	MB_SLOT_ID1	GND	COM_TX	GND	PCIE_RX_DP2	G
PCIE_TX_DP2	PCIE_TX_DN5	PCIE_TX_DP7	MB_SLOT_ID2	PCIE_CLK_100M_DP	COM_RX	PCIE_RX_DP5	PCIE_RX_DN2	H
PCIE_TX_DN2	GND	PCIE_TX_DN7	PCIE_PERST_N	PCIE_CLK_100M_DN	MB_ON	PCIE_RX_DN5	GND	I

Table 22 Pin description for AirMax® 3x8 signal

Signal	Type	Description
GND	Ground	Ground return
FAN_PWM[1..0]	Output	FAN PWM output from motherboard or uServer sled to mid-plane; OD output from motherboard.
FAN_TACH[3..0]	Input	FAN TACH input from mid-plane to motherboard or uServer sled. OD output at mid-plane. PU at motherboard or uServer sled needed.
PMBUS_DATA	Bi-direction	PMBus data line; 5V_AUX level.
PMBUS_CLK	Output	PMBus clock line; 5V_AUX level.
PMBUS_ALERT_N	Input	PMBus alert line; OD from mid-plane; low active; need PU at motherboard or uServer sled.
PCIE_PERST_N	Output	PCIe reset signal from motherboard to mid-plane. Low active. 3.3V push pull from motherboard.
PCIE_TX_DP/N[7..0]	Output	PCIe transmit signal from motherboard to mid-plane; AC decoupling at motherboard side.
PCIE_RX_DP/N[7..0]	Input	PCIe receive signal from mid-plane to motherboard; AC decoupling at mid-plane side.

PCIE_CLK_100M_DP/N	Output	100MHz PCIe clock from motherboard to mid-plane
MNG_TX_DP/N	Output	Management SGMII/KX transmit
MNG_RX_DP/N	Input	management SGMII/KX receive
MB_SLOT_ID[2..0]	Input	Slot location from mid-plane to motherboard or uServer sled. PD 100ohm or open at mid-plane to indicate different slot locations.
MB_ON	Output	Motherboard on indication from motherboard to mid-plane; Push Pull P3V3_STBY output at motherboard
COM_TX	Output	3.3V UART console TX from motherboard or uServer sled to mid-plane
COM_RX	Output	3.3V UART console RX from mid-plane to or uServer sled
MATED_IN_N	Input	Mated detection pin. E4 is a short, last mate pin on 3x6 and 3x8 AirMax® connector to indicate fully mating of sled. Follow Figure 12-7 for implementation at motherboard or uServer sled side and at mid-plane side. Fully mating of sled enables hot swap controller on motherboard or uServer sled. This action also notice mid-plane the presence of sleds.

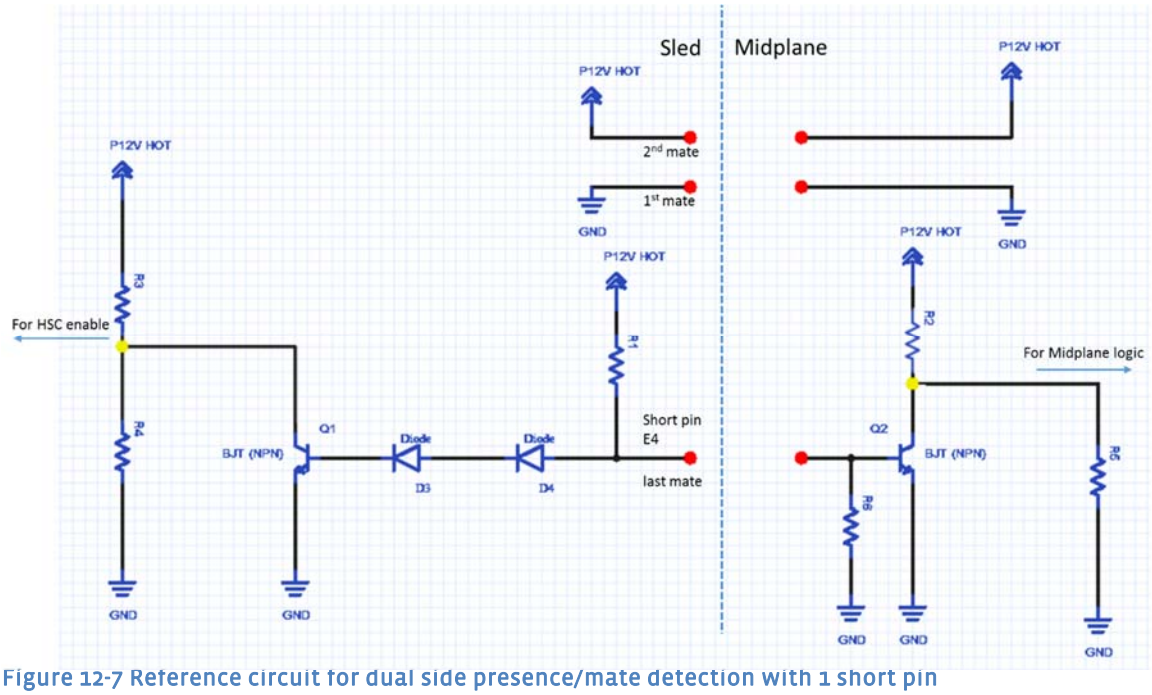


Figure 12-7 Reference circuit for dual side presence/mate detection with 1 short pin

12.2.5 Pressfit Cable

A pressfit cable is enabled for the use case of ORv2 and Cubby chassis. A side view is shown in Figure 14-8.

One side of the Pressfit cable is a pressfit power connector. The pressfit power connector is installed on motherboard with pressfit process, and secured by a screw for added strength. Pressfit power connector shares the same footprint and pin define as FCI/10124648-001LF.

The other side of the pressfit cable is a panel mount connector. It is installed on a sheet metal panel with tool-less install and removal.

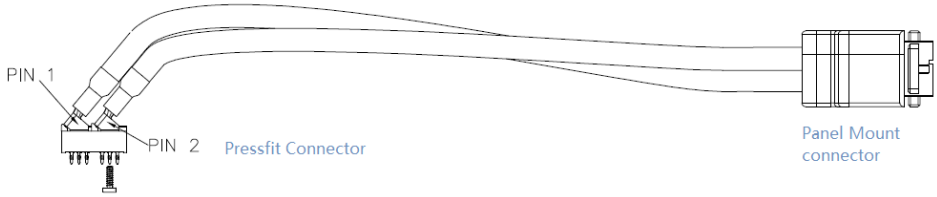
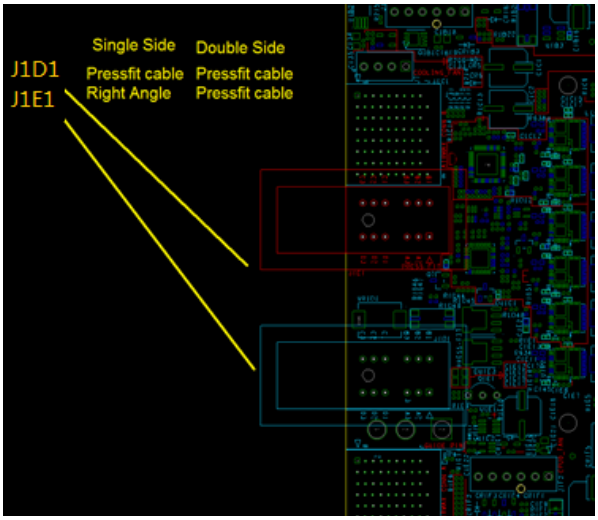


Figure 12-8: Pressfit Cable drawing-Side View

Pressfit cable is part of motherboard PCBA as a FRU; Pressfit cable cannot be replaced in the field.

Single side will have one pressfit cable on board while double side will have 2 pressfit cables to support more power. Vendor shall follow below BOM option for EVT and DVT.



12.3 Mid-plane

The motherboard design can support high speed mid-plane. The vendor shall design high speed mid-plane to validate the connection, signal integrity, power delivery and hot-swap. Mid-plane support is not Tioga Pass POR feature.

High speed mid-plane is mid-plane with power delivery, plus high speed interconnect on mid-plane. One or both of AirMax® 3x8 and AirMax® 3x6 need to be populated for this

use case. Mid-plane can have one of these optional active components: high speed signal switch, high speed signal repeater/buffer. Mid-plane can have no active components as well, if mid-plane only provide high speed interconnect with PCB trace.

Mid-plane provides mechanical and electric interface to DC power cable assembly. Each of the two slugs of DC power cable assembly is fixed to mid-plane through 2x screws. There is a notch feature on lug of DC power cable. Mid-plane should design a key feature to mate with this notch to provide foolproof design.

2x 80mm FANs are directly attached to 2x fan connectors on motherboard. The design of mid-plane should allow the replace of mid-plane without removing motherboard from tray.

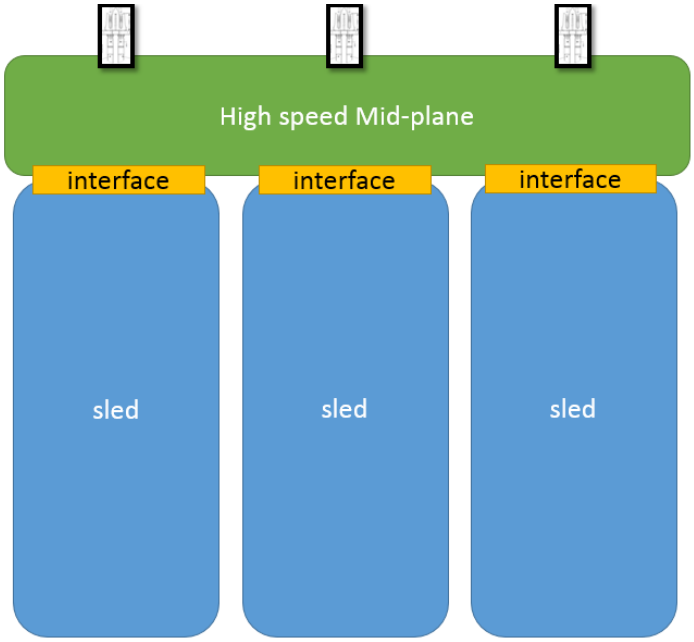


Figure 12-9 Sled with high speed mid-plane Layout

13 ORv2 Implementation

13.1 Cubby for ORv2

Figure 13-1 shows Cubby enclosure for Tioga Pass-ORv2 sled. Vendor should refer to 3D for more detail. Cubby serves as the mechanical and power delivery interface between ORv2 and Tioga Pass-ORv2 sled.

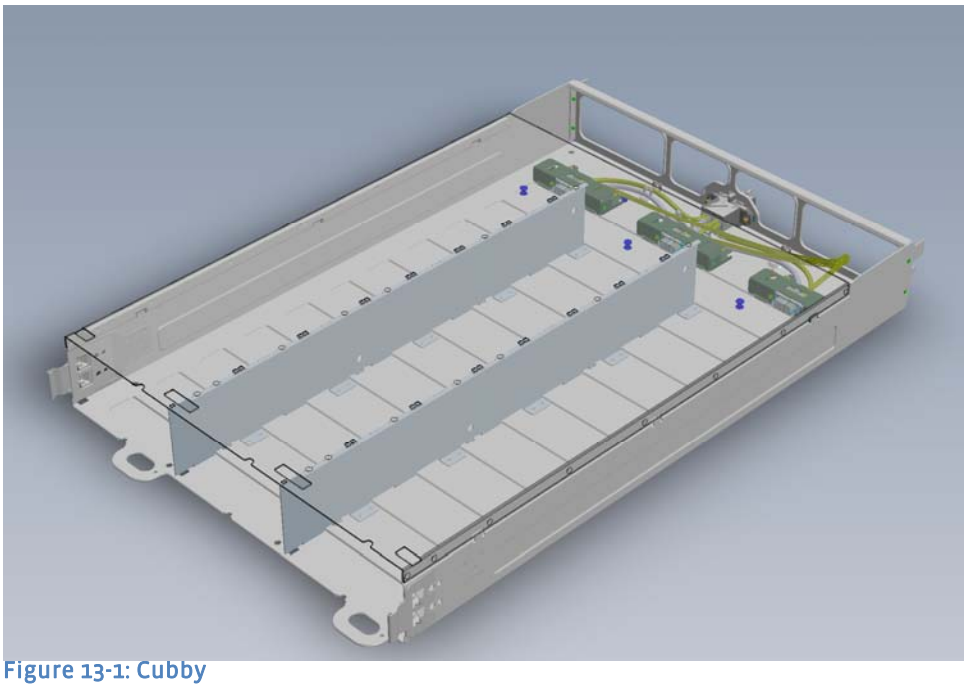


Figure 13-1: Cubby

13.2 Tioga Pass-ORv2 Power Delivery

There is one bus bar in each power zone of ORv2. There are up to 3x sleds in each cubby enclosure.

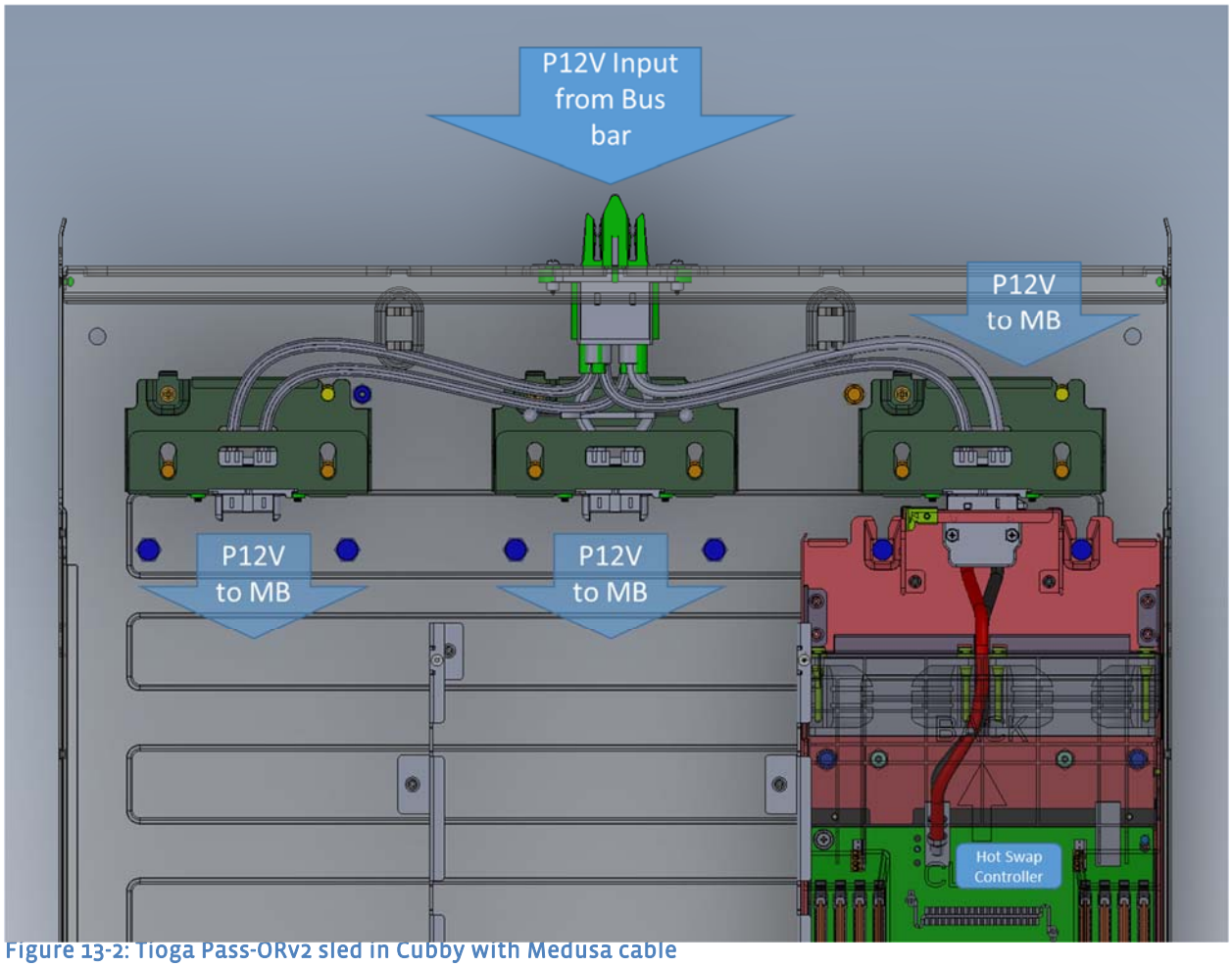
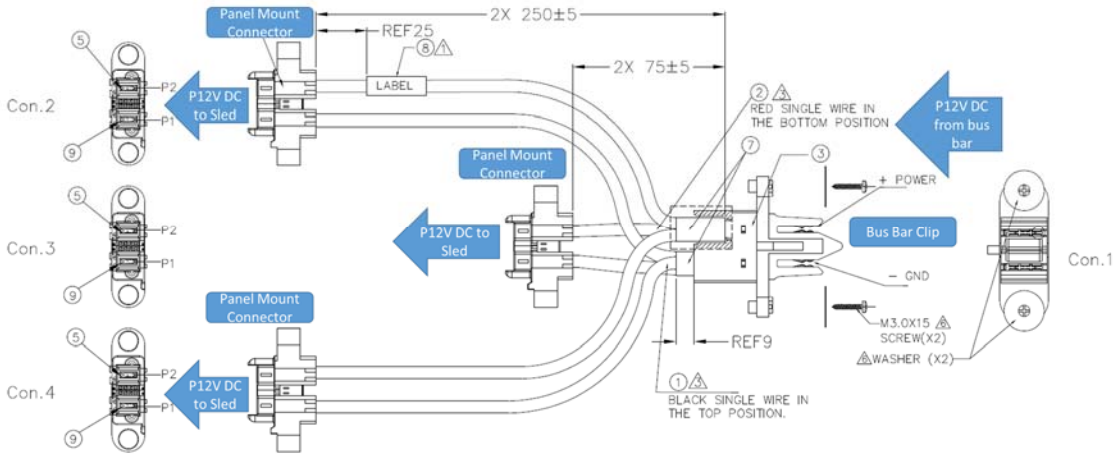


Figure 13-2: Tioga Pass-ORv2 sled in Cubby with Medusa cable

A Medusa cable (TE/2820303-2) is created to take DC power input from bus bar and delivery to each of the 3x sleds. Medusa cable delivers up to 40A to each node continuously in 65C local ambient considering pre-heating from sleds. One side of Medusa cable has a bus bar clip assembly to interface with bus bar. The other side of Medusa cable has 3 spited panel mount connectors with built in mechanical floating feature.

Figure 13-3: Medusa Cable drawing top view



Each motherboard for ORv2 sled has one or two pressfit cables installed on motherboard. Detail of the pressfit cable is described in section 12.2.5. One side of the pressfit cable is a pressfit connector that installed on the motherboard; the other side of the pressfit cable is a panel mount connector (shown in Figure 15-4) that interfaces with one of the panel mount connectors on medusa cable. Panel mount connector can be removed from the sled without using tools.

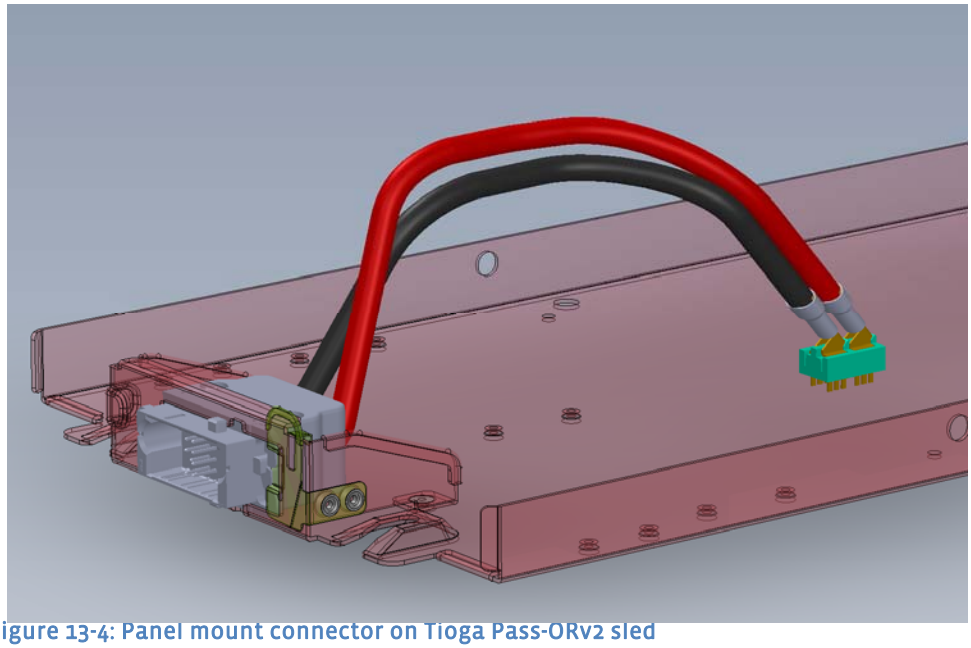


Figure 13-4: Panel mount connector on Tioga Pass-ORv2 sled

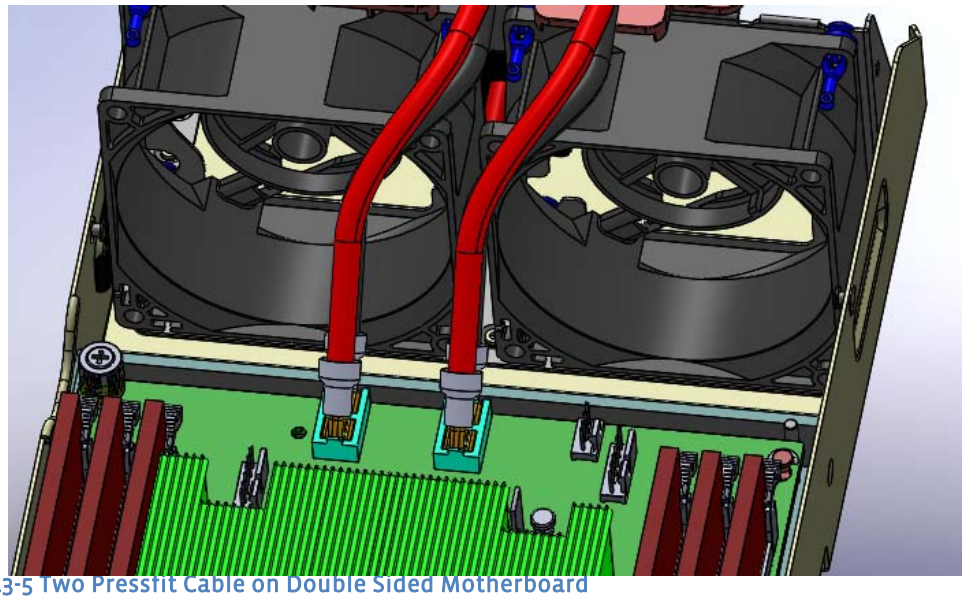


Figure 13-5 Two Pressfit Cable on Double Sided Motherboard

13.3 Tioga Pass-ORv2 single side sled

Single side Sled supports up to 12x DIMM slots. All DIMM slots are on the component side of motherboard.

A sheet metal tray serves as mechanical interface between the motherboard and Cubby. It provides mechanical retention for the components inside the tray, such as pressfit cable, fan, riser card, PCIe cards, Hard Drive, Mezzanine card. The combination of tray, motherboard with pressfit cable and the other components assembled in the tray is an Tioga Pass-ORv2 sled. Vendor should refer to 3D for more detail.

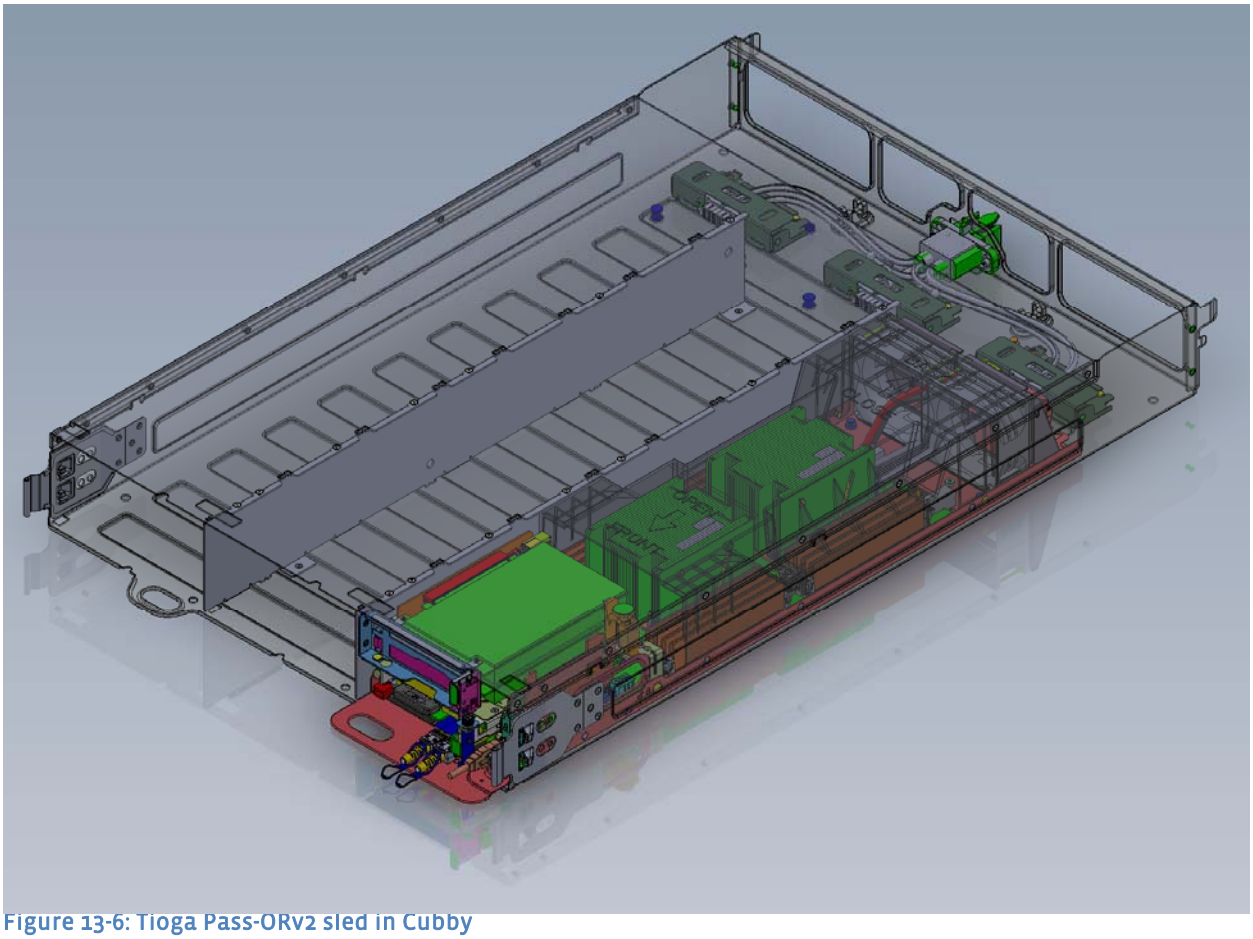


Figure 13-6: Tioga Pass-ORv2 sled in Cubby

13.4 Tioga Pass- ORv2 Double Side Sled

Double side Sled supports up to 24x DIMM slots. 12x DIMM slots are on the component side of motherboard and 12x DIMM slots are on the solder side of motherboard.

The sled provides mechanical support of motherboard and all other components in system. The sled allows DIMMs on both side of the system being serviced without removal of motherboard from sled. Vendor should refer to 3D for more detail.

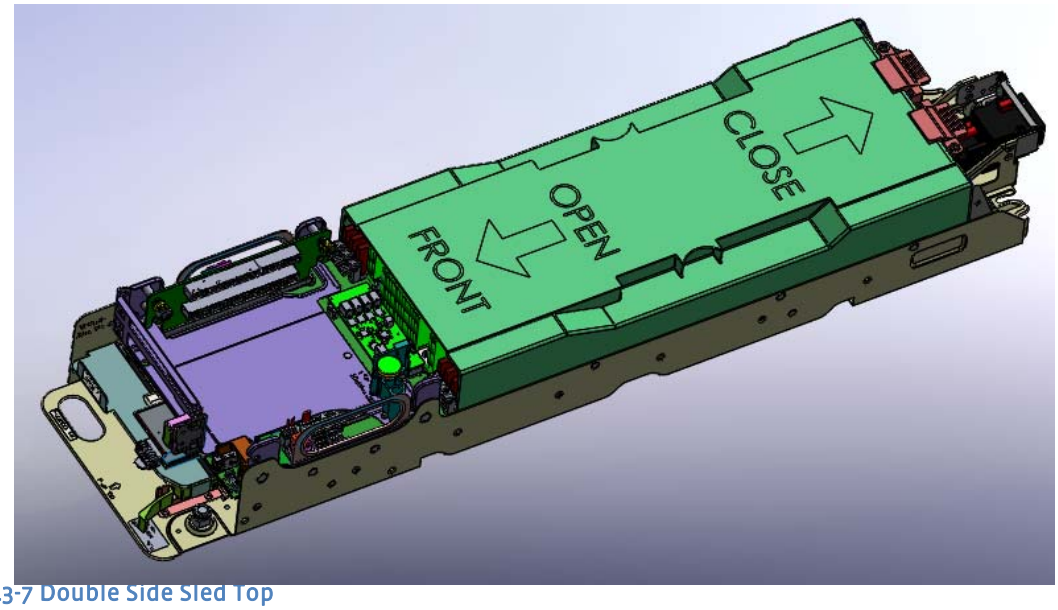


Figure 13-7 Double Side Sled Top

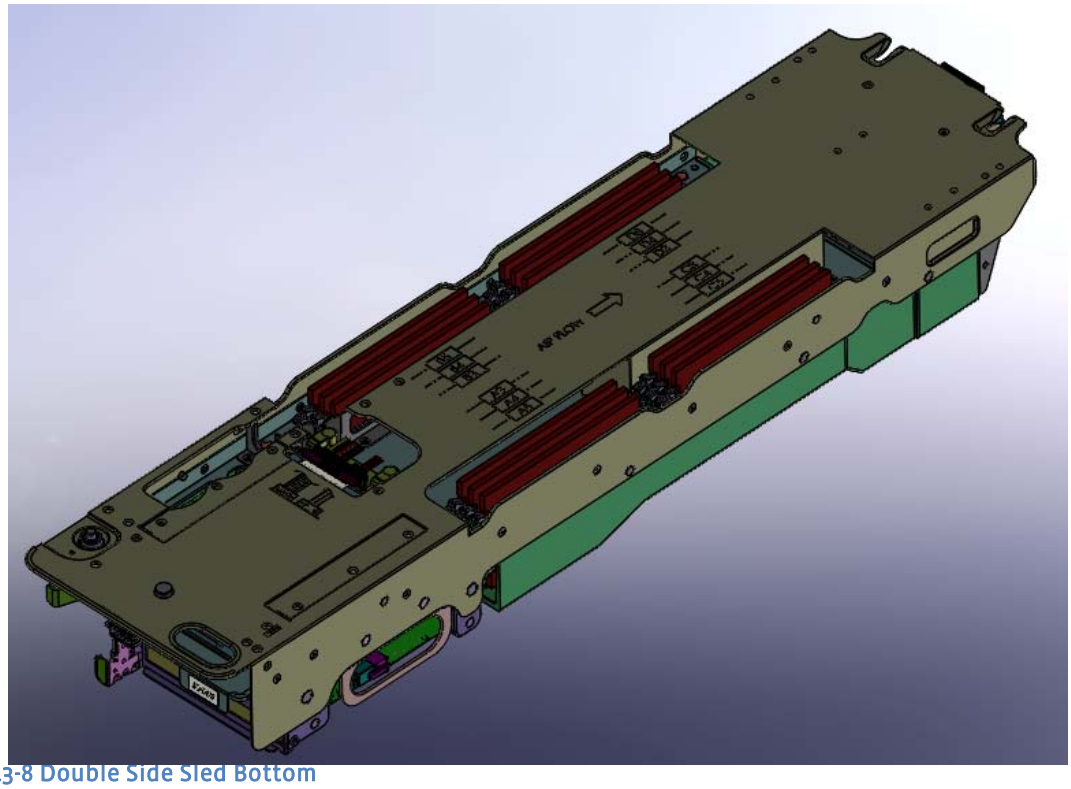


Figure 13-8 Double Side Sled Bottom

14 Mechanical

Tioga Pass-ORv2 sled should work with Open Rack V1 and ORv2 mechanically and the implementation guide provided in Chapter 13.

14.1 Single Side Sled mechanical

For Single side SKU, Tioga Pass shares most mechanical design and tooling as Leopard in ORv2 sled. The only difference to mechanical design and tooling is the standoffs' shoulder gap shall be adjusted to accommodate Tioga Pass PCB thickness (92mil).

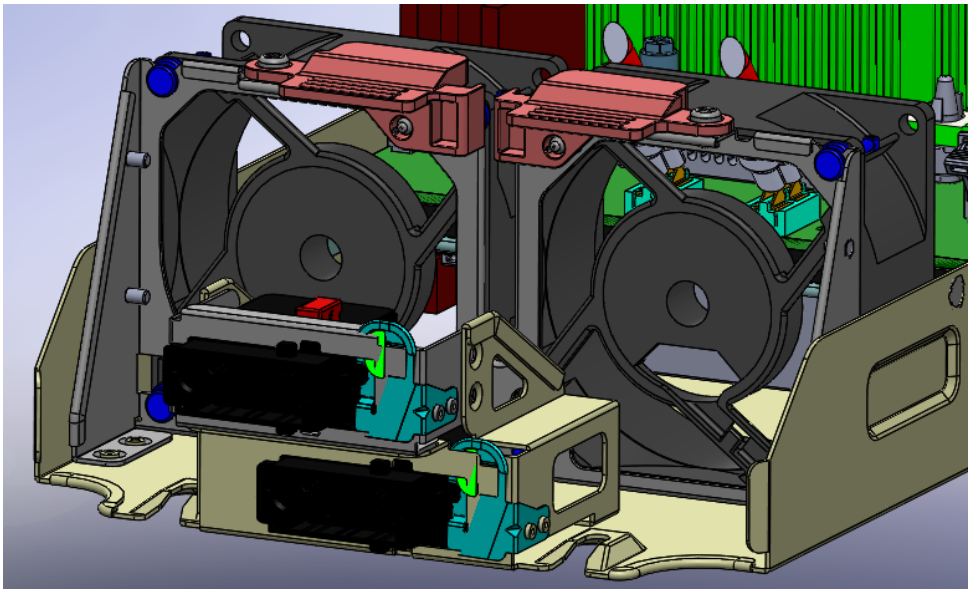
14.1.1 PCIe and HDD bracket

There is a metal bracket near I/O side of the tray to provide mechanical support for two full-height PCIe cards and a 3.5" hard drive or three full-height PCIe cards.

14.2 Double Side Sled mechanical

Double side sled for ORv2 is a new design. Motherboard sits in middle of the sled and has DIMM socket on both side. Please refer to 15.4 for details.

Double side motherboard has two pressfit cables , as a result, there shall be two connector holders stacked up to hold the pressfit panel mount connectors.



14.3 Fixed Locations

Refer to mechanical DXF file for fixed locations of mounting hole, PCIe x16 slot and power connector.

14.4 PCB Thickness

To ensure proper alignment of the motherboard and mid-plane interface within its mechanical enclosure, the motherboard should follow PCB stack up in Table 5-6 to have 92mil (2.34mm) PCB thickness. And mid-plane PCB thickness should also be 92mil (2.34mm). Mezzanine card and riser card PCB thickness should be 62mil (≈1.57mm).

14.5 Heat Sinks and ILM

The motherboard shall support heat sinks that are mounted according to the Intel® thermal mechanical specification and design guide. The vendor shall comply with all keep out zones defined by Intel® in the above referenced specification.

14.6 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 are listed in section 16.3.

- CPU0 / CPU1
- DIMM slot numbering, as described in Figure 11-11
- LEDs as defined in 11.9.2
- Switches as PWR and RST

14.7 DIMM Connector Color

Colored DIMM connectors shall be used to indicate the first DIMM of each memory channel. This first DIMM on each channel is defined as the DIMM placed physically furthest from its associated CPU. This DIMM connector shall be populated first when the memory is only partially populated. The First DIMM connector shall be a different color than the remaining DIMM connectors on the same memory channel.

14.8 PCB Color

Different PCB colors shall be used to help identify the motherboards revision. Table below indicates the PCB color to be used for each development revision.

Table 23 PCB Color

Revision	PCB Color
EVT	Red
DVT	Yellow
PVT	Green
MP	Green

15 Motherboard Power system

15.1 Input Voltage

15.1.1 Input voltage Level

The nominal input voltage delivered by the power supply is 12.5 VDC nominal at light loading with a range of 11V to 13V. The motherboard shall accept and operate normally with input voltage tolerance range between 10.8V and 13.2V when all under voltage related throttling features are disabled in section 15.2.

Motherboard's under-voltage protection level should be less than 10.1V.

15.1.2 Capacitive Load

Previous server generations required a maximum capacitive load of 4,000 uF. This requirement does not apply to Tioga Pass design. The motherboard design requires greater than 10,000 uF capacitive loading on P12V_AUX for supplying surge current from CPU VR, and reduce the slew rate of P12V_AUX decaying for NVDIMM feature at surprising power fail. Hot Swap Controller design should limit the inrush current to node during soft-start to less or equal to 10A.

15.1.3 P12V as AUX rail

There is only one 12V rail delivered to the motherboard as auxiliary power rail. Caution need to be taken to provide proper isolation to PCIe device, HDD, FAN, and all other devices in system, to meet voltage and timing requirement during running time and power on/off. The isolation circuit should have a soft start to avoid inrush current to P12V Aux rail, and prevent SOA damage of isolation MOSFET.

15.1.4 P1V_PSU to GND clearance

Due to P12V_PSU is without over current protection of Hotswap controller, modify P12V_PSU to GND and other shape based on these requirements.

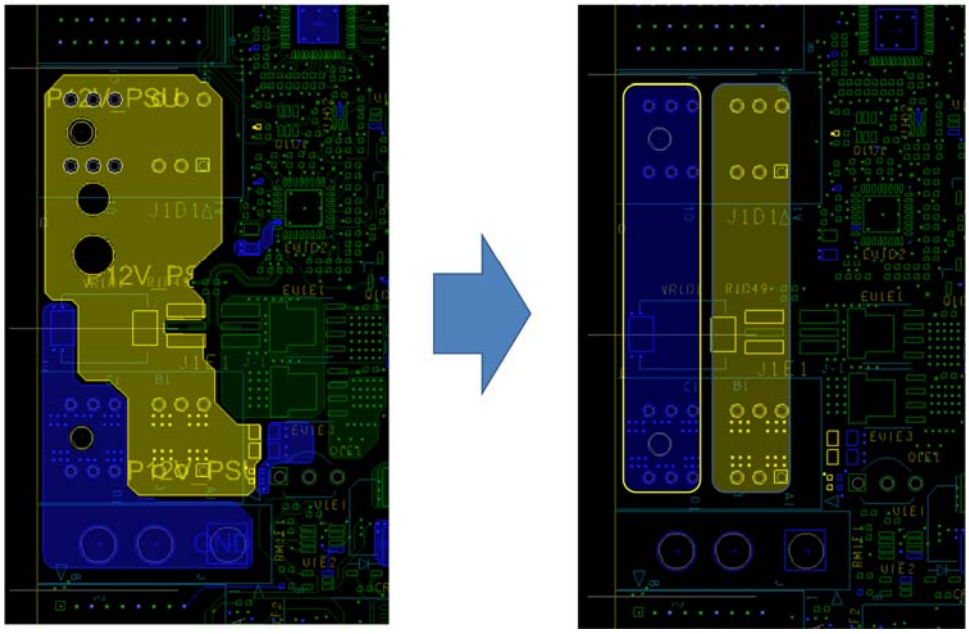
Minimal requirements:

- On same layer and adjacent layers, P12V_PSU shape to all other nets, including GND ≥ 40mil
- On different layers, from P12V_PSU shape to all other nets, including GND ≥ 2 layers of dielectrics if overlapping

Refereed practice if power delivery and board space allows:

- On same layer and adjacent layers, P12V_PSU shape to any other nets ≥ 80mil
- On different layers, from P12V_PSU shape to other nets has no overlap

1. 40mil min gap on same layer from P12V_PSU to other net
2. No adjacent layers have overlap of P12V_PSU and GND



P12V_PSU trace is needed to provide biasing for Hot swap controller and related circuit. Such trace must be ≤20mil, and has 40mil clearance to other signal on same layer. On adjacent layer, it is preferred to generate void in plane to provide clearance to P12V_PSU where there is no other tradeoff.

15.2 Hot Swap Controller (HSC) Circuit

In order to have a better control of 12.5V DC power input to each motherboard, one HSC (ADI/ADM1278) is used on the motherboard. HSC circuit provides the following functions:

- Inrush current control when motherboard is inserted and powered up.
- Current limiting protection for over current and short circuit. Over current trip point should be able to set to 53.9A and 41.4A with Iset jumper setting for single side motherboard; default is 53.9A. Over current trip point should be able to set to 71.5 and 78.9A with Iset jumper setting for double side motherboard; default is 71.5A.
- HSC UV protection shall be set to 10V~10.1V and OVP shall be set to 14.3V~14.4V.
- SOA protection during MOSFET turning on and off.
- HSC fault protection is set to latch off (default) with retry as stuff option.
- PMBUS interface to enable Intel® Manageability Engine and BMC following actions
    - Report server input power and log event if it triggers upper critical threshold.
    - Report input voltage (up to 1 decimal point) and log event if it triggers either lower or upper critical threshold.
    - Log status event based on hot swap controller's status register.
- Use HSC or external circuit to provide fast (<20us) over current sense alert to trigger system throttling and CPU fast PROCHOT#; Over current based fast PROCHOT# shall be controlled by HSC Iset jumper. Fast PROCHOT# threshold shall be slightly lower than HSC DC OCP set point to be useful. Feature can be disabled by BMC GPIO directly. BIOS has a setting to control Enable/Disable/ [no change]. No change is the default. This means follow the BMC initial setting. BMC sets it to disable as the default. Before BMC is ready, the hardware POR state is enable.
- Use HSC or external circuit to provide fast (<20us) under-voltage alert to trigger system throttling and CPU fast PROCHOT#. This feature is enabled by default with resistor option to disable. The threshold is set to 11.5V by default and with option to set it 11V. A jumper for UV_HIGH_SET is implemented together with BMC GPIOAA5(AST2500 pin T20) to control under voltage FPH trip point. When the jumper is at pin 1 and pin 3, trip point is 11.5v or follow BMC; when jumper is at pin 3 and 5, trip point is 11.0v.
- Use HSC or external circuit to provide fast (<20us) under-voltage alert to trigger system FAN throttling. This feature is disabled by default with resistor option to enable.
- Use HSC or external circuit to provide HSC timer alert to trigger system throttling before HSC OCP happens.
- Please refer to Table 24 for setting requirements of System, CPU, and memory sub-system throttling.

Table 24 Entry point of System, CPU, and Memory Sub-system Throttling

Condition	Threshold	Action	Enable control	Default

Board input power over current limit	>41.4A or 53.9A[Default] by jumper setting for SS >71.5A[Default] or 76.8A for DS	Trigger throttle to system in < 20us	BMC GPIO	Disable
Board input power under voltage	<11.5V/11V	Trigger throttle to system in < 20us	Resistor option	Enable
Board input power under voltage	<11.5V/11V	Disable MOSFET between P12V_AUX to P12V_FAN	Resistor option	Disable
Board input power under voltage	<10.5V	Disable MOSFET between P12V_AUX to P12V_FAN	Resistor option	Enable
HSC Timer Alert	>400mV[11]	Trigger throttle to system in < 20us	Resistor option	Enable
CPU VR hot	Determined by CPU VR design	Trigger throttle to PROCHOT in < 20us	N/A	Enable (always)
Memory VR hot	Determined by Memory VR design	Trigger throttle to MEM_HOT in < 20us	N/A	Enable (always)

- The voltage drop on HSC current sense resistor should be less or equal to 25mV at full loading. Hot swap controllers should have SMBUS address set to 0x45 (7bit format) with 0.25mohm Rsen on
- The power reporting of hot swap controller needs to be better than +/-2% from 50W to full loading in room temperature as a minimal requirement. Vendor shall optimize HSC power reporting by taking measurement on multiple samples and using firmware to apply different offset based on system loading and temperature. Example as below:

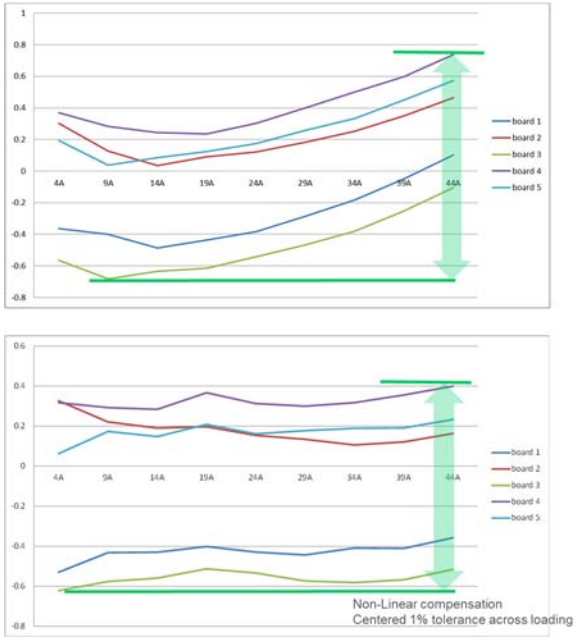


Figure 15-1 HSC power reporting

[11] Based on ADM1278 Timer threshold=1V for over-current protection

15.3 CPU VR

15.3.1 CPU Maximum Power

The motherboard shall be designed to handle a processor with a maximum TDP of 165W for Intel® Xeon® Scalable processor family (aka Skylake-SP). Support for processors of higher TDP is not required. As a result the vendor shall optimize the CPU VR accordingly.

15.3.2 CPU VR Optimizations

CPU VR optimizations shall be implemented to remove cost and increase the efficiency of the power conversion system. Vendors shall only use the minimum number of total phases to support the maximum CPU power defined in 15.3.1. CPU VR should have auto phase dropping feature, and run at optimized phase count among 1, 2, 3,..., and maximum phase count. CPU VR should support all Power States to allow the VRM to operate at its peak efficiency at light loading.

CPU VR should be compliant to latest VR13 specification and validation method and pass test with margin.

15.3.3 CPU VRM Efficiency

For CPU efficiency measurement:

- VID is set to 1.8 and 1.6 V (2x tests)
- Vin is set to 12.5V
- Efficiency is measured from input inductor to socket
- Driver and controller loss should be included
- Output voltage is gathered from Vsense of socket
- No additional air flow shall be supplied to the VR area other than the air flow caused by VRTT tool FAN
- Test is done in room temperature(20°C~25°C)
- Voltage measurement shall be done by tool and method with 0.05% accuracy or better
- Current measurement shall be done by tool and method with 0.25% accuracy or better
- Efficiency should be over 92% at 10A-100A load

Below is the efficiency test set up process and efficiency curve for reference:

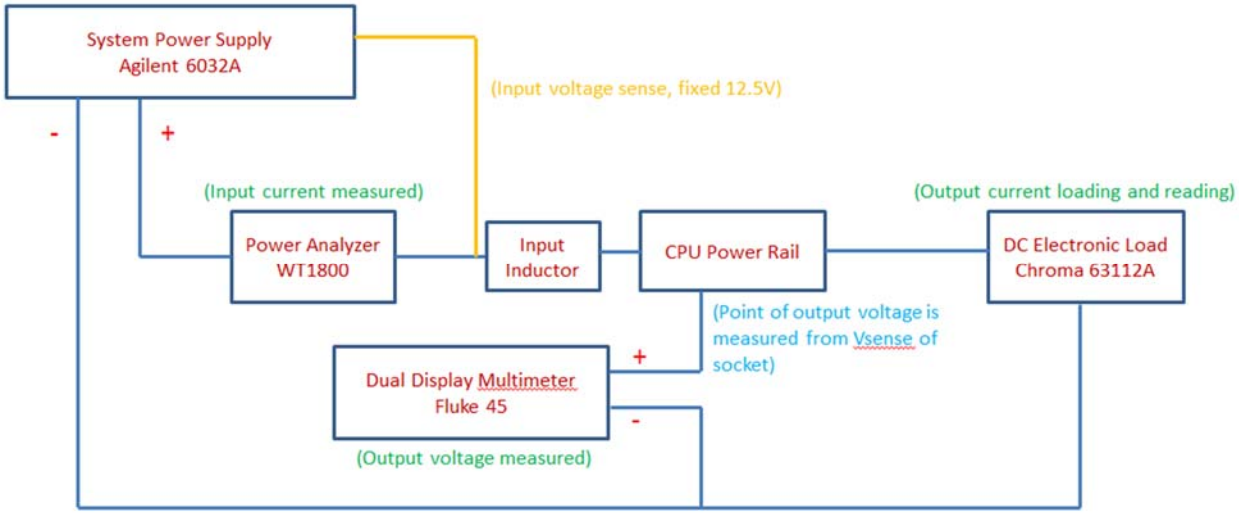


Figure 15-2 CPU VRM Efficiency test set up

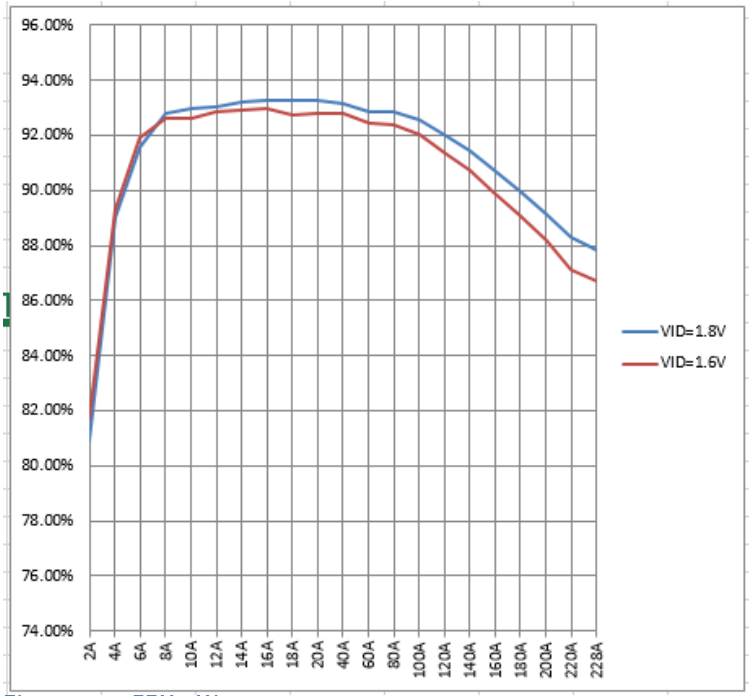


Figure 15-3 CPU effiency curve

Vendors are encouraged to exceed the above efficiency requirement and may propose higher efficiency VRMs that may come at additional cost. Power efficiency measured from 12.5V input to CPU socket should also be analyzed and improved.

15.3.4 CPU core VR configuration

Vendor should use CPU core VR solution with all configurations stored in NVRAM without any external resistor strapping. Vendor should provide utility under CentOS to perform VR configuration change. Configuration change should take effect without AC cycling node. The guaranteed rewrite count of NVRAM should be greater or equal to 15.

15.4 DIMM VR

15.4.1 DIMM Maximum Power

The motherboard has a DIMM configuration of 2 CPU sockets, 6 channels per socket, and 2 Slots per channel. Vendor should follow memory controller vendor's guideline to design and validate DIMM power rail to support maximum power needed for this configuration, and support 1.2V DDR4 DIMM.

15.4.2 DIMM VR Optimizations

DIMM VR should support auto phase dropping for high efficiency across loading. DIMM VR should be compliant to latest VR12.5 specification and memory controller vendor's updated validation guideline, and pass test with margin.

Vendor shall have different BOM options in VR area to optimize for Single side board with 12x DIMM slots and Double side board with 24x DIMM slots.

15.4.3 DIMM VR Efficiency

For DIMM VR efficiency measurement

- VID is set to 1.20V
- Vin is set to 12.5V
- Efficiency is measured from input inductor to PCB near DIMM sockets
- Driver and controller loss should be included
- Output voltage is gathered from PCB at middle of the 3rd and 4th of 6x DIMM slots
- No additional air flow shall be supplied to the VR area
- Test is done in room temperature(20°C~25°C)
- Voltage measurement shall be done by tool and method with 0.05% accuracy or better
- Current measurement shall be done by tool and method with 0.25% accuracy or better
- Efficiency curve shall be higher than the envelope defined below

Below is memory VDDQ power rail efficiency test set up process and efficiency curve for reference:

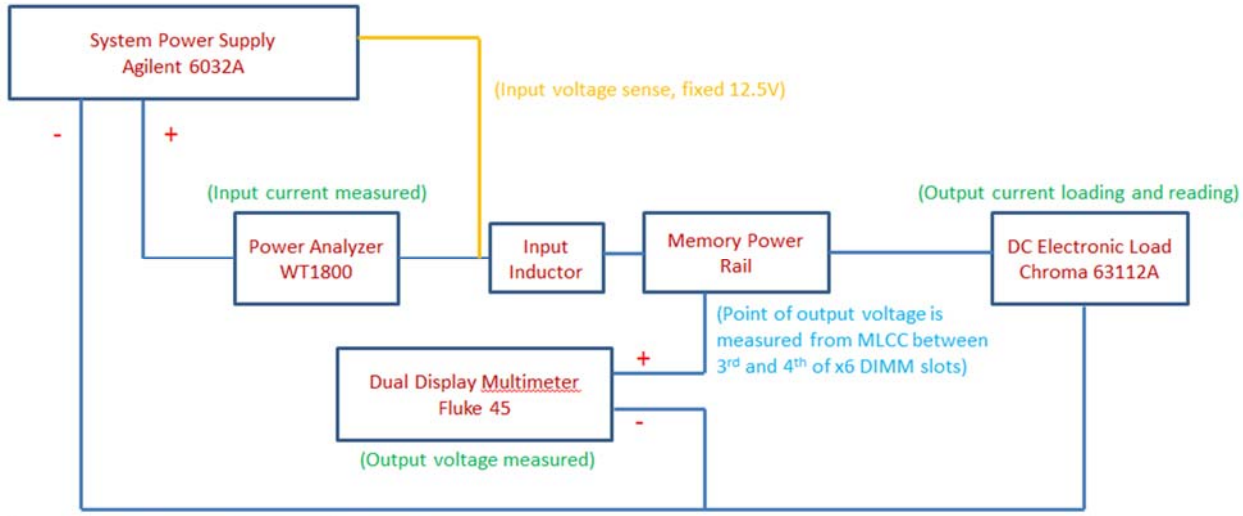


Figure 15-4 Memory VDDQ Efficiency Test setup

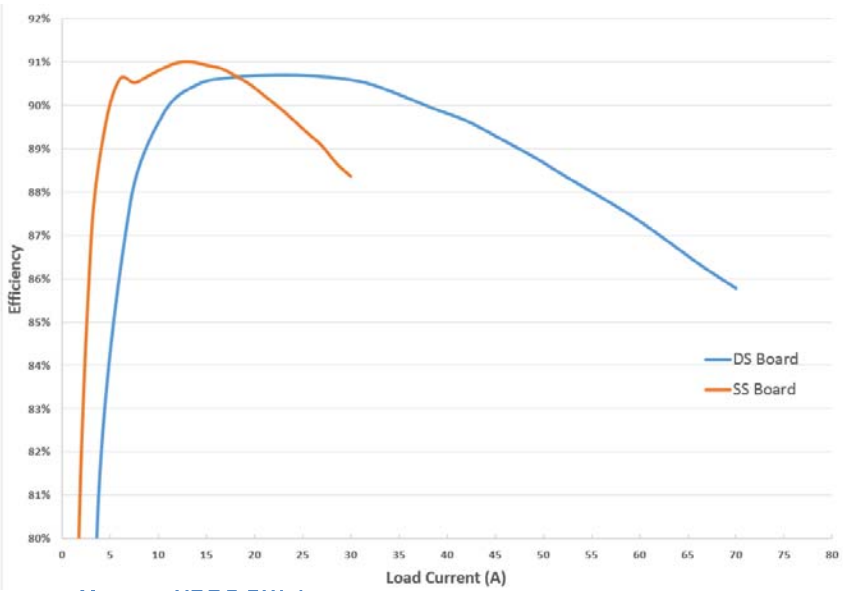


Figure 15-5 Memory VDDQ Efficiency curve

15.4.4 DIMM VR configuration

DIMM VR has same configuration requirement as CPU VR, listed in 15.3.4.

15.4.5 DIMM VR MLCC Co-Layout

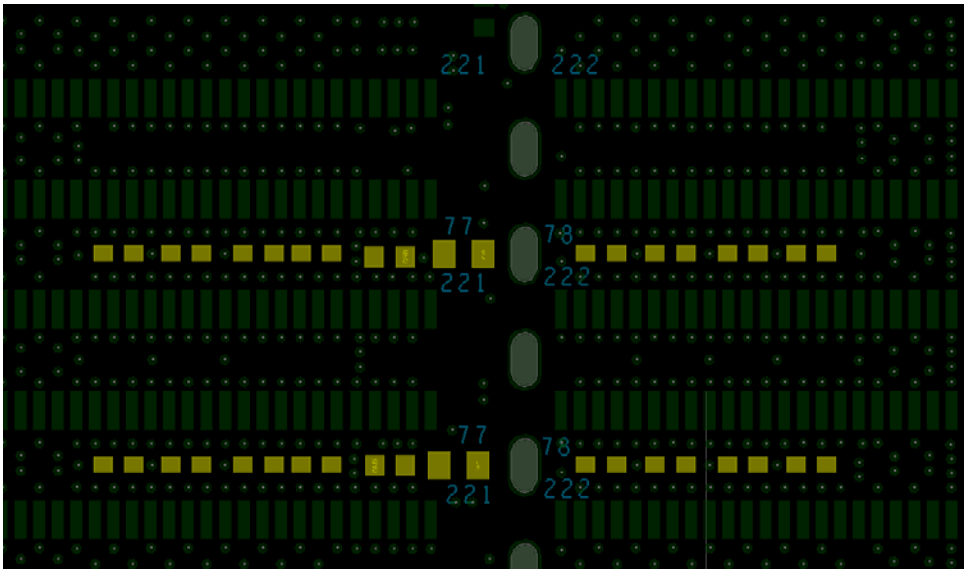
There are 18x C0603 and 2x C0805 for each DDR4 Memory VR on top layer of board between DIMM fields. It is shown in the picture below.

This change is to add 18x C0603 and 2x C0805 MLCC footprints for each DDR4 Memory VR, and place on bottom side. The added MLCC shall be placed on the exact same location as the corresponding MLCC on top layer, using same footprint.

The added MLCC will only be populated for single sided board since they conflict with bottom side SMT DIMM sockets.

Vendor shall check DFM for both cases:

1) Double side board, populate 24 DIMM sockets on both sides. No added MLCC specified in this section is populated
2) Single side board, populate 12 DIMM sockets on top side. Added MLCC specified in this section are populated



15.5 MCP (Multi Chip Package) VRM

There is 1x Voltage Regulator Module for each CPU socket to supply the power rails needed for MCP Power.

For board that does not need MCP support and Intel® Omni-Path Architecture (Intel® OPA) support, MCP VRM and the connector for MCP VRM is not installed. All system functions other than MCP and Intel® Omni-Path Architecture shall still be supported.

15.6 VRM design guideline

For VRM, vendor should list current budget for each power rail based on worst case loading case in all possible operation conditions. General requirements for VR component selection and VR design should meet 150% of this budget, and OCP should set to 200% of this budget. Vendors should do design check, inform purchasers about the actual OCP setting chosen for VRM and explain the reason if it cannot meet this general requirement above.

For VRM which requires firmware or power code or configuration file, vendors should maintain version control to track all the releases and changes between each version, and provide a method to retrieve version through application software during system run time. This software method should run under CentOS 7.x 64-bit with updated Kernel specified by customer.

All switching VRs should reserve testing hook for bode plot measurement.

CPU, DIMM and PCH VR power stages are listed as below:

VR rail	# of phases	Tioga Pass AVL1	Tioga Pass AVL2
PVCCIN_CPU0	5	TDA21470	FDMF3180
PVCCIN_CPU1	5	TDA21470	FDMF3180
PVSA_CPU0	1	TDA21460	FDMF3172
PVSA_CPU1	1	TDA21460	FDMF3172
PVCCIO_CPU0	1	TDA21460	FDMF3172

PVCCIO_CPU1	1	TDA21460	FDMF3172
PVDDQ_ABC	1 for SS/2 for DS	TDA21470	FDMF3180
PVDDQ_DEF	1 for SS/2 for DS	TDA21470	FDMF3180
PVDDQ_GHJ	1 for SS/2 for DS	TDA21470	FDMF3180
PVDDQ_KLM	1 for SS/2 for DS	TDA21470	FDMF3180
PVNN_PCH_STBY	1	TDA21460	FDMF3172
P1V05_PCH_STBY	1	TDA21460	FDMF3172

- Different BOM options and VR firmware are allowed to accommodate AVL with exceptions. Vendor's manufacture process shall be able to handle different BOM, matching AVL of Power Stage.
- Different BOM options are limited to these components
  - RC snubber in switching node
  - Pull low resistor at Pin 37 OCSET
  - 1K/0.1%ohm serial resistor and 22pF decoupling cap between pin 38 IOUT and pin 39 REFIN
  - 0.1uF decoupling cap between Pin 39 REFIN and GND
  - 1000pF decoupling cap at Pin 36 TOUT_FLT
  - 0Ohm Boost resistor at pin 32 PHASE
- Different BOM options do not apply to all other components
  - Input/Output Inductor/Capacitor shall not have BOM dependency to power stage AVL
- If different VR firmware is required to support different power stages, VR firmware shall have unique ID in user specific area matching each power stage.

15.7 Hard Drive Power

The motherboard shall supply power to all possible 9 hard drives connected. This means to support 1A continuous per HDD on 12.5VDC power rail, and 0.75A continuous per HDD on 5VDC power rail. In-rush current required to spin up the drive must also be considered in power delivery design. Both 12V and 5V disk output power rails shall protect against shorts and overload conditions.

For 1 individual SATA ports, power will be delivered through a combo SATA connector, ALLTOP C18625-1133-L or equivalent. The pin assignment shall follow description in Table 25.

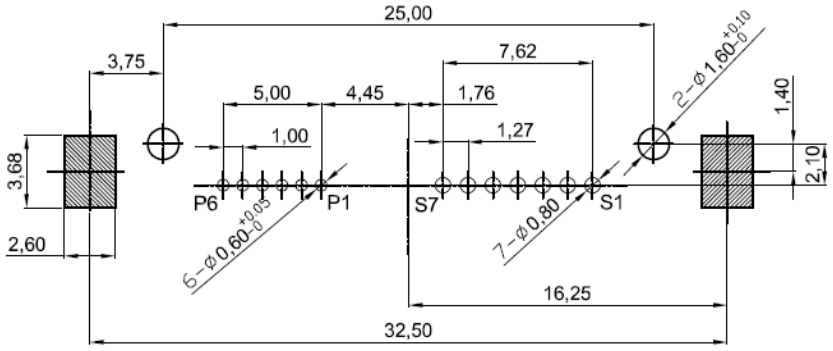


Figure 15-6 13-pin HDD power/signal combination connector

Table 25 13-pin HDD power/signal combination connector pin definition

Pin	Description
P1	+5VDC
P2	+5VDC
P3	GND
P4	GND
P5	+12VDC
P6	+12VDC
S1	GND
S2	SATA_TX_P
S3	SATA_TX_N
S4	GND
S5	SATA_RX_N
S6	SATA_RX_P
S7	GND

15.7.1 Spin-up Delay

When hard drive spins up after power on, it draws excessive current on both 12V and 5V. The peak current may reach 1.5A ~ 2A range in 12V. System may have up to 9 hard drives installed, so there is need to spin up hard drive in sequence. BIOS should implement 5 seconds delay between each hard drive spinning up. In order to do this, SATA hard drive's power cable should have pin 11 as NC (No Connection) to enable hard drive's spin-up delay function.

15.8 System VRM efficiency

Vendors shall supply high efficiency VRMs for all other voltage regulators over 20W not defined in this specification. All other voltage regulation modules shall be 91% efficiency over the 30% to 90% load range. Vendors are encouraged to deliver systems with higher efficiencies. If higher efficiencies are available at additional cost vendors shall present those options.

15.9 Power On

Motherboard should be set to restore last power state during AC on/off. This means that, when AC does on/off cycle, motherboard should power on automatically without requiring power button. Only when motherboard is powered off on purpose, then motherboard should be kept power off through AC on/off.

15.10 High power use case

High power use case means system power is between 480W (40A@12V) and 960W (80A@12V). This is not a typical FB use case, but need to have support to enable testing of such configuration. Typically it is caused by fully populating NVDIMM, high TDP CPU, and heavy load on PCIe slots, or a combination of the above.

Motherboard design and power delivery shall allow such use case with BOM change below:

- Populate both Power connectors with pressfit cables
- Change Rsen of HSC from 2x 1mOhm to 2x 0.5mOhm

Vendor shall perform simulation during design, and testing during validation for high power kit.

16 Environmental and Regulations

16.1 Environmental Requirements

The motherboard shall meet the following environmental requirements:

- Gaseous Contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Ambient operating temperature range: -5C to +45C
- Operating and Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40C to +70C
- Transportation temperature range: -55C to +85C (short-term storage)

The full system shall meet the following environmental requirements:

- Gaseous Contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Ambient operating temperature range: -5C to +35C
- Operating and Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40C to +70C
- Transportation temperature range: -55C to +85C (short-term storage)
- Operating altitude with no de-ratings: 1000m (3300 feet)

System would be deployed into datacenter with following environment.

Site 1 as

- Temperature: 65F to 85F
- Humidity: 30% to 85%
- Altitude: 1000m (3300 feet)

Site 2 as

- Temperature: 65F to 85F
- Humidity: 30% to 85%
- Altitude: 300m (1000 feet)

16.2 Vibration & Shock

The motherboard shall meet shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) & IEC721-3-(*) Standard & Levels, the testing requirements are listed in Table 26. The motherboard shall exhibit fully compliance to the specification without any electrical discontinuities during the operating vibration and shock tests. No physical damage or limitation of functional capabilities (as defined in this specification) shall occur to the motherboard during the non-operational vibration and shock tests.

Table 26 Vibration and Shock Requirements

	Operating	Non-Operating

Vibration	0.5g acceleration, 1.5mm amplitude, 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, 3mm amplitude, 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

16.3 Regulations

Vendor need to provide CB reports of the motherboard and tray in component level. These documents are needed to have rack level CE. The sled should be compliant with RoHS and WEEE. The motherboard PCB should have UL 94V-0 certificate. The vendor should design an EMI panel kit and pass FCC Class A.

17 Labels and Markings

The motherboard shall include the following labels on the component side of the motherboard. The labels shall not be placed in a way, which may cause them to disrupt the functionality or the air flow path of the motherboard.

Description	Type	Barcode Required?
MAC Address. One per network interface[12]	Adhesive label	Yes
Vendor P/N, S/N, REV (Revision would increment for any approved changes)	Adhesive label	Yes
Vendor Logo, Name & Country of Origin	Silk Screen	No
PCB vendor Logo, Name	Silk Screen	No
Purchaser P/N	Adhesive label	Yes
Date Code (Industry Standard: WEEK / YEAR)	Adhesive label	Yes
RoHS compliance	Silk Screen	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 recycle 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.	Silk Screen	No
CE Marking	Silkscreen	No
UL Marking	Silkscreen	No
Vendor Asset Tag[13]	Adhesive label	Yes

[12] MAC label for LOM is on motherboard; MAC label for NIC is on NIC.

[13] Work with purchaser to determine proper placement (if an asset tag is necessary)

18 Prescribed Materials

18.1 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

18.2 Capacitors & Inductors

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

- Only Aluminum Organic Polymer Capacitors shall all be used they must be rated 105C, and shall be selected only from Japanese Manufacturers.
- All capacitors will have a predicted life of at least 50,000 hours at 35C inlet air temperature, under worst conditions.
- Tantalum capacitor using manganese dioxide cathode is not allowed.
- SMT Ceramic Capacitors with case size > 1206 are not preferred. Vendor shall discuss with Facebook before using MLCC > 1206 case by case. Size 1206 still allowed when installed far from PCB edge, and with a correct orientation that minimizes risks of cracks.
- X7R Ceramics material shall be used for SMT capacitors by default. COG or NP0 type should be used in critical portions of the design. X6S can be used in CPU Cage area. Vendor shall discuss with Facebook before using X5R with evaluation of worst case temperature of the location.
- Only SMT inductors may be used. The use of through-hole inductors is disallowed.

18.3 Component De-rating

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

19 Reliability and Quality

19.1 Specification Compliance

Vendors must ensure that the motherboard meets these specifications as a stand-alone unit and while functioning in a complete server system. The vendor is ultimately responsible for assuring that the production motherboards conform to this Specification with no deviations. The Vendor shall exceed the quality standards demonstrated during the pilot build (PVT) while the motherboard is in mass production. Customer must be notified if any changes are made which may impact product quality.

19.2 Change Orders

Vendors must notify customer any time a change is made to the motherboard. A Specification Compliance Matrix will be submitted to customer for each revision of the motherboard including prototype samples.

19.3 Failure Analysis

Vendors shall perform failure analysis on defective units, which are returned to the vendor. Feedback shall be provided to customer with a Corrective Action plan within two weeks from the date, which the units were received at Vendor's Facility.

19.4 Warranty

The Vendor shall warrant the motherboard against defects and workmanship for a period of two years from the date of initial deployment at customer's facility. The warranty is fully transferable to any end user.

19.5 MTBF Requirements

The motherboard shall have a minimum calculated MTBF of 300K hours at 90% confidence level at 45C ambient temperature. The motherboard shall also demonstrate the MTBF requirement above by running at full load and 50% of time and performing AC cycling test 50% of time at 45C. Typical alternation period is 1 week for stress test and 1 week for AC cycling test. This MTBF demonstration shall finish prior to First Customer Shipment (Pilots samples, Mass Production units).

The motherboard shall have a minimum Service Life of 5 years (24 Hours / day, Full Load, at 45C ambient temperature).

Vendors shall provide a calculated MTBF number based on expected component life.

19.6 Quality Control

Below is a list of manufacturing requirements to ensure ongoing product quality:

- Incoming product must have less than 0.1% rejections
- PCBA yield(ICT &BFT) Target is 93.0% for MP 1-6months and vendor should continue to improve it to 94.0% for MP 6-12months and 95.5% after MP over 1 year
- L10 system yield target is 95.1% for MP 1-6 months and vendor should continue to improve it to 96.5% for MP 6-12months and 97.5% after MP over 1 year
- Vendors will implement a quality control procedure during Production, by sampling motherboards at random from the production line and running full test to prove ongoing compliance to the requirements. This process shall be documented and submitted prior to Production. The relative reports shall be submitted on an ongoing basis.
- Vendors will conduct an ongoing burn-In procedure for use in Production (Production will not start without an agreement on some sort of burn-in procedure). Vendors shall submit documentation detailing the burn in procedure.

19.7 Change Authorization and Revision Control

After the motherboard is released to mass production, no design changes, AVL changes, manufacturing process or materials changes are allowed without prior written authorization from customer. The AVL (Approved Vendor List) is defined by the list of components specified in the BOM (Bill of Materials).

Any request for changes must be submitted to customer with proper documentation showing details of the changes, and reason for the changes. This includes changes

affecting form, fit, function, safety, or serviceability of the product. Major changes in the product (or in the manufacturing process) will require re-qualification and/or re-certification to the Product. A new set of First Article Samples may be required to complete the ECO process. Any modifications after approval shall phase-in during production without causing any delays or shift of the current production schedule. Vendors shall provide enough advance notice to customer to prevent any discontinuation of production.

All changes beginning with the pilot run must go through a formal ECO process. The revision number (on the motherboard label) will increment accordingly. Revision Control: copies of all ECOs affecting the product will be provided to customer for approval.

19.8 PCB Tests

Server ODM should arrange Independent 3rd party lab testing on Delta-L, IST, and IPC-6012D for each motherboard, riser card and midplane PCB from every PCB vendors.

Server ODM cannot use the PCB vendor for these tests. Server ODM should submit reports for review and approval before a PCB vendor can be used in mass production. The testing lots should be manufactured at the same facility of a PCB vendor with same process that planned to be used by mass production.

Delta-L requires 3 different PCB fabrication lots from the PCB vendor, >10pcs coupons each time. Environmental shipping, packaging, and handling of this board is vital to test success; overnight shipping direct from PCB vendor to Delta-L independent lab is recommended.

IST is done once. It is recommended to IST test during DVT stage. It is required to be tested on a board manufactured at the same time as a board that completely passes Delta-L. (Run Delta-L, if it passes then ask the IST lab to run IST on the board they receive.) IST test profile is 3x cycles to 250°C and up to 1000x cycle to 150°C. Passing criteria is 150x cycles' average, and 100x cycles minimum for 35x coupons.

IPC-6012C is done when 2x of the 5x Delta-L tests passing from a PCB vendor (Passing at the independent test lab).

Server ODM should also request each PCB vendor to provide >10pcs Impedance coupon measurements and X-section check reports for each stage.

ODM should work with PCB house to implement Impedance, IST and Delta-L coupon to break off panel without increasing unit cost of PCB. These coupons will be on a breakoff on motherboard or working panel for riser cards.

19.9 Secondary Component

Secondary component planning should start from EVT and reach 80% of total number of BOM items in PCBA BOM in EVT. The rest of secondary component should be included in DVT.

It is recommended that PCB is planned with 3 vendors at EVT. EVT and DVT build plan should cover all possible combinations of key components of DC to DC VR including output inductor, MOSFETs and driver.

ODM should provide 2nd source plan and specification compare before each build stage.

20 Deliverables

20.1 OS Support

Motherboard shall support CentOS 7.x 64-bit with updated Kernel specified by customer, and pass Red Hat certification tests.

20.2 Accessories

All motherboard related accessories, including heat sink, back-plate and CPU socket protectors, should be provided and installed at the vendor's factory. All accessory boards including debug card, PCIe riser card, should be provided by the vendor.

20.3 Documentation

The vendor shall supply the following documentation to customer:

- Projection Action Tracker
- Bug Tracker
- Testing Status Tracker
- Design documents
    - Schematics for EVT, DVT and PVT(Cadence and PDF)
    - Board Layout EVT, DVT and PVT (Cadence and Gerber RS-274)
    - Board Design Support Documents:
        - System Block Diagram
        - Power distribution Diagram
        - Power and Reset Sequence Diagram
        - High Speed Signal Integrity Simulation, especially for DDR4 memory
        - Power Integrity Simulation, for important power rails such as CPU and DDR4 memory
        - SMBUS and JTAG Topology
        - GPIO Table for BMC and PCH
        - Hardware Monitor Topology
        - Clock Topology
        - Error Management Block Diagram
- BIOS Version plan, Version Tracker, and specification
- BMC Version plan, Version Tracker, and specification
- BMC Sensor Table
- Mechanical 2D Drawings (DXF and PDF)
- Mechanical 3D model (IGS or STEP, and EASM)
- BOM with MFG name, MFG P/N, Quantity, Reference Designators, Cost
- BOM in customer's defined format, whose definition is provided in separate file.
- Validation documents
    - Server Hardware Validation Items: Test Plan and Report
    - FAI test plan and Report
    - VR test Plan and Report
    - Signal Integrity Test Plan and Report

- Functional Test Report
- MTBF Test Plan and Report, including calculation
- System AVL(CPU, DIMM, PCIe cards, Mezzanine Cards, SSD) Qualification Test Plan and Report
- Reliability Test Plan and Report
- De-rating Report (worst conditions)
- 2nd source component Plan and Test Report
- Thermal Test Plan and Report (with indication of critical de-ratings, if any)
- Mechanical Test Plan and Report

20.4 Mass Production First Article Samples

Prior to final project release and mass production, the Vendor will submit the following samples and documentation:

- All the pertinent documentation described in section 20.3 and any other documents and reports, necessary for customer to release the product to mass Production.
- Pilot samples which are built in the allocated Facility for mass production.
- A full Specification Compliance Matrix
- A full Test/Validation Report
- Production line final Test 'PASS' tickets
- Samples which have passed the production burn-in process
- Samples shipped using the approved for production-shipping box described in section 21.

21 Shipping

The motherboard shall be shipped using a custom packaging containing multiple motherboards in each package. The quality of the packing assembly will be such that the motherboard will not get damaged during transportation. The units shall arrive in optimum condition and will be suitable for immediate use. Shock Test for the shipping box shall be conducted by the Vendor and submitted to customer for audit and approval.

22 Appendix

22.1 Appendix: Commonly Used Acronyms

This section provides definitions of acronyms used in the system specifications.

ANSI – American National Standards Institute

BIOS – basic input/output system

BMC – baseboard management controller

CFM – cubic feet per minute (measure of volume flow rate)

CMOS – complementary metal–oxide–semiconductor

DCMI – Data Center Manageability Interface

DDR4 – double data rate type 4

DHCP – dynamic host configuration protocol

DIMM – dual inline memory module

DPC - DIMMs per memory channel

DRAM – dynamic random access memory

ECC – error-correcting code

EEPROM - electrically erasable programmable read-only memory

EMI – electromagnetic interference

FRU – field replaceable unit

GPIO – general purpose input output

I²C – inter-integrated circuit

IPMI – Intelligent plattform management interface

KCS – keyboard controller style

LAN – local area network

LPC – low pin count

LUN – logical unit number

MAC – media access control

MTBF – mean time between failures

MUX – multiplexer

NIC – network interface card

OOB – out of band

ORv1 – Open Rack Version One

ORv2 – Open Rack Version Two

OU – Open Compute Rack Unit (48mm)

PCB – printed circuit board

PCIe – peripheral component interconnect express

PCH – platform control hub

POST – power-on self-test

PSU – power supply unit

PWM – pulse-width modulation

PXE – preboot execution environment

UPI – Intel® Ultra Path Interconnect

QSFP – Quad small form-factor pluggable

RU – rack unit (1.75")

SAS – serial-attached small computer system interface (SCSI)

SATA – serial AT attachment

SCK – serial clock

SDA – serial data signal

SDR – sensor data record

SFP - small form-factor pluggable

SMBUS – systems management bus

SMBIOS – systems management BIOS

SOL – serial over LAN

SPI – serial peripheral interface

SSD – solid-state drive

SSH – Secure Shell

TDP – thermal design power

TOR – top of rack

TPM – trusted platform module

U – Rack unit

UART – universal asynchronous receiver/transmitter

UEFI – unified extensible firmware interface

UL – Underwriters Laboratories

22.2 Mechanical drawings

Following mechanical drawings are provided

- 3D CAD for Tioga Pass-ORv2 Single Side sled (preliminary, subject to design change)
- 3D CAD for Tioga Pass-ORv2 Double Side sled

22.3 SMBIOS FRU mapping table

Tioga Pass SMBIOS and FRU Mapping v01								
SMBIOS Type 1~4 and 11				Map	FRU			Note
Type	Offset (0-base)	Field	BIOS Default		Area	Field	FRU Default(xxx.txt)	
System Information (Type 1)	04h	Manufacturer	[ODM name]		Product Info Area	Manufacturer Name	[ODM name]	
	05h	Product Name	[Tioga Pass Single Side/Tioga Pass Double Side]+Product PN/Model Name		Product Info Area	Product Name	[Tioga Pass Single Side/Tioga Pass Double Side]	
					Product Info Area	Part Number/Model Name	' '	
	06h	Version	To be filled by O.E.M.	←	Product Info Area	Product Version	' '	
	07h	Serial Number	To be filled by O.E.M.	←	Product Info Area	Product Serial Number	' '	
	SMBIOS type 1 didn't support Asset Tag.				Product Info Area	Asset Tag	' '	
	08h	UUID generated by AMI DMIEDIT utility.			N/A for FRU			BIOS send UUID to Management Controller during POST
Base Board Information (Type 2)	04h	Manufacturer	[ODM name]		Board Info Area	Board Manufacturer	[ODM name]	
	05h	Product	[Tioga Pass Single Side/Tioga Pass Double Side]	←	Board Info Area	Board Product Name	[Tioga Pass Single Side/Tioga Pass Double Side]	
	06h	Version	To be filled by O.E.M.	←	Board Info Area	Board Part Number	[board part number]	<ODM MFG filled>
	07h	Serial Number	To be filled by O.E.M.	←	Board Info Area	Board Serial Number	M1 ODM_DEFINE	<ODM MFG filled>
	SMBIOS type 2 didn't support this field.				Board Info Area	FRU File ID	[FRU file version, example Ver 0.01]	
	08h	Asset Tag	To Be Filled By O.E.M.	←	Product Info Area	Asset Tag	' '	
System Enclosure or Chassis (Type 3)	04h	Manufacturer	[ODM name]		Product Info Area	Manufacturer Name	[ODM name]	
	05h	Type	17	←	Chassis Info Area	Chassis Type	17	
	06h	Version	To be filled by O.E.M.	←	Chassis Info Area	Chassis Part Number	' '	
	07h	Serial Number	To be filled by O.E.M.	←	Chassis Info Area	Chassis Serial Number	M3 ODM_DEFINE	<ODM MFG filled>
	08h	Asset Tag	To be filled by O.E.M.	←	Product Info Area	Asset Tag	' '	
Processor Information (Type 4)	20h	Serial Number	To Be Filled By O.E.M.	←	Chassis Info Area	Chassis Extra	M3 ODM_DEFINE	<ODM MFG filled>
Processor Information (Type 4)	20h	Serial Number	To Be Filled By O.E.M.	←	Chassis Info Area	Chassis Extra	M3 ODM_DEFINE	<ODM MFG filled> <Present if 2nd CPU installed>
OEM Strings (Type 11)		String 1	To Be Filled By O.E.M.	←	Board Info Area	Board Extra: FB PCBA part number	[Facebook PCBA part number]	<ODM MFG filled>
		String 2	To Be Filled By O.E.M.	←	Product Info Area	Product Extra: FB L10 part	[Facebook L10 part number]	<ODM MFG filled>
		String 3	To Be Filled By O.E.M.	←	Product Info Area	Product Extra: Product Build	[Such as EVT1/EVT2/DVT]	<ODM MFG filled>
		String 4	To Be Filled By O.E.M.	←	Product Info Area	Product Extra: L10 build time	[Generate L10 build time, mfg. site if	<ODM MFG filled>
		String 5	Ppin Value		N/A for FRU			depend on each CPU
		String 6	Ppin Value		N/A for FRU			depend on each CPU
		String 7	PCH SKU		Board Info Area	PCH SKU	[PCH-4/PCH-T/PCH-x]	
		String 8 ~ 16	To Be Filled By User		N/A for FRU			"Name of Save as User Defaults" "CRC of Setup" <exists if user create it>