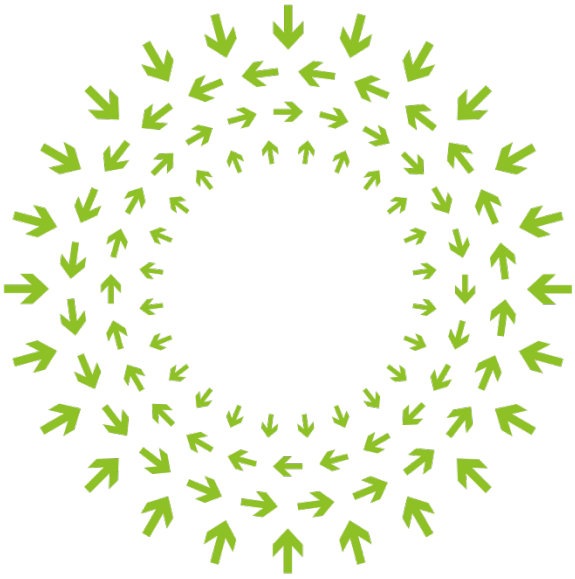


OPEN

Compute Project

Facebook Server
Intel Motherboard v3.0

Author: Jia Ning, Engineer, Facebook

Contents

Contents 2
1 Scope 5
2 Overview 5
2.1 License 5
3 Efficient Performance Motherboard v3 Features 6
3.1 Block Diagram 6
3.2 Placement and Form Factor 6
3.3 CPU and Memory 7
3.4 Platform Controller Hub 8
3.5 Printed Circuit Board Stackup (PCB) 8
4 Basic Input Output System (BIOS) 10
4.1 BIOS Chip 10
4.2 BIOS Source Code 10
4.3 BIOS Feature Requirements 10
5 Baseboard Management Controller (BMC) 14
5.1 Management Network Interface 14
5.2 Local Serial Console and Serial-Over-LAN (SOL) 14
5.3 Graphic and GUI 15
5.4 Remote Power Control and Power Policy 15
5.5 Port 80 POST 15
5.6 Power and System Identification LED 15
5.7 Platform Environment Control Interface (PECI) 16
5.8 Power and Thermal Monitoring and Power Limiting 16
5.9 Sensors 16
5.10 System Event Log (SEL) 21
5.11 Fan Speed Control in BMC 23
5.12 BMC Firmware Update 24
5.13 BMC Update: Dual PCH flash 24
5.14 BMC Update Complex Programmable Logic Device (CPLD) 26
6 Thermal Design Requirements 26

6.1 Data Center Environmental Conditions 26
6.2 Server Operational Conditions 27
6.3 Thermal Kit Requirements 28
7 Midplane Interface and Midplane 29
7.1 Connector Modules 29
7.2 Pin Definition 29
7.3 Population options 33
7.4 Midplane for Open Rack V1 34
8 Motherboard Power system 35
8.1 DC Cable Assembly 35
8.2 Input Voltage 35
8.3 Hot Swap Controller (HSC) Circuit 35
8.4 CPU Voltage Regulator (VR) 36
8.5 DIMM Voltage Regulator 37
8.6 Voltage Regulator Module Design Guideline 37
8.7 Hard Drive Power 38
8.8 System VRM Efficiency 39
8.9 Power On 39
9 I/O System 40
9.1 PCIe* x24 Slot Connector/Riser Card 40
9.2 DIMM Connector 41
9.3 PCI-E Mezzanine Card 41
9.4 Network 43
9.5 USB 43
9.6 SATA 44
9.7 Next Generation Form Factor (NGFF) Connectors 44
9.8 Debug Header 44
9.9 Switches and LEDs 47
9.10 Fan Connector 49
9.11 TPM Connector and Module 49
10 Mechanical 51
10.1 Shelf 51
10.2 Tray and Sled 51
10.3 PCIe and HDD bracket 52
10.4 Fixed Locations 52

10.5 PCB Thickness 52
10.6 Heat Sinks and Independent Loading Mechanism (ILM) 52
10.7 Silk Screen 52
10.8 DIMM Connector Color 52
10.9 PCB Color 52
11 Environmental Requirements 53
11.1 Vibration and Shock 53
12 Prescribed Materials 54
12.1 Disallowed Components 54
12.2 Capacitors and Inductors 54
12.3 Component De-rating 54

1 Scope

This document defines the technical specifications for the Intel motherboard v3 used in Open Compute Project servers.

2 Overview

When data center design and hardware design move in concert, they can improve efficiency and reduce power consumption. To this end, the Open Compute Project is a set of technologies that reduces energy consumption and cost, increases reliability and choice in the marketplace, and simplifies operations and maintenance. One key objective is openness—the project is starting with the opening of the specifications and mechanical designs for the major components of a data center, and the efficiency results achieved at facilities using Open Compute technologies.

One component of this project is a custom motherboard. This document describes the Open Compute Project Efficient Performance Intel motherboard. The motherboard is power-optimized and bare bones, designed to provide the lowest capital and operating costs. Many features found in traditional motherboards have been removed from the design.

2.1 License

As of April 7, 2011, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0:
Facebook, Inc.

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at http://opencompute.org/licensing/, which may also include additional parties to those listed above.

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 entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. 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.

3 Efficient Performance Motherboard v3 Features

3.1 Block Diagram

The figure below illustrates the functional block diagram of the Efficient Performance Motherboard v3.

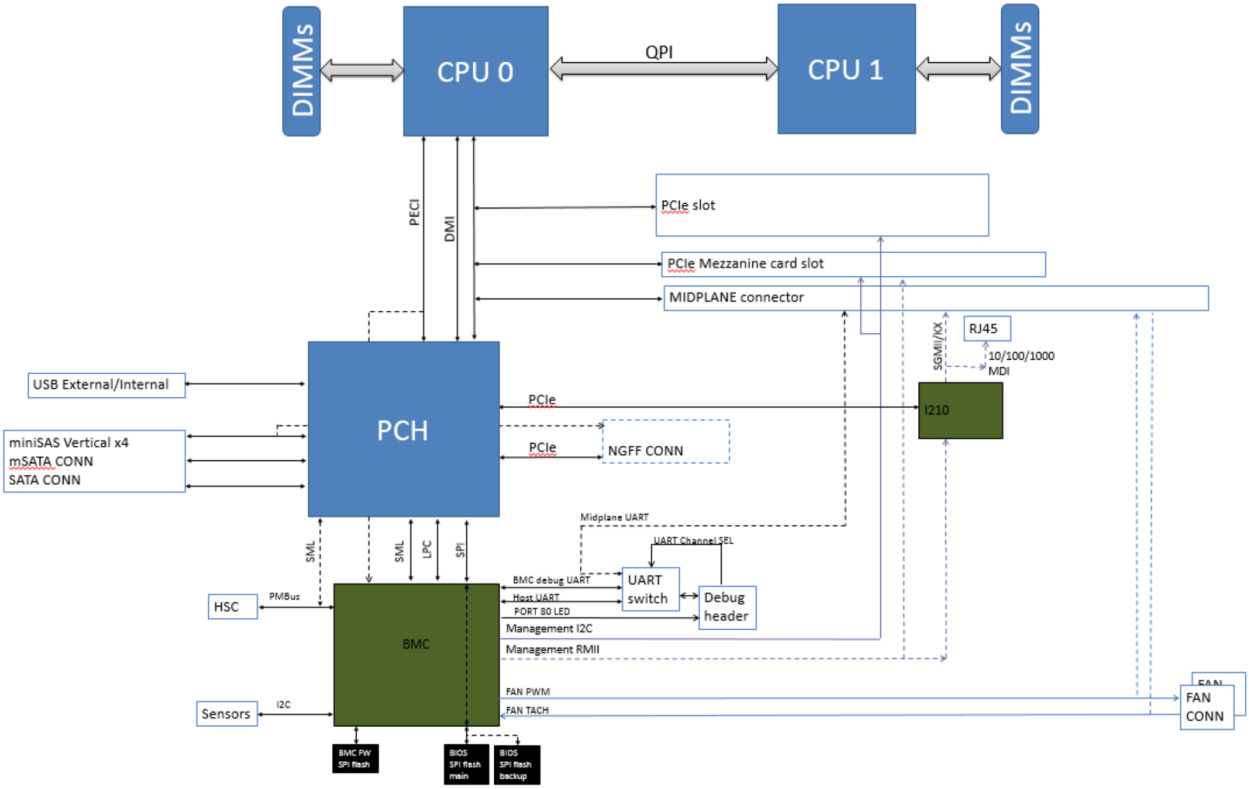


Figure 1: Efficient Performance Motherboard v3 Functional Block Diagram

3.2 Placement and Form Factor

The motherboard's form factor is 6.5 inches x 20 inches. Figure 2 illustrates board placement. The placement shows the relative positions of key components. Exact dimension and position information is available in the mechanical drawing interchange file (DXF). Strictly follow the form factor, CPU, and DIMM socket location, PCI-E slot position, front I/O ports' position, Power and Reset buttons, PCI-E Mezzanine Card connector position, midplane connectors, battery connector, and mounting hole locations. Other components can be shifted based on layout routing as long as relative position is maintained.

The following internal connectors are placed as close as possible to front of the board in order to have easy frontal access:

- 1x vertical SATA signal connector
- 1x SATA power connector
- 1x Debug card header
- 1x mSATA connector(dual layout with one next-generation form factor [NGFF] connector)
- 1x RJ45 connector

The following midplane connectors should be placed at the back of the board to provide the midplane connection co-layout:

- 1x Airmax VS2 3x8
- 1x Airmax Guide
- 2x Airmax VS Power 2x2

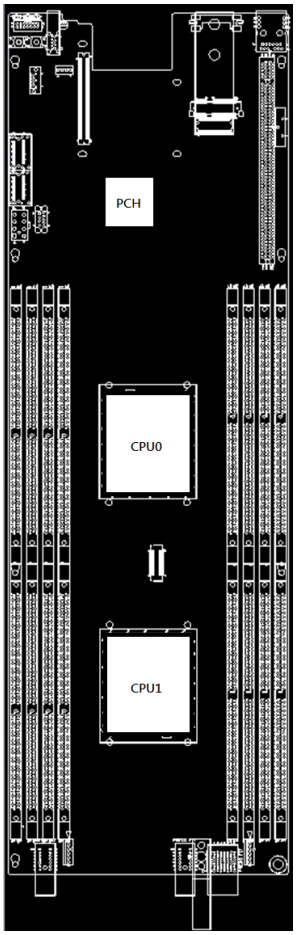


Figure 2: Efficient Performance Board v3 Placement

3.3 CPU and Memory

3.3.1 CPU

The motherboard uses next generation Intel® Xeon® CPUs with a TDP (thermal design power) up to 145W. The motherboard supports these features:

- 2 next generation Intel® Xeon® processors
- Single Processor Mode
- DDR4 direct attached memory support on cpu0 and cpu1

3.3.2 Non-Volatile DIMM (NVDIMM)

Besides traditional DDR4 DIMM, the motherboard needs to support NVDIMM on all DIMM slots. A power failure detection circuit must be implemented to initiate three actions related to data transferring:

1. CPU cache flush

2. Memory controller write pending queue flush
3. Joint Electron Device Engineering Council (JEDEC) standard DIMM support to store DRAM memory into flash (non-volatile) memory

Due to system energy storing and timing requirement limitations, the logic of action 1) is disabled by default with the resistor option set to Enable. The logic of action 2) and 3) is enabled by default with the resistor option set to Disable. The original design manufacturer (ODM) will work with an NVDIMM vendor to implement a BIOS design.

Action 2) and 3) should be supported and validated in a DDR4 SKU board.

The under-voltage based power failure detection circuit should also trigger separate CPU/DIMM/FAN throttling with separate resistor options for Enable and Disable. The default is set to Enable. This reduces power consumption and slows down voltage decay on P12V.

3.4 Platform Controller Hub

The Efficient Performance board v3 uses the next generation Intel® platform controller hub (PCH), which supports the following features:

- USB ports
- SATA ports connected to NGFF connector, mSATA connector, standard SATA port and miniSAS ports
- PCIe ports to the NGFF connector
- SPI interface, connect to BMC to enable BMC the capability to perform BIOS upgrade and recovery
- SMBUS interface (master and slave)

3.5 Printed Circuit Board Stackup (PCB)

The Efficient Performance board v3's PCB stack-up and impedance control are defined in the table below.

Table 1: Efficient Performance Board PCB Stackup

Layer	Plane Description		Copper (oz)	Thickness (mil)	Dielectric (er)
		Solder Mask		0.5	3.8
L1	TOP	Signal	0.5+1.0	1.9	
		PrePreg		2.7	3.5
L2	GND1	Ground	1.0	1.3	
		Core		4.0	3.6
L3	IN1	Signal	1.0	1.3	
		PrePreg		7.7	4.0
L4	GND2	Ground	1.0	1.3	
		Core		4.0	3.6
L5	IN2	Signal	1.0	1.3	
		PrePreg		12.0	4.3
L6	VCC1	Power	2.0	2.6	
		Core		4.0	3.6
L7	VCC2	Power	2.0	2.6	

		PrePreg		12.0	4.3
L8	IN3	Signal	1.0	1.3	
		Core		4.0	3.6
L9	GND3	Ground	1.0	1.3	
		PrePreg		7.7	4.0
L10	IN4	Signal	1.0	1.3	
		Core		4.0	3.6
L11	GND4	Ground	1.0	1.3	
		PrePreg		2.7	3.5
L12	BOT	Signal	0.5+1.0	1.9	
		Solder Mask		0.5	3.8
		Total		85.2	Tolerance: +/-8mil

Table 2 Efficient Performance Board PCB Impedance Control

Trace Width (mil)	Air Gap Spacing (mil)	Impedance Type	Layer	Impedance Target (ohm)	Tolerance (+/- %)
4.0		Single	1,12	50	15.0
6.5		Single	1,12	40	15.0
3.9	4.1	Differential	1,12	83	17.5
3.8	8.2	Differential	1,12	95	17.5
4.9	5.1	Differential	1,12	85	17.5
4.0		Single	3,5,8,10	53	10.0
4.5		Single	3, 5,8,10	50	10.0
8.0		Single	3, 5,8,10	40	10.0
4.0	4.0	Differential	3, 5,8,10	83	12.0
4.0	8.0	Differential	3, 5,8,10	95	12.0
5.0	7.0	Differential	3, 5,8,10	85	12.0

4 Basic Input Output System (BIOS)

The ODM is responsible for supplying and customizing a BIOS for the motherboard. This section outlines the specific BIOS requirements.

4.1 BIOS Chip

The BIOS chip should use PCH's SPI interface through a baseboard management controller (BMC) controlled multiplexer (MUX) for the BMC to perform offline BIOS updates or recovery. The vendor is responsible for selecting a specific BIOS chip. The chip should fit the required functionality and be able to support potential additional features that may be required in the future. Considering the space needed for both BIOS and PCH, we recommend a 16MByte size. The vendor should provide a flash region plan for different code and current used size for each region to justify the sizing of the serial peripheral interface (SPI) flash.

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

A secondary identical BIOS chip is designed to share the same SPI bus with multiplexed chip select pin. The MUX is controlled by BMC. More detail of this scheme is described in Section 5.13.

4.2 BIOS Source Code

The BIOS source code comes from the American Megatrends Inc. (AMI) extensible firmware interface (EFI). The ODM is responsible for maintaining the BIOS source code to make sure it has the latest version from AMI and Intel.

4.3 BIOS Feature Requirements

4.3.1 Optimization

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

- Disabling unused devices, including PCI-E lanes, USB ports, SATA/SAS ports, clock generator and buffer ports
- Tuning CPU/chipset settings to reach minimized power consumption and best performance
- Open Turbo Mode tuning option for power limits, short and long time duration
- Using SPEC Power guidelines as guidance for the ODM to validate BIOS tuning results

4.3.2 Setup Menu

The ODM must provide a BIOS specification, which includes the complete BIOS, setup menu, and default settings. The setup menu allows its options to be configured before the operating system loads. The configuration options available through the boot menu include the following:

- Setting for adjusting memory speed, quick path interconnect (QPI) speed, Speed-step/Turbo mode, and CPU idle power state. The default follows the CPU and chipset vendor's POR unless otherwise mentioned.
- Setting for enable different Turbo Mode tuning settings based on CPU SKU and Memory configuration. The default is Turbo enable with the CPU vendor's POR unless otherwise mentioned.
- Setting for power feature after AC power failure. The default is set to Restore Last Power State.
- Setting for Console redirection. There are three selectable options:
    - "Serial over LAN" with baud rate 57600, no flow control, and terminal type VT100
    - "Local Console" with baud rate 57600, no flow control, and terminal type VT100
    - "Auto", means that Serial-Over-LAN port is enabled by default, and BIOS will switch to local console automatically depending on hardware strapping. The default option is "Auto".
- Setting for altitude of server deployment location. The default is 300 meters.
- Setting for watchdog timer. The default is Enabled. The timeout value is 15 minutes. The watchdog timer is disabled after POST.
- Event log viewing and clearing.
- Setting for the error correcting code (ECC) memory error threshold. Available settings are 1, 4, 10 and 1000. The default setting for the engineering validation test (EVT), design verification test (DVT), and production validation test (PVT) is 4. The default setting for mass production is 1000.
- Setting for the ECC memory error event log threshold. Available settings are Disabled, 10, 50, 100. The default setting is 10.
- If a CMOS checksum error occurs (for example, caused by BIOS update), the BIOS should load the system default loading program automatically after showing a text message in the console for 5 seconds and then rebooting system to apply them. It should not wait for user input to load the system defaults.
- Setting to disable all "wait for keyboard input to continue" types of features. The default is "not to wait for keyboard input to continue".
- Calculate the checksum of the BIOS setting, display it in the BIOS setup menu, and output it to the SMBIOS table.
- Setting to save and load 10 different user default sets.
- Setting of the unified extensible firmware interface (UEFI) and Legacy boot options. The default is UEFI.

4.3.3 PXE Boot

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

- 1st: USB device if available
- 2nd: Network (search through all available network interfaces)
- 3rd: SATA HDD
- 4th: SATA-CDROM
- 5th: Other removable device

If there is no bootable device found, the BIOS should keep loop searching for a bootable device. The BIOS should support UEFI and legacy boot mode options. The default is UEFI. UEFI and legacy boot mode have an independent boot loop.

Boot mode and boot order can be displayed and changed from the BMC.

4.3.4 Board ID

The motherboard should provide 4 strapping pins to be used as BOARD_ID [3:0], so the BIOS can do correct board initialization based on different board configurations. The board ID is also routed to the BMC to be accessed by BMC firmware.

4.3.5 FAB Revision ID

The motherboard should provide 4 strapping pins to be used as FAB_REVISION_ID [3:0], so the BIOS can differentiate correct board FAB versions. FAB revision ID is also routed to BMC to be accessed by BMC firmware.

4.3.6 Remote BIOS Update

Vendors should provide tool(s) to implement remote BIOS update function. Vendors must validate update tools on each BIOS release during development and production. Any tool(s) provided should support three 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 (as in Scenario 2). Output must include a detailed value-meaning description for each setting. Settings must include pre-production setup menus/options too.
- Scenario 2: Update BIOS with pre-configured set of BIOS settings
  - Update/change multiple BIOS settings. Settings must include pre-production setup menus/options. Tool(s) should provide detailed value-meaning descriptions for each setting.
  - Reboot.
- Scenario 3: BIOS/firmware update with a new revision
  - Load new BIOS/firmware on machine and update, retaining the current BIOS settings.
  - Reboot.
- Scenario 4: Use BMC to update BIOS in PCH flash (also described in Section 5.13)
  - Update BIOS flash from BMC. The update time should be less than five minutes including uploading the BIOS binary and updating/verifying flash.
  - Update with a command line script in the Linux environment from a remote server. A graphical user interface (GUI) is not acceptable.

Additionally, the update tools must have the following capabilities:

- Ability to update from the operating system (OS) through ssh; the current OS is CentOS 5.2 (64-bit)
- Require no more than one reset cycle to the system to complete the BIOS update or the BIOS setup option change
- No user interaction (e.g. prompts)

- BIOS updates and option changes do not take longer than five minutes to complete
- Can be scripted and propagated to multiple machines

4.3.7 Event Log

The BIOS should do event logging through BMC system event logs (SEL). The combination of BIOS and BMC should meet the SEL log requirements in Section 5.10.

4.3.8 BIOS Error Codes

BIOS fatal error codes listed in following table should be enabled for POST-code output. The major and minor codes alternately display.

Table 3 BIOS Error Codes

Fatal Errors	Major Code	Minor Code	Error Description

4.3.9 POST Code for Option ROM Entry and Exit

Special BIOS post-codes are assigned to indicate the entry and exit options for ROM. A two-byte sequence is assigned for the entry and the exit to avoid same sequence used on other BIOS post-codes. For example, AA-C0 indicates entry, and BB-C1 indicates exit. These two sequences should be avoided in other processes that use post-codes.

5 Baseboard Management Controller (BMC)

The motherboard uses a BMC for various platform management services and interfaces with hardware, BIOS, PCH.

The BMC should be a standalone system in parallel to the host. The health status of the host system should not affect the normal operation and network connectivity of the BMC. The BMC cannot share memory with the host system. BMC management connectivity should work independently from the host. If using a shared NIC, there should be no NIC driver dependency for out-of-band (OOB) communication.

5.1 Management Network Interface

The BMC should have both I2C port and RMII/NCSI port for OOB access.

Three OOB access methods should be supported. Option 2 and Option 3 share same device footprint as the co-layout.

- Option 1: Shared-NIC uses I2C or RMII/NCSI interfaces to pass management traffic on the data network of the Mezzanine 10GE NIC
- Option 2: Shared-NIC uses RMII/NCSI interfaces to pass management traffic on data network of I210-AS. I210-AS has SGMII/KX interface to midplane
- Option 3: Shared-NIC uses RMII/NCSI interfaces to pass management traffic on data network of I210-AT. I210-AT has 10/100/1000 MDI interface to RJ45

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

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

5.2 Local Serial Console and Serial-Over-LAN (SOL)

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

- A local serial console on debug header, described in section 9.8.2
- A remote console, also known as Serial-Over-LAN (SOL) through the management network described in section 5.1.

It is preferred that both interfaces are functional at all stages of system operation.

The BMC has the control of switching the console input and output between SOL and the local serial console on the fly. When there is a legacy limitation that only one interface is functional, the default is set to Local. Upon SOL activation, the BMC switches the console to SOL automatically in order to establish a SOL session at any time. When the universal asynchronous receiver/transmitter (UART) console switch described in 9.8.3 is changed to channel 01, the BMC needs to identify this action and switch the console back to Local to re-enable local console access.

While the system boots, POST codes will be sent to port 80 and will be decoded by BMC to drive the LED display described in section 5.5. The POST codes should be displayed in the SOL console during system POST. Before the system displays the first screen, the POST codes are dumped to and displayed in the SOL console in sequence. For example, "[00] [01] [02] [E0]…" etc. After the system displays the first screen in the SOL console, the last POST code received on port 80 is displayed in the lower right corner of the SOL console screen.

A serial console buffer feature is required. The buffer needs to save, at minimum, the last five screens of local console output and at least the last five screens of remote console output. It must save 80 columns x 24 rows of data 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 internal or external SDRAM of the BMC. Buffer data is cleared within five seconds after the loss of standby power. The Buffer should NOT be stored in any non-volatile media.

5.3 Graphic and GUI

Because the GUI does not scale, the motherboard does not require the BMC to support graphic, KVM and GUI features. All BMC features need to be available via command line mode, by In-band and OOB Intelligent Platform Management Interface (IPMI) command, or SOL.

5.4 Remote Power Control and Power Policy

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

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

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

5.5 Port 80 POST

The vendor should implement BMC to support a port 80 POST code display to drive the 8-bit HEX GPIO to debug header described in section 9.8. The BMC post function must be ready before the system BIOS starts to send the first POST code to port 80. The POST code should also be sent to SOL as mentioned in section 5.2.

The BMC should have access to POST code and record up to 20 POST codes. The OOB raw command can be used to retrieve the last 20 POST codes from the BMC.

5.6 Power and System Identification LED

The motherboard must combine the Power LED and the System Identification LED into a single blue LED at the front side.

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

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

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

Blinking the Power LED blinking is used as a system identifier. The on time is different during power on and power off.

5.7 Platform Environment Control Interface (PECI)

The BMC should access the platform environment control interface (PECI) through PCH by default. The PECI connection implementation should follow Intel guidelines. The BMC should be able to execute PECI raw commands.

The vendor should implement the board design to connect the CPU PECI interface to the PCH PECI. It should be accessed by PCH directly by default. As an option it should reserve the ability to establish a direct connection from the BMC to the CPU PECI interface.

5.8 Power and Thermal Monitoring and Power Limiting

The vendor should implement BMC firmware to support platform power monitoring. Power limiting for processor, memory, and platform, PCH is required. Access to this function must be available through In-Band and Out-of-Band.

The vendor should implement BMC FW to support thermal monitoring, including processor, memory, chipset, and Inlet/outlet air temperatures. The inlet and outlet air temperature sensors should be compatible with the PCH supported device list. Using the TI TMP421 with external transistors is preferred to detect inlet and outlet temperature. Take caution when implementing inlet air sensors. Taking care will help you avoid preheating nearby components, and will reduce the amount of heat conducted through PCB.

5.9 Sensors

5.9.1 Analog Sensors

The BMC has access to all analog sensors on the motherboard either directly or through the PCH. All analog sensors need to be displayed in the sensor data record (SDR) repository.
The analog sensors required are listed in the table below. The lower and upper critical threshold is listed for system event logging purpose. Please refer to section 5.10 for logging requirements.

Table 4: Analog Sensor Table with Lower and Upper Critical Values

Sensor Name	Sensor Number	Lower Critical	Upper Critical
Outlet Temp	0x01		75
Inlet Temp	0x07		40
PCH Temp	0x08		85
P0 Temp	0x09		PROCHOT -2
P1 Temp	0x0A		PROCHOT -2
P0 VR Temp	0x02		75(preliminary)
P1 VR Temp	0x03		75(preliminary)
P0 DIMM VR0 Temp	0x0B		Per VR spec
P0 DIMM VR1 Temp	0x0C		Per VR spec
P1 DIMM VR0 Temp	0x0D		Per VR spec
P1 DIMM VR1 Temp	0x0E		Per VR spec
HSC Temp	0x0F		TBD
P0 CH0 DIMM0 Temp	0x11		82
P0 CH0 DIMM1 Temp	0x12		82
P0 CH1 DIMM0 Temp	0x13		82
P0 CH1 DIMM1 Temp	0x14		82
P0 CH2 DIMM0 Temp	0x15		82
P0 CH2 DIMM0 Temp	0x16		82
P0 CH3 DIMM0 Temp	0x18		82
P0 CH3 DIMM1 Temp	0x19		82
P1 CH0 DIMM0 Temp	0x1A		82
P1 CH0 DIMM1 Temp	0x1B		82
P1 CH1 DIMM0 Temp	0x1C		82
P1 CH1 DIMM1 Temp	0x1D		82
P1 CH2 DIMM0 Temp	0x1E		82
P1 CH2 DIMM0 Temp	0x1F		82
P1 CH3 DIMM0 Temp	0x20		82
P1 CH3 DIMM1 Temp	0x21		82
HSC Input Power	0x29		350
HSC Output Current	0x28		47
HSC Input Voltage	0x2A	10.8	13.2
P0 Package Power(PECI)	0x2C		N/A
P1 Package Power(PECI)	0x2D		N/A
P0 core VR Power	0x22		290
P0 core VR Current	0x23		188
P0 core VR Voltage	0x24	1.47	1.85
P1 core VR Power	0x25		290
P1 core VR Current	0x26		188
P1 core VR Voltage	0x27	1.47	1.85
P0 DIMM VR0 Power	0x32		47
P0 DIMM VR0 Current	0x33		76
P0 DIMM VR0 Voltage	0x34	1.15	1.25
P0 DIMM VR1 Power	0x35		47
P0 DIMM VR1 Current	0x36		76
P0 DIMM VR1 Voltage	0x37	1.15	1.25

P1 DIMM VR0 Power	0x38		47
P1 DIMM VR0 Current	0x39		76
P1 DIMM VR0 Voltage	0x3A	1.15	1.25
P1 DIMM VR1 Power	0x3C		47
P1 DIMM VR1 Current	0x3D		76
P1 DIMM VR1 Voltage	0x3E	1.15	1.25
SYS FAN0	0x46	500	14000
SYS FAN1	0x47	500	14000
P0 PROCHOT	0x30		
P1 PROCHOT	0x31		
Airflow	0x4A	N/A	N/A
P3V3	0xD0	3.00	3.60
P5V	0xD1	4.50	5.50
P12V	0xD2	11.0	13.2
P1V05_STBY	0xD3	0.95	1.15
P1V8_AUX	0xD4	1.70	1.90
P3V3_AUX	0xD5	3.00	3.60
P5V_AUX	0xD6	4.50	5.50
P3V_BAT	0xD7	2.70	3.30

5.9.2 Discrete Sensors

The vendor should implement BMC firmware access and display discrete sensors in SDR. The BMC should log abnormal sensor readings to SEL.

The discrete sensors required and the SEL format is listed in the table below.

Table 5: Discrete Sensors

Sensor Name	Sensor #	Event Data1	Event Data2	Event Data3
System status	0x10	[0]=1b, CPU0 Socket Occupied [1]=1b, CPU1 Socket Occupied [2]=1b, CPU0 Thermal trip [3]=1b, CPU1 Thermal trip [6]=1b, CPU CATERR [7]=1b, System Throttle		
Pwr Thresh Evt	0x3B	[1]=1b, Limit Exceeded		
Battery Mon	0x45	[0]=1b, Low Asserted		
SEL Status	0x5F	[1]=1b, SEL Clear [8]=1b, SEL Rollover		
DCMI Watchdog	0x60	[0]=1b, Timer Expired [1]=1b, Hard Reset [2]=1b, Power Down [3]=1b, Power Cycle		

		[8]=1b, Timer Interrupt
Processor Fail	0x65	[4]=1b, FRB3
TSOD SMBus Sts	0x6E	0x00~0x03
Chassis Pwr Sts	0x70	[0]=1b, Power Off [1]=1b, Power Cycle [2]=1b, Power On [4]=1b, AC Lost
Thermal Limit 1	0x7D	[1]=1b, Limit Exceeded
System booting status	0x7E	[0]=1b, SLP S4 [1]=1b, SLP S3 [2]=0b, PCH PWROK [3]=0b, SYS PWROK [4]=1b, Platform Reset [5]=0b, BIOS Post Complete
VR HOT	0x82	[0]=1b, CPU0 Core VR Hot [1]=1b, CPU1 Core VR Hot [2]=1b, CPU0 DIMM VR0 HOT [3]=1b, CPU0 DIMM VR1 HOT [4]=1b, CPU1 DIMM VR0 HOT [5]=1b, CPU1 DIMM VR1 HOT
CPU/DIMM HOT	0xB3	[0]=1b, CPU0 PROCHOT [1]=1b, CPU1 PROCHOT [2]=1b, CPU0 CH0/1 MEMHOT [3]=1b, CPU0 CH2/3 MEMHOT [4]=1b, CPU1 CH0/1 MEMHOT [5]=1b, CPU1 CH2/3 MEMHOT
PMBus Status Word Low	0x28	[0]=1b,None of the Above [1]=1b,CML [2]=1b, Temperature [3]=1b, VIN UV FAULT [4]=1b, IOUT OC FAULT [5]=1b, VOUT OV FAULT [6]=1b, HSC OFF [7]=1b, BUSY
PMBus Status Word High	0x42	[0]=1b, Unknown [1]=1b, Other [2]=1b, Fans [3]=1b, Power Good# [4]=1b, MFR Specific

		[5]=1b, Input [6]=1b, Iout/Pout [7]=1b, Vout
PMBus Status MFR	0x5E	[0]=1b, MFR Defined [1]=1b, MFR Defined [2]=1b, MFR Defined [3]=1b, MFR Defined [4]=1b, MFR Defined [5]=1b, MFR Defined [6]=1b, MFR Defined [7]=1b, MFR Defined
PMBus Status Input	0x9F	[0]=1b, Pin OP WARNING [1]=1b, Iin OC WARNING [2]=1b, Iin OC FAULT [3]=1b, Unit Off For Low Input Voltage [4]=1b,Vin UV Fault [5]=1b, Vin UV Warning [6]=1b, Vin OV Warning [7]=1b, Vin OV Fault
CPU0 Error	0x91	[0]=CPU VR Fault [1]=CPU CATERR [2]=Machine Check Error [3]=Thermal Trip
CPU1 Error	0x92	[0]=CPU VR Fault [1]=CPU CATERR [2]=Machine Check Error [3]=Thermal Trip

5.9.3 Event Only Sensors

The event only sensors are not shown in SDR. These sensors trigger SELs if an abnormal value is detected. The required event only discrete sensors and SEL format is listed in the table below.

Table 6: Event Only Sensors

Sensor Name	Sensor #	Event Data1	Event Data2	Event Data3
PCH Health	0x17	[7,6]=10b [5,4]=10b [3..0]=00h, Firmware Status	=0, Forced GPIO Recovery =1, Image Execution Failed =2, Flash Erase Error =3, Flash Corrupted =4, Internal Error = 5,BMC No Respond =6, Direct Flash Update by BIOS	<Extended Error Code>

			=7, Mfg. Error =8, Flash Integrity Error	
POST Err	0x2B	[7:6]=10b or 11b [5:4]=10b [3:0]=00b,Firmware Error	ED1[7:6]= 10b, LSB of POST Error Code ED1[7:6]= 11b, Per IPMI Spec	MSB of POST Error Code
Power Error	0x56	[0]=1b, SYS_PWROK Failure [1]=1b, PCH_PWROK Failure		
ProcHot Ext	0x3C	[7:6]=10b [5:4]=10b [3:0]=0Ah, Processor Thermal Throttling Offset	[7:2] Reserved [1:0] 0h = Native, 1h = External (VR)	[7:5] CPU/VR Number [4:0] Reserved
MemHot Ext	0x3D	[7:6]=10b [5:4]=10b [3:0]=09h, Memory Thermal Throttling Offset	[7:2] Reserved [1:0] 0h = Native, 1h = External (VR)	[7:5] CPU/VR # [4:3] Channel # [2:0] DIMM # [4:0] Reserved for VR HOT
Machine Chk Err	0x40	[7:6]=10b [5:4]=10b [3:0]=0Bh, Uncorrectable Or 0Ch, Correctable	Error Code ID	[7:5] CPU # [4:3] Source 00b = QPI 01b = LLC [2:0], if QPI, (0b = QPI 0; 1b = QPI 1), if LLC, Core #
PCIe Err	0x41	[7:6]=10b [5:4]=10b [3:0]=07h, Correctable Or 08h, Uncorrectable Or 0Ah, Bus Fatal	[2:0] Function # [7:3] Device #	[7:0] Bus #
Other IIO Err	0x43	[7:6]=10b [5:4]=10b [3:0]=00h, Other IIO	Error Code ID	[7:5] CPU # [4:3] Reserved [2:0] Source 000b = IRP0 001b = IRP1 010b = IIO-Core 011b = VT-d
Memory ECC Error	0x63	[7:6]=10b [5:4]=10b [3:0]=00h, Correctable Or 01h, Uncorrectable	[1:0] Logical Rank [7:2] Reserved	[7:5] CPU # [4:3] Channel # [2:0] DIMM #

5.10 System Event Log (SEL)

The vendor should implement BMC to support the SEL.

5.10.1 Errors to be Logged

- **CPU error**
  Both correctable ECC and uncorrectable ECC errors should be logged in the event log. Error categories include Link and L3 Cache.
- **Memory error**
  Both correctable ECC and uncorrectable ECC errors should be logged in the event log. The error log should indicate the location of DIMM, CPU socket#, channel # and slot #. Memory error reporting needs to be tested by both debug tool error injection and the reworked ECC DIMM.
- **QPI error**
  All errors that have status, register should be logged into the Event Log. Fatal or non-fatal classification follows the chipset vendor's recommendation.
- **PCI-E error**
  All errors that have status, register, should be logged in the Event Log. This includes root complex, endpoint device, and any switch upstream/downstream ports if available. Link disable on errors should also be logged. The error classifications (fatal, non-fatal, or correctable) follow the chipset vendor's recommendation.
- **POST error**
  All POST errors, which are detected by BIOS during POST, should be logged in the Event Log.
- **Power error**
  Two kinds of power errors should be logged. One is a 12.5V DC input power failure that causes all power rails on the motherboard to lose power, including standby power. The other is an unexpected system shutdown during system S0/S1 while 12.5V DC input is still valid.
- **MEMHOT# PROCHOT# and VRHOT#**
  Memory hot errors and processor hot errors should be logged. The error log should identify whether the error source is internal or external. Internal errors come from the processor or from memory. External errors come from voltage regulators, over-current-triggered throttling, or under-voltage-triggered throttling.
- **FAN Failure**
  FAN failure error should be logged if the fan speed is out of expected ranges between the lower and upper critical threshold. The Error log should also identify which fan fails.
- **PMBus status error**
  The PMBus status sensors check PMBus controller health status and log errors if an abnormal value is detected. The PMBus controller can be a DC Hot Swap Controller (HSC) or a PMBus AC to DC power supply unit (PSU).

Based on needs, each of the above errors can be enabled or disabled separately.

5.10.2 Error Threshold Setting

Enable error threshold settings for both correctable and uncorrectable errors. Once a programmed threshold is reached, the system should trigger an event and log it.

- Memory correctable ECC: The suggested threshold value for the MP stage is 1000. The suggested threshold value for EVT/DVT/PVT is 4. Each of these stages should also have options of 1, 4, 10, and 1000. When a threshold is reached, the BIOS should log the event, including DIMM location information and the output DIMM location code through the debug card.

- ECC error event log threshold: Defines the maximum number of correctable DIMM ECC events logged in the same boot. The default value is 10. Disable, 10, 50, and 100 must be available as options.
- QPI Error: Follow the chipset vendor's recommendation.
- PCIE Error: Follow the chipset vendor's recommendation.

5.10.3 Critical SEL Filter

Original equipment manufacturer (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 an option to log only critical SELs that need service or indicate a power cycle state change, and SEL clear and overflow.

5.11 Fan Speed Control in BMC

The vendor should enable fan speed control (FSC) on BMC. The BMC samples thermal related analog sensors in real time. The FSC algorism processes these inputs and drives two pulse width modulation (PWM) outputs in optimized speed.

5.11.1 Data Gathering for FSC

The BMC needs to gather data as input of FSC. The data to gather appears in the following table.

Table 7: FSC Data

Type of data	Data used be 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
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 ≥ **TBD** sample/s.

5.11.2 Fan Speed Control in BMC

The BMC should support FSC implementation to deliver optimal FAN speed control for data center environments. The BMC should support both In-Band and OOB FSC configuration update. Updates should take effect immediately without rebooting. The BMC should support fan boost during fan failure.

5.11.3 Fan Connection

The motherboard has 2 fan headers. The motherboard and midplane interfaces also have optional FAN tachometer and PWM connections.

5.12 BMC Firmware Update

Vendors should provide tool(s) to implement a remote BMC firmware update that will not require any physical input at the system level. We define a remote update as an update either through OOB by management network or through logging into the local OS (CentOS) by the data network.

A remote BMC firmware update may take a maximum of 5 minutes to complete. The BMC firmware update process and BMC reset process should require no reboot or power down of the host system. It should also have no impact on the normal operations of the host system. The BMC needs to be fully functional, with updated firmware after the update and it must reset without further configuration.

By default, the update should recovery the BMC to its factory default settings. There must also be options to preserve SEL and configuration. The MAC address should not be cleared with the BMC firmware update.

5.13 BMC Update: Dual PCH flash

5.13.1 Hardware Scheme

The vendor should implement BMC to be able to access the host system's PCH flash. It should also be able to remotely recover the host system's PCH flash from corruption. The PCH flash stores BIOS and PCH codes, and both BIOS and PCH regions should be updated. The dual PCH flash chip's hardware scheme is designed for this purpose. The hardware scheme contains two multiplexers controlled by a BMC's general purpose output (GPO). GPO_A controls PCH or the BMC has access to flash. GPO_B controls primary or backup flash and is accessed by CS# signal. Primary flash is the default flash PCH uses to access the BIOS and PCH code. Backup flash is stored with a known good image and is not usually modified. You can modify the backup flash through the command line.

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

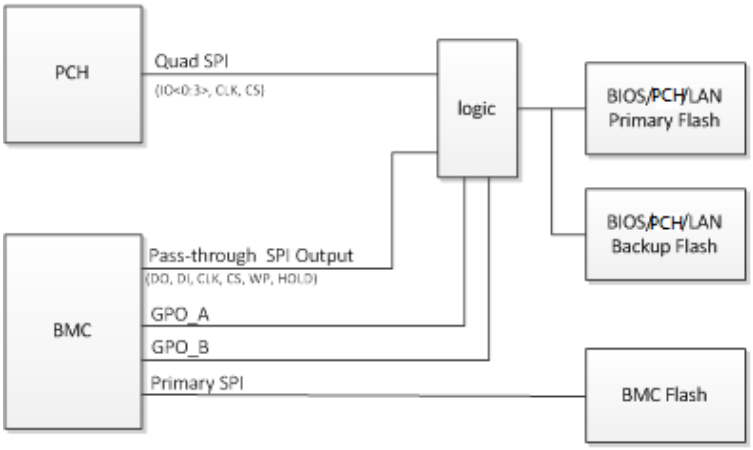


Figure 3: BMC and Dual PCH flash diagram

5.13.2 Software Scheme

The motherboard must support the following software features:

1. Auto BIOS recovery from backup flash to primary flash:
   When the 1st boot up after the BIOS FW update fails and the BMC watchdog timer (WDT) times out, the BMC determines that the BIOS needs to be recovered. It then takes over control of flash access and duplicates the backup flash content to the primary flash storage.
2. Manual BIOS recovery from backup flash to primary flash:
   Use the IPMI OEM command to trigger the BMC to control flash access and duplicate backup flash content to primary flash.
3. Manual BIOS save from primary flash to backup flash:
   Use IPMI OEM command to trigger the BMC to control flash access and duplicate the primary flash memory to the backup flash memory.
4. Manual Mux control:
   User the IPMI OEM command to trigger the BMC to change the state of two multiplexers.
5. Manual BIOS recovery from network to flash:
   User the IPMI OEM command to trigger a PCH flash binary to be transferred to the BMC and written to the primary or secondary flash memory, depending on the state of the multiplexer.
6. Enable/Disable Auto BIOS recovery
   User the IPMI OEM command to Enable and Disable feature 1 (above).
7. Get dual BIOS feature status:
   User the IPMI OEM command to get the current state of two multiplexers and the status of the Auto recovery feature (Enable/Disable).
8. Get dual BIOS versions:

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

5.14 BMC Update Complex Programmable Logic Device (CPLD)

The vendor should implement BMC to access the motherboard complex programmable logic device (CPLD) and perform a CPLD code upload. It should also be able to update and verify from the OOB command line mode.

6 Thermal Design Requirements

To meet thermal reliability requirements, the thermal and cooling solution should dissipate heat from the components when the system is operating at its maximum thermal power. Find the thermal solution by setting a high power target for the initial design in order to avoid redesigning the cooling solution. The final thermal solution of the system delivered should be the most optimized and energy efficient with the lowest capital and operating costs. The thermal solution should not allow components in the motherboard to overheat. Under following environmental conditions, neither the CPU nor the memory should throttle due to any thermal issue:

- Inlet temperature lower than or equal to 35°C, and 0 inch H^2O data center pressure with all fans functioning in each thermal zone
- Inlet temperature lower than or equal to 35°C, and 0.01 inch H^2O data center pressure with one fan (or one rotor) failure in each thermal zone

6.1 Data Center Environmental Conditions

The thermal design for the motherboard needs to satisfy the operational conditions at data centers. Conditions listed below use Facebook's Data Center as reference.

6.1.1 Location of Data Center/Altitude

Data centers may be located 1000m above sea level (at maximum). Any variation of air properties or environmental difference due to the high altitude needs to be accounted for in the thermal design.

6.1.2 Cold Aisle Temperature

Data centers keep 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 may fluctuate minutely depending on the outside air temperature. Every component must be cooled and maintained below its maximum specification temperature in any cold aisle temperature.

6.1.3 Cold Aisle Pressurization

Data centers maintain cold aisle pressure between 0 inches H^2O and 0.05 inches H^2O. The thermal design should factor in the worst operational pressurization in the data center, which is 0 inches H^2O and 0.01 inches H^2O with a single fan (or rotor) failure.

6.1.4 Relative Humidity

Data centers maintain relative humidity between 20% and 85%. Environmental condition changes due to the high altitude may not need to be considered if the thermal design can meet the requirement with maximum relative humidity (85%).

6.2 Server Operational Conditions

6.2.1 System Loading

The power consumption of individual components in the motherboard vary with the application the server is running or the motherboard's SKU. The total power consumption of the system may also vary with the usage or with the number of PCI-E cards on the system. Please see the summary below:

- System loading: idle to 100%
- Number of PCI-E full height or half height cards installed: 0 - 2
- Number of PCI-E Mezzanine card installed: 0 - 1
- Number of 3.5" HDD: 0 - 1

The system can consume up to 300W per node on a DC bus bar with a 14 node per power zone configuration in an Open Rack V1 4.2KW power zone. The real system configuration of each sled may consume more than 300W per node. The plan of record (worst-case configuration) for the thermal and power delivery design is two CPUs with 16x 8GB DIMMs.

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

6.2.2 DDR DIMM DRAM Operation

The thermal design should meet DIMM maximum operating temperatures (85°C) with a single refresh rate. A thermal test should be done based on DIMM module approved vendor list (AVL). The BIOS and memory subsystem should have an optimized refresh rate. It should use the optional DIMM auto-self-refresh (ASR) based on the DIMM's temperature. The implementation should follow the updated DDR4 memory controller and the DIMM vendor's specification.

6.2.3 Inlet Temperature

Inlet air temperatures will vary. The cooling system should cover the following inlet temperatures (20°C, 25°C, 30°C, and 35°C). Cooling above 30°C is beyond the system's operating specification. It is used during validation, however, to demonstrate design margin. CPU throttling is not allowed to activate over the validation range of 0°C – 35°C.

6.2.4 Pressurization

The thermal solution should not consider the availability of extra airflow from data center cooling fans. If (and only if) one rotor in server fan fails, the negative or positive data center pressurization can be considered in the thermal solution in the hot aisle or in the cold aisle.

6.2.5 Fan Redundancy

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

6.2.6 System Airflow or Volumetric Flow

The unit of airflow (or volumetric flow) used for this specification is cubic feet per minute (CFM). The maximum allowable airflow per watt is 0.16. The desired airflow per watt is 0.1 or lower at the mean temperature (plus or minus standard deviation). See section 6.1.2 for temperature definitions.

6.2.7 Delta T

The Delta T is the air temperature difference across the system or the temperature difference between the outlet air temperature and the inlet air temperature of the system. The Delta T must be greater than 11.7°C (21°F). The desired Delta T is 20°C (36°F) when the inlet air temperature is lower than 30°C.

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

6.3 Thermal Kit Requirements

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

6.3.1 Heat Sink

The heat sink design should have the most optimized design with the lowest cost. The heat sink design should be reliable and meet all the conditions described above. The heat sink can have no more than three heat pipes. Both CPU should use same CPU heat sink design. The ODM can always propose different heat sink types if there is alternative way to provide cost benefits. The heat sink should be without complex installation guidance (for example, air-flow direction).

6.3.2 System Fan

The system fan must be highly power-efficient with dual bearings. The design should minimize the propagation of vibration caused by the fan's rotation. The fan's minimum frame size is 60 mm x 60 mm. The maximum frame size is 80 mm x 80 mm. The ODM can propose a larger frame size if and only if there are cost benefits. The maximum thickness of the fan should be less than 38mm. Each rotor in the fan should have a maximum of 5 wires. Power consumption should not exceed 5% of total system power excluding the fan power (unless a rotor fails).

Back rush current is not acceptable. The 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% overshoot or no overshoot for the input current. The system should stay within its power envelope (300W for Open Rack V1 configuration) for all fan operations.

6.3.3 Air Duct

The air duct must be part of the motherboard tray cover. The design must be energy efficient, simple, and easily serviceable. It should also be unified for all system SKUs. It is preferred that the manufacturer uses highly green materials or reusable materials for the air duct.

6.3.4 Thermal Sensor

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

7 Midplane Interface and Midplane

The midplane interface of the motherboard described in this section contains options to support two use cases. The section also contains a midplane design that enables a motherboard to be used in Open Rack V1.

7.1 Connector Modules

There are 3 types of connector modules used in the motherboard interface. The AVL is shown in Table 8. The placement of connector modules is shown in the Figure 4 below.

Table 8 Connector Types

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 Small Press-Fit/E4 Short	FCI/10124755-111LF	FCI/10124756-101LF

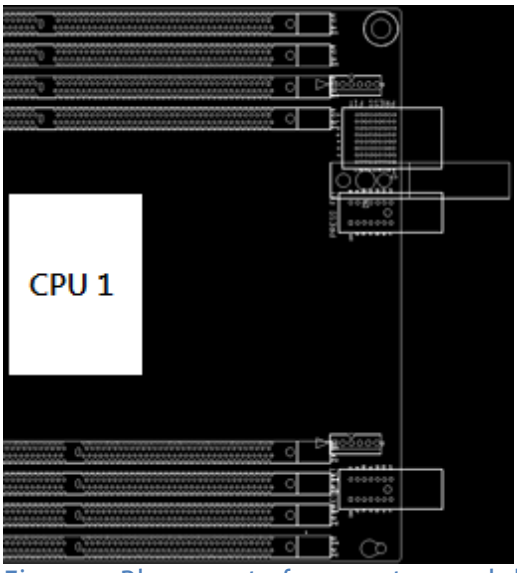


Figure 4: Placement of connector modules

7.2 Pin Definition

7.2.1 Airmax Power 2x2

In use cases with less or equal to **TBD** A[1] on 12 volts DC, one pair of Airmax power 2x2 R/A low profile headers and receptacles are populated. Two pairs of these connectors support up to **TBD** A on 12 volts DC.

[1] In 65°C ambient

7.2.1.1 Sled Side

Up to two Airmax power 2x2 R/A low profile headers are used at sled side, as shown in Figure 5. Pin definitions appear in the following table.

Table 9: 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

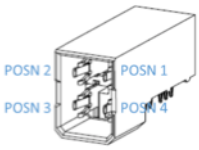


Figure 5: Power connector at sled side

7.2.1.2 Midplane Side

Up to two Airmax power 2x2 R/A receptacles for co-planar application are used at the midplane side, as shown in Figure 6. The receptacle has a long pin and a short pin to control the mating sequence. The part number with an S-S-L-S pattern is used to ensure that at least one ground pin mates before any power pin mates. Refer to the table below for detail.

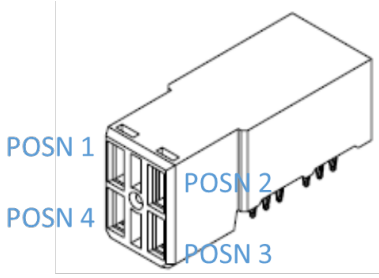


Figure 6: Power connector at midplane side

Table 10: 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

7.2.2 Airmax Guide

The motherboard and midplane uses one pair of Airmax 7.2mm R/A guides ONLY when blind mating is needed.

7.2.2.1 *Sled Side*

The sled side uses one Airmax 7.2mm R/A guide blade, as seen below.

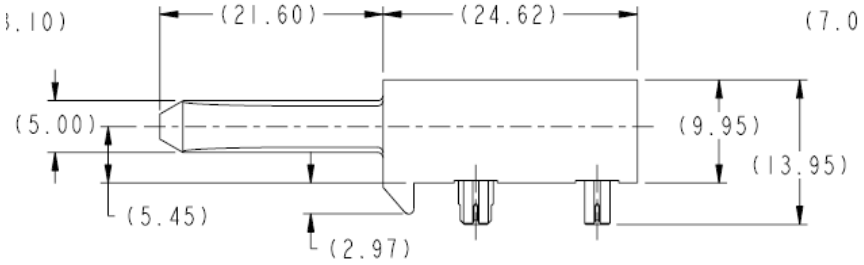


Figure 7: Airmax 7.2mm R/A guide blade

7.2.2.2 *Midplane Side*

The midplane side uses one Airmax 7.2mm R/A guide socket, as seen below.

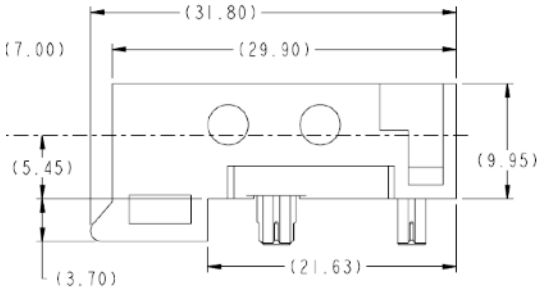


Figure 8: Airmax 7.2mm R/A guide socket

7.2.3 Airmax 3x8 Signal

One Airmax VS/VS2 3x8 connector follows the definition below on the sled side. The signal description is shown in Table 11. It supports up to x8 PCI-E plus side-band signals and management interfaces. This connector is optional to support a high-speed midplane with a PCI-E interface.

Table 11: Pin Definition for Airmax 3x8 Signal

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_N	PCIE_RX_DN5	GND	I

Table 12: 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_N	Output	Motherboard on indication from motherboard to midplane; OD output at motherboard and needs PU at midplane.
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 the 3x6 and 3x8 Airmax connector to indicate full mating of the sled. Follow for implementation at motherboard or uServer sled side and at midplane side. Fully mating the sled enables a hotswap controller on the motherboard or the uServer sled. This action also notices the presence of sleds in the

midplane.

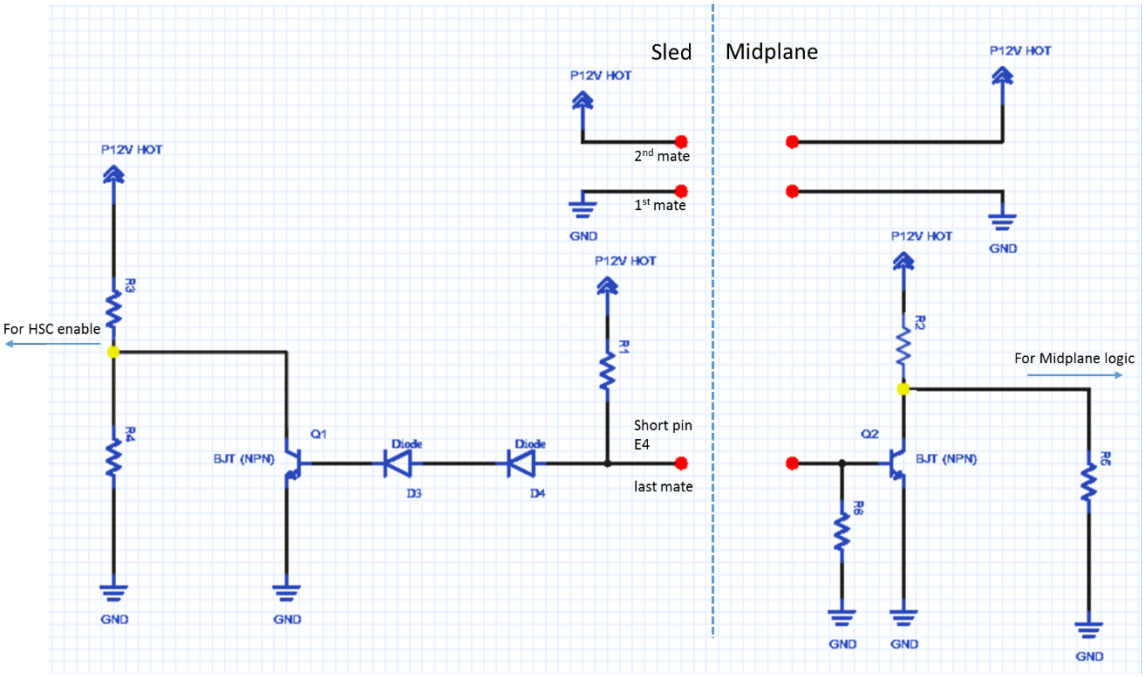
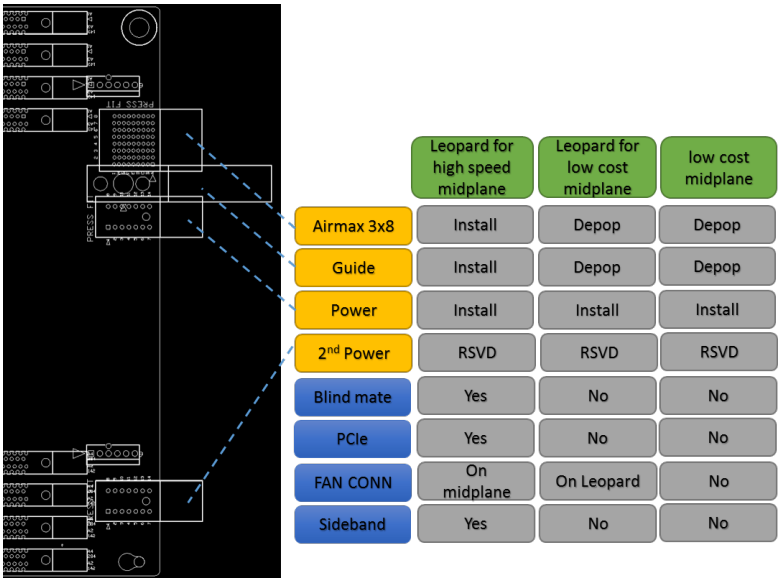


Figure 9: Reference circuit for dual side presence/mate detection with 1 short pin

7.3 Population options

The motherboard has two population options for the midplane connectors to support different use cases. The motherboard works with a different midplane in each use case. The population options and major differences are listed in Figure 10. The high-speed midplane is not covered in this document. The low-cost midplane is described in section 7.4.



	Leopard for high speed midplane	Leopard for low cost midplane	low cost midplane
Airmax 3x8	Install	Depop	Depop
Guide	Install	Depop	Depop
Power	Install	Install	Install
2nd Power	RSVD	RSVD	RSVD
Blind mate	Yes	No	No
PCIe	Yes	No	No
FAN CONN	On midplane	On Leopard	No
Sideband	Yes	No	No

Figure 10: Motherboard midplane connector population options

7.4 Midplane for Open Rack V1

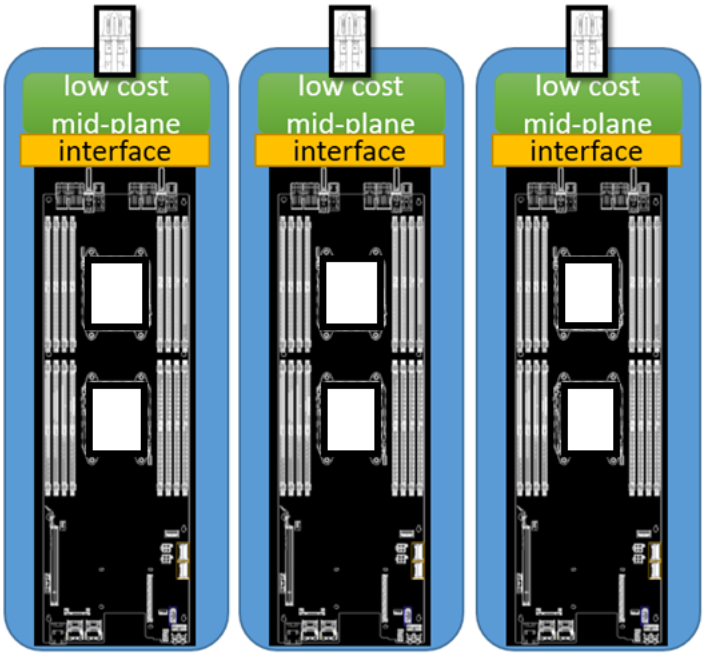


Figure 11: Motherboard to be used with Open Rack V1

For the use case shown in Figure 11, the motherboard is placed in a sheet metal tray similar to an OCP Open Rack tray[2] to form the sled. Each sled includes a low-cost midplane to bypass power from the bus bar clip to the motherboard. This midplane is completely passive. The hotswap controller is on the motherboard. The interface between the motherboard and the midplane does not involve blind mating or hotswapping. The hotswap interface is between the bus bar clip and bus bar.

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

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

[2] http://www.opencompute.org/wp/wp-content/uploads/2013/01/Open_Compute_Project_Intel_Server_Open_Rack_Specification_v0.3.pdf

8 Motherboard Power system

8.1 DC Cable Assembly

The DC cable assembly includes the bus bar clip assembly, DC cables, slug with screw hole and notch, and a common mode choke. Methode is the assembly vendor. The, part number is 5313-07450-00107 shown in Figure 12. The DC cable has two wires, a red insulator for power and a black insulator for ground to supply 12.5 Volt DC power from the bus bar clip to the midplane. The DC cable is AWG6. The DC cable assembly includes the bus bar clip, the DC cable, and the slug, which is rated for 50A in the worst operating conditions.

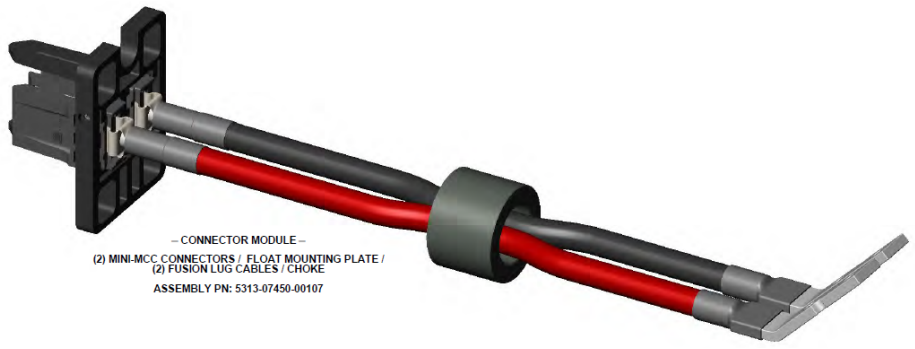


Figure 12: DC Cable Assembly

8.2 Input Voltage

8.2.1 Input Voltage Level

The nominal input voltage delivered by the power supply is 12.5 volts DC 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. The motherboard's under-voltage protection level should be less than 10V.

8.2.2 Capacitive Load

To ensure compatibility with the system power supply, the motherboard may not have a capacitive load greater than 4000µF. The capacitive load of the motherboard should not exceed the maximum value of 4000µF under any operating condition listed in Section 11.

8.2.3 P12 Voltage as AUX Rail

There is only one 12V rail delivered to the motherboard as an auxiliary power rail. Ensure that the PCIe device, HDD, and all other devices have the proper insulation, to meet voltage and timing requirements when running and powering on/off.

8.3 Hot Swap Controller (HSC) Circuit

In order to better control the 12.5V DC power input to each motherboard, one HSC (ADI/ADM1276/ADM1278; TI/LM25066I/LM25062/LM25063) is used on the motherboard. The HSC circuit provides the following functions:

- Inrush current control when the motherboard is inserted and powered up.
- Current limiting protection for over currents and short circuits. Users should be able to set the over current trip point to 33A and 46A with a jumper setting. The default is 46A.

- SOA protection during MOSFET turning on and off
- HSC fault protection is set to latch off (Default) with retry as a stuff option.
- PMBUS interface to enable PCH and BMC with the following actions:
    - Report server input power and log an event if it triggers an upper critical threshold
    - Report input voltage (up to 1 decimal point) and log an event if it triggers either a lower or upper critical threshold
    - Log status events based on the hot swap controller's status register
- Use HSC or an external circuit to provide a fast (<20us) over the current sense alert to trigger system throttling and CPU PROCHOT#. This feature needs to be controlled by BMC GPIO directly. The BIOS has a setting to control Enable/Disable/[No Change]. No Change is the default setting. It means follow the BMC's initial setting. The BMC sets it to Disable as the default. Before the BMC is ready, the hardware's POR state is Enable.
- Use the HSC or external circuit to provide a fast (<20us) under-voltage alert to trigger system throttling and CPU PROCHOT#. This feature is enabled by default with a resistor option to Disable.
- Use HSC or external circuit to provide fast (<20us) under-voltage alert to trigger system fan throttling. This feature is disabled by default with a resistor option to Enable.
- Please refer to the table below or **preliminary** setting requirements of System, CPU, and memory sub-system throttling.
- The voltage drop on the HSC current sense resistor should be less or equal to 25mV at full loading. The hot swap controllers should have the SMBUS address set to 0x11 (7-bit format).
- The power reporting of the hotswap controller needs to be better than 2% from 50W to full loading at room temperature.

Table 13: Entry Point of System, CPU, and Memory Sub-system Throttling

Condition	Threshold	Action	Enable control	Default
Board input power over current limit	>TBD A or TBD A by jumper setting	Trigger throttle to system in < 20us	BMC GPIO	Disable
Board input power under voltage	<11.5V (preliminary)	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)

8.4 CPU Voltage Regulator (VR)

8.4.1 CPU Maximum Power

The motherboard shall be designed to handle a processor with a maximum TDP of 145W CPU. Support for processors of higher TDP is not required. As a result, the vendor shall optimize the CPU voltage regulator (VR) accordingly.

8.4.2 CPU VR Optimizations

CPU VR optimizations shall be implemented to reduce 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 section 8.4.1. The CPU VR should have an auto-phase-dropping feature, and run at optimized phase count among 1, 2, 3, … and maximum phase count. The CPU VR should support all power states to allow the VRM to operate at its peak efficiency at light loading.

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

8.4.3 CPU VRM Efficiency

The minimum efficiency for the CPU VRM efficiency is 91% over the 30% to 90% load range and 93% over the 50% to 70% load range for TDP power of CPU, measured from the 12.5V input to the VRM output. Vendors are encouraged to exceed the above efficiency requirement and may propose higher efficiency VRMs that may come at additional cost. The power efficiency measured from 12.5V input to the CPU socket should also be analyzed and improved.

8.4.4 CPU Core VR Configuration

The vendor should use a CPU core VR solution. All configurations shall be stored in NVRAM without any external resistor strapping. The vendor should provide a utility under CentOS to perform a VR configuration change. Any configuration change should take effect without the AC cycling node. The guaranteed rewrite count of NVRAM should be greater or equal to 15.

8.5 DIMM Voltage Regulator

8.5.1 DIMM Power Rails

The motherboard design should have a DIMM Power Rails option for DDR4 on DDR4 SKU.

8.5.2 DIMM Maximum Power

The motherboard has a DIMM configuration for 16x slots of two CPU sockets. The vendor should follow the memory controller vendor's guideline to design and validate the DIMM power rail to support the maximum power needed for this configuration, and support a 1.2V DDR4 DIMM.

8.5.3 DIMM VR Optimizations

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

8.5.4 DIMM VR Configuration

The DIMM VR has same configuration requirement as the CPU VR, listed in section 8.4.4.

8.6 Voltage Regulator Module Design Guideline

For the voltage regulator module (VRM), the vendor should list the current budget for each power rail based on worst case loading scenarios in all possible operation conditions. The 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 a design check, get the actual OCP setting chosen for VRM, and explain the reason if it cannot meet this general requirement.

For a VRM that requires firmware or power code or a configuration file, vendors should maintain version control to track all the releases and changes between each version. Vendors

must also provide a method to retrieve versions via software during system run time. This software method should run under CentOS.

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

8.7 Hard Drive Power

The motherboard shall supply power to all possible connected hard drives (nine maximum). Hard drives require 12 volt DC and 5 volt DC power sources. For individual SATA ports, 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 the industry standard conventions described in the table below.

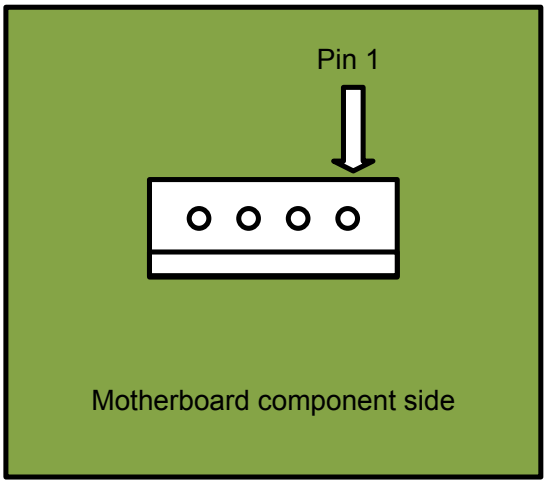


Figure 13: 4-pin floppy disk power connector

Table 14: Pin Description

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

For SATA ports inside miniSAS connectors, power will be delivered through the 8-pin (2x4) ATX power connector shown in the figure below. The connector is **TBD** or equivalent with mating connector **TBD** or equivalent. The pin definition is described in the table below. It fans out to eight standard SATA power cables. The power design should also allow this connector to supply P12V on the PCI-E riser.

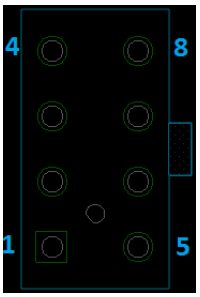


Figure 14: 4pin ATX power connector

Table 15: Pin Description

Pin	Description
1,2,3,4	GND
5,7	+5VDC
6,8	+12VDC

8.7.1 HDD Power Requirements

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

8.7.2 Output Protection

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

8.7.3 Spin-up Delay

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

8.8 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.

8.9 Power On

The motherboard should be set to restore the last power state during AC on/off. This means when the AC cycles on/off, the motherboard should power on automatically without someone pressing the power button. When the motherboard is powered off on purpose, it should be kept off through AC on/off.

9 I/O System

This section describes the motherboard's I/O features.

9.1 PCIe* x24 Slot Connector/Riser Card

The motherboard has one PCI-E x24 slot connector to support PCI-E Gen 3 signals. The slot location must follow the mechanical requirement delivered in DXF format.

The vendor should also provide a PCI-E riser card so two full-height PCI-E cards could be plugged in horizontally and locked in position. The top slot has x16 PCI-E connector and supports a full-height, full-length PCI-E card. The bottom slot has x8 PCI-E connector. It is open ended and supports a full-height half-length PCI-E card. Other riser configurations are provisioned. For example, a riser with 3 x8 PCI-E connectors. The motherboard uses SLOT CONFIG[0..1] to identify the different types of riser.

The PCI-E riser card should fit to a 2-OpenU-height tray. Replacing the PCI-E riser should not require the removal of any other components.

Table 16 contains information on how the JTAG and the reserved pins on the PCI-E x24 slot connector on the motherboard. These support two PCI-E connectors on the riser card. Follow Table 17 to use the reserved pins on PCI-E x8 connector (open end slot 0 [low]) on the riser card. Follow Table 18 to use the reserved pins on PCI-E x16 connector (slot 1 [high]) on the riser card.

Table 16: PCI-E x24 Slot Connector Reserved Pin Usage on MB

Pin	Pin Defined	Description
A5	CLK_100M_P	Extra 100MHz clock for the PCI-E slot on riser card
A6	CLK_100M_N	
A7	GND	Ground reference for the clock
A8	LAN_SMB_DAT	SMBUS DATA from Management I2C port of BMC
B9	LAN_SMB_CLK	SMBUS CLOCK from Management I2C port of BMC
B12	LAN_SMB_ALERT_N	SMBUS Alert signal to Management I2C port of BMC
A32	CLK_100M_P	Extra 100MHz clock for the PCI-E slot on riser card
A33	CLK_100M_N	
B82	GND	Ground reference for the PCIe signal
A50	SLOT_CONFIG0	Riser slot configure 0
A83	SLOT_CONFIG1	Riser slot configure 1
B17	SLOT_PRSNT2-1	SLOT PRESENT1
B31	SLOT_PRSNT2-2	SLOT PRESENT2
B48	SLOT_PRSNT2-3	SLOT PRESENT3
B81	SLOT_PRSNT2-4	SLOT PRESENT4
B114	SLOT_PRSNT2-5	SLOT PRESENT5
B115	SLOT_PRSNT2-6	SLOT PRESENT6

Table 17: PCI-E x8 Connector Open End Slot 0 (low) Reserved Pin Usage on Riser

Pin	Pin Defined	Description

A19	SLOT0_ID_0	Slot0 ID bit 0. Strapping to PCIe card for slot number identification; pull down on riser card slot 0 with 100ohm
A32	LAN_SMB_CLK	SMBUS CLOCK from Management I2C port of BMC
A33	LAN_SMB_DAT	SMBUS DATA from Management I2C port of BMC
B12	LAN_SMB_ALERT_N	SMBUS Alert signal to Management I2C port of BMC
B30	SLOT0_ID_1	Slot0 ID bit 1. Strapping to PCIe card for slot number identification; pull down on riser card slot 0 with 100ohm
B31	SLOT_PRSNT2-1	SLOT PRESENT1
B48	SLOT_PRSNT2-2	SLOT PRESENT2

Table 18: PCI-E x16 Connector Slot 1 (high) Reserved Pin Usage on Riser

Pin	Pin Defined	Description
A19	SLOT1_ID_0	Slot1 ID bit 0. Strapping to PCIe card for slot number identification; Left open on riser card slot 1 with 100Ohm pull down option (not installed).
A32	LAN_SMB_CLK	SMBUS CLOCK from Management I2C port of BMC
A33	LAN_SMB_DAT	SMBUS DATA from Management I2C port of BMC
B12	LAN_SMB_ALERT_N	SMBUS Alert signal to Management I2C port of BMC
B30	SLOT1_ID_1	Slot1 ID bit 1. Strapping to PCIe card for slot number identification; pull down on riser card slot 1 with 100ohm
B31	SLOT_PRSNT2-3	SLOT PRESENT3
B48	SLOT_PRSNT2-4	SLOT PRESENT4
B81	SLOT_PRSNT2-6	SLOT PRESENT6

Table 19: SLOT CONFIG Definition

SLOT_CONFIG0	SLOT_CONFIG1	MB x24 slot mapping to riser slots
0	0	Lane 0-7 (Riser lower x8 slot), Lane 8~23 (Riser upper x16 slot) (POR)
0	1	Lane 0-7->Riser lower x8 slot/Lane 8~15-> Riser middle x8 slot/Lane 17~23-> Riser upper x8 slot
1	0	RSVD
1	1	RSVD

9.2 DIMM Connector

The motherboard requires a 15u" gold contact for the DDR4 DIMM connectors. The connector uses black a housing. It uses a white/nature latch for the far-end DIMM in a DDR channel, and a black latch for the near-end DIMM in a DDR channel.

9.3 PCI-E Mezzanine Card

The motherboard has one PCI-E x8 Mezzanine card connector that holds x8 PCI-E Gen 3 signals. The Mezzanine card connector on the motherboard should use (FCI 61082-121402LF) or an equivalent connector, with mating (FCI 61083-124402LF) or an equivalent connector on the Mezzanine card.

Table 20: Mezzanine Connector Pin Definition

Signal	Description	Pin	Pin	Signal	Description
P12V_AUX	Aux Power	61	1	MEZZ_PRSNT1_N	Present pin1, Short to Pin120 on Mezz Card
P12V_AUX	Aux Power	62	2	P5V_AUX	Aux Power
P12V_AUX	Aux Power	63	3	P5V_AUX	Aux Power
GND	Ground	64	4	P5V_AUX	Aux Power
GND	Ground	65	5	GND	Ground
P3V3_AUX	Aux Power	66	6	GND	Ground
GND	Ground	67	7	P3V3_AUX	Aux Power
GND	Ground	68	8	GND	Ground
P3V3	Power	69	9	GND	Ground
P3V3	Power	70	10	P3V3	Power
P3V3	Power	71	11	P3V3	Power
P3V3	Power	72	12	P3V3	Power
GND	Ground	73	13	P3V3	Power
LAN_3V3STB_ALERT_N	SMBus Alert for OOB	74	14	**NCSI_RCSDV**	BMC NCSI
SMB_LAN_3V3STB_CLK	SMBus Clock for OOB	75	15	**NCSI_RCLK**	BMC NCSI
SMB_LAN_3V3STB_DAT	SMBus Data for OOB	76	16	NCSI_TXEN	BMC NCSI
PCIE_WAKE_N	PCIE Wake Up	77	17	RST_PLT_MEZZ_N	PCIe Reset Signal
NCSI_RXER	BMC NCSI	78	18	RSVD (MEZZ_SMCLK)	Reserved(PCIe Slot SMBus Clock)
GND	Ground	79	19	RSVD (MEZZ_SMDATA)	Reserved(PCIe Slot SMBus Data)
NCSI_TXD0	BMC NCSI	80	20	GND	Ground
NCSI_TXD1	BMC NCSI	81	21	GND	Ground
GND	Ground	82	22	NCSI_RXD0	BMC NCSI
GND	Ground	83	23	NCSI_RXD1	BMC NCSI
CLK_100M_MEZZ1_DP	100MHz PCIe Clock	84	24	GND	Ground
CLK_100M_MEZZ1_DN	100MHz PCIe Clock	85	25	GND	Ground
GND	Ground	86	26	RSVD(CLK_100M_MEZZ2_DP)	Reserved(2nd Set of 100MHz PCIe Clock)
GND	Ground	87	27	RSVD(CLK_100M_MEZZ2_DN)	Reserved(2nd Set of 100MHz PCIe Clock)
MEZZ_TX_DP_C<0>	PCIE TX Signal	88	28	GND	Ground
MEZZ_TX_DN_C<0>	PCIE TX Signal	89	29	GND	Ground
GND	Ground	90	30	MEZZ_RX_DP<0>	PCIE RX Signal
GND	Ground	91	31	MEZZ_RX_DN<0>	PCIE RX Signal
MEZZ_TX_DP_C<1>	PCIE TX Signal	92	32	GND	Ground
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 Mezzanine Card	120	60	GND	Ground

9.4 Network

9.4.1 Data Network

The motherboard uses a SFP+ 10G Mezzanine card[3] as its primary data network interface at the I/O side. There is option of a single port or a dual port. There is an x8 PCI-E port routed to the midplane connector to provide an option for future data network connections.

One I210 1 GB NIC is on board to provide an optional 10/100/1000 data connection.

9.4.2 Management Network

The motherboard has three management network interface options for the BMC's connection. The management network shares the data network's physical interface. The management connection should be independent from the data traffic, and OS/driver conditions. The information below echoes the requirement in section 5.1, with an emphasis on the physical connection.

- SFP+ shared-NIC from the Mezzanine 10 GB NIC or the PCI-E NIC, driven by the BMC through RMII/NC-SI or I2C. I2C is the default.
- SGMII/KX shared-NIC connected to the midplane interface from I210-AS, driven by the BMC through the RMII/NCSI.
- 10/100/1000 MDI shared-NIC connected to RJ45 from I210-AT (co-layout with I210-AS), driven by the BMC through the RMII/NCSI.

9.4.3 IPv4/IPv6 Support

The motherboard needs to be deployed in both IPv4 and IPv6 network environments. All data networks and management networks should have this capability. This includes, but is not limited to:

- DHCP and static IP settings
- PXE booting capability
- NIC and BMC firmware support
- OS drivers
- Utilities in both IPv4 and IPv6.

9.5 USB

The motherboard requires one external USB port at the front and one internal Type A USB connector. The BIOS should support the following devices on the motherboard's USB ports:

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

[3] http://opencompute.org/wp/wp-content/uploads/2012/10/Intel_Mezzanine_Card_Design_Specification_v0.5.pdf

- USB optical drive (bootable)

The placement of internal Type A USB connectors should enable a P5V to 1x SATA NGFF adaptor with 2280 size SATA NGFF card support.

9.6 SATA

The motherboard has PCH on board. PCH has a SATA controller support for SATA3 ports, and an sSATA controller to support SATA3 ports. These ports will be connected to several connectors, including miniSAS, mSATA, and standard SATA.

Both the mSATA connector and the vertical SATA connector need to be placed near the I/O side of the motherboard for easy access. The HDDs are attached to all SATA connectors and must follow the spin-up delay requirement described in section 8.7.3.

The placement of the vertical SATA connector should consider providing the SATA signal to one SATA NGFF adaptor with 2280 size SATA NGFF card support.

9.7 Next Generation Form Factor (NGFF) Connectors

The motherboard has one optional onboard NGFF connector (Key ID=M and H4.2 Type). The NGFF connector is populated only when the mSATA connector is not populated as a co-layout. The NGFF connector has an optional connection to the PCIe x4 or SATA Port from the PCH. The on board NGFF connector supports either the 2242 or the 2260 card form factor. Sideband signals such as WAKE# and SATA_ACTIVITY should be connected when they are available.

For NGFF 2280, the design should support one SATA to NGFF adaptor card. The SATA signal is from the vertical SATA connector on the motherboard and power is from the internal USB Type A connector.

9.8 Debug Header

The motherboard has one debug header at the front. The debug card can be plugged into this header directly or through a cable. This debug header should support hot plugging. Through this debug 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 a console redirection function. The two 7-segment LED displays should show BIOS POST code and DIMM error information. One reset button will trigger a system reset when pressed. The UART channel selection button sends a negative pulse to the motherboard to select and rotate the UART console in a loop of host console->BMC debug console->midplane console. The default after the debug card is plugged in or the system powers up is host console. A jumper is used to determine the power supply to the debug header pin 14 as P5V or P5V_AUX. P5V is the default and P5V_AUX is optional and used only for connecting the BMC debug console and the midplane console at S5. The jumper is for engineering use only. It is located close to the front of motherboard.

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

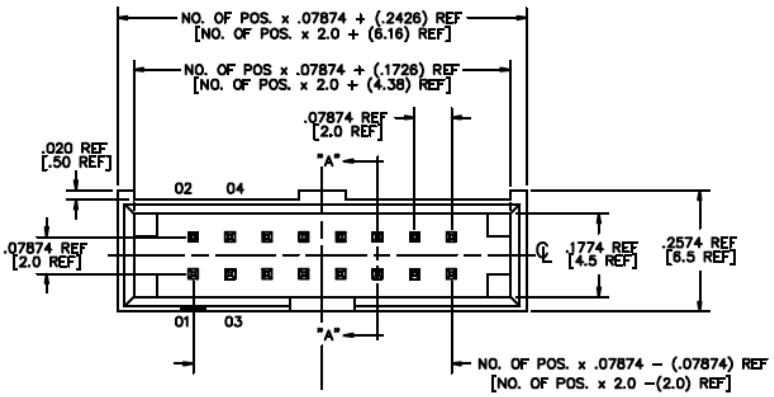


Figure 15: Debug header

Table 21: Debug Header Pin Definitions

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)
10	Serial Receive (motherboard receive)
11	System Reset
12	UART Channel Selection
13	GND
14	P5V(default)/P5V_AUX

9.8.1 Post Codes

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

During POST, the BIOS should also output POST codes to BMC SOL. When a SOL session is available during POST, a remote console should show the POST codes as mentioned in Section 5.2.

During the boot sequence, the BIOS shall initialize and test each DIMM module. If a module does not initialize or does not pass the BIOS test, post codes should flash on the debug card to indicate which DIMM has failed. The first hexadecimal character indicates which CPU interfaces with the DIMM module. The second hexadecimal character indicates the number of the DIMM module. The POST code will also display the error major code and error minor code from the

Intel memory reference code. The display sequence will be “00”, DIMM location, error major code and error minor code with a one second delay between every code displayed. The BIOS shall repeat the display sequence indefinitely to allow time for a technician to service the system. The DIMM location code table is TBD. The DIMM number count starts from the furthest DIMM from the CPU.

9.8.2 Serial Console

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

Table 22: Debug Card Mini-USB UART Pins

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 with a 0.1” pitch header (FCI/ 68685-306LF or equivalent) to connect to an optional Class 2 Bluetooth module (Roving Networks/RN42SM[4] or equivalent).

Table 23: Bluetooth Header Pins

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

9.8.3 UART Channel Selection

When the debug card is plugged in, pin 12 shall be used to notify the motherboard’s BMC GPIO of its presence. When someone presses the UART Channel selection button on the debug card the same pin is used to send pulses to the motherboard to trigger a UART connection change among the host console (Default) -> BMC debug console->midplane console and loop. The falling edge triggers the change.

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

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

The motherboard side should implement logic to detect the debug card's presence when 10K or stronger pulls down presents on Pin 12.

The motherboard side should implement logic to trigger the UART connection change when someone presses the UART Channel selection button. The motherboard should stop its original POST-code display for one second when a falling edge to ground is detected, and display the channel number for one second to the debug port POST code to indicate a UART channel change. If the system POST code does not change within one second, the motherboard outputs the original display. If the system POST code does change within one second, the latest POST code displays.

Table 24: UART Channel And Connection

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

9.8.4 Other Debug Use Design Requirements on the Motherboard

An XDP header is required for BIOS debug and should be populated in both EVT and DVT samples. Access to the XDP header should not be mechanically blocked by the CPU heatsink or other components.

The SMBUS debug header should be inserted for the SMBUS on the 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 requirements, it should be added and populated in both EVT and DVT samples.

9.9 Switches and LEDs

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

9.9.1 Switches

Vertical tactile switches must be placed behind the debug header. The push button actuator has a minimum 2.5mm diameter and must protrude 9mm+/-1mm from the top of the actuator to the PCB surface. The system power button must be red and must be located on the left. The systems reset button must be black and must be located on the right.

If a user depresses the power switch for less than four seconds, the system shall issue a power management event. This event must indicate that a user triggered the switch. If the user depresses the power switch 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.

A silkscreened label on the motherboard shall indicate the functionality of each switch. The labels PWR and RST are acceptable.

9.9.2 LEDs

The table below indicates the color and function of each LED on the motherboard. The motherboard's silkscreened label shall indicate the functionality of each LED. Looking from I/O towards LED, from right to left, the sequence is Blue, Green, and Yellow.

Table 25: LED Functionality

LED Color	Function	Silk Screen Label
Blue	Power LED. This LED shall have the same functionality of a traditional PC power LED. It shall only illuminate if the motherboard is in the power on state.	PWR
Green	Hard drive activity. This LED shall illuminate when there is activity on the motherboard's SATA hard drive interfaces, or the onboard mSATA and NGFF connector interface.	HDD
Yellow	This LED replaces the functionality of the PC speaker. The motherboard shall cause the LED to light to illuminate for the same duration and sequence as the PC speaker would normally beep. The LED allows easier diagnosis in a noisy data center environment.	BEEP

The Beep Error LED codes appear in the table below.

Table 26: Beep Error LED Pattern Table

Error Description	LED Patterns						
PEI memory error	On (2s)	Off (0.25s)	On (2s)	Off (0.25s)	On (2s)	Off (3s)	…(repeat)
PEI DXE not found	On (2s)	Off (0.25s)	On (2s)	Off (0.25s)	On (0.25s)	Off (3s)	…(repeat)
DXE No Console	On (0.25s)	Off (0.25s)	On (0.25s)	Off (0.25s)	On (2s)	Off (3s)	…(repeat)

DXE flash update fail	On (2s)	Off (0.25s)	On (0.25s)	Off (0.25s)	On (0.25s)	Off (3s)	…(repeat)
DXE PCI requirement not met	On (0.25s)	Off (0.25s)	On (0.25s)	Off (0.25s)	On (0.25s)	Off (3s)	…(repeat)

9.10 Fan Connector

The motherboard has two fan connectors. Fan connector signals should follow the "4-Wire Pulse Width Modulation (PWM) Controlled Fans Specification" (Revision 1.3 September 2005) published by Intel. Each fan has six pins, which is compatible with a standard 4-wire fan connector, and also can be used to support a dual rotor fan that shares a PWM control signal but has separate tachometer signals. The fan connector pin definitions appear in the table below Use LOTES YBA-WAF-016 or its equivalent as a fan connector. Its mating part is a LOTES GAP-ABA-WAF-038 (or equivalent). Connect the fan power at the downstream side of the hot swap controller (HSC). The fan power needs to be turned off during S5.

Table 27: FAN Connector Pin Definitions

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

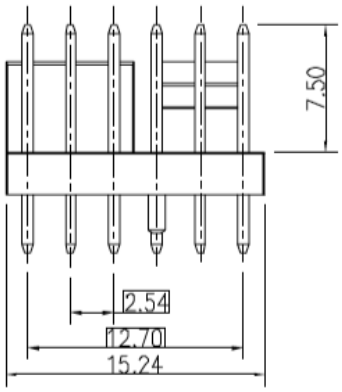


Figure 16: Fan connector diagram

9.11 TPM Connector and Module

A 15pin vertical receptacle connector (Figure 18) is defined on the motherboard for the LPC TPM module. The connector pin definition on the motherboard side is shown in the table below. Use an FCI 91911-31115 receptacle or its equivalent on the motherboard.

The TPM module is a 23mm(L)x 13mm(W) x 1.57mm(T) PCB with an FCI/91911-31515 header or its equivalent in the center of the module and a **TBD** TPM chip.

Table 28: TPM Header Pin Definitions

Pin	Description	Pin	Description
1	LPC_CLK_33M	9	P3V3
2	LPC_RST_N	10	TPM_PRSNT_N
3	LPC_LAD0	11	NC
4	LPC_LAD1	12	LPC_SERIRQ
5	LPC_LAD2	13	GND
6	LPC_LAD3	14	GND
7	LPC_FRAME_N	15	NC
8	NC		

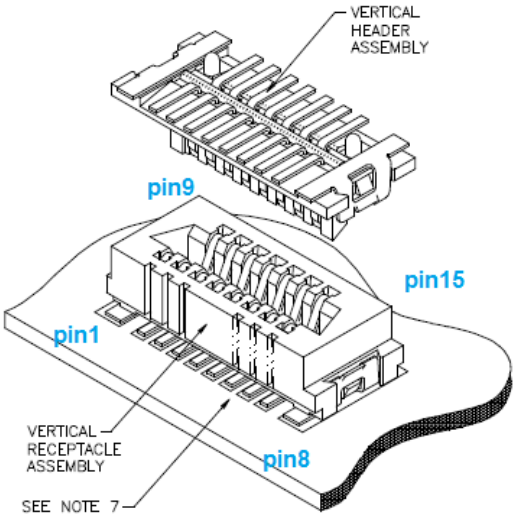


Figure 17: TPM header diagram

10 Mechanical

Mechanically, the motherboard should work with Open Rack V1. Open Rack V1 has three power zones (10 Open U each). Each power zone can hold five 2-OpenU shelves. Each shelf can hold up to three sleds. Each sled includes a sheet metal tray, one motherboard, and other components.

10.1 Shelf

Instead of depending on the figure below, the vendor should refer to a mechanical DXF file for more detail.

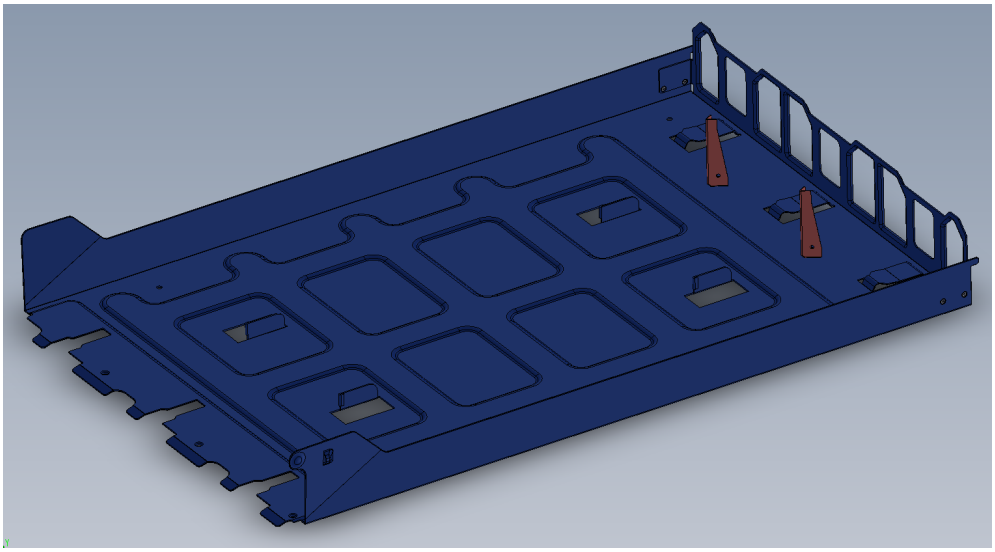


Figure 18: Sled

10.2 Tray and Sled

A sheet metal tray serves as the mechanical interface between the motherboard and the shelf. It also provides mechanical retention for the components inside the tray. This includes the bus bar cable assembly, the fan, the midplane, the riser card, PCIe cards, a hard drive, and the mezzanine card. The combination of tray, motherboard and the other components assembled in the tray is a sled.

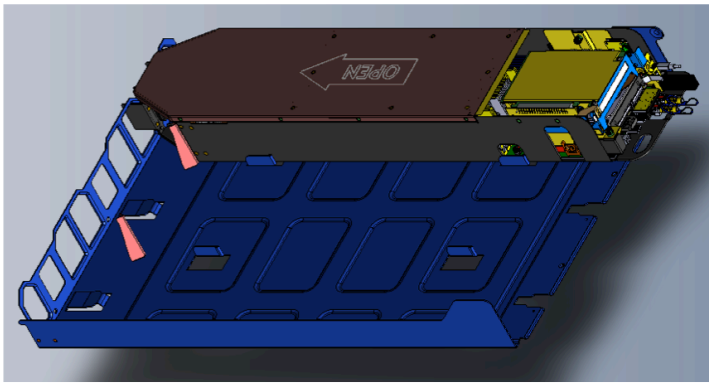


Figure 19: Sled on shelf

10.3 PCIe and HDD bracket

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

10.4 Fixed Locations

Refer to the mechanical DXF file for the fixed locations of the mounting holes, the 24 PCI-E slots and the power connector.

10.5 PCB Thickness

To ensure proper alignment of the motherboard and the midplane interface within mechanical enclosure, the motherboard should follow PCB stack up in Table 1 to have 85 mil (2.16mm) PCB thickness. The midplane PCB thickness should also be 85 mil (2.16mm). Mezzanine card and riser card PCB thickness should be 62 mil (≈1.6mm).

10.6 Heat Sinks and Independent Loading Mechanism (ILM)

The motherboard shall support heat sinks that are mounted according to the Intel CPU thermal mechanical specification and design guide. The vendor shall comply with all keep out zones defined by Intel in the above referenced specification. A standard heat sink and Narrow independent loading mechanism (ILM) solution is preferred. Only when the vendor's CPU/VR placement doesn't allow a standard Intel heat sink and Narrow ILM to be used, the vendor may choose non-standard heat sink. Approval is required.

10.7 Silk Screen

The silk screen shall be white in color and shall include labels for the components listed below. Additional items required on the silk screen are listed below.

- CPU0/CPU1
- DIMM slot numbering, as described in section 9.8.1
- LEDs as defined in Table 25
- Power and reset switches as PWR and RST.

10.8 DIMM Connector Color

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

10.9 PCB Color

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

Table 29: PCB Color

Revision	PCB Color
EVT	Red
DVT	Yellow

PVT	Green

11 Environmental Requirements

The motherboard must 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 OCP system also meets these requirements. In addition, the full system has an operating altitude with no de-ratings of 1000m (3300 feet).

11.1 Vibration and Shock

The motherboard meets shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) and IEC721-3-(*) Standard & Levels. The testing requirements are listed in Table below.

Table 30: Vibration and Shock Requirements

	Operating	Non-Operating
Vibration	0.5g acceleration, 1.5mm amplitude, 5 to 500 Hz, 10 sweeps at 1 octave / minute for 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 for each of the three axes (one sweep is 5 to 500 to 5 Hz)
Shock	6g, half-sine 11mS, 5 shocks for each of the three axes	12g, half-sine 11mS, 10 shocks for each of the three axes

12 Prescribed Materials

12.1 Disallowed Components

The following components cannot be used in the motherboard's design:

- Components disallowed by the European Union's Restriction of Hazardous Substances Directive (RoHS 6)
- Trimmers and/or potentiometers
- Dip switches

12.2 Capacitors and Inductors

The following limitations apply to the use of capacitors:

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

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

12.3 Component De-rating

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