Facebook Server Intel Motherboard V4.0

Project Tioga Pass

Rev 0.30

Author:

Whitney Zhao, Hardware Engineer, Facebook

Jia Ning, Hardware Engineer, Facebook

1 Revision History

Table 1-1

Date	Name	Description
6/3/2015	Jia Ning	- Version 0.1 release
7/29/2016	Whitney Zhao	- Version 0.2 release - Updated Figure and Tables References - Update contents
1/31/2017	Whitney Zhao	- Version 0.3 release for OCP summit - Update contents - Minor corrections

© 2016 Facebook.

As of July 26, 2016, the following persons or entities have made this Specification available under the Open Compute Project Hardware License (Permissive) Version 1.0 (OCPHL-P), which is available at http://www.opencompute.org/.../spec-submission-process/.

Facebook, Inc.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, non-infringement, fitness for a particular purpose, or title, related to the Specification. The Specification implementer and user assume the entire risk as to implementing or otherwise using the Specification. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE."

2 Scope

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

3 Contents

1 Revision History ... 2
2 Scope ... 4
3 Contents ... 4
4 Overview ... 8
  4.1 Overview ... 8
  4.2 Open Rack Introduction ... 8
5 Physical Specifications ... 9
  5.1 Block Diagram ... 9
  5.2 Placement and Form Factor ... 9
  5.3 CPU and Memory ... 10
  5.4 Platform Controller Hub (PCH) ... 11
  5.5 PCIe Usage ... 12
  5.6 PCB Stack-Up ... 12
6 BIOS ... 16
  6.1 BIOS Chip ... 16
  6.2 BIOS Source Code ... 16
  6.3 BIOS Feature Requirements ... 16
7 PCH Intel® SPS Firmware Plan of Record ... 22
8 BMC ... 23
  8.1 Management Network Interface ... 23
  8.2 Local Serial Console and SOL ... 23
  8.3 Graphic and GUI ... 24
  8.4 Remote Power Control and Power policy ... 24
  8.5 Port 80 POST ... 24
  8.6 Power and System Identification LED ... 24

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

8.7 Platform Environment Control Interface (PECI) 25
8.8 Power and Thermal Monitoring and power limiting 25
8.9 SMBUS Diagram 25
8.10 Sensors/Events 25
8.11 SEL 33
8.12 FSC in BMC 34
8.13 OEM commands 35
8.14 BMC FW chip and Firmware Update 37
8.15 BMC Update Dual PCH flash 37
8.16 BMC Update and Access CPLD 39
8.17 BMC Time Sync 39
8.18 PCIe and Mezzanine card Thermal monitoring 39
8.19 BMC PPIN Implementation 40
8.20 BMC Average Power Reporting 40
8.21 BMC Access and Update VR Firmware 40
8.22 BMC MSR Dump from OOB 40
8.23 BMC Network Status 41
8.24 BMC Secure boot 41
9 Thermal Design Requirements 43
9.1 Data Center Environmental Conditions 43
9.2 Server operational condition 43
9.3 Thermal Kit Requirements 45
10 I/O System 47
10.1 PCIe x32 Slot/Riser Card 47
10.2 DIMM Sockets 54
10.3 Mezzanine Card 54
10.4 Network 59
10.5 USB 60
10.6 SATA 61
10.7 M.2 61
10.8 Debug Header 62
10.9 Switches and LEDs 66
10.10 Fan connector 67

10.11 TPM Connector and Module 68
10.12 Sideband Connector 69
10.13 VGA header 69
11 Rear Side Power, I/O and Midplane 70
11.1 Overview of Footprint and Population Options 70
11.2 Rear Side Connectors 70
11.3 Midplane 75
12 ORv2 Implementation 77
12.1 Cubby for ORv2 77
12.2 Intel Motherboard V4.0-ORv2 Power Delivery 77
12.3 Intel Motherboard V4.0-ORv2 Single-Side Sled 79
12.4 Intel Motherboard V4.0- ORv2 Double Side Sled 80
13 Mechanical 81
13.1 Single Side Sled mechanical 81
13.2 Double Side Sled Mechanical 81
13.3 Fixed Locations 81
13.4 PCB Thickness 81
13.5 Heat Sinks and ILM 81
13.6 Silk Screen 81
13.7 DIMM Connector Color 81
13.8 PCB Color 82
14 Motherboard Power System 83
14.1 Input Voltage 83
14.2 Hot-Swap Controller (HSC) Circuit 84
14.3 CPU VR 86
14.4 DIMM VR 87
14.5 MCP (Multi Core Package) VRM 89
14.6 VRM design guideline 89
14.7 Hard Drive Power 90
14.8 System VRM efficiency 91
14.9 Power On 92
14.10 High power use case 92

15 Environmental and Regulations 93
15.1 Environmental Requirements 93
15.2 Vibration & Shock 93
15.3 Regulations 94
16 Labels and Markings 95
17 Prescribed Materials 96
17.1 Disallowed Components 96
17.2 Capacitors & Inductors 96
17.3 Component De-rating 96
18 Reliability and Quality 97
18.1 Specification Compliance 97
18.2 Change Orders 97
18.3 Failure Analysis 97
18.4 Warranty 97
18.5 MTBF Requirements 97
18.6 Quality Control 97
18.7 Change Authorization and Revision Control 98
18.8 PCB Tests 98
18.9 Secondary Component 99
19 Deliverables 100
19.1 OS Support 100
19.2 Accessories 100
19.3 Documentation 100
19.4 Mass Production First Article Samples 101
20 Shipping 102
21 Appendix 103
21.1 Appendix: Commonly Used Acronyms 103
21.2 Mechanical drawings 104
21.3 SMBIOS FRU mapping table 104
21.4 Add-on-Card Thermal Interface Spec for Intel Motherboard V4.0 104

4 Overview

4.1 Overview

Intel Motherboard V4.0 (also referred to "motherboard" or "Tioga Pass" in this document, unless noted otherwise) is a derivative design of Intel's Cloud Reference board based on the next generation Intel® Xeon® processor (codenamed Skylake) microarchitecture. The motherboard supports double sideboard Stock Keeping Unit (SKU) and single side sled designs in Open Rack Version 2 (ORv2).

ORV2 is the 2013 OCP rack design, and the Intel Motherboard V4.0-ORv2 sled is the ORV2 compatible compute sled.

4.2 Open Rack Introduction

This chapter provides background information of ORV2. Details of the motherboard's electrical and mechanical interfaces to ORV2 are described in Chapter 14and Chapter 0.

4.2.1 Open Rack V2 Introduction

ORV2 has two 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 Power Supply Units (PSUs), 3x Battery Backup Units (BBU), and provides 6.3KW[2] continuous max loading through the single bus bar that the power zone it is attached to.

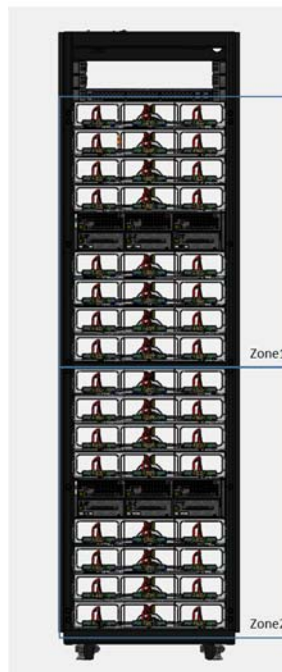


Figure 4-1 Open Rack V2 with 2x power zones

[2] 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 Intel Motherboard V4.0. The dashed lines are for reserved connections, dual layout, and the high-speed mid-plane option.

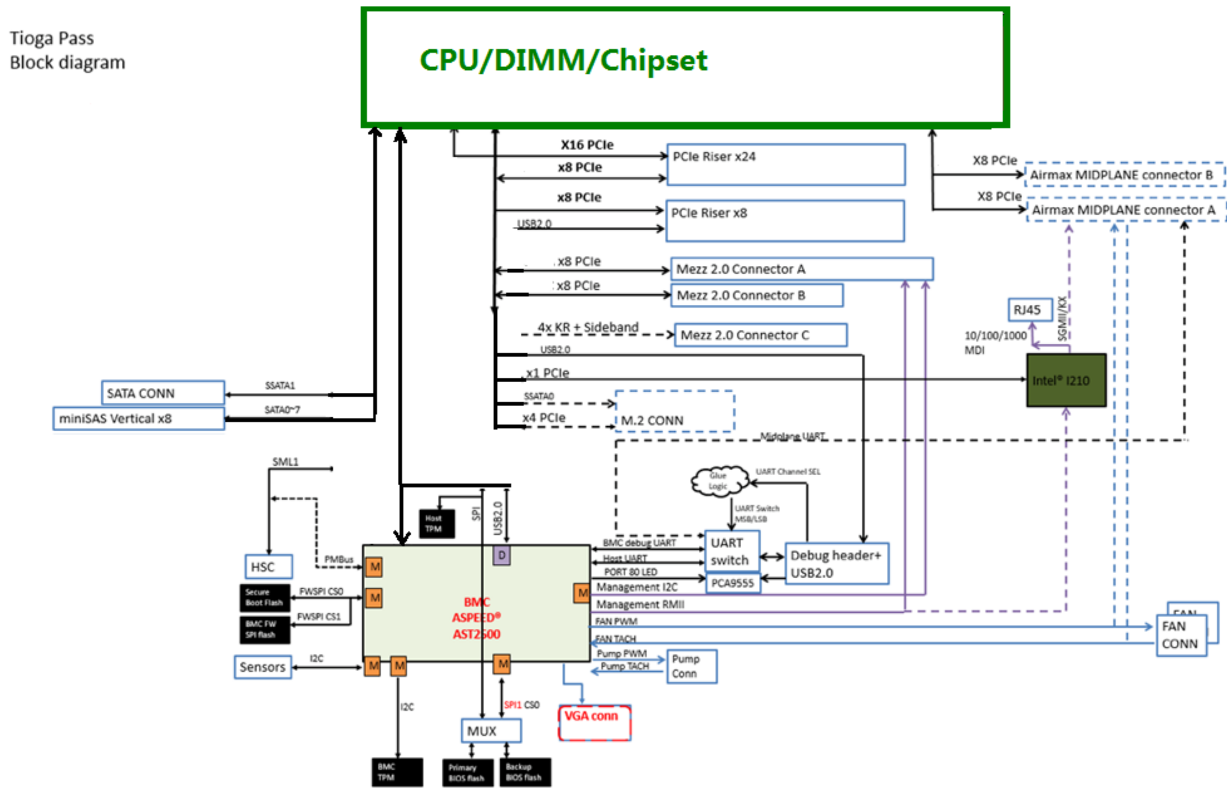


Figure 5-1 Intel Motherboard V4.0 Block Diagram

5.2 Placement and Form Factor

The board form factor is 6.5 inch by 20 inch (6.5"x20"). Figure 5-2 illustrates board placement and is meant to highlight the relative position of key components. Exact dimension and position information is available in DXF format for layout and 3D modeling. The vendor is responsible for completing component placement.

The layout and placement of the following components shall be followed strictly to specification:

- Form Factor
- Central Processing Unit (CPU) and DIMM socket location
- Peripheral Component Interconnect Express (PCI-e) slot position
- Front I/O port' position
- Power and reset button
- PCI-e Mezzanine card connector position
- Mid-plane connectors

- Battery connector
- Mounting holes

Other components can be shifted based on layout routing so long as the relative position is maintained.

The following internal connectors shall be placed as close as possible to the front of the board in order to allow easy access:

- (1) Vertical combo Serial ATA (SATA) signal and power connector
- (1) 14-pin Debug Card header
- (1) USB 3.0 Type-A connector, right angle
- (1) M.2 connector with 2280 and 22110 support
- (1) RJ45
- (1) USB Type-C
- (1) Proprietary VGA connector

The following midplane connector footprints shall be placed at the backside of the board to provide midplane connection co-layout:

- (2) AirMax VS2® 3x8
- (1) AirMax® Guide
- (2) AirMax® VS Power 2x2

Placeholder for placement

Figure 5-2 Intel Motherboard V4.0 Placement

Refer to DXF file for critical component placement.

5.3 CPU and Memory

5.3.1 CPU

The motherboard shall support the next generation Intel® Xeon® processor product family and provision the support of all future CPUs unless noted otherwise. The features listed below shall be supported by the motherboard:

- Support two next generation Intel® Xeon® processor product family processors
- Two, full-width Intel process links
- Single processor support mode

5.3.2 DIMM

The motherboard shall have a DIMM subsystem designed as described below:

- DDR4 direct-attach memory support for CPU0 and CPU1
- DDR4 registered memory interface on each CPU
- The farthest DIMM slot of each channel is on the component side of the Printed Circuit Board (PCB)
- The nearest DIMM slot of each channel is on the solder side of the PCB

- Follow the updated JEDEC DDR4 specification with 288-pin DIMM socket
- The board design shall allow Intel® Xeon® Processors to operate at maximum POR memory speeds

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 an NVDIMM to implement the Basic Input/Output System (BIOS) design.

The undervoltage 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 Platform Controller Hub (PCH)

The motherboard shall support following features:

- USB 2.0 ports:
    - (1) Type-A, front connector
    - (1) Type-C, front connector
    - (1) for in-band BMC firmware update
    - (1) for x32 riser connector
- Drive connectors:
    - (1) M.2 connector
    - (1) SATA 6Gbps port
    - (1) MiniSAS HD x8 port
    - (1) MiniSAS HD x4 port
- (1) PCIe x4 port to M.2 connector, co-layout with a SATA port to the M.2 connector
- Serial Peripheral Interface (SPI) mux with BMC to allow the capability for BIOS upgrade and recovery

- SPI interface for the Trusted Platform Module (TPM) header
- System Management Bus (SMBus) interface (master & slave)
- Intel® Server Platform Services (SPS) firmware with Intel® Node Manager
    - Platform Environment Control Interface (PECI) access
    - SMLink0 to connect to the BMC
    - Intel Management Engine (ME) to obtain Hot-Swap Controller (HSC) PMBus related information
    - Power capping capability
    - Temperature sensor reading capability from the BMC
    - Support for all PCH SKUs in terms of power delivery and thermal design

5.5 PCIe Usage

5.5.1 PCIe Hot Plug

- The x16 PCIe in Airmax connectors from CPU1, and the x32 PCIe in riser slot shall support the standard hot-swap PCIe signal.
- Motherboard design shall connect PE_HP_SCL/SDA of CPU0 and CPU1 to either the Airmax connector or the BMC in order for the CPU to obtain access to the expender logic directly through Airmax connector or BMC. Refer to the SMBus block diagram for connection.

PCIe power hot-swap and the expander logic is not in the scope of the motherboard design.

5.6 PCB Stack-Up

The 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 5-1 Motherboard PCB Stack Up

Material: IT170GRA1				
Layer	Layer Type	Material Requirement	Thickness (mil)	Thickness Tolerarance (mil)
	Solder Mask		0.50	
1	Top	0.5oz+plating	1.80	
	prepeg	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	2116 50.2% x 2	10.00	
	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: The vendor must follow the below items and bring issues to Facebook if difficulties arise:

- Total Board Thickness: 92 ± 9 mil
- Keep thickness between (L2 & L3), be target 3x of the thickness between (L3 & L4), or absolute ration > 2.46 minimum
- Keep thickness between (L4 & L5) be at least 3x of the thickness between (L5 & L6), or absolute ratio > 2.46 minimum
- Keep the thickness between (L10 & L11) be at least 3x of the thickness between (L11 & L12), or an absolute ratio > 2.46 minimum
- Must meet width design ±20% tolerance with ±10 impedance control for signal-end and differential trace
- The prepreq between L1 & L2 and L13 & L14 should be controlled with ± 0.4 mil tolerance
- The prepreq between L7 & L8 should be controlled thickness min. > 5 mil

Table 5-2 Motherboard PCB Impedance Control

Layer	Target Design Single Track Break Out Geometry				Target Design SE 40Ω				Target Design SE 50Ω				Target Design Diff Breakout Geometry			
	Width	Ohms	Ref. Plane	Sim. Zo	Width	Ohms	Ref. Plane	Sim. Zo	Width	Ohms	Ref. Plane	Sim. Zo	Width	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	

Table 5-3 Motherboard PCB Impedance Control with Insertion Loss

Layer	Target Design DIFF 85Ω				Insertion Loss Spec (db/inch)		Target Design DIFF 100Ω				Insertion Loss Spec (db/inch)	
	Width / Space	Ohms	Ref. Plane	Sim. Zo	@ 4GHz	@ 8GHz	Width / Space	Ohms	Ref. Plane	Sim. Zo	@ 4GHz	@ 8GHz
1	5.7/7.0 & 4.5/5	85± 10%	L2	87.72/87.49	-0.69	-1.38	3.95/14.05	100±10%	L2	104.07	-0.69	-1.38
				85.28/ 84.62			4.25/4.875			100.59		
2												

3	4.5/7	85± 10%	L2/4	90.79	-0.65	-1.25	3.99/12.52	100±10%	L2/4	101.01	-0.65	-1.25
				84.19			4.125 /12.385			99.4		
4												
5	4.5/7	85± 10%	L4/6	91.55	-0.65	-1.25	3.9912.52	100±10%	L4/6	101.93	-0.65	-1.25
				84.92			4.125 /12.385			100.31		
6												
7												
8												
9												
10	4.5/7	85± 10%	L9/11	91.55	-0.65	-1.25	3.99/12.52	100±10%	L9/11	101.93	-0.65	-1.25
	0	0		84.92			4.125 /12.385			100.31		
11												
12	4.5/7	85± 10%	L11/13	91.55	-0.65	-1.25	3.99/12.52	100±10%	L11/13	101.93	-0.65	-1.25
	0	0		84.19			4.125 /12.385			99.4		
13												
				85.78 /85.21			4.125 /13.875			100.59		
14		85± 10%			-0.69	-1.38		100±10%			-0.69	-1.38

6 BIOS

Vendors shall be responsible for supplying and customizing the BIOS for the motherboard. 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 the PCH's SPI interface through BMC controlled MUX for the 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. 32 Megabyte (32MB) size is recommended considering the space required for both BIOS and Intel® ME firmware. The vendor should provide a 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 the BIOS chip so that the 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 be able to fit major available SPI flash vendors' package drawing.

A secondary, identical BIOS chip is designed to share the same SPI bus with the multiplexed CS pin (the MUX is controlled by the BMC). More detail of this scheme is described in section 8.15.

6.2 BIOS Source Code

A United Extensible Firmware Interface (UEFI) BIOS firmware should be used. The vendor shall be responsible for maintaining BIOS source code to ensure that it has the latest code release from Intel and codebase vendors. The 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. BIOS tuning includes:

- Disabling unused devices, such as PCI and PCIe ports, USB ports, SATA/SAS ports, clock generator, buffer ports, etc
- CPU/Chipset settings to attain minimum power consumption and the best performance for a data center environment
- Enabling the 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 tuned BIOS settings

6.3.2 Setup Menu

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

- Settings for adjusting memory speed, QPI speed, speed-step/turbo mode, and the following CPU C-state power states. 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 follows the CPU and chipset 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 Serial-Over-LAN (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 the 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 300 meters
- 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 system reset once the timer expires. The watchdog timer is always disabled after Power-On-Self-Test (POST).
- Available settings for the ECC error threshold are 1,4, 10, and 1000. The default setting is 1 for EVT/DVT/PVT, and 1000 for MP.
- Setting for the ECC error event log threshold – available settings are 1, 4, 10, and 1000. The default setting is 1 for EVT/DVT/PVT and 1000 for MP.
- Setting for ECC error event log threshold. Available settings are disabled, 10, 50, and 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 System Event Log (SEL).
- The default setting to disable all "wait for keyboard input to continue" features is "not to wait for keyboard input to continue".
- Calculate checksum of BIOS setting, display it in the BIOS setup menu, and Ouput
- 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 it in the BIOS setup menu, and output to System Management (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 Protected Processor Inventory Number (PPIN) 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 Preboot eXecution Environment (PXE) boot capability in IPv4 and IPv6 environments at the same time and boot from the SATA/SAS and USB interface. The BIOS should provide boot option selection capability. The default boot device priority is:

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

If a bootable device is not found, the BIOS should loop and search for a bootable device. The BIOS should support UEFI and legacy boot mode operations, with UEFI being the default. UEFI and legacy boot mode have independent boot loop.

The boot mode and boot order can be displayed and changed from the BMC with the 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 that the BIOS can perform correct board initialization based on different board configurations. The board ID is also routed to the BMC for BMC firmware access.

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 ID1 is to identify M.2 Type, 0=SATA, 1=PCIe
- Board SKU ID0 is to identify ODM.0=Vendor1, 1=Vendor2

Table 6-1: Board SKU ID

SKU_ID[4:0]					Description
0	0	X	0	0	Single Side/Fairchild/ M.2=SATA
0	0	X	1	0	Single Side/Fairchild/M.2=PCIe
0	1	X	0	0	Single Side/Infineon/ M.2=SATA
0	1	X	1	0	Single Side/Infineon/ M.2=PCIe
1	0	X	0	0	Double Side/Fairchild/ M.2=SATA
1	0	X	1	0	Double Side/Fairchild/ M.2=PCIe
1	1	X	0	0	Double Side/Infineon/ M.2=SATA
1	1	X	1	0	Double Side/Infineon/ M.2=PCIe

The motherboard has a 1Kbit EEPROM (addres 0xA2 in 8-bit format) for soft-strap board ID to be accessed by BIOS on the host SMBus. The definition of the soft-strap is to be decided. The vendor shall keep this EEPROM blank until the definition is provided by

Facebook. If no definition is defined during DVT, then the vendor shall remove it from the Bill Of Materials (BOM).

6.3.5 FAB Revision ID

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

Table 6-2: FAB Revision ID

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 a remote BIOS update function. The vendor must validate update tools on each BIOS release during development and production. Tool(s) provided should support the following 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 8.15)
  - 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 Secure Shell (ssh); the current OS based is CentOS 6.4 64-bit with a kernel version specified by the 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 the system
- BIOS update or BIOS setup option change should not take more than five minutes to complete
- BIOS update procedure can be scripted and propagated to multiple machines

6.3.7 Event log requirement

BIOS should perform event logging through BMC SEL with Generator ID 0x0001 and the combination of BIOS and BMC should meet the SEL log requirements in section 8.11.

6.3.8 BIOS Error Code Requirement

BIOS fatal error codes listed in Table 6-3 should be enabled for POST code output. The vendor should display major and minor code alternatively as specified in 10.8.2.

Table 6-3 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 byte sequences are assigned for the entry and the exit to avoid the same sequence used on other BIOS POST codes. 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 code process.

6.3.10 PPIN BIOS Implementation

The BIOS shall support Protected Processor Inventory Number (PPIN) with its default setting set to [Unlock and Enable]. There are two ways that the user can access the PPIN:

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

The BIOS shall perform 2x actions to synchronize the PPIN value to the 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® SPS Firmware Plan of Record

Intel Motherboard V4.0 uses an Intel® chipset. Its Management Engine (Intel® ME) runs Intel® Server Platform Services (Intel® SPS) firmware. Intel® SPS firmware is required for system operation.

Intel® SPS 4.0 Firmware (FW) consists of two parts:

- Intel® SPS Silicon Enabling FW – required to boot the platform and have a fully functional platform.
- Intel® SPS 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 assistance from a BMC. Both parts are required for Tioga Pass.

The Intel® SPS firmware is stored in PCH flash. The vendor should provide a utility to update the Intel® SPS firmware in CentOS 6.4 64-bit with a kernel version specified by the customer through ssh. The utility should support updating Intel® SPS firmware and BIOS either together or separately, and provide the option to update only Intel® SPS Firmware's operation region or the entire Intel® SPS firmware region. Vendor should also implement a BMC to update PCH flash where Intel® SPS firmware is located as described in section 8.15.

8 BMC

Intel Motherboard V4.0 uses the ASPEED® AST2500 BMC with one, x16 4Gb DDR4 DRAM for various platform management services. It interfaces with hardware, BIOS, and the Intel® SPS firmware. The following outlines BMC requirements:

- The BMC should be a standalone system that operates in parallel with the host (dual processor x86)
- The health status of the host system should not affect normal operation or network connectivity of the BMC
- The BMC cannot share memory with the host system
- The BMC management connectivity should work independently of the host, and not have a Network Interface Card (NIC) driver dependency for Out-Of-Band (OOB) communication if using a shared NIC

8.1 Management Network Interface

The following outlines requirements for the BMC's management network interface:

- The BMC should have both the I²C port and RMII/NC-SI port for OOB access
- Three OOB access options should be supported. Option 2 and Option 3 share the same device footprint as co-layout.
    - Option 1: The shared NIC uses I²C or RMII/NCSI interfaces to pass management traffic on the data network of the Mezzanine 25GE NIC
    - Option 2: The shared NIC uses RMII/NCSI interfaces to pass management traffic on theIntel® I210-AS data network. The Intel® I210-AS has SGMII/KX interface to midplane.
    - Option 3: The shared NIC uses RMII/NCSI interface to pass management traffic on the Intel® I210-AT data network. The Intel® I210-AT has a 10/100/1000 MDI interface to RJ45.
- The BMC firmware needs to flexible about which interface and device to activate by hardware strapping, or a preset priority policy. The BMC needs to ensure that unused interfaces and devices are disabled and do not interfere with the activated management interface and device.
- The OOB MAC address should use the NIC's data network MAC with an offset defined by NIC vendors.
- The BMC management network firmware and utility must support all features defined in this specification for both IPv4 and IPv6 network environments.

8.2 Local Serial Console and SOL

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

- A local serial console on the debug header (described in section 10.8)
- Remote console, also known as Serial-Over-LAN (SOL) through the management network (described in section 0).

It's preferred that both of these interfaces are functional at all stages of system operation.

During system boot-up, POST codes will be send to port 80 and decoded by the BMC to drive the LED display as described in section 8.5. POST codes should be displayed through SOL console during system POST. Before the system has the first screen, POST codes are dumped 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 on the lower right corner of the SOL console screen.

A serial console buffer feature is required. The buffer needs to save, at least, the last 5x screens of local and 5x screens of remote console output with 80x24 (80 columns by 24 rows) for each screen. The OOB raw command is used to extract and display the buffer. The buffer has to be stored in volatile media, such as an internal or external SDRAM of BMC. The SOL buffer data is cleared within five seconds of the removal of standby power. The SOL SHOULD NOT be stored in any non-volatile media for security and privacy reasons. The SOL buffer implementation shall allow the SOL buffer being dumped by script with OEM command to file (for scaling of data collection).

8.3 Graphic and GUI

Because the Graphic User Interface (GUI) is not scalable, all of the BMC features need to be available in command line model with In-band and OOB IPMI commands, or SOL. Intel Motherboard V4.0 adds support of GUI and KVM on hardware level to accommodate the OCP customers whose environment requires using of VGA and KVM.

8.4 Remote Power Control and Power policy

The vendor should implement BMC firmware to support remote system power on/off/cycle and warm reboot through In-Band or Out-of-Band IPMI commands.

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 by IPMI and take effect without a BMC firmware cold reset or system reboot.

It should take less than three seconds from AC on for the BMC to process the power button signal and power up the system for POST. A long waiting period for this process IS NOT allowed.

8.5 Port 80 POST

The vendor should implement BMC support for port 80 POST code display to drive an 8-bit HEX General-Purpose Input/Output (GPIO) to debug header. The BMC POST function needs to be ready before the system BIOS sends the 1st POST code to port 80. The POST code should also be sent to SOL as mentioned in section 8.2.

The BMC should have access to 256x POST codes and records. The OOB OEM command command can be used to retrieve the last 256x POST codes from the BMC.

8.6 Power and System Identification LED

The motherboard combines the Power LED and the System Identification LED in a single, blue LED on the front.

- Power LED – ON: used to show readiness of major runtime power rails (P12V, p5V, and P3V3), but NOT the readiness of all runtime power rails (the CPU core power rail, for example)
- Power LED – Blinking: used for system identification. The on time is different during power on and power off

The power and system identification LED has four unique behaviors/states to identify the system power state and chassis status:

Table 8-1: Power and System LED States

State	Behavior
Power Off, Chassis Identify Off	LED Off
Power Off, Chassis Identify On	LED On for 0.1s, LED OFF for 0.9s, and loop (1 Hz, 10% duty cycle)
Power On, Chassis Identify OFF	LED Off
Power On, Chassis Identify On	LED On for 0.9s, LED OFF for 0.1s, and loop (1 Hz, 90% duty cycle)

8.7 Platform Environment Control Interface (PECI)

The BMC should access the PECI through PCH SMLink by default

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

The vendor should implement board design to connect CPU PECI interface to PCH PECI or BMC PECI by adding an analog switch controlled by BMC GPIOAB2(ASPEED 2500 Pin T22). CPU PECI is accessed by Intel® ME firmware by default when the GPIO is low and it will switch to BMC access PECI when ME is not responsive.

8.8 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® SPS-NM is required. This function should be available through In-Band and Out-Of-Band connectivity.

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 that temperature reporting is accurate, the TI TMP421 with external transistor is the preferred component for detection of Inlet and Outlet temperature. Caution could be taken for inlet air sensor implementation to avoid preheating nearby components and heat conducted through the PCB. The airflow sensor is not a physical sensor; it's calculated based on system FAN PWM.

8.9 SMBUS Diagram

Placeholder for SMBUS diagram

Figure 8-1: SMBUS Diagram

8.10 Sensors/Events

This chapter describes analog, discrete, and events. The list includes all of the sensors/events required. It does not include all of the detailed requirements. Refer to the BMC specification for more detailed requirements.

8.10.1 Analog sensors

The BMC has access to all analog sensors on the motherboard directly or through the PCH Management Engine. All analog sensors need to be displayed in the Sensor Data Record (SDR) repository.

The analog sensors required are listed in Table 8-2. Refer to section 8.11 for logging requirements.

Table 8-2 Analog Sensor Table with Lower and Upper Critical

Sensor name	Sensor#	Lower Critical	Upper Critical
Outlet Temp	0x01	na	75
P0 VR Temp	0x02	na	85
P1 VR Temp	0x03	na	85
P0 Temp	0x05	na	DTSMax-2
P1 Temp	0x06	na	DTSMax-2
Inlet Temp	0x07	na	40
PCH Temp	0x08	na	66
P0 Therm Margin	0x09	na	-2
P1 Therm Margin	0x0A	na	-2
P0 DIMM VR0 Temp	0x0B	na	72
P0 DIMM VR1 Temp	0x0C	na	72
P1 DIMM VR0 Temp	0x0D	na	72
P1 DIMM VR1 Temp	0x0E	na	72
HSC Temp	0x0F	na	75
P0 core VR PIN	0x11	na	240
P1 core VR PIN	0x12	na	240
P0 DIMM VR0 PIN	0x13	na	47
P0 DIMM VR1 PIN	0x14	na	47
P1 DIMM VR0 PIN	0x15	na	47
P1 DIMM VR1 PIN	0x16	na	47
P0 core VR POUT	0x22	na	255
P0 core VR Curr	0x23	na	98
P0 core VR Vol	0x24	1.35	1.96
P1 core VR POUT	0x25	na	255
P1 core VR Curr	0x26	na	98
P1 core VR Vol	0x27	1.35	1.96
HSC Output Curr	0x28	na	47.8
HSC Input Power	0x29	na	501
HSC Input Volt	0x2A	11.3	13.2
P0 Package Power	0x2C	na	na
P1 Package Power	0x2D	na	na
P0 DTSmax	0x30	na	na
P1 DTSmax	0x31	na	na
P0 DIMM VR0 POUT	0x32	na	47
P0 DIMM VR0 Curr	0x33	na	76
P0 DIMM VR0 Vol	0x34	1.15	1.25
P0 DIMM VR1 POUT	0x35	na	47
P0 DIMM VR1 Curr	0x36	na	76

P0 DIMM VR1 Vol	0x37	1.15	1.25
P1 DIMM VR0 POUT	0x38	na	47
P1 DIMM VR0 Curr	0x39	na	76
P1 DIMM VR0 Vol	0x3A	1.15	1.25
P1 DIMM VR1 POUT	0x3C	na	47
P1 DIMM VR1 Curr	0x3D	na	76
P1 DIMM VR1 Vol	0x3E	1.15	1.25
Riser card slot2 P12V power	TBD	na	75
Riser card slot2 P12V Curr	TBD	na	
Riser card slot2 P12V Vol	TBD	na	
Riser card slot3 P12V power	TBD	na	
Riser card slot3 P12V Curr	TBD	na	
Riser card slot3 P12V Vol	TBD	na	
Riser card slot4 P12V power	TBD	na	
Riser card slot4 P12V Curr	TBD	na	
Riser card slot4 P12V Vol	TBD	na	
P1 DIMM VR1 Curr	0x3D	na	76
P1 DIMM VR1 Vol	0x3E	1.15	1.25
SYS FAN0	0x46	500	9000 for SS, DD TBD
SYS FAN1	0x47	500	9000 for SS, DD TBD
Airflow	0x4A	na	na
C1 Local Temp	0x53	na	Based on AVL
C1 Remote Temp	0x54	na	Based on AVL
C2 Local Temp	0x4B	na	Based on AVL
C2 Remote Temp	0x4C	na	Based on AVL
C3 Local Temp	0x4D	na	Based on AVL
C3 Remote Temp	0x4E	na	Based on AVL
C4 Local Temp	0x4F	na	Based on AVL
C4 Remote Temp	0x50	na	Based on AVL
P0 DIMM0 Temp	0xB4	na	81
P0 DIMM1 Temp	0xB5	na	81
P1 DIMM0 Temp	0xB6	na	81
P1 DIMM1 Temp	0xB7	na	81
P3V3	0xD0	3	3.6
P5V	0xD1	4.5	5.5
P12V_STBY	0xD2	11.3	13.2
P1V05_PCH_STBY	0xD3	0.95	1.15
PVNN_PCH_STBY	0xD4	0.8	1.1
P3V3_AUX	0xD5	3	3.6
P5V_AUX	0xD6	4.5	5.5
P3V_BAT	0xD7	2.7	3.6

8.10.2 Discrete sensors

The vendor should implement BMC firmware access and display discrete sensors in SDR. The BMC should log abnormal sensor readings to the SEL. The discrete sensors required and the SEL format is listed in Table 8-3 for error decoding.

Table 8-3: Discrete Sensor Table with Sensor Number, Offset, and Event Date 1/2/3

Sensor name	Sensor#	Sensor Offset	ED1	ED2	ED3
System Status	0x10	[0]=1b, CPU0 socket occupied(A,S,R) [1]=1b, CPU1 socket occupied(A,S,R) [2]=1b, CPU0 Thermal trip(S,R) [3]=1b, CPU1 Thermal trip(S,R) [4]=1b, CPU0 FIVR FAULT(S,R) [5]=1b, CPU1 FIVR FAULT(S,R) [6]=1b, CPU CATERR(S,R) [7]=1b, System throttle(A,D,S,R)[3]	Trigger Sensor Offset	0xFF	0xFF
HSC Sts Low	0x2E	[0]=1b,None of the Above(A,D,S,R) [1]=1b,CML(A,D,S,R) [2]=1b, Temperature(A,D,S,R) [3]=1b, VIN UV FAULT(A,D,S,R) [4]=1b, IOUT OC FAULT(A,D,S,R) [5]=1b, VOUT OV FAULT(A,D,S,R) [6]=1b, HSC OFF(A,D,S,R) [7]=1b, BUSY(A,D,S,R)	Trigger Sensor Offset (Temperature represents 82h)	0xFF (Temperature represents result of request STATUS_TEMPERATURE command)	0xFF
HSC Sts High	0x2F	[0]=1b, Unknown(A,D,S,R) [1]=1b, Other(A,D,S,R) [2]=1b, Fans(A,D,S,R) [3]=1b, Power Good(A,D,S,R) [4]=1b, MFR Specific(A,D,S,R) [5]=1b, Input(A,D,S,R) [6]=1b, Iout/Pout(A,D,S,R) [7]=1b, Vout(A,D,S,R)	Trigger Sensor Offset (Temperature represents 82h) (Iout/Pout represents 86h) (Vout represents 87h)	0xFF (Temperature represents result of request STATUS_TEMPERATURE command) (Iout/Pout represents result of request STATUS_IOUT command) (Vout represents result of request STATUS_VOUT command)	0xFF
SEL Status	0x5F	[1]=1b, SEL Clear(A,S,R) [8]=1b, SEL Rollover(A,S,R)	Trigger Sensor Offset	0xFF	0xFF
DCMI Watchdog	0x60	[0]=1b, Timer expired(A) [1]=1b, Hard Reset(A) [2]=1b, Power Down(A) [3]=1b, Power Cycle(A) [8]=1b, Timer interrupt(A)	[7:6]=11b [5:4]=00b [3:0] Trigger Sensor Offset	[7:4] Reserve for timer action as 00 [3:0] Timer use 00h=reserved 01h=BIOS FRB2 02h=BIOS/POST 03h=OS Load 04h=SMS/OS 05h=OEM	0xFF
Processor Fail	0x65	[4]: FRB3/Processor Startup/Initialization failure(A,D,S,R) (CPU didn't start)	Trigger Sensor Offset	0xFF	0xFF

[3] A=Assertion, D=De-assertion, S=Threshold settable, R=Threshold is readable

Chassis Pwr Sts	0x70	[0]=1b, Power Off/Power Down(A,S,R) [1]=1b, Power Cycle(A,S,R) [2]=1b, Power On(A,S,R) [3]=1b, Soft-Shutdown(A,S,R) [4]=1b, AC Lost(A,S,R) [5]=1b, Hard Reset(A,S,R)	Trigger Sensor Offset	0xFF	0xFF

Sys booting sts	0x7E	[0]=1b, SLP S4 #N(S,R)
		[1]=1b, SLP S3 #N(S,R)
		[2]=0b, PCH PWROK(S,R)
		[3]=0b, SYS PWROK(S,R)
		[4]=1b, Platform reset #N(S,R)
		[5]=0b, BIOS post complete #N(S,R)
			none	none	none

CPU0 Error	0x91	[0]=1b, CPU FIVR Fault(A,R) [3]=1b, Thermal Trip(A,R)	Trigger Sensor Offset	none	none
CPU1 Error	0x92	[0]=1b, CPU FIVR Fault(A,R) [3]=1b, Thermal Trip(A,R)	Trigger Sensor Offset	none	none
VR HOT	0xB2	[0]=1b, CPU0 core VR hot(A,D,S,R) [1]=1b, CPU1 core VR hot(A,D,S,R) [2]=1b, CPU0 DIMM VR0 HOT(A,D,S,R) [3]=1b, CPU0 DIMM VR1 HOT(A,D,S,R) [4]=1b, CPU1 DIMM VR0 HOT(A,D,S,R) [5]=1b, CPU1 DIMM VR1 HOT(A,D,S,R)	Trigger Sensor Offset	0xFF	0xFF
CPU_DIMM HOT	0xB3	[0]=1b, CPU0 PROCHOT(A,D,S,R) [1]=1b, CPU1 PROCHOT(A,D,S,R) [2]=1b, CPU0 CH0/1 MEMHOT(A,D,S,R) [3]=1b, CPU0 CH2/3 MEMHOT(A,D,S,R) [4]=1b, CPU1 CH0/1 MEMHOT(A,D,S,R) [5]=1b, CPU1 CH2/3 MEMHOT(A,D,S,R)	Trigger Sensor Offset	0xFF	0xFF
NTP Status	0xED	[0]=1b, NTP date / time sync failed (A, D, R)	00h, NTP date / time sync failed	0xFF	0xFF

8.10.3 Events

Events are triggered if a GPIO transition is detected. The Event only discrete sensors that are required and the Event Data format is provided in Table 8-4.

Generator ID in the event log shows which piece of firmware generates the log:

- 0x602C = Intel® SPS ME Firmware
- 0x0001 = BIOS/UEFI system Firmware
- 0x0020 = BMC Firmware

Table 8-4: Event only Sensor

Sensor name	Sensor#	Generator ID	ED1	ED2	ED3
SPS FW Health	0x17	0x602c	[7,6]=10b – OEM code in byte 2 [5,4]=10b – OEM code in byte 3 [3..0] – Health Event Type =00h –Firmware Status	Follow the Intel® SPS FW specification	Follow the Intel® SPS FW specification

NM Exception	0x18	0x602c	[0:2]-Reserved. [3]=1b, Policy Correction Time Exceeded [4:5]=10b-OEM code in byte 3. [6:7]=10b-OEM code in byte 2.	00h: Entire platform 01h: CPU subsystem 02h: Memory subsystem 03h: HW Protection 04h: High Power I/O subsystem.	<Policy ID>
NM Health	0x19	0x602c	[0:3]=0010b, Sensor Intel® Node Manager [4:5] =10b-OEM code in byte 3. [6:7] =10b-OEM code in byte 2.	Follow the Intel® SPS FW specification	Follow the Intel® SPS FW specification
NM Capabilities	0x1a	0x602c	[0] – Policy interface capability = 0 – Not Available. = 1 – Available. [1] – Monitoring capability = 0 – Not Available. = 1 – Available. [2] – Power limiting capability = 0 – Not Available. = 1 – Available. [4:7] – Reserved.	none	none
NM Threshold	0x1b	0x602c	[0:1]-Threshold Number. [2]=reserved [3]=0b, Threshold exceeded =1b, Policy Correction Time Exceeded [4:5]=10b-OEM code in byte 3. [6:7]=10b-OEM code in byte 2.	00h: Entire platform 01h: CPU subsystem 02h: Memory subsystem 03h: HW Protection 04h: High Power I/O subsystem.	<Policy ID>
CPU0 Therm Statu	0x1c	0x602c	0h - CPU Critical Temperature. Indicates whether CPU temperature is above critical temperature point. 1h - PROCHOT# Assertions. Indicates whether PROCHOT# signal is asserted. 2h - TCC Activation. Indicates whether CPU thermal throttling functionality is activated due to CPU temperature being above Thermal Circuit Control Activation point.	none	none
CPU1 Therm Statu	0x1d	0x602c	0h - CPU Critical Temperature. Indicates whether CPU temperature is above critical temperature point. 1h - PROCHOT# Assertions. Indicates whether PROCHOT# signal is asserted. 2h - TCC Activation. Indicates whether CPU thermal throttling functionality is activated due to CPU temperature being above Thermal Circuit Control Activation point.	none	none
POST Error	0x2B	0x0001	[7:6] 10b or 11b [5:4] 10b [3:0] Offset 0x00 (System Firmware Error)	If ED1[7:6]= 10b, LSB of OEM POST Error Code If ED1[7:6]= 11b, Per IPMI Spec	MSB of OEM POST Error Code
Pwr Thresh Evt	0x3b	0x602c	01h, Limit Exceeded	none	none
Machine Chk Err	0x40	0x0001	[7:6]=10b [5:4]=10b [3:0]=0Bh, Uncorrectable Or 0Ch, Correctable	Machine Check bank Number (Any one of 0 to 21)	[7:5] CPU Number [4:0] Core Number

PCIe Error	0x41	0x0001	[7:6]=10b [5:4]=10b [3:0]= 04h = PCI PERR 05h = PCI SERR 07h, correctable 08h, uncorrectable 0Ah, Bus Fatal	[7:3] Device Number [2:0] Func Number	[7:0] Bus No
Other IIO Err	0x43	0x0001	[7:6] 10b [5:4] 10b [3:0] Offset 0x00 (Other IIO)	[7:0] Error ID [Refer to Intel® Xeon® Processor E5 v3 Product Family Datasheet, Vol. 1 Sec 11.1.7 IIO module error codes]	[7:5] CPU # [4:3] Reserved [2:0] Source 000b = IRP0 001b = IRP1 010b = IIO-Core 011b = VT-d 100b = TBD 101b = Misc Others = Reserved
ProcHot Ext	0x51	0x0020	[7:6]=10b [5:4]=10b [3:0]=0Ah, Processor thermal throttling offset	[7:2] Reserved [1:0] 0h = Native, 1h= External (VR), 2h = External(Throttle)	[7:5] CPU/VR Number [4:0] Reserved
MemHot Ext	0x52	0x0020	[7:6]=10b [5:4]=10b [3:0] Memory thermal throttling offset (09h)	[7:2] Reserved [1:0] 0h = Native, 1h = External (VR), 2h = External(Throttle)	[7:5] CPU/VR Number [4:3] Channel Number [2:0] DIMM Number [4:0] Reserved for VR HOT
Power Error	0x56	0x0020	[7:6]=00b [5:4]=00b [3:0]=Event offset: 01h, SYS_PWROK Failure 02h, PCH_PWROK Failure	0xFF	0xFF
Memory ECC Error	0x63	0x0001	[7:6]=10b [5:4]=10b [3:0]=00h, correctable 01h, uncorrectable 05h, Correctable ECC error Logging Limit Reached.	[7:4] Reserved [3:2] 00b=All info available 01b=DIMM info not valid 10b=CHN info not valid 11b=CPU info not valid [1:0] Logical Rank	[7:5] CPU Number [4:2] Channel Number [1:0] DIMM Number
Software NMI	0x90	0x0001	[7:6] = unspecified byte 2 [5:4] = unspecified byte 3 [3:0] = Software NMI offset (03h)	0xFF	0xFF
Button	0xAA	0x0020	[7:4] 0h [3:0] 0h: Power button pressed 2h: Reset button pressed	0xFF	0xFF
Power State	0xAB	0x0020	[7:4] 0h [3:0] 0h: Transition to running 2h: Transition to power off	0xFF	0xFF
Power Policy	0xAC	0x0020	[7:6]=00b [5:4]=00b [3:0]=Event offset: 05h: Soft-power control failure	0xFF	0xFF
ME Status	0xAE	0x0020	[7:6]=00b [5:4]=00b [3:0]=Event offset: 01h: Controller access degraded or unavailable 03h: Management controller unavailable	0xFF	0xFF

Network Status	0xB1	0x0020	[7:6]=00b [5:4]=00b [3:0]=Event offset: 00h, After BMC has IP assigned, and Both IPv4 and IPv6 network cannot ping gateway. BMC in disconnection state 01h, Either IPv4 or IPv6 can ping gateway after disconnection state	0xFF	0xFF
PCH Thermal Trip	0xBF	0x0020	01h, State Asserted	0xFF	0xFF
ME Gl Reset Warn	0xc5	0x602c	A0h	If state is asserted: Time for which Intel® ME will delay Global Platform Reset. =00h – FEh – time in unites specified in Event Data 3 =FFh – Infinite delay. For debug purposes, you could configure the delay time as infinity. In this case, the BMC is not required to respond to the event – global reset is suppressed unconditionally.	If State is Asserted: Time Units for which Intel® ME will delay Global Platform Reset. =00h – reserved =01h – minutes =02h – FFh – reserved

System Event	0xE9	0x0020	[7:0]=E5h, Timestamp Clock Synch.
			[7:0]=C4h, PEF Action.
				if ED1 = E5h:
				0x00: event is first of pair
				0x80: event is second of pair
				if ED1 = C4h:
				0x1: PEF Action
					if ED1 = E5h:
					Cause of time changed:
					00h: NTP
					01h: Host RTC
					02h: Set SEL time Command
					03h: Set SEL time UTC offset Command
					FFh: Unknown
					if ED1 = C4h:
					FFh

Critical IRQ	0xEA	0x0020	00h, Front Panel NMI / Diagnostic Interrupt	0xFF	0xFF
CATERR/MSMI	0xEB	0x0020	00h: IERR (CATERR_N Hold Low) 01h: MSMI_N Hold Low 0Bh: MCERR (CATERR_N 16 BCLK pulse) 0Ch: MSMI_N 16 BCLK Pulse	0xFF	0xFF

Dual BIOS Up Sts	0xEf	0x0020	01h: Auto Recovery 02h: Manual Recovery 03h: OOB Directly 04h: Auto Detect 05h: BIOS Crash by SLP_S3_N cycling Recovery	if ED1 = 01h 01h: FRB2 WDT timeout 02h: BIOS Good de-assert (GPION2) 07h: Watchdog not enable if ED1 = 02h 03h: Recovery from Gold to Primary 04h: Recovery from Primary to Gold if ED1 = 03h 05h: Primary directly 06h: Gold directly if ED1 = 04h 08h: BMC Self test failed 09h: Unexpected power off 0Ah: BMC ready pin de-assert (GPIOQ4) If ED1 = 05h 03h: Recovery from Gold to Primary	if ED2 = 01h or 02h 01h: Start the progress for recovery 02h: End the progress for recovery 03h: Checksum compare failed 04h: Primary BIOS is not present 05h: Gold BIOS is not present

8.11 SEL

The vendor should implement BMC support for the System Event Log (SEL).

8.11.1 Error to be logged

- Analog Sensors – An event is logged when analog sensor triggers an upper or lower threshold.
- Discrete Sensors – Sensor with an “A” and “D” note in the offset column indicates that the related event will trigger an assertion and de-assertion event.
- Events – The BIOS and BMC sensors defined in Table 8-4 are capable of triggering the SEL.
- Machine Check Error (MCE) – MCE shall be logged during runtime when MCE generates SMI#. The BIOS shall check MCE banks after a warm reboot and log the error before clearing the MCE.
- Memory Error – Both correctable and uncorrectable ECC errors should be logged into the SEL. Each log entry should indicate the location of the DIMM by CPU socket number, channel number, and slot number. The memory error reporting needs to be tested by both XDP injection and reworked ECC DIMM.
- Error – All errors that have status registers should be logged in the event log. Fatal or non-fatal classification follows the chipset vendor’s recommendation.
- PCIe error – All errors that have status registers should be logged in the event log, including root complex endpoint devices and any switch upstream/downstream ports, if available. Link disable on errors should also be logged. Fatal, non-fatal, or correctable classification follow the chipset vendor recommendations.
- POST error – All POST errors that are detected by the BIOS during system POST should be logged into the event log.
- Power error – There are two kinds of power errors that 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 errors and processor hot errors should be logged. The error log should identify the internal or external source of the error (either processor, memory, voltage regulator, over-current triggered throttling, or under-voltage triggered throttling).
- **FAN Failure** – FAN failure errors should be logged if the FAN speed is outside of an expected range between the lower and upper critical thresholds. The error log should also identify which fan has failed.
- **PMBus status error** – PMBus status sensors check the PMBus controller health status and logs an error if an abnormal value is detected. The PMBus controller can be a DC How Swap Controller (HSC) or PMBus AC to DC power supply unit.
- **Intel® SPS FW related Error logging** –Error logging can be enabled or disabled based on needs.

8.11.2 Error Threshold Setting

- Error threshold settings for correctable and uncorrectable errors should be enabled. After a threshold is reached, an event should be triggered and logged.
- The Memory Correctable ECC should be set to [4] with the option for [1, 4, 10, 1000] for evaluation, development, and pilot run stage. The Memory Correctable ECC should be set to [1000] for MP. When the threshold has been reached, the BIOS should log the event, including DIMM location information with output for the debug card.
- The ECC error event log threshold defines the maximum number of correctable DIMM ECC is logged in the same boot. The default value is 10, with Disable, 10, 50, and 100 as options.
- PCIe Errors follow the chipset vendor's suggestion.

8.11.3 Critical SEL Filter

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

8.12 FSC in BMC

The vendor should enable Fan Speed Control (FSC) on the 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.

8.12.1 Data gathering for FSC

The BMC needs to get data as possible input for FSC. The data that should be obtained is shown in Table 8-5.

Table 8-5: Data for FSC Input

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	4 FAN tachometer inputs
Airflow	Airflow sensor

The sampling rate of the data should be ≥ 1 sample/s.

8.12.2 FSC in BMC

The vendor shall follow and implement FSC and an FSC update interface per the *Facebook Server Fan Speed Control Interface*[4]. The BMC should support both In-Band and Out-of-Band FSC configuration updates. Updates should take effect immediately without requiring a reboot.

8.12.3 Fan Connection

The motherboard has 2x FAN and 1X Pump headers on motherboard. The motherboard and midplane interface both have optional FAN tachometer and PWM connections.

8.13 OEM commands

The vendor shall implement OEM features with the OEM commands listed in the Table 8-6:

Table 8-6: OEM Commands

Command	NF/Lun	Cmd	Function Description
Set DIMM Information	0x30/00	0x1C	Write DIMM Type to BMC; Typically used by BIOS
Get DIMM Information	0x30/00	0x1D	Read DIMM Type from BMC
Set PHY Reset Status	0x30/00	0x30	Change setting to enable/disable shared NIC Phy reset of all shared NICs during system warm reboot and DC cycle
Get PHY Reset Status	0x30/00	0x31	Read setting of Keep NIC Phy Link up feature of all shared NICs
Set First Time BIOS Boot Flag	0x30/00	0x40	For inband BIOS update utility to set Flag for first time boot after an inband BIOS update action
Perform BIOS Recovery	0x30/00	0x41	Copy backup BIOS image to primary BIOS image
Perform BIOS Backup	0x30/00	0x42	Copy primary BIOS image to backup BIOS image
Set dual BIOS Mux	0x30/00	0x43	Change Mux setting of CS# between Primary BIOS and backup BIOS
Set dual BIOS Recovery	0x30/00	0x44	Setting of enable/disable Auto BIOS Recovery after 1st time boot after a BIOS update
Get dual BIOS Status	0x30/00	0x45	Read setting of Dual BIOS recovery, Primary/backup BIOS CS# Mux status, and SPI Host Mux status
Set Log Filter	0x30/00	0x46	Enable/disable log filtering to show critical log only or show full log

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

Get Log Filter	0x30/00	0x47	Read setting of log filter
Replay SOL Buffer	0x30/00	0x48	Replay last 16KB of SOL screen in SOL console
Get 80 Port Record	0x30/00	0x49	Replay the last port 80 post codes(up to 256x POST records)
Get Dual BIOS Version	0x30/00	0x50	Read BIOS version of Primary and backup image from BMC DRAM
Get VR FW Version	0x30/00	0x51	Read CPU and DDR VR FW version
Set BIOS Boot Order	0x30/00	0x52	Change setting of default BIOS boot order
Get BIOS Boot Order	0x30/00	0x53	Read order of default BIOS boot order
Set Dual BIOS Version Obtain	0x30/00	0x54	Request BMC to read BIOS version from physical Primary and backup flash device, and save in BMC DRAM. Get Dual BIOS version reads the same variable in BMC DRAM.
Get BIOS Flash Info	0x30/00	0x55	Read BIOS SPI device ID and status register
Set Post Start	0x30/00	0x73	BIOS to inform BMC post starts
Set Post End	0x30/00	0x74	BIOS to inform BMC post ends
Set PPIN	0x30/00	0x77	BIOS to write PPIN to BMC
Get PPIN	0x30/00	0x78	Read PPIN from BMC
Get BIOS Version	0x30/00	0x83	Read BIOS version that BIOS passed to BMC during BIOS POST
Set Network Sequence	0x30/00	0xB0	Change setting of OOB interface priority between LOM and Mezzanine card
Get Network Sequence	0x30/00	0xB1	Read setting of OOB interface priority
Get Fast PROCHOT	0x30/00	0xD0	Read setting of Current Based FAST_PROCHOT_N
Set Fast PROCHOT	0x30/00	0xD1	Change setting of Current Based FAST_PROCHOT_N
Set VR Monitor Enable	0x30/00	0xD2	Change setting of Enable/Disable CPU Vcore and Memory VDDQ VR sensor polling
Store VR version To BMC	0x30/00	0xD3	Request BMC to read VR version and save to DRAM
Get VR Monitor Enable	0x30/00	0xD4	Read setting of Enable/Disable CPU Vcore and Memory VDDQ VR sensor polling
Clear Memory Status	0x30/00	0xD6	Clear status of "Px_CHxDIMMx_Sts"
MSR Dump	0x30/00	0xD7	Command to request BMC to start fetch MSR dump log task, read MSR dump log, and get MSR dump status
Set MEZZ Protocol Priority	0x30/00	0xD8	Change setting of Mezz OOB interface priority to be NC-SI first, or I2C first
Get MEZZ Protocol Priority	0x30/00	0xD9	Read setting of Mezz OOB interface priority
Set Power Capping	0x30/00	0xDA	Experimental feature not covered by this specification
Get Power Capping	0x30/00	0xDB	Experimental feature not covered by this specification
Set EIN Collection	0x30/00	0xDC	Experimental feature not covered by this specification
Get EIN Collection	0x30/00	0xDD	Experimental feature not covered by this specification
SOL Dump	0x30/00	0xDE	SOL dump configuration and get SOL dump up to 128KB (Default is 64KB)
Get PIN	0x30/00	0xDF	Read average PIN in any duration between 0.1 to 60 seconds
Set GPIO	0x30/00	0xE0	Set GPIO status
Get GPIO	0x30/00	0xE1	Read GPIO status
Set NTP Server	0x30/00	0xE4	Set NTP server IP address and sync policy
Get NTP Server	0x30/00	0xE5	Read NTP server IP address and sync policy
Restore Factory Default	0x32/00	0x66	Restore Factory Default

Set Preserve Configuration	0x32/00	0x83	Set BMC configuration to be preserved or not after BMC FW update
Get Preserve Configuration	0x32/00	0x84	Read settings of BMC configuration to be preserved or not after BMC FW update
Get CPLD Info	0x32/00	0x88	Read CPLD FW checksum, CPLD device ID, and FW version

8.14 BMC FW chip and Firmware Update

The BMC FW flash chip should use the BMC's SPI interface. The vendor is responsible for selecting a specific flash chip that fits the required functionality and potential additional features that may be required in the future. 32MB size is recommended.

Vendors should provide tool(s) to implement remote BMC firmware updates that will not require any physical access to the system. Remote update includes Out-Of-Band via management network or through In-Band via logging into the local OS (CentOS) with the data network. These tool(s) shall support CentOS 6.4 64-bit with updated kernel specified by customer.

A remote BMC firmware update may take a maximum of five minutes to complete. I^2C sideband has a bottleneck to achieve this requirement; the NC-SI interface is needed. The BMC firmware update process and BMC reset process require no reboot or power down of the host system and should have no impact to normal operation of the host system. The BMC needs to be fully functional with updated firmware after the update and reset without further configuration.

In-Band BMC firmware updates can be performed through KCS or USB. USB is the preferred interface due to higher update speed.

The default update should recover the BMC to factory default configuration; an option to preserve the SEL and configuration is required. The MAC address is based on the NIC MAC, so it should not be cleared with a BMC firmware update.

8.15 BMC Update Dual PCH flash

8.15.1 Hardware Scheme

The vendor should implement functionality for the BMC to access the host system PCH flash and remotely recover a host system's PCH from corruption. The PCH flash stores BIOS and Intel® SPS firmware code; both the BIOS and Intel® SPS firmware region should be updated. A dual PCH flash chip hardware scheme is designed for this purpose. The hardware scheme contains two multiplexers controlled by BMC GPO. GPO_A controls PCH or BMC access to flash. GPO_B controls the primary backup flash that is accessed by the CS# signal. The primary flash is the default flash that PCH uses to access BIOS and Intel® SPS firmware code. The backup flash can be modified via command line.

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 ensure PCH has access to the Primary flash by default.
- The BMC shall check the system power status and not change GPO_A and GPO_B status when the system is in S0.
- The BMC shall set Intel® ME to recovery mode before changing GPO_A and GPO_B status from [0,0] to another value.

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

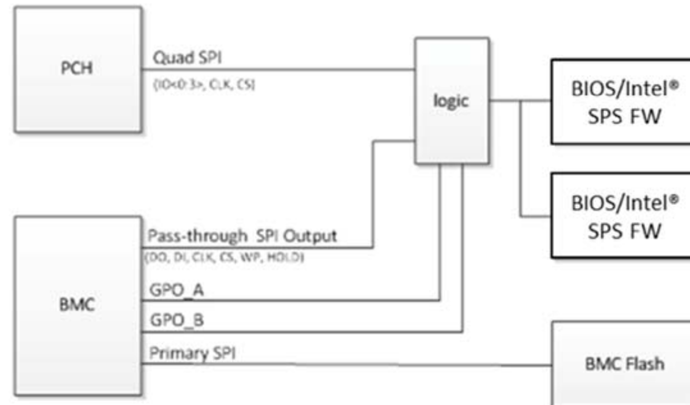


Figure 8-2 BMC and Dual PCH flash diagram

8.15.2 Software Scheme (To be updated, DVT will not have 2nd bios popped)

The following software features must be supported:

1) Auto BIOS recovery from backup flash to primary flash
   - The first boot after BIOS update failure and BMC watchdog timer timeout, the BMC determines that the BIOS recovery needs to recover. The BMC takes over control of flash access and duplicates backup flash content to the primary flash. The BMC will then return flash access to PCH and power cycle the system. This recovery attempt is executed only once after a BIOS update failure. The BMC shall generate SEL for each recovery attempt.

2) Manual BIOS recovery from backup flash to primary flash
   - The user uses the IPMI OEM command to trigger BMC control of flash access and duplication of backup of flash content to primary flash.

3) Manual BIOS save from primary flash to backup flash
   - The user uses the IPMI OEM command to trigger BMC control of flash access and duplication of primary flash to backup flash.

4) Manual Mux control
   - The 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
   - The user uses the IPMI OEM command to trigger a PCH flash binary to be transferred to the BMC and written to the primary or secondary flash, depending

on the state of the multiplexer.

6) Enable/Disable Auto BIOS recovery
   - The user uses the IPMI OEM command to Enable and Disable the Auto BIOS recovery feature. The default is Disable for EVT/DVT/PVT, and Enable for MP.

7) Get dual BIOS feature status
   - The user uses the IPMI OEM command to receive the current state of two multiplexers and the the auto recovery Enable/Disable status.

8) Get dual BIOS version
   - The user uses the IPMI OEM command to access and return the BIOS version for both the primary and secondary flash.

The vendor shall provide an update utility that supports CentOS 6.4 64-bit with an updated kernel specified by the customer.

8.16 BMC Update and Access CPLD

The vendor should implement a way for the BMC to access motherboard CPLD through JTAG and perform:

- CPLD code upload from a remote control server to the BMC
- Update code from the BMC to the CPLD through the CPLD JTAG interface
- Verification of CPLD code

All of these steps shall be done from the OOB command line mode. The vendor shall provide an update utility that supports CentOS 6.4 64-bit with an updated kernel specified by customer.

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

8.17 BMC Time Sync

During BMC initialization:

- If it's the 1st time being powered on via an AC source (G3 exit), the BMC should receive its initial RTC time from PCH RTC. Otherwise, the BMC should keep its own RTC.

During BMC running time:

- If the system is in S0 and the BIOS POST has completed, the BMC should sync with PCH RTC every hour
- If the system is in S5, the BMC should retrieve the NTP UTC following the OEM command setting of NTP periodical sync interval (default =2x per hour)
- Log NTP status time sync failed event if a mismatch of > ±5 seconds is found, or if BMC cannot obtain NTP time after 3x retries. Retry interval is set to 30 seconds by default.

8.18 PCIe and Mezzanine card Thermal monitoring

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

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 detailed requirement and implementation.

8.19 BMC PPIN Implementation

- The BMC shall send SMI# to PCH after BMC reset.
- The BIOS shall serve the request by using Set_PPIN (OEM command) to write PPIN of CPU0 and CPU1 to the BMC. The BIOS shall also Set_PPIN when BIOS POST completes.
- The BMC shall return the PPIN to user with Get_PPIN (OEM_Command).

8.20 BMC Average Power Reporting

- The BMC shall record READ_EIN_EXT every 0.1 seconds in a ring buffer.
- The BMC shall support the OEM command Get_PIN to calculate and report the average power between current and a duration defined in the GET_PIN command.
- The ring buffer size shall support any duration from 0.1 seconds to 60 seconds, in 0.1 second increments.
- The return shall have a resolution of 0.1W.

8.21 BMC Access and Update VR Firmware

- The vendor should implement a way for the BMC to access motherboard CPU VCCIN and memory VDDQ VR controller's firmware.
- The vendor shall implement a script or utility through the OOB master write-read command to perform a VR firmware code update and verify it from the OOB command line mode.
- The script or utility shall support CentOS 6.4 64-bit with updated kernel specified by the customer.
- During the VR firmware update, the 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.
- The vendor shall implement an OEM command to read the VR FW version.

8.22 BMC MSR Dump from OOB

The BMC MSR Dump from OOB feature is a debug feature to allow the user access to 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.

- The vendor shall implement a way for the BMC to dump the MSR from both CPUs through the Intel® ME and PECI.
- The vendor shall provide a utility that supports CentOS 6.4 64-bit with a kernel version specified by the customer.

- The BMC firmware shall apply MSR dumps automatically when either an IERR or MCERR have been asserted.
- The BMC shall store the MSR dump in the BMC flash.
- During the dump, the BMC shall reject chassis power related commands to avoid interrupting the dump.

8.23 BMC Network Status

The vendor shall implement a way for the BMC to log a disconnection in both IPv4 and IPv6 OOB networks when the logging criteria is met. The logging criteria for sensor #0xB1, Event Data 1=00h is:

- BMC has been successfully assigned an IP address through DHCP with either IPv4 or IPv6
- And, the BMC cannot get a response by pinging the Default Gateway for IPv4 for 10 consecutive attempts; every attempt is 10 seconds apart
- And, the BMC cannot get a response by pinging the Default Gateway for IPv6 for 10 consecutive attempts; every attempt is 10 seconds apart
- The BMC shall not log the sensor #0xB1, Event Data 1=00h when either of the IPv4 or IPv6 gateways have at least one response from the gateway within 10 consecutive pings.
- Only one network status log shall be logged by the BMC until the next ping gets response on the IPv4 or IPv6 interface
- The BMC shall log sensor #0xB1, Event Data 1=01h once the next ping receives a response on either the IPv4 or IPv6 interface.

8.24 BMC Secure boot

The vendor shall implement hardware and firmware to support BMC secure boot.

8.24.1 BMC Secure boot hardware

The Tioga Pass has one TPM module with two independent TPM1.2 integrated circuits on the module. The SLB9645 is for the BMC and the SLB9670 is for system firmware. Pin 8 has a special implementation to allow it to be used as both a physical presence pin and a as a reset pin for SLB9645.

The design is based on a few key points:

- The BMC secure boot SPI flash device (flash size TBD) can be set to read-only by hardware strapping and the 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 the secure boot SPI flash device
- The BMC firmware flash image shall not have privilege to reset SLB9645 hardware without resetting the BMC processor itself and trigger code loading from CS0 with Secure Boot image
- SPI0.0 is the secure flash, or ROM;
- SPI0.1 is the field-upgradable flash, containing U-boot, env, the kernel, etc.

8.24.2 BMC Secure boot Firmware flow

The BMC secure boot firmware flow is depicted in Figure 8-4. The vendor shall implement flow option 1 and 2 in U-boot. The BMC secure boot firmware shall utilize hardware hash and crypto engine in AST2500 for SHA-256 hash generation to short the BMC booting time.

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

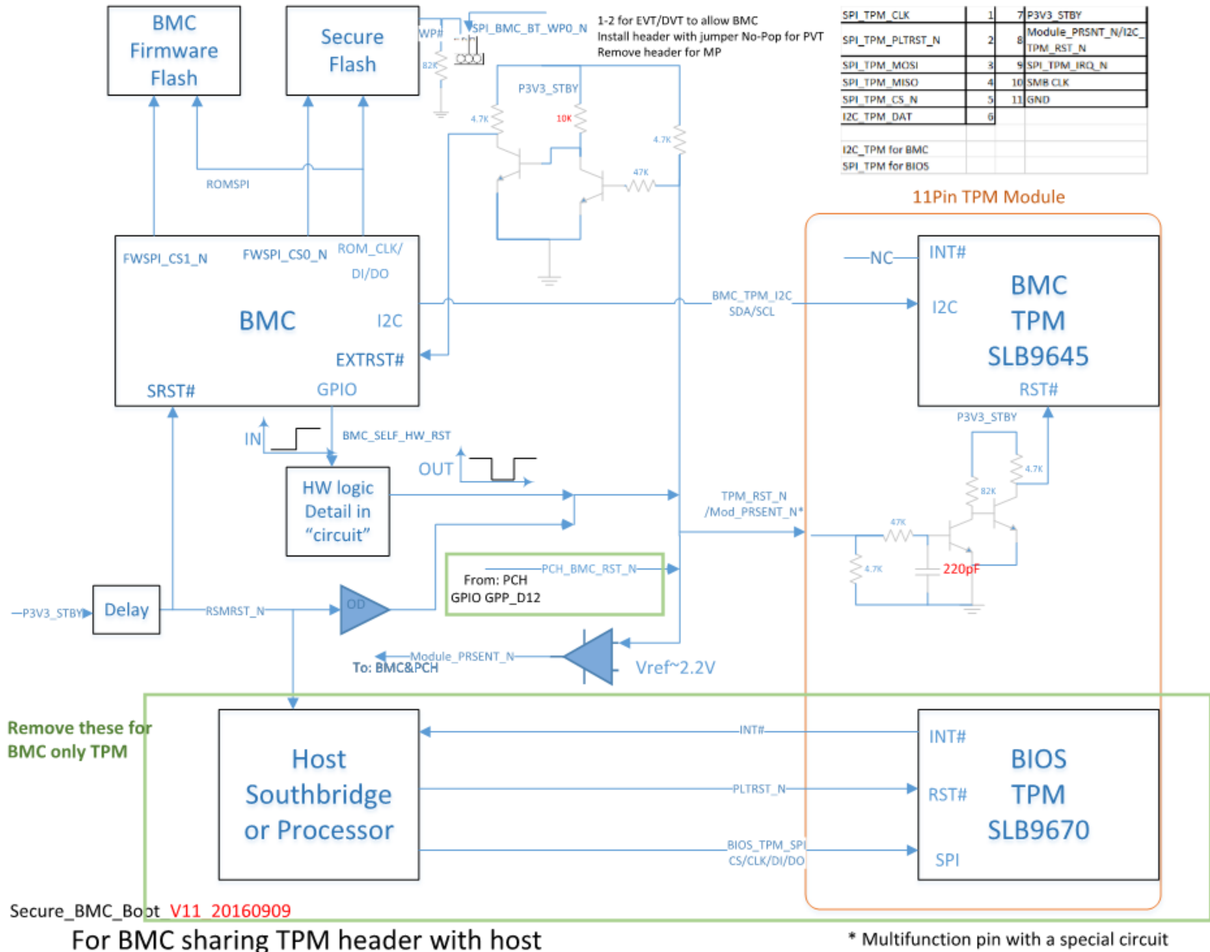


Figure 8-3: BMC Architecture

Placeholder for Secure Boot Flow

Figure 8-4: Firmware Secure Boot Flow

9 Thermal Design Requirements

To meet thermal reliability requirements, the thermal and cooling solution should dissipate heat from the components when the 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. The thermal solution should not allow any overheating issues for any components in the system. CPU or memory should not throttle due to any thermal issue under the 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" H_2O datacenter pressure with one FAN (or one rotor) in each thermal zone failed

9.1 Data Center Environmental Conditions

The thermal design for the Intel Motherboard V4.0 needs to satisfy the data center operational conditions as described below.

9.1.1 Location of Data Center/Altitude

Data centers may be located 6000 meters above sea level or higher. Any variation of air properties or environmental difference due to the high altitude needs to be considered when creating the thermal design.

9.1.2 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 the data center. Every component in the system must be cooled and maintained below its maximum specified temperature in any cold aisle temperature in a data center.

9.1.3 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.

9.1.4 R.H

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

9.2 Server operational condition

9.2.1 System Loading

The power consumption of individual components on the system motherboard will vary by application or by motherboard SKU. The total system power consumption may vary with use or with the number of existence of PCIe cards on the system:

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

A unified thermal solution that can cover 100% of 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.

9.2.2 DDR DIMM DRAM Operation

Thermal design should meet DIMM maximum 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.

9.2.3 Inlet Temperature

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

9.2.4 Pressurization

Except for the condition of a single rotor in a server fan failing, 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.

9.2.5 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.

9.2.6 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 9.1.2 for the temperature definitions.

9.2.7 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.

9.2.8 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.

9.3 Thermal Kit Requirements

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

9.3.1 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.

9.3.2 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 a maximum of five 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 (300W for Open Rack V1 configure) in all condition of fan operation.

9.3.3 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.

9.3.4 Thermal sensor

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

10 I/O System

This section describes the motherboard I/O requirements.

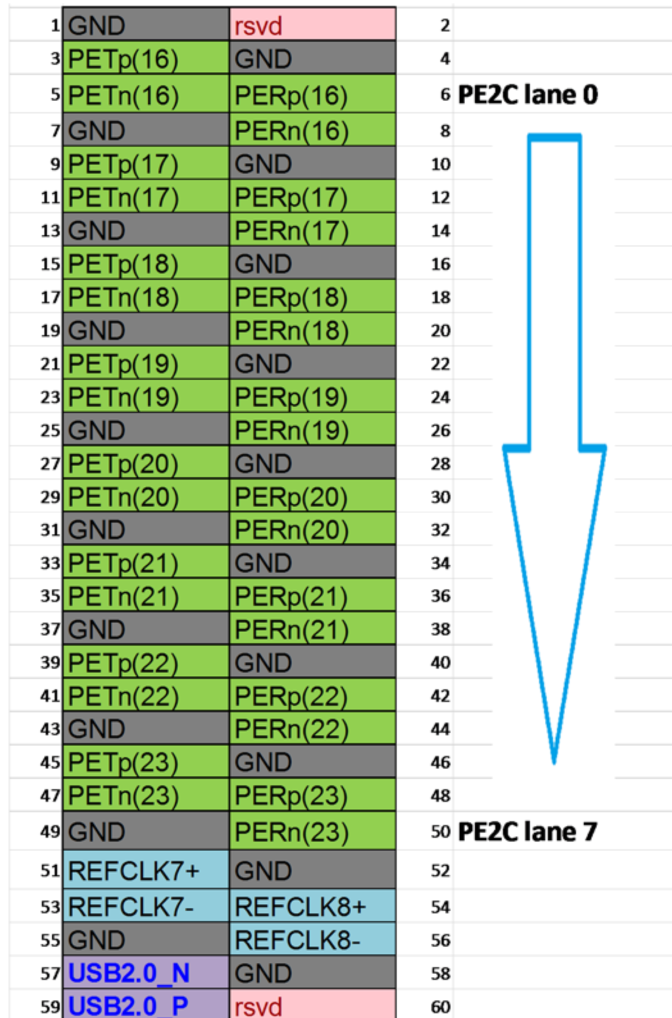
10.1 PCIe x32 Slot/Riser Card

10.1.1 Riser Slot Interface Between Riser Card and Motherboard

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

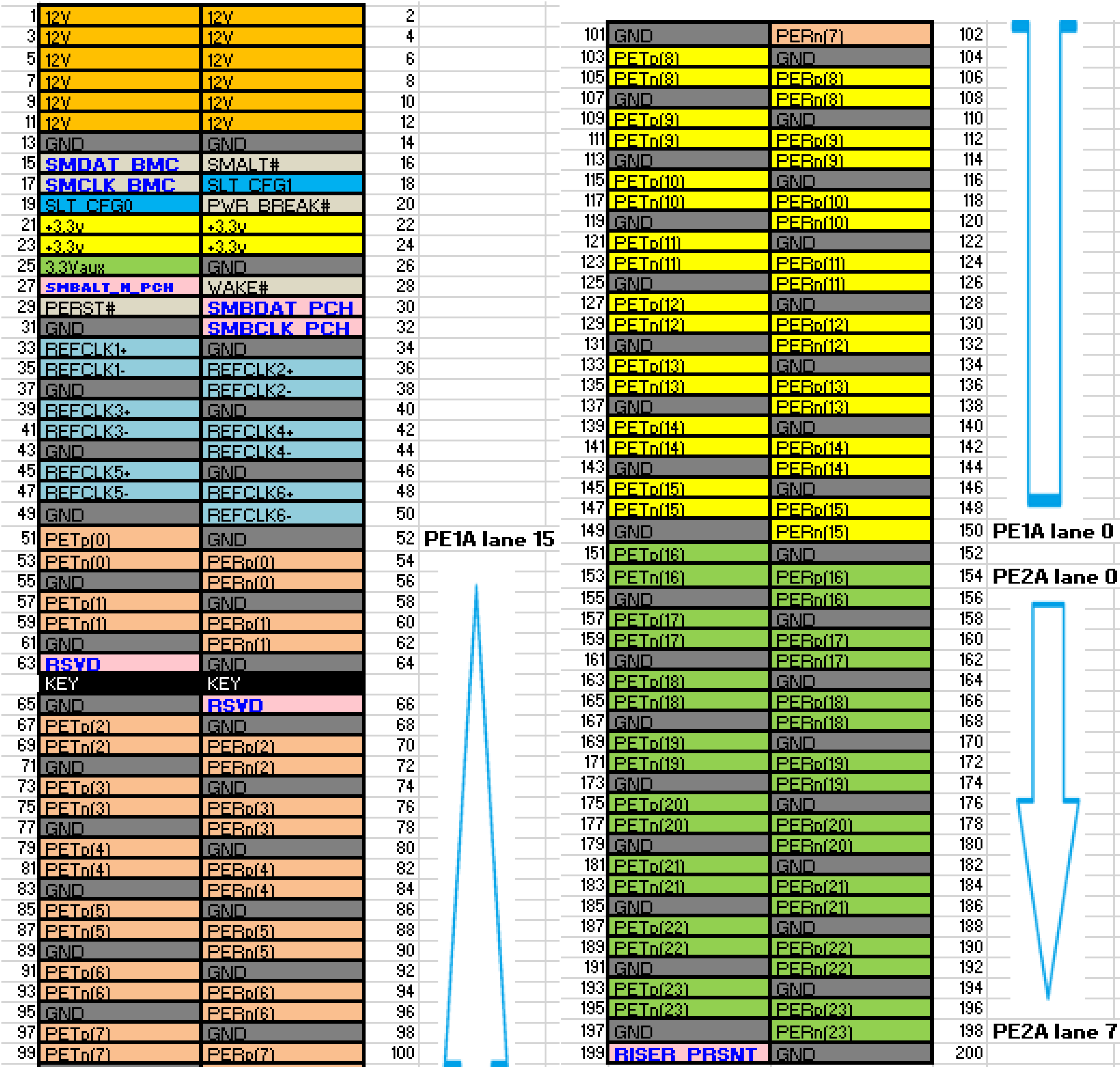
- x24 slot Samtec/HSEC8-1100-01-L-DV-A-K (200-pin) – used for x24 PCIe and power delivery
- x8 slot Samtec/HSEC8-130-01-L-DV-A-TR(60-pin) – used for x8 PCIe, and USB signal

Slot location must follow mechanical requirement that outlined in the DXF document. 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).



Pin	Signal	Signal	Pin	
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	REFCLK7+	GND	52	
53	REFCLK7-	REFCLK8+	54	
55	GND	REFCLK8-	56	
57	USB2.0_N	GND	58	
59	USB2.0_P	rsvd	60	

Figure 10-1



Pin	Signal	Signal	Pin	
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_N_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	PE1A lane 15
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	

Pin	Signal	Signal	Pin	
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	PE1A lane 0
151	PETp(16)	GND	152	
153	PETn(16)	PERp(16)	154	PE2A lane 0
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	PE2A lane 7
199	RISER_PRSNT	GND	200	

Figure 10-2

10.1.2 Riser card types

The vendor shall enable the 3x riser cards described below for single-side and double-side use cases in DVT.

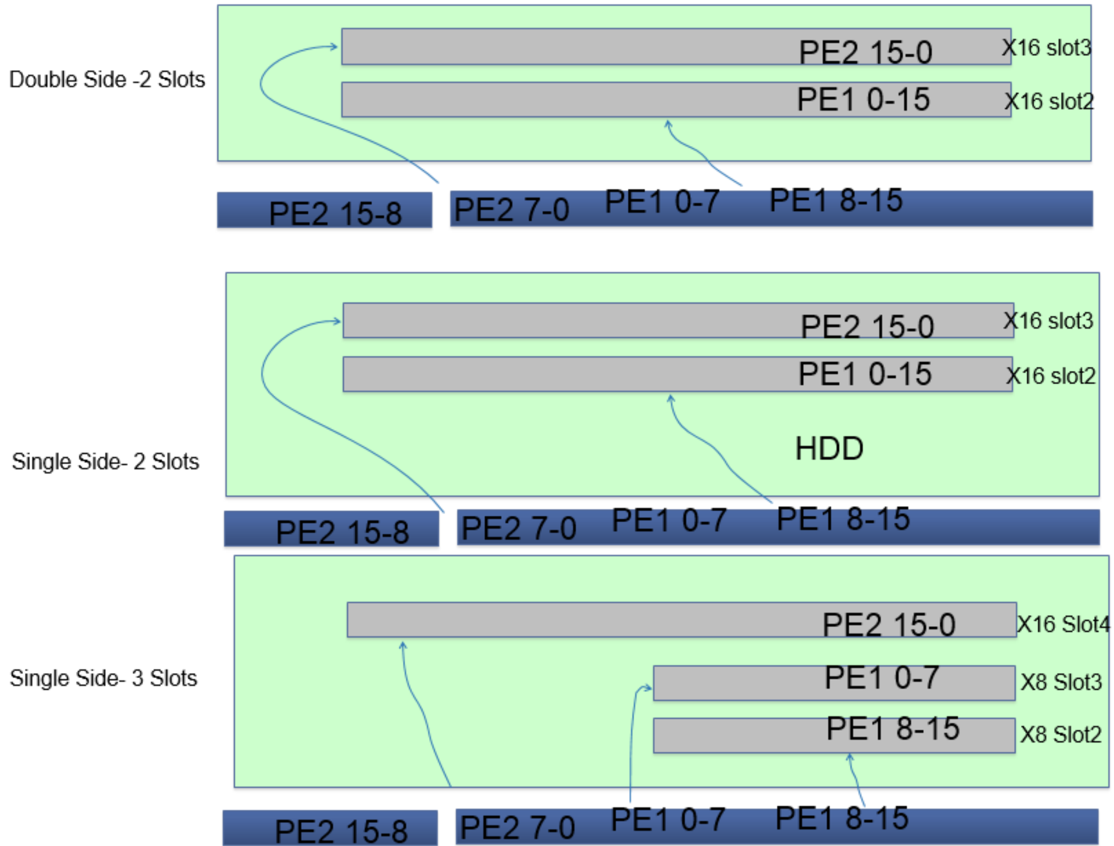


Figure 10-3

Table 10-4, Table 10-5, Table 10-6, Table 10-7, and Table 10-12 describe the pin information for each PCIe slot:

Table 10-1 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 10-2 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 10-3 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 10-4 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 10-5 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 10-6 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

The riser card should implement an SMBus Mux to avoid address confliction of PCIe cards. Reference Figure 10-4 and Figure 10-5 for implementation and slave address assignment in 8-bit.

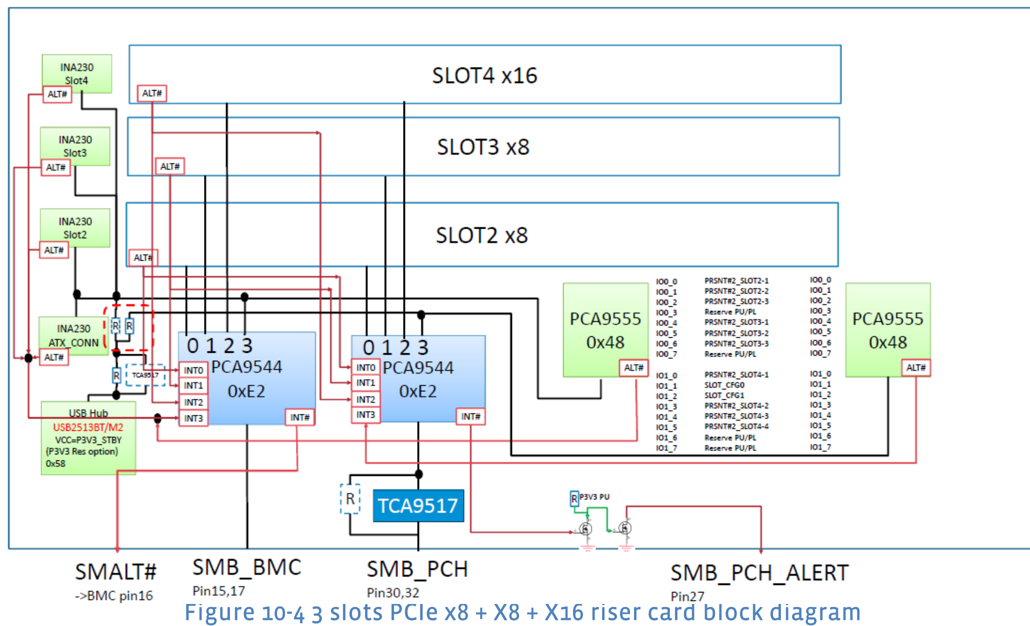


Figure 10-4 3 slots PCIe x8 + X8 + X16 riser card block diagram

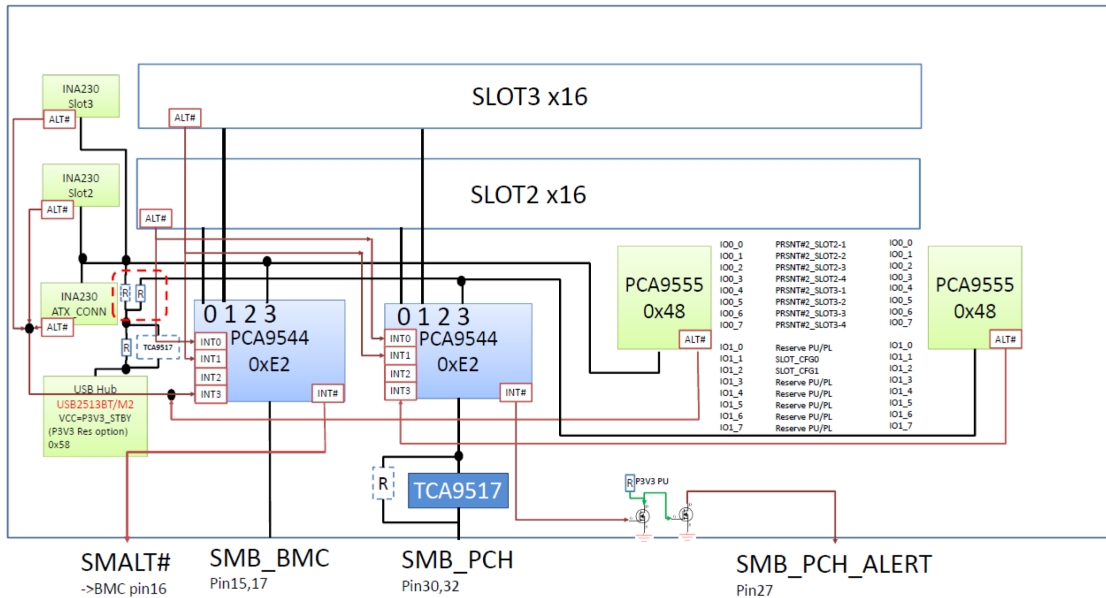


Figure 10-5 2 Slots PCIe X16 Riser Card Block Diagram

The riser card should implement I²C to GPIO expander (PCA9555) to be accessed by both BIOS and BMC on motherboard to tell which AICs are currently on each slot. The BIOS shall follow 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 10-1: 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 10-2: 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

The 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 the motherboard.

The slave addresses are shown in Table 10-3.

Table 10-3: Riser Card Slave Addresses

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

10.1.3 Riser Card Power outlet

A 2x2 right angle power connector (Molex/46991-1004 or equivalent) is required for power delivery from the riser to PCIe cards that need more power than PCIe slot allows.

Table 10-4 Riser Power Connector Pin Define

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

10.1.4 USB on Riser Card

There is a USB 2.0 port from the motherboard's PCH USB 2.0 Port 2 (BW61/CA61) connected to the motherboard's riser card interface. A USB hub is required to fan-out the USB 2.0 to two to three PCIe slots on the riser card. The hub P/N is Microchip USB2513BT/M2 for both the 2-slot riser and 3-slot riser.

10.1.5 PCIe Super Slots Optional Implementation

This section describes an optional implementation of PCIe super slots in Tioga Pass. This implementation adds a 2x footprint on the bottom of x24 and x8 HSEC8 connectors. The purpose of this implementation is extend a functional riser card slot on the bottom of the motherboard.

The vendor shall perform a careful evaluation of SI risk with this implementation.

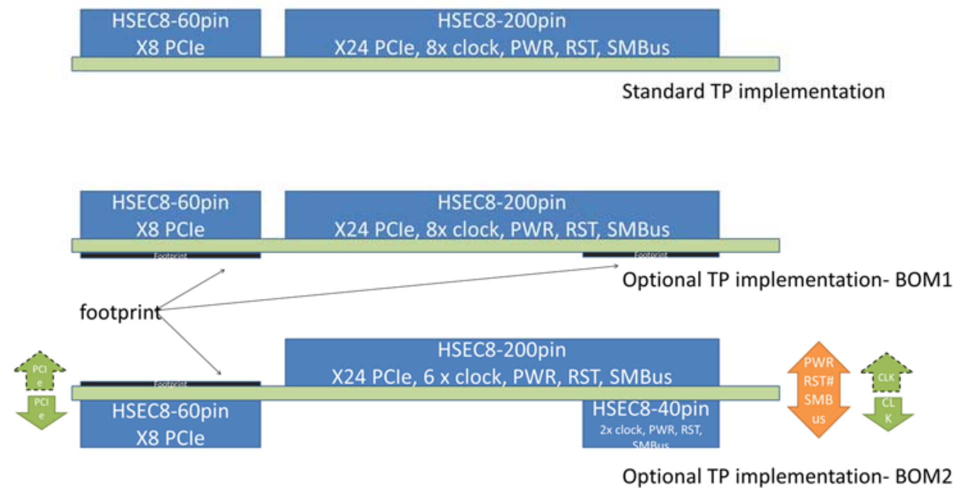


Figure 10-6: Optional Tioga Pass Super Slot Implementation

10.2 DIMM Sockets

The motherboard requires a 15u" gold contact for the DDR4 SMT DIMM socket. This socket incorporates yellow housing with a white/nature latch for DIMM in a far DDR channel, and black housing with a white/nature latch for DIMM in a near DDR channel. The vendor shall announce if the color selection will increase the cost of the DIMM Socket.

10.3 Mezzanine Card

Intel Motherboard v4.0 is compatible with *Mezzanine Card for Intel v2.0 Motherboard*[5] and *OCP Mezzanine Card 2.0 design specifications*[6].

The motherboard has OCP Mezz 2.0 connector A and connector B to provide up to x16 PCIe Gen3 connection to Mezzanine card. The motherboard also has an OCP Mezz 2.0 Connector C to provide up to x4 KR. Connector C can be used independently on the Mezzanine card side.

Motherboard mezzanine connector A is named as Slot1 in system.

10.3.1 Connector A

Table 10-5 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

[5] http://files.opencompute.org/oc/public.php?service=files&t=2047e49112f6109c0f7e595cc93af8ae&download
[6] http://files.opencompute.org/oc/public.php?service=files&t=b9b9b1892b8584c52aeaf53bf8706ce0&download

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
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.

10.3.2 Connector B

Table 10-6 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

Signal	Description	Pin	Pin	Signal	Description
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	

10.3.3 Connector C

Table 10-7 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

10.3.4 Baseboard ID

The baseboard ID allows the Mezzanine card to have awareness of various connected baseboard types. The baseboard ID only applies connector A and connector B. The implementation of the baseboard ID circuit on Tioga Pass is shown in Figure 10-7. R1 is 10KΩ and R2 is 887Ω on both connector A and connector B.

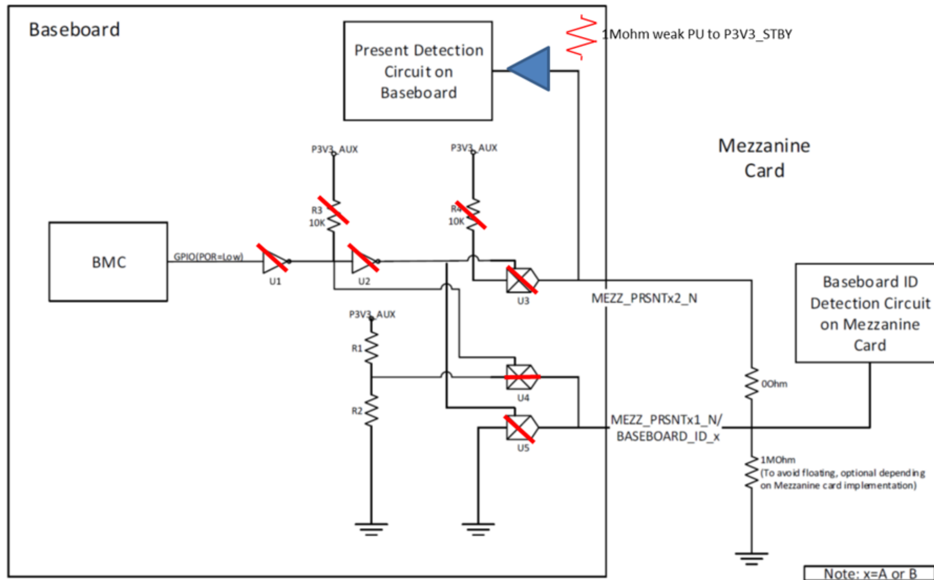


Figure 10-7: Baseboard Circuit Diagram

The Mezzanine card identifies different baseboard types using the resistor pair R1/R2 based on the resistance shown in the tables below:

Table 10-8: 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 10-9: 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

10.3.5 Mezzanine present pin

Tioga pass' Mezzanine pin connection to the BMC and PCH is shown in **Error! Reference source not found.**:

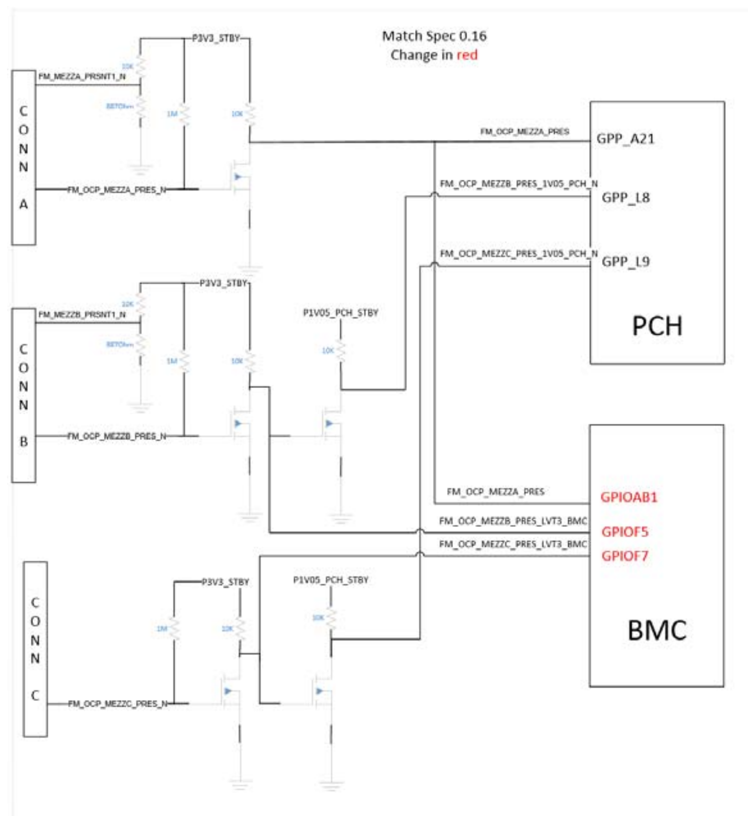


Table 10-10: Mezzanine Pin Schematic Diagram

10.4 Network

10.4.1 Data network

The motherboard uses the OCP Mezzanine 2.0 card as the primary data network interface on the I/O side. This card provides an option for single port or dual port.

Once Intel® I210 1G NIC is installed on the board to provide an optional 10/100/1000 data connection.

10.4.2 Management Network

The motherboard has three options of management network interfaces for the BMC connection. The management network shares the data network's physical interface. The management connection should be independent from data traffic and OS/driver condition.

- SFP+ shared-NIC from Mezzanine 10G NIC or PCIe NIC, driven by BMC through RMII/NC-SI or I^2C. I^2C being default
- SGMII/KX shared-NIC connected to midplane 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

10.4.3 IPv4/IPv6 Support

The system needs to have deployment capability in both IPv4 and IPv6 network environments. All data and management networks should have this capability. This includes, but is not limited to, DHCP and static IP setting, PXE booting capability, NIC and BMC firmware support, OS driver, and utility in both IPv4 and IPv6.

10.5 USB

The motherboard has one external Type-A, right angle USB 2.0/3.0 port and one USB 3.0 Type-C port located in front of the motherboard. The BIOS should support the following USB devices:

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

The Type-C USB port is reserved for Intel debug port.

The external USB 2.0/3.0 port shall have the 5x USB3.0 signals remapped with the debug interface. Refer to section 10.8.2 for more information.

USB power enabling is controlled by the CPLD to allow the debug port to be powered in S5 to perform BMC related troubleshooting.

Table 10-11

Channel Status	System Status	FM_CPLD_DBG_PWER_EN_N

00 (Host)	S5	Disable
00 (HOST)	S0	Enable
01 (Debug)	S0 or S5	Enable
10 (IPMI)	S0 or S5	Enable

10.6 SATA

SATA port 0~7 can be connected to one vertical mini-SAS HD 8 ports connector. sSATA ports 2~5 can be connected to one mini-SAS HD 4 ports connector.

sSata port 0 is connected to 1x M.2 connector on the motherboard with dual layout. PCIe x4 to M.2 is the default population

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

sSATA port 1 is connected to the 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. The connector is an Alltop/C18625-11331-L.

The vendor shall use the DXF document for SATA port placement.

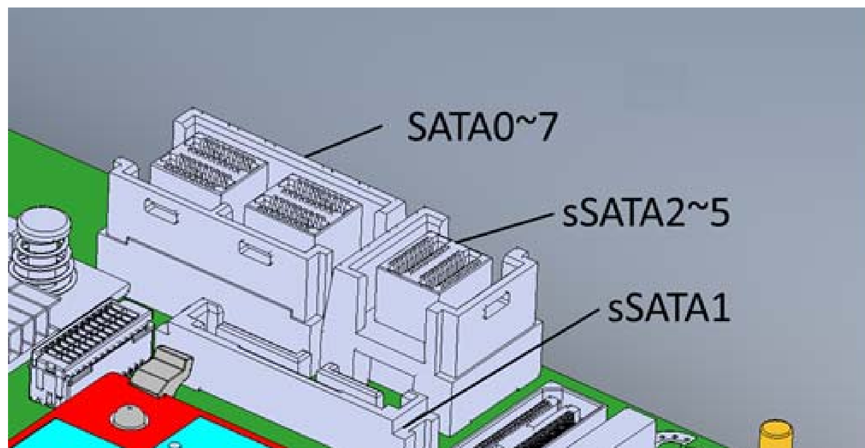


Figure 10-8: SATA and sSATA

Both the mSATA connector and the vertical SATA connector need to be placed near the I/O side of the motherboard for easy access. HDDs attached to all SATA connectors need to follow spin-up delay requirement described in section 14.7.3.

10.7 M.2

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

The sideband signals such as WAKE# and SATA_ACTIVITY should be connected when it applies. The PERST# signal shall go active before the power on M.2 connector is removed per the PCI CEM specification.

The vendor shall add SMBUS and Alert connections to the BMC base on the latest M.2 specification. Please be aware of the M.2 SMBUS is at 1.8v level. The vendor shall use a

shunt regulator to create 1.8V for SMBUS pull up and add level shift to connect to BMC SMB3.

10.8 Debug Header

The motherboard has two debug headers to work with two types of debug cards – a 14-pin debug card and a USB 3.0 debug card.

10.8.1 Debug Port Mechanical and Electrical

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

The UART channel selection button sends a negative pulse to the motherboard CPLD to select and rotate the UART console in the following loop:

Host Console → BMC Debug Console → Midplane Console

The Host Console is the default state after the debug card is connected or the system powers on. CPLD counts the negative pulse and turns the state machine states in a loop of 00-->01->10->11 to control the Most Significant Bit (MSB) and Least Significant Bit (LSB). MSB and LSB contol the output of FSA3357 to switch between the host, the BMC, and the midplane console.

The connector for the debug header is a 14-pin, shrouded, vertical, 2mm pitch connector. Figure 10-9 is an illustration of the headers. The debug card should have a key to match with the notch to avoid pin shift when plugging in.

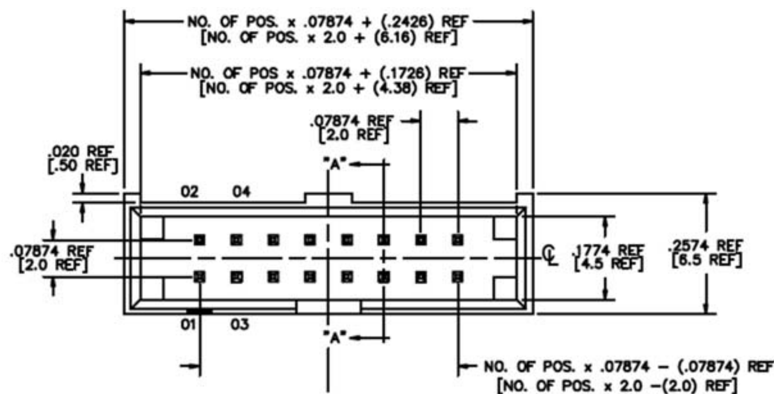


Figure 10-9 Debug Header

Table 10-12 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

10.8.2 Debug Port Dual Layout to USB 3.0 Connector

The 2nd debug header is for remapping the 5x USB 3.0 signals to UART signals to pass console data, and I²C signals to access POST codes through an I²C GPIO expander on the motherboard. The block diagram and pin definitions is shown in Figure 10-10. Please refer to “debug_card_v2_fb_sch_v06_20161127.pdf” in the OCP spec package for a detailed schematic.

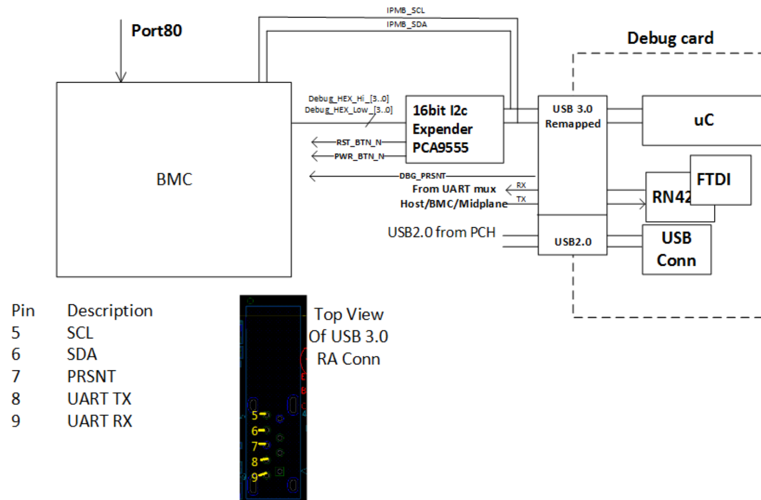


Figure 10-10: Debug Port Dual Layout to USB 3.0 Connector

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

The vendor shall follow the diagram above to implement the SMBus address. The BMC uses 0x20h, the PCA9555 uses 0x4Eh, and the microcontroller on the debug card uses 0x60h. PCA9555 GPIO mapping is shown in Table 10-13:

Table 10-13: PCA9555 GPIO Mapping

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

10.8.3 Debug Port Power Policy

Placeholder to define debug port power status in different system status (S0 /S5) and channel status, on 14x pins debug port and USB3.0.

10.8.4 POST Codes

POST codes are sent to the 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, the BIOS should also output POST codes to the BMC SOL. When the SOL session is available during POST, the remote console should show POST code as mentioned in section 8.2.

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 codes will also display error major code and minor code from Intel's memory reference code.

The display sequence will be:

"00" → DIMM location → Major code → Minor code

...with a 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 shown in Table 10-14. The DIMM number count starts with the furthest DIMM from the CPU.

10.8.5 Placeholder for DIMM Error Code Table

Table 10-14 DIMM Error Code Table

Code	Result

Placeholder for silkscreens

10.8.6 Serial Console

The output stage of the system's 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 have a mini-USB type connector with pin definition shown in Table 10-15. A separate convertor is required to provide a RS-232 transceiver and DB9 connector.

Table 10-15 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[7] or equivalent).

Table 10-16 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

10.8.7 UART Channel Selection

When the debug card is connected, pin-12 shall be used to for detecting the motherboard BMC GPIO. When UART Channel selection button on debug card is pressed, the same pin is used to send pulses to the motherboard to trigger UART connection change among host console (Default) -> BMC debug console-> midplane console (falling edge triggers).

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

The debug card has a 10KΩ pull-down resistor for pin-12, with a white UART channel selection button between pin-12 and ground.

The motherboard side should implement logic to detect debug card presence when 10K or stronger pull-down exists on pin-12.

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

The motherboard should stop its original POST code display for 1 sec when a falling edge to ground is detected, and give a 1 sec display of channel number to debug port POST code as an indication of UART channel change. If the system POST code does not change within this 1 sec, the motherboard outputs the original display. If the system POST code is changed within this 1 sec, the latest POST code should be displayed. The motherboard also has two LEDs for displaying of UART connection status as described in Table 10-17.

Table 10-17 UART channel and connection

Channel	UART Connection
00	Host console
01	BMC debug console
02	Midplane debug console

10.8.8 Other Debug Use Design Requirements on Motherboard

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

The 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.

10.9 Switches and LEDs

The motherboard shall include a power switch, reset switch, power LED, HDD activity LED, and Beep error LED.

10.9.1 Switches

Vertical tactile switches are placed behind the debug header. The push button actuator has a minimum 2.5mm diameter and protrudes 9mm+/-1mm from the top of the actuator to the PCB surface. The system power button is red and on the left. The system's 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 depressed 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 firmware or have any 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.

10.9.2 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 10-18. Looking from the I/O towards the LEDs, from right to left, the sequence is Blue, Green, Yellow, Green, Green.

Table 10-18 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]

10.10 Fan connector

The motherboard has two 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 six pins, which is compatible with standard 4-Wire FAN connector, and also can be used to support a dual rotor fan that shares PWM control signal but has separate tachometer (TACH) signals. The fan connector pin definition is available in Table 10-19. LOTES APHD0019-P001A or equivalent shall be used as a 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 a pump header to support water cooling. Pump header has same form factor as the fan connectors. The pump header will have a cap to avoid connection to a system fan.

Table 10-19 FAN Connector Pin Definition

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

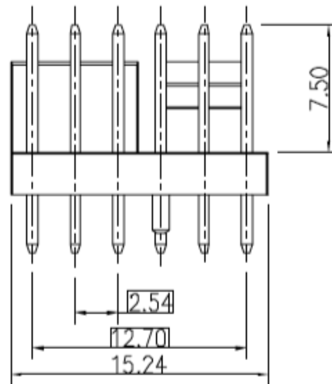


Figure 10-11 Fan Connector

10.11 TPM Connector and Module

An 11-pin vertical receptacle connector is defined on MB for SPI and I^{2}C TPM module. The connector pin definition on the motherboard side is shown in **Error! Reference source not found.**.
FCI/91931-31111LF receptacle or equivalent should be used on motherboard.

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

Table 10-20: TPM Header Pin Definition

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			

Placeholder for Picture

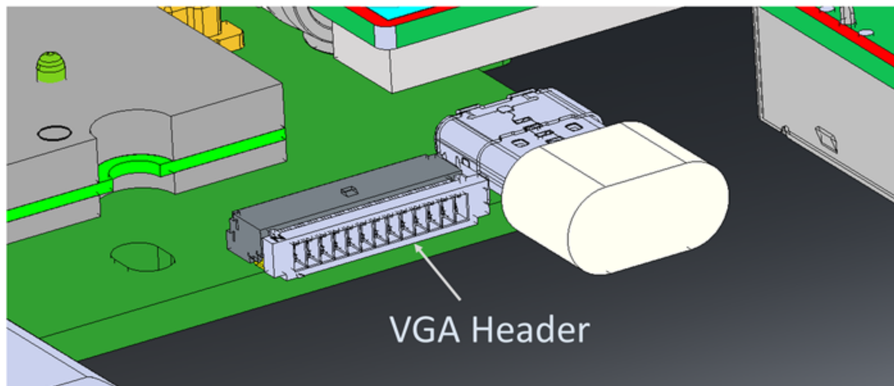
10.12 Sideband Connector

An 8-pin 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

10.13 VGA header

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



The vendor shall refer to the DXF file for VGA header placement. Table 10-21 shows the signal assignment.

Table 10-21: VGA Connector Signal Assignment

#	Signal
1	RED
2	RED_RTN(GND)
3	Green
4	Green_RTN (GND)
5	Blue
6	Blue_RTN (GND)
7	V-Sync
8	GND
9	H-Sync

10	GND (H-sync)
11	SDA
12	SCL
13	PWR

11 Rear Side Power, I/O and Midplane

11.1 Overview of Footprint and Population Options

There are four footprints at rear side of the motherboard to provide power to the motherboard and I/O to a midplane. The population of the footprints is flexible to fit the need of different use cases.

Population options and major differences are listed in **Error! Reference source not found.**. The high speed midplane is not covered in this document. The ORv2 implementation is described in Chapter 12.

Placeholder for midplane connector population options

Motherboard midplane connector population options

11.2 Rear Side Connectors

11.2.1 Footprints and Connectors/Pressfit Cable

There are three types of connector footprints. A design can install a combination of three types of connectors and one type of Pressfit cable to the footprints. The AVL is listed in Table 11-1. The placement of connector footprints is shown in **Error! Reference source not found.**.

Table 11-1

Connector Type	Sled side P/N	Midplane 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

Placeholder for Placement picture

Placement of connector modules

11.2.2 AirMax® power 2x2

In a use case with less or equal to 49 A[8] on 12VDC, one pair of AirMax® power 2x2 R/A low profile headers and receptacles are populated. Two pairs of this connector support up to 98A on 12VDC.

11.2.2.1 Sled Side

Up to two AirMax® power 2x2 R/A low profile headers are used at sled side and is shown in **Error! Reference source not found.**. Pin definition is as Table 11-2.

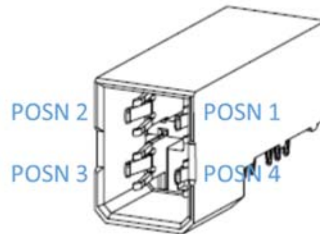


Figure 11-1: Power Connector at Sled Side

Table 11-2 Pin definition on both sled and midplane side

Position	Type	Description
1,2	Power	P12V_AUX power rail from midplane 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

11.2.2.2 Midplane side

Up to two AirMax® power 2x2 R/A receptacles for co-planar applications are used on the midplane side, and shown in Figure 11-2. This receptacle has long and short pins 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 11-3 for detail.

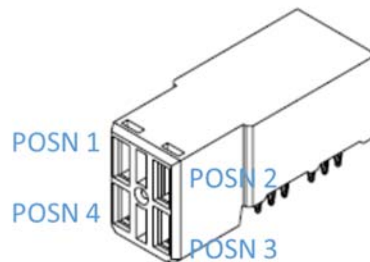


Figure 11-2: Power Connector at Midplane Side

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

Table 11-3 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

11.2.3 AirMax® Guide

One pair of AirMax® 7.2mm R/A guides is used on the motherboard and midplane ONLY in the use cases that blind mate is needed.

11.2.3.1 Sled Side

The sled side uses one AirMax® 7.2mm R/A guide blade as Figure 11-3**Error! Reference source not found.**.

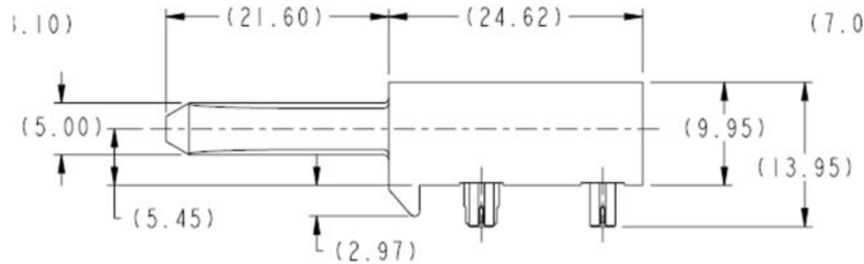


Figure 11-3: AirMax® 7.2mm R/A Guide Blade

11.2.3.2 Midplane Side

The midplane side uses one AirMax® 7.2mm R/A guide socket as Figure 11-4.

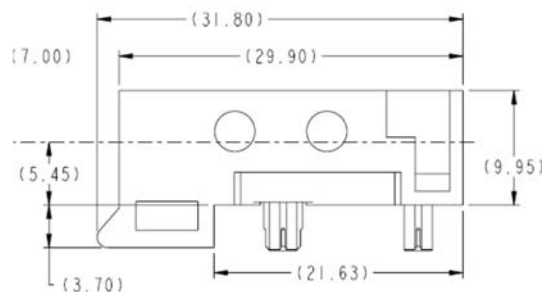


Figure 11-4: AirMax® 7.2mm R/A Guide Socket

11.2.4 AirMax® 3x8 Signal

One AirMax® VS/VS2 3x8 connector follows the definition in Table 11-4 in the perspective of the sled side. The signal description is shown in Table 11-5. It supports up to x8 PCIe plus side band signals and management interface. This connector is optional to support

a high speed midplane with PCIe interface.

Table 11-4 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 11-5 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 midplane; OD output from motherboard.
FAN_TACH[3..0]	Input	FAN TACH input from midplane to motherboard or uServer sled. OD output at midplane. 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 midplane; low active; need PU at motherboard or uServer sled.
PCIE_PERST_N	Output	PCIe reset signal from motherboard to midplane. Low active. 3.3V push pull from motherboard.
PCIE_TX_DP/N[7..0]	Output	PCIe transmit signal from motherboard to midplane; AC decoupling at motherboard side.
PCIE_RX_DP/N[7..0]	Input	PCIe receive signal from midplane to motherboard; AC decoupling at midplane side.
PCIE_CLK_100M_DP/N	Output	100MHz PCIe clock from motherboard to midplane
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 midplane to motherboard or uServer sled. PD 100ohm or open at midplane to indicate different slot locations.

MB_ON	Output	Motherboard on indication from motherboard to midplane; Push Pull P3V3_STBY output at motherboard
COM_TX	Output	3.3V UART console TX from motherboard or uServer sled to midplane
COM_RX	Output	3.3V UART console RX from midplane 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 11-5 for implementation at motherboard or uServer sled side and at midplane side. Fully mating of sled enables hot-swap controller on motherboard or uServer sled. This action also notice midplane the presence of sleds.

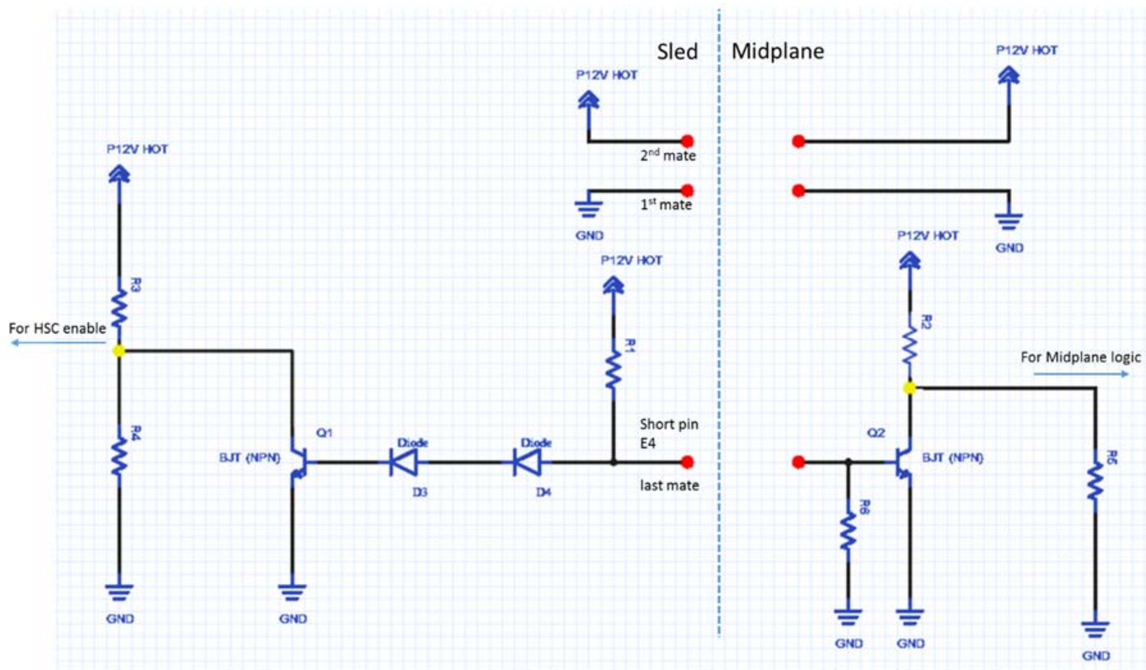


Figure 11-5 Reference Circuit for Dual Side Presence/Mate Detection with 1 Short Pin

11.2.5 Pressfit Cable

A pressfit cable is enabled for Orv2 Cubby chassis installations. A side view is depicted in Figure 11-6. A pressfit cable is enabled for the use case of ORv2 and Cubby chassis. A side view is shown in Figure 11-6.

One side of the Pressfit cable is a pressfit power connector. The pressfit power connector is installed on the motherboard with Pressfit process, and secured by a screw for added strength. The 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 installation and removal.

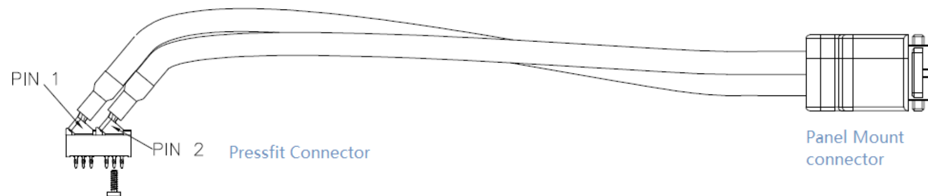
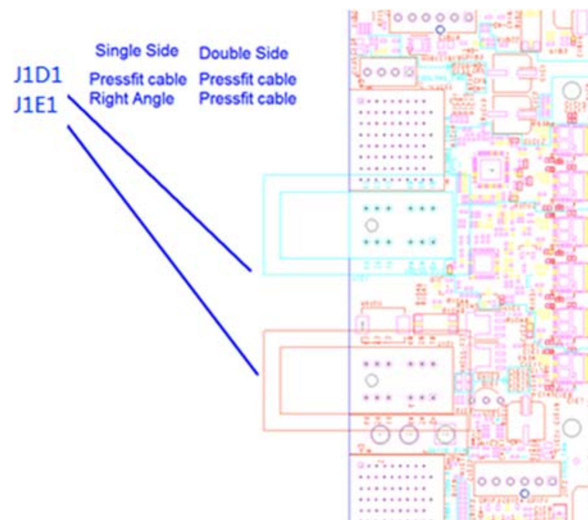


Figure 11-6: Pressfit Cable drawing-Side View

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

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



11.3 Midplane

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

The high speed midplane is a midplane with power delivery and a high speed interconnect. One or both of the AirMax® 3x8 and AirMax® 3x6 must to be populated for this use case. The midplane can have one of these optional active components – high speed signal switch, high speed signal repeater/buffer. The midplane cannot have any active components if the midplane only provides high speed interconnect with PCB trace.

The midplane provides a mechanical and electrical interface for the DC power cable assembly. Each of the two slugs of DC power cable assembly is fixed to the midplane through two screws. There is a notch feature on the lug of DC power cable. The midplane should design a key feature to mate with this notch to provide foolproof design.

Two 80mm fans are directly attached to two fan connectors on the motherboard. The design of the midplane should allow the replace of midplane without removing the motherboard from the tray.

12 ORv2 Implementation

12.1 Cubby for ORv2

The Cubby serves as the mechanical and power delivery interface between ORv2 and Intel Motherboard V4.0-ORv2 sled. Figure 12-1 shows the Cubby enclosure for the Intel Motherboard V4.0-ORv2 sled. The vendor should refer to the 3D design file for more detailed information.

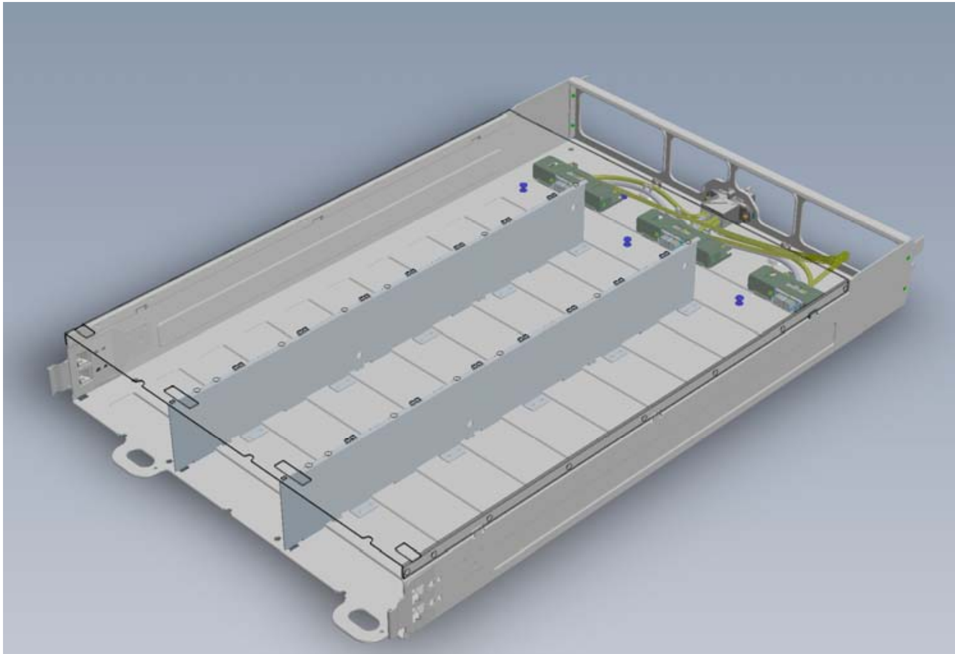


Figure 12-1: Cubby

12.2 Intel Motherboard V4.0-ORv2 Power Delivery

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

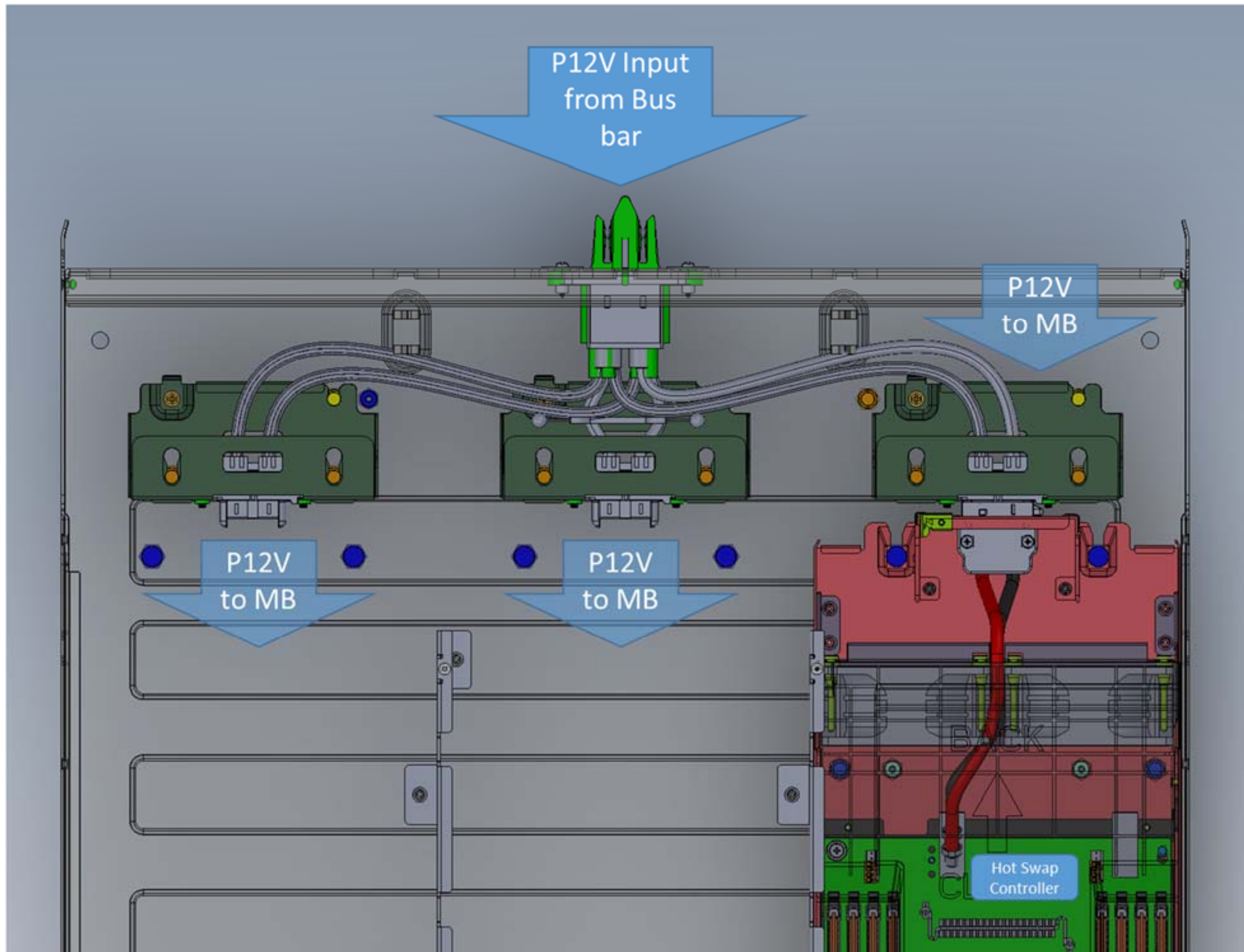


Figure 12-2: Intel Motherboard V4.0-ORv2 sled in Cubby with Medusa cable

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

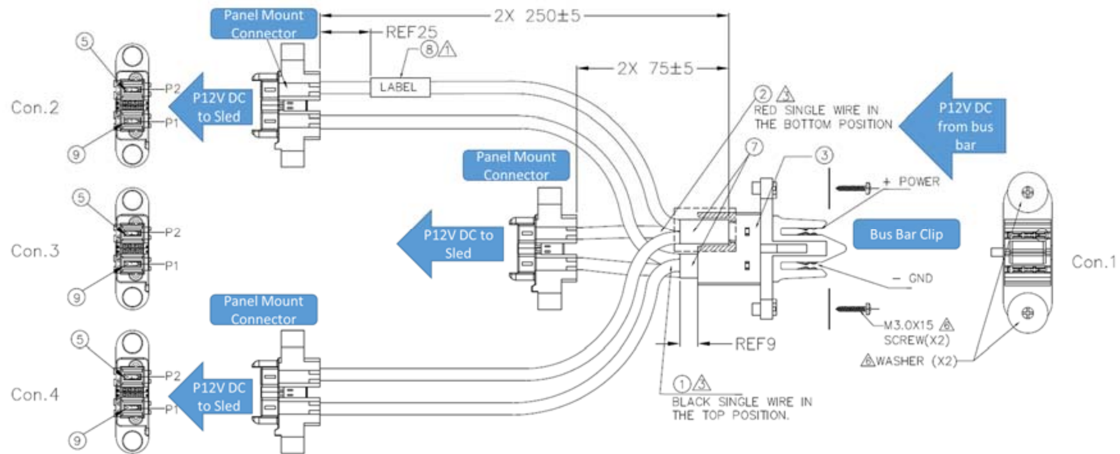


Figure 12-3: Medusa Cable drawing top view

Each motherboard for the ORv2 sled has a Pressfit cable installed on the motherboard. 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 **Error! Reference source not found.**) that interfaces with one of the panel mount connectors on medusa cable. The panel mount connector can be removed from the sled without using tools.

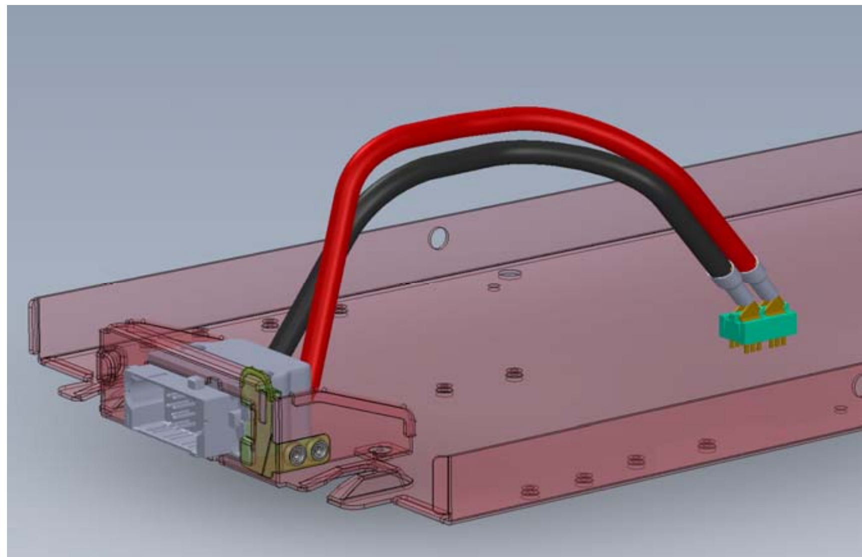


Figure 12-4: Panel Mount Connector on Intel Motherboard V4.0-ORv2 Sled

12.3 Intel Motherboard V4.0-ORv2 Single-Side Sled

A sheet metal tray serves as the mechanical interface between the motherboard and Cubby. It provides mechanical retention for the components inside the tray, such as the pressfit cable, fan, riser card, PCIe cards, hard drive, and Mezzanine card. The combination of tray, motherboard with pressfit cable and the other components assembled in the tray is an Intel Motherboard V4.0-ORv2 sled. Vendor should refer to 3D for more detail.

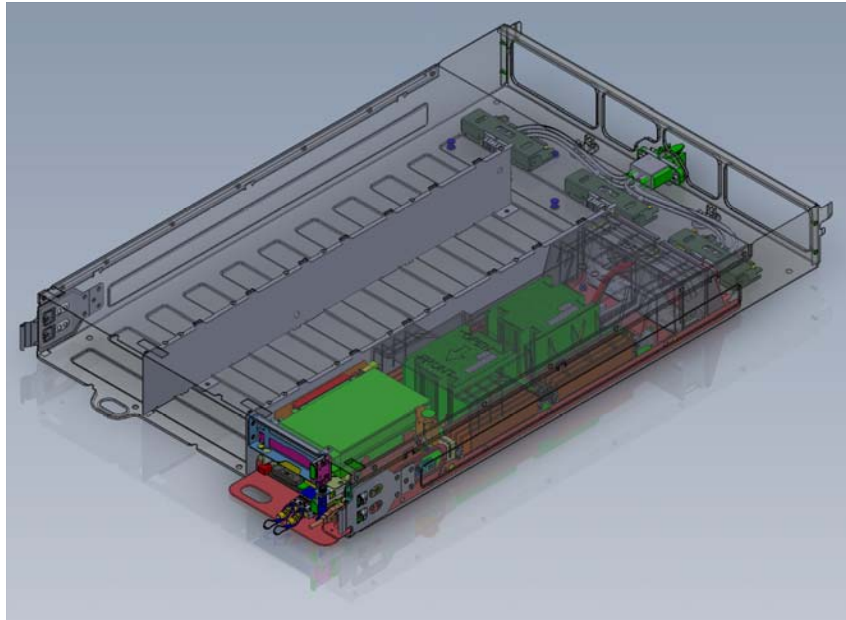


Figure 12-5: Intel Motherboard V4.0-ORv2 sled in Cubby

12.4 Intel Motherboard V4.0- ORv2 Double Side Sled

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

13 Mechanical

The Intel Motherboard V4.0-ORv2 sled should work with Open Rack V1 and ORv2 mechanically and the implementation guide provided in Chapter 12.

13.1 Single Side Sled mechanical

For the 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 (89.44mil).

13.1.1 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.

13.2 Double Side Sled Mechanical

Double side sled for ORv2 is a new design.

Placeholder: Use this section for describing double side sled mechanical design.

13.3 Fixed Locations

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

13.4 PCB Thickness

To ensure proper alignment of the motherboard and midplane interface within its mechanical enclosure, the motherboard should follow PCB stack up in Table 5-1 to have 89.44mil (2.27mm) PCB thickness. The midplane PCB thickness should also be 89.4mil (2.27mm). The Mezzanine card and riser card PCB thickness should be 62mil (≈1.57mm).

13.5 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 referenced specification.

13.6 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 0.

- CPU0 / CPU1
- DIMM slot numbering, as described in **Error! Reference source not found.**
- LEDs as defined in 10.9.2
- Switches as PWR and RST

13.7 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 that is the furthest

placed 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.

13.8 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 13-1 PCB Color

Revision	PCB Color
EVT	Red
DVT	Yellow
PVT	Green

14 Motherboard Power System

14.1 Input Voltage

14.1.1 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 14.2.

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

14.1.2 Capacitive Load

Previous server generations required a maximum capacitive load of 4,000 uF. This requirement does not apply to Intel Motherboard V4.0 design. The motherboard design requires greater than 10,000 uF capacitive loading on P12V_AUX for supplying surge current from CPU VR, and slew rate reduction of P12V_AUX decaying for NVDIMM feature at surprising power fail. The hot-swap controller design should limit the inrush current to the node during soft-start to less or equal to 10A.

14.1.3 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.

14.1.4 P1V_PSU to GND clearance

Due to P12V_PSU is without over current protection of hot-swap 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

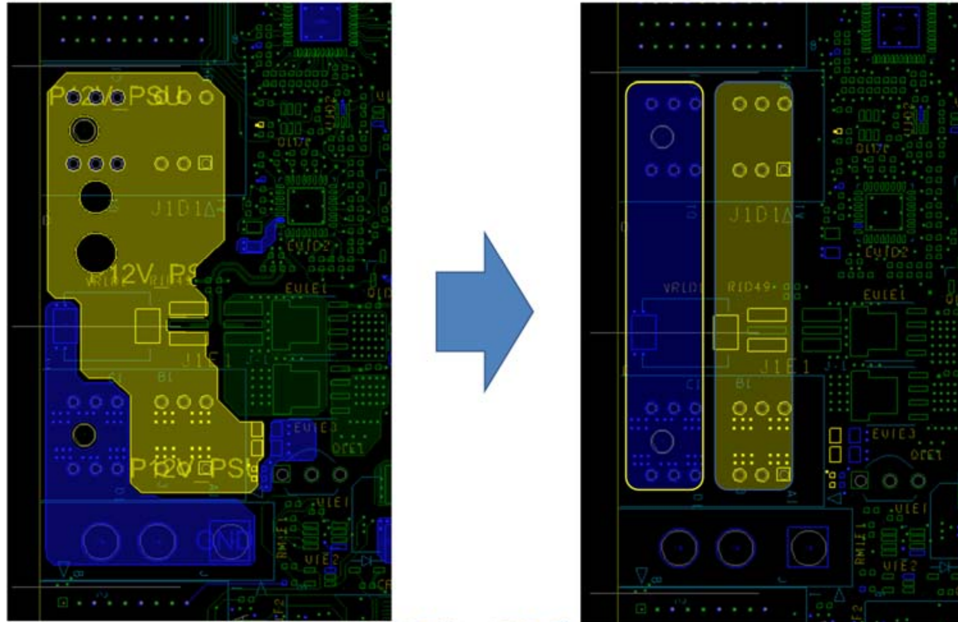


Figure 14-1: P1V_PSU to GND Clearance

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.

14.2 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. The HSC circuit provides the following functions:

- In-rush current control when the motherboard is inserted and powered on.
- Current limiting protection for over current and short circuit conditions. Over current trip point should be able to set to 48.8A and 41.4A with Iset jumper setting for single-side motherboard; default is 48.8A. 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
- 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 PCH Intel® ME 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 of theHSC 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. This feature can be disabled by BMC GPIO directly. The 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 the BMC is ready, the hardware POR state is enable.
- Use of the 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, the trip point is 11.5v or follow BMCs. When the jumper is at pin-3 and pin-5, the trip point is 11.0v.
- Use of the 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 enabled.
- Use of the HSC or external circuit to provide HSC timer alert to trigger system throttling before HSC OCP happens.
- Refer to Table 14-1 for setting requirements of System, CPU, and memory sub-system throttling.

Table 14-1 Entry point of System, CPU, and Memory Sub-system Throttling

Condition	Threshold	Action	Enable control	Default
Board input power over current limit	>40.6A or 47.9A[Default] by jumper setting for SS >70.8A[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[9]	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 the HSC current sense resistor should be less or equal to 25mV at full loading. Hot-swap controllers should have SMBUS address set to 0x11 (7-bit

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

format) on single side with 0.5mΩ Rsen, 0x45 (7bit format) on double side with 0.25mohm Rsen

- The power reporting of the hot-swap controller needs to be better than +/-2% from 50W to full loading in room temperature as a minimal requirement. The 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:

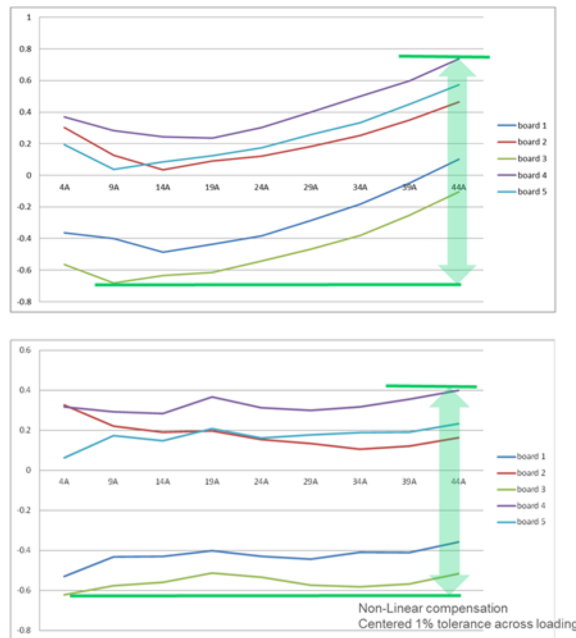


Figure 14-2

14.3 CPU VR

14.3.1 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 **Error! Reference source not found.**. 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 VR specification and validation method and pass test with margin.

14.3.2 CPU VRM Efficiency

For CPU efficiency measurement:

- VID is set to 1.8V and 1.6V (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 curve shall be higher than the envelope defined below

Placeholder for VR Eff

Figure 14-3 Efficiency envelope requirement of CPU VCCIN VR

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.

14.3.3 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.

14.4 DIMM VR

14.4.1 DIMM Maximum Power

The motherboard has a DIMM configuration of two CPU sockets, per socket, and two slots per channel. The vendor should follow the vendor's memory controller guidelines to design and validate DIMM power rail to support maximum power needed for this configuration, and support 1.2V DDR4 DIMM.

14.4.2 DIMM VR Optimizations

DIMM VR should support auto phase dropping for high efficiency across loading. DIMM VR should be compliant to latest VR 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 DIMM slots and Double side board DIMM slots.

14.4.3 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 furthest two DIMM slots from CPU
- 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

(Placeholder)

Figure 14-4 Efficiency envelope requirement of Memory VDDQ VR

14.4.4 DIMM VR configuration

DIMM VR has same configuration requirement as CPU VR, listed in 14.3.3.

14.4.5 DIMM VR MLCC Co-Layout

There are 18 C0603 and 2 C0805 for each DDR4 Memory VR on the top layer of the board between DIMM fields. It is shown in Figure 14-5.

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

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

The 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-sideboard, populate 12 DIMM sockets on top side. Added MLCC specified in this section are populated

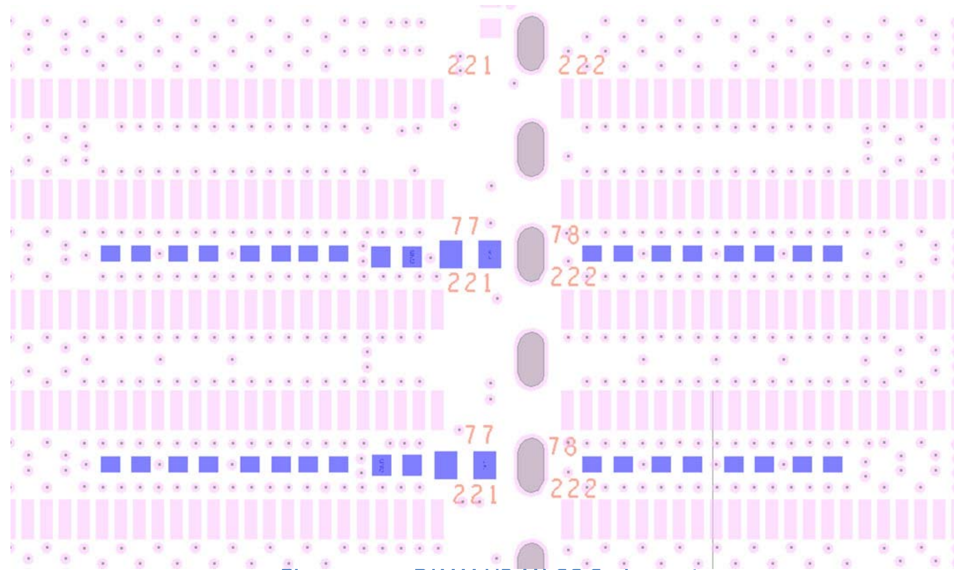


Figure 14-5: DIMM VR MLCC Co-Layout

14.5 MCP (Multi Core 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 be still be supported.

14.6 VRM design guideline

For VRM, the vendor should list the 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 that require firmware, 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 6.4 64-bit with a kernel version specified by the customer.

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

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

Table 14-2: CPU, DIMM, and PCH VR

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_	2	TDA21470	FDMF3180
PVDDQ_	2	TDA21470	FDMF3180
PVDDQ_	2	TDA21470	FDMF3180
PVDDQ_	2	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. The 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.

14.7 Hard Drive Power

The motherboard shall supply power to all possible nine hard drives connected. Drives require 12VDC and 5VDC power sources. For single individual SATA ports, the power will be delivered through a traditional 4-pin floppy disk power connector, Tyco 171825-4 or equivalent. The mating connector is a Tyco 171822-4. The pin assignment shall follow industry standard convention described in
Table 14-3.

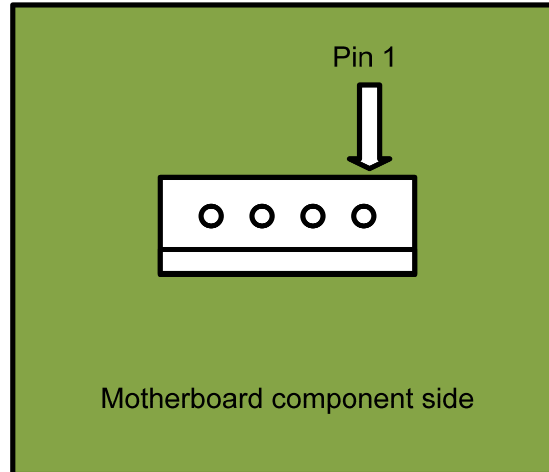


Table 14-3 4-pin floppy disk power connector

Pin	Description
1	+5VDC
2	GND
3	GND
4	+12VDC

14.7.1 HDD Power Requirements

The motherboard must provide enough power delivery on 12.5VDC and 5VDC to support up to nine HDDs this platform supports. This means supporting 1A continuous current per HDD on a 12.5VDC power rail, and 0.75A continuous current per HDD on a 5VDC power rail. In-rush current required to spin up the drive must also be considered in power delivery design.

14.7.2 Output Protection

Both 12V and 5V disk output power rails shall protect against shorts and overload conditions.

14.7.3 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 nine hard drives installed, so there is a need to spin up hard drives in sequence. BIOS should implement five seconds delay between each hard drive spinning up. In order to do this, the SATA hard drive's power cable should have pin-11 as NC (No Connection) to enable hard drive's spin-up delay function.

14.8 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 make those options known.

14.9 Power On

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

14.10 High power use case

High power use case means the the system power is between 480W (40A@12V) and 960W (80A@12V). This is not a typical FB use case, but support is required to enable testing of such configuration. Typically it is caused by fully populating Non-Volatile DIMM, 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 1mΩ to 2x 0.5mΩ

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

15 Environmental and Regulations

15.1 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: -5°C to +45°C
- Operating and Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C
- Transportation temperature range: -55°C to +85°C (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: -5°C to +35°C
- Operating and Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C
- Transportation temperature range: -55°C to +85°C (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: 65°F to 85°F
  - Humidity: 30% to 85%
  - Altitude: 1000m (3300 feet)

Site 2 as:

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

15.2 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 15-1. The motherboard shall exhibit full 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 15-1 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)

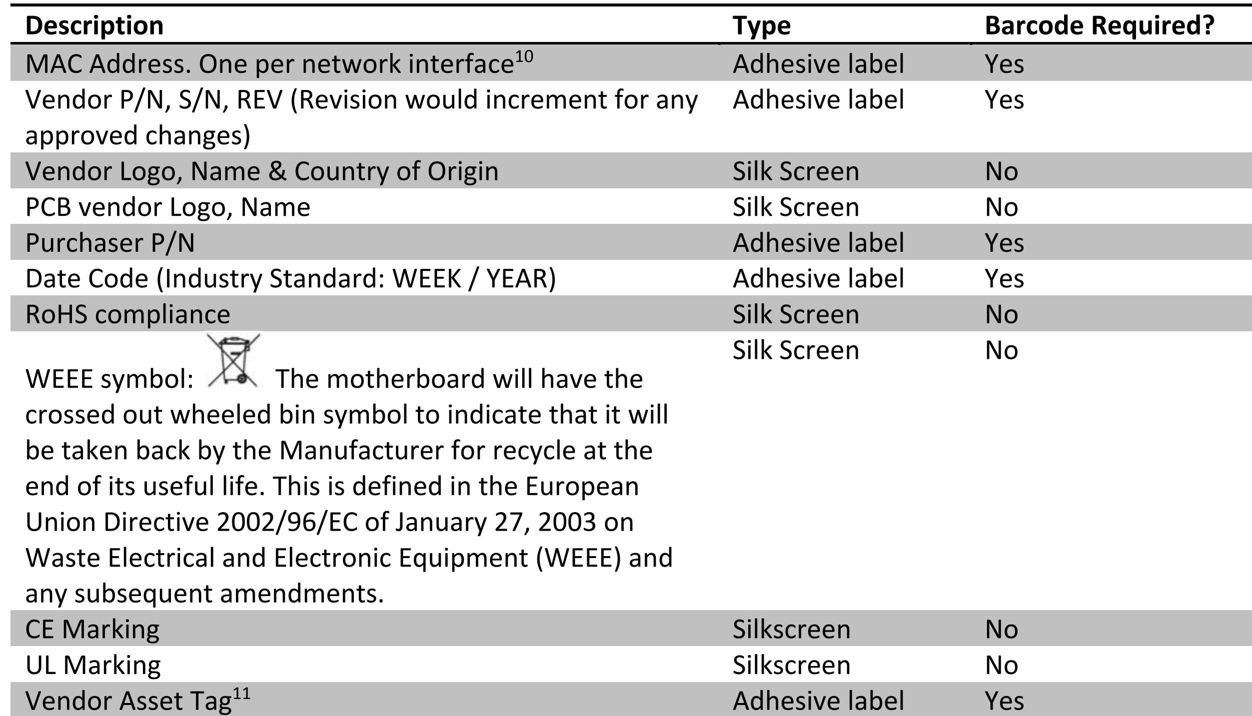
Shock	6g, half-sine 11mS, 5 shocks per each of the three axes	12g, half-sine 11mS, 10 shocks per each of the three axes

15.3 Regulations

The vendor needs 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.

16 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 airflow path of the motherboard.



Description	Type	Barcode Required?
MAC Address. One per network interface[10]	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[11]	Adhesive label	Yes

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

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

17 Prescribed Materials

17.1 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

17.2 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 105°C, and shall be selected only from Japanese Manufacturers.
- All capacitors will have a predicted life of at least 50,000 hours at 45C 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. The 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.

17.3 Component De-rating

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

18 Reliability and Quality

18.1 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. The customer must be notified if any changes are made which may impact product quality.

18.2 Change Orders

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

18.3 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 the vendor's facility.

18.4 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.

18.5 MTBF Requirements

The motherboard shall have a minimum calculated MTBF of 300K hours at 90% confidence level at 45°C 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 one 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 45°C ambient temperature).

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

18.6 Quality Control

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

- Incoming product must have less than 0.1% rejections
- Cpk values will exceed 1.33 (Pilot Build & Production)
- 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.

18.7 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.

18.8 PCB Tests

Server ODM should arrange Independent 3rd party lab testing on Delta-L, IST, and IPC-6012C 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 five 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 working panel for riser cards or midplane board.

18.9 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 components should be included in DVT.

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

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

19 Deliverables

19.1 OS Support

Motherboard shall support CentOS 6.4 64-bit with updated Kernel specified by customer, and pass Red Hat certification tests.

19.2 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.

19.3 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

19.4 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 19.3 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 0.

20 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. A shock test for the shipping box shall be conducted by the vendor and submitted to customer for audit and approval.

21 Appendix

21.1 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

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

21.2 Mechanical drawings

Following mechanical drawings are provided

- 3D CAD for Intel Motherboard V4.0-ORv2 Single Side sled (preliminary, subject to design change)
- 3D CAD for Intel Motherboard V4.0-ORv2 Double Side sled

21.3 SMBIOS FRU mapping table

Tioga Pass SMBIOS and FRU Mapping v01

SMBIOS Type 1~4 and 11				Ma	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 Info Area	Product Name	**[Tioga Pass Single Side/Tioga Pass Double Side]**	
	SMBIOS type 1 didn't support Model Name.				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			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 Number	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 number	[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 applicable]	<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>

21.4 *Add-on-Card Thermal Interface Spec for* Intel Motherboard V4.0