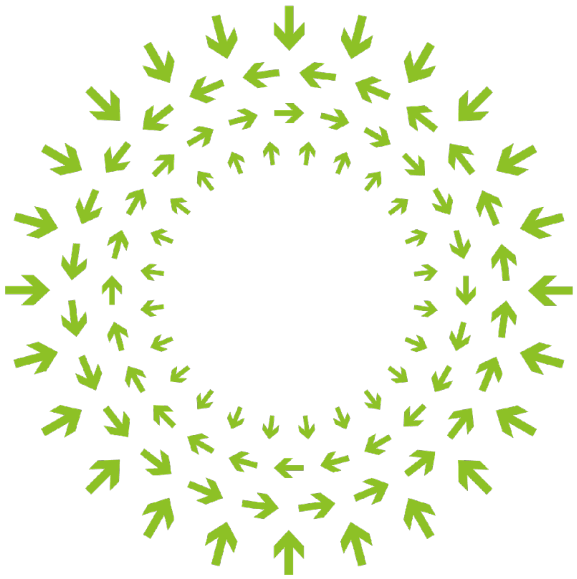


OPEN

Compute Project

Facebook Angels Landing System Specification

1.0

Author:
Hao Shen, Hardware Engineer, Facebook
Garnett Thompson, Hardware Engineer, Facebook
Anthony Chan, Power Engineer, Facebook
Michael Haken, Mechanical Engineer, Facebook
Yin Hang, Thermal Engineer, Facebook
Harsha Bojja, Manufacturing Quality Engineer, Facebook

Table of Contents

1. License

PLEASE PICK EITHER THE OCP CLA OPTION OR THE OWF OPTION. ONLY ONE CAN BE USED. DELETE THE ONE NOT USED.

1.1. OPTION A: OCP CLA

Contributions to this Specification are made under the terms and conditions set forth in Open Compute Project Contribution License Agreement ("OCP CLA") ("Contribution License") by:

Facebook

You can review the signed copies of the applicable Contributor License(s) for this Specification on the OCP website at https://www.opencompute.org/legal-documents
Usage of this Specification is governed by the terms and conditions set forth in
Open Web Foundation Final Specification Agreement ("OWFa 1.0")
also known as a "Specification License".

Notes:

1) The following clarifications, which distinguish technology licensed in the Contribution License and/or Specification License from those technologies merely referenced (but not licensed), were accepted by the Incubation Committee of the OCP:

None

2) The above license does not apply to the Appendix or Appendices. The information in the Appendix or Appendices is for reference only and non-normative in nature.

NOTWITHSTANDING THE FOREGOING LICENSES, THIS SPECIFICATION IS PROVIDED BY OCP "AS IS" AND OCP EXPRESSLY DISCLAIMS ANY WARRANTIES (EXPRESS, IMPLIED, OR OTHERWISE), INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, FITNESS FOR A PARTICULAR PURPOSE, OR TITLE, RELATED TO THE SPECIFICATION. NOTICE IS HEREBY GIVEN, THAT OTHER RIGHTS NOT GRANTED AS SET FORTH ABOVE, INCLUDING WITHOUT LIMITATION, RIGHTS OF THIRD PARTIES WHO DID NOT EXECUTE THE ABOVE LICENSES, MAY BE IMPLICATED BY THE IMPLEMENTATION OF OR COMPLIANCE WITH THIS SPECIFICATION. OCP IS NOT RESPONSIBLE FOR IDENTIFYING RIGHTS FOR WHICH A LICENSE MAY BE REQUIRED IN ORDER TO IMPLEMENT THIS SPECIFICATION. THE ENTIRE RISK AS TO IMPLEMENTING OR OTHERWISE USING THE SPECIFICATION IS ASSUMED BY YOU. IN NO EVENT WILL OCP BE LIABLE TO YOU FOR ANY MONETARY DAMAGES WITH RESPECT TO ANY CLAIMS RELATED TO, OR ARISING OUT OF YOUR USE OF THIS SPECIFICATION, INCLUDING BUT NOT LIMITED TO ANY LIABILITY FOR LOST PROFITS OR ANY CONSEQUENTIAL, INCIDENTAL, INDIRECT, SPECIAL OR PUNITIVE DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO

THIS SPECIFICATION, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND EVEN IF OCP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

1.2 Acknowledgements

The Contributors of this Specification would like to acknowledge the following companies for their feedback:

Intel
Quanta

List all companies or individuals who may have assisted you with the specification by providing feedback and suggestions but did not provide any IP.

2. OCP Tenets Compliance

[Note to author of this specification: This section is optional if the supplier chooses to highlight additional tenet compliance that was not covered in the baseline specification. Refer to the baseline specification template for details on each tenet]

2.1. Openness

Angels Landing in Zion Design follows the OAI concept. The whole Zion system implements OAM, OpenBMC technology to support ML/AI workload. Facebook contributes the whole Zion design specification in OCP.

2.2. Efficiency

Angels Landing serves as a CPU box in Zion system. It is designed with a modular concept where we can support different configurations for different use cases to optimize TCO.

2.3. Impact

Zion system is the system that is designed to support large scale AI workloads with OAMs. It has been implemented in a large-scale DC environment with huge improvement of production efficiency.

2.4. Scale

Zion System is designed to optimize TCO for AI workload. It also implements OpenBMC in design to enable large scale monitoring and telemetry.

3. Revision Table

Date	Revision #	Author	Description
07-21-2020	0.1	Facebook	Initial Draft
04-19-2021	0.2	Facebook	Update spec and drawing
07-28-2021	1.0	Facebook	Release

4. Scope

This document describes technical specifications of Facebook's Angels Landing server node used in OCP. Angels Landing is part of Zion System which has a modular design that contains server box (Angels Landing), interconnection box (Clear Creek) and GPU box (Emerald Pools).

4.1 Terminology

Term	Description
2S-MB	2 Socket Motherboard with Intel® Cooper Lake CPU
Angels Landing 2S	2 Socket Server
Angels Landing 4S	4 socket Server
BIC	Bridge IC
BMC	Baseboard management Controller
CPX-6S	Intel® Cooper Lake, 14nm processor
EDP/TDP	Electrical/Thermal Design Power
Emerald Pools	8 Accelerator System
IMC	Integrated Memory Controller
Intel® PFR	Intel® Platform Firmware Resiliency
ME	Intel® Manageability Engine
PCH	Platform Controller Hub – Intel® Lewisburg C620
PCIe	PCI Express high-speed serial bus
PECI	Platform Environment Control Interface
SMBUS/PMBUS/I2C	Low-speed serial protocols for low level management and monitoring tasks
TPM	Trusted Platform Module
UDIMM/RDIMM/LRDIMM	Unbuffered/Registered/Load Reduced DIMM
UPI	Intel® Ultra Path Interconnect
Zion	Facebook's AI Platform

5. Zion Overview

AI is a difficult problem. First, it is demanding. The nature of AI algorithms requires a tremendous amount of compute power, large amounts of memory and massive IO throughput. To receive good results, AI models could become very big and traditional server architecture can no longer handle them nicely anymore. Second, the algorithm evolves fast. The model size and complexity can easily go 10x each 18 month, which is purely driven by the business needs. Hardware needs to catch up this trend by providing more compute power and memory. The Zion System is designed to meet this need.

Zion system includes three boxes: Angels Landing, Clear Creek and Emerald Pools. Angels Landing is the CPU box; Clear Creek is the interconnect box; Emerald Pools is the accelerator box. These three boxes are connected as the Figure below shows:

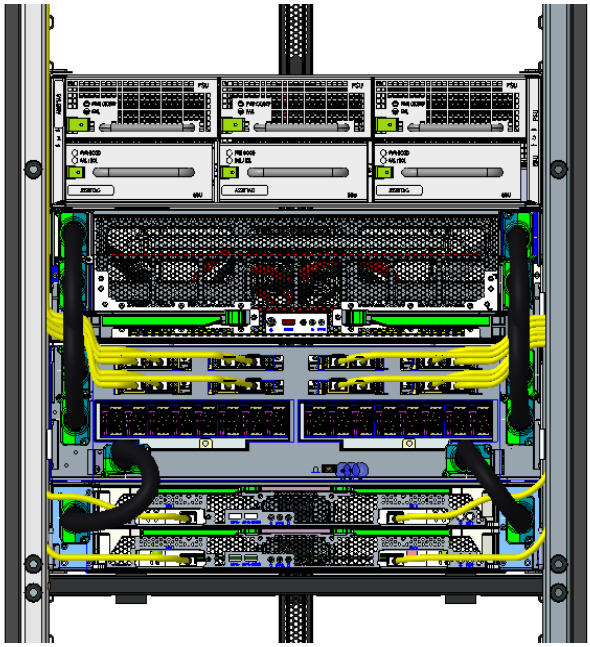


Figure 1: Zion System View

In this document we will focus on the CPU box: Angels Landing. Please refer to Zion system spec and Emerald Pools system spec for more information.

6. Angels Landing Overview

Angels Landing is a 4-socket system based on Intel® CPX-6S processor architecture. The system contains 4x CPUs which are interconnected through Intel® UPI bus in a Fully Connected configuration. This system supports up to 48x DIMMs @3200MT/S and 4x NICs. The CPU box also provided the connectivity to the GPU and interconnect box using PCIe interface. Angels landing can be configured as a 2-socket system also based on the modular design concept we implemented.

6.1 system introduction

Figure. 2 below illustrates the Angels Landing box with the top cover removed. There are 2 trays in the system's front side. Each tray occupies 1 OU space and carries one 2S-MB - dual-socket motherboard. Each tray can be serviced individually.

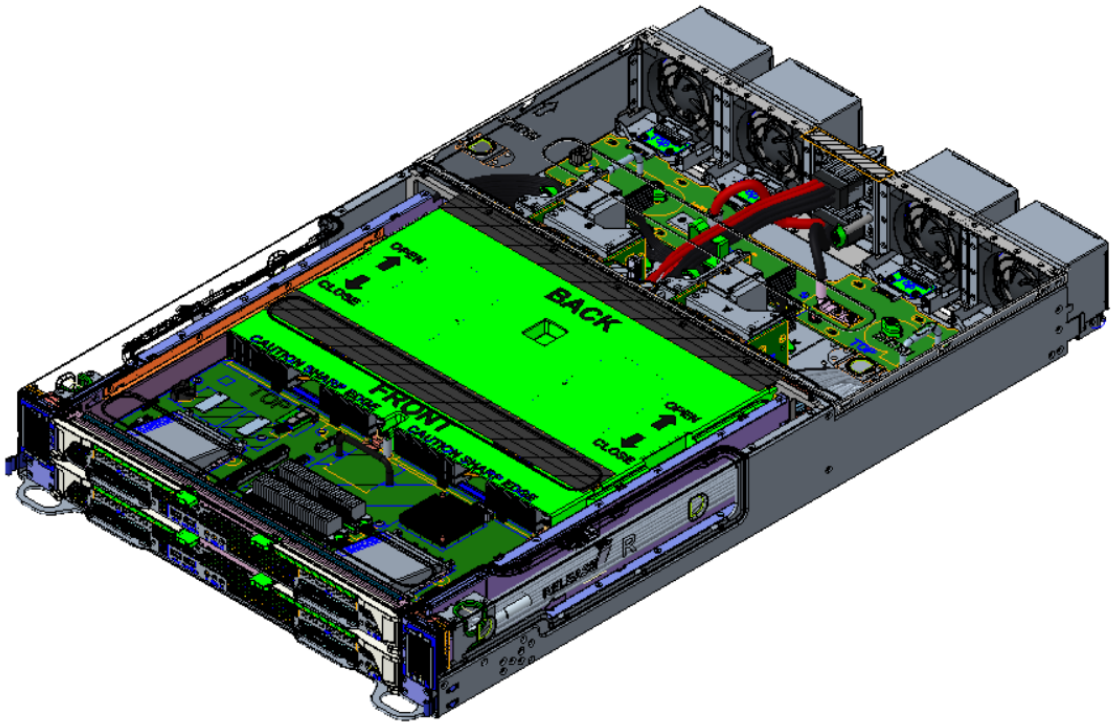


Figure 2: System View, Top cover removed

Figure 3 shows more details about the backplane connection idea. The backplane connection has been separated into two groups: The UPI system interconnect is implemented with 4x individual UPI BP boards which are used to connect the 2 * 2S MB together. Each motherboard has two PCIe connectors with 1 on each side that connect Angels Landing to Emerald Pools.

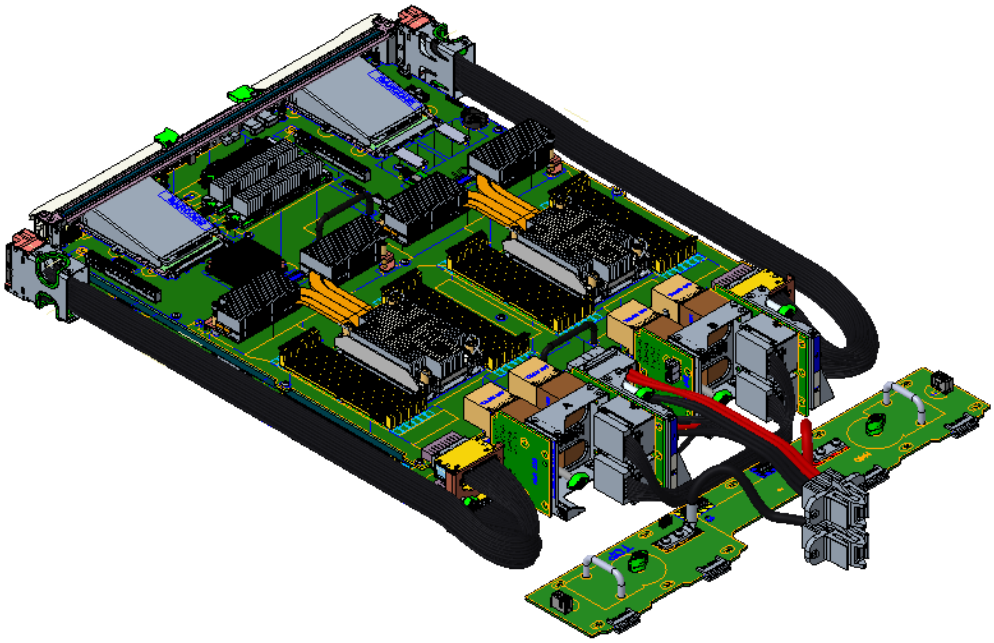


Figure 3: Angels Landing PCB back view

The Block diagram of Angles landing is listed in Figure 4. Each motherboard has two CPUs, PCH and BMC's. Two MB are connected through an UPI network. Only one motherboard (MB0) can be treated as a master board. The PCH on the slave board will be disabled. Each motherboard has the same design and either one of them can be configured as master MB.

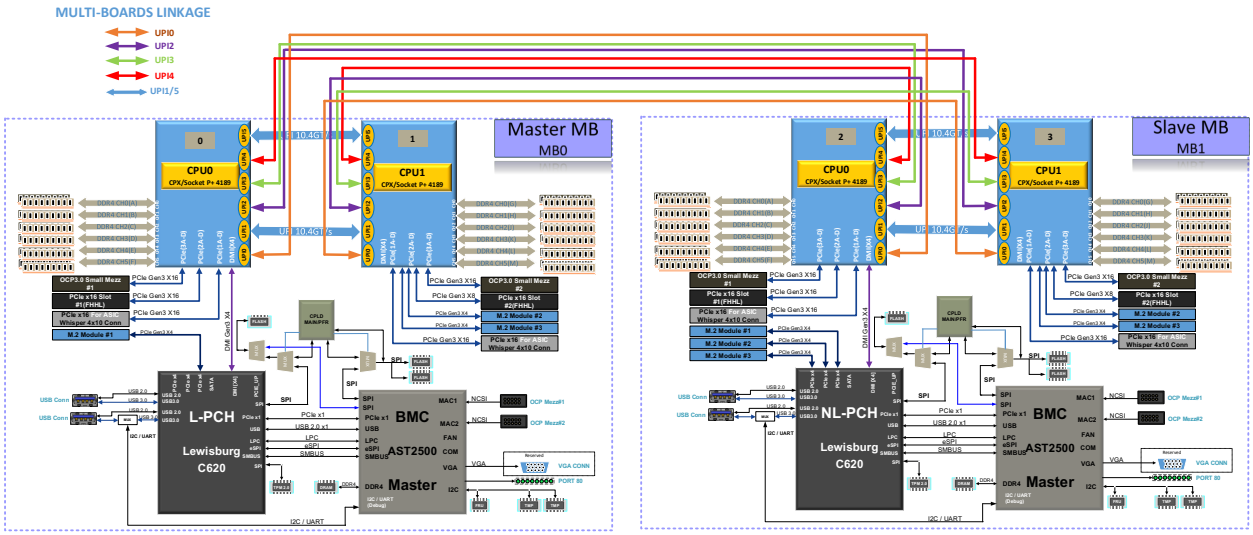


Figure 4: Angels Landing Block Diagram

Angel Landing can be configured as a 2S server by removing the MB1 motherboard and the UPI backplane boards. Our mechanical design can also support those two skus at the same time.

Figure 5 shows a typical rack configuration of the Zion system. Each rack will contain 2x Zion systems.

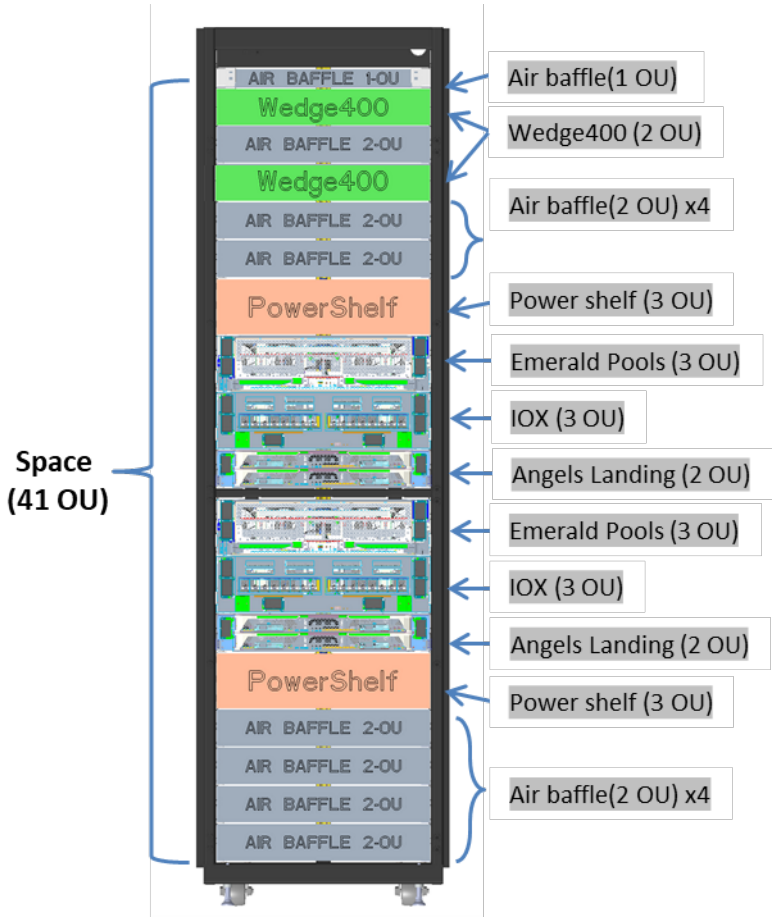


Figure 5: Angels Landing Rack Level Configuration

6.2 Angels Landing Components

6.2.1 Component list

Table 6.2-1 below lists the major components in the Angels Landing system.

Table 6.2-1: Angels Landing Major Component List

Item	Description
2S-MB	2* 2-socket module MBs. Each MB sits in a tray which contains: ● 2* Cooperlake-SP,24 Core, 8339HC 1.8GHz ● 24* DDR4 DIMM slots ● 2* OCP NIC3.0 slots up to 100G NICs

	• 1* NVMe SSD as boot drive 2* 2TB NVMe for storage
UPI Backplanes	4*UPI backplanes are used for 4* CPUs connectivity
UPI cable	2* UPI cables that support fully connected topology
Whisper Cables	Internal Whisper cables are used to connect 2* 2S MB cards to External systems such as Emerald Pools
PDB card	Power Distribution Board (w/ optional 48V support)
Fan boards	4* Fan boards with gold finger connection to PDB
Power Supply	Open Rack V2 Central PSU
Chassis	2OU Chassis for Open Rack v2

6.2.2 Component Placement and Form Factor

2S-MB form factor is 431mm x 466mm, Figure 3 illustrates board to board connection in Angels Landing system. Figure 6 shows the location of CPU, DIMMs, Heatsink, M.2 connector and NIC cards. The placement is meant to show key components' relative positions, while exact dimension and position information would be exchanged by DXF format for layout and 3D model for mechanical.

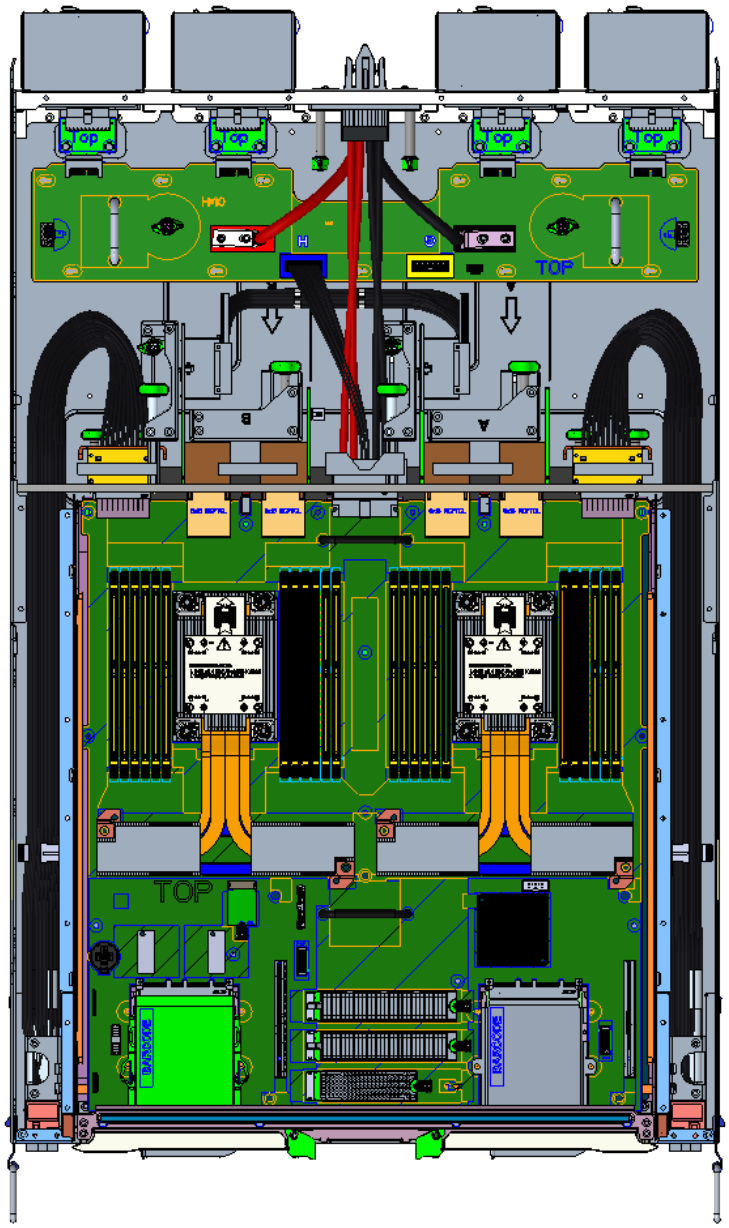


Figure 6: Angels Landing connectivity, top view

Figure 7 below shows major Angels Landing components breakdown:

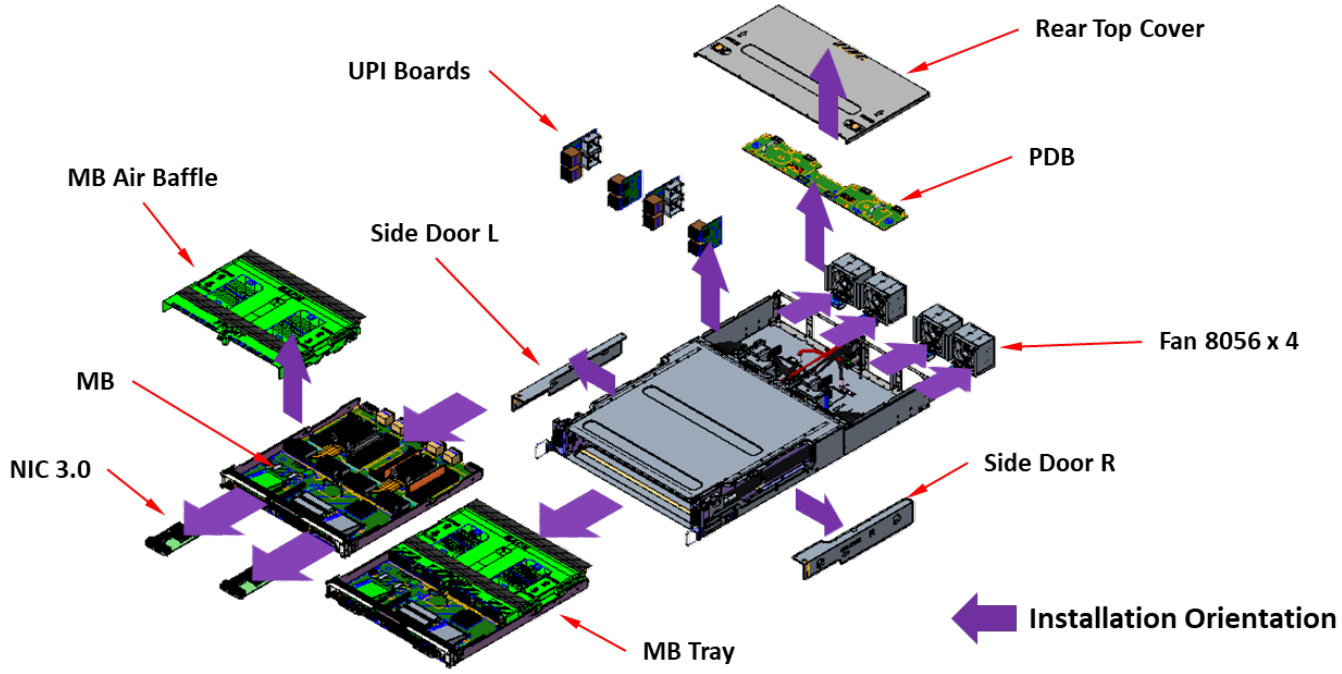


Figure 7: System Level components Breakdown

System contains the following PCBs:

- 2x 2-Socket CPU motherboards. These two boards are same in design
- 4x UPI boards to connect for Motherboards through backplane. Those UPI boards are all different.
- 1x Power Distribution Board (PDB) to support power to all four MBs.
- 2x Fan boards to support 4x fans
- 2x busbar connectors

The following backplane connector has been designed at back of 2S module MB to provide power, UPI backplane and PCIe cables connections:

- 4x 6x16 Amphenol ExaMax Connector(10131770-101LF) for UPI
- 2x 4x10 TE Whisper for PCIe
- 1x FCI Power Blade Ultra for power and side band signal

Figure 8 gives a back view about where those connectors are located.

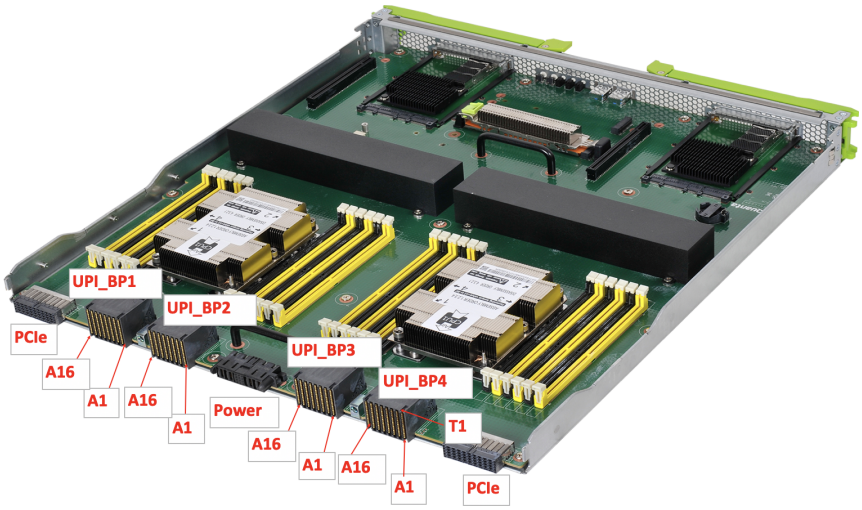


Figure 8: AL backplane connectors on MB

The following internal connectors has been placed at the front of 2S module MB for easy front access:

- 1x SMD switch to enable/disable Intel® ASD (At Scale Debug) (MFG: Pinrex, MFG_PN: 210-HP1-030BR1)
- 3x M.2 connectors with 2280 and 22110 support
- 2x USB type A
- 1X customized VGA connector (MFG: Foxconn, MFG_PN: DZ11A51-H8R1-4F).

Each MB support the following connector:

- 3x M.2 connectors with 2280 and 22110 support
- 2x OCP NIC 3.0 connectors to support 2x 100G NIC. Each NIC connects to one CPU.
- 1x BSM connector to support Facebook's BSM modules
- No PCIe CEM connector is supported in Angels Landing design.

Please refer to DXF for critical component placement.

6.3 System Connectivity

This section illustrates the UPI and PCIe interconnection. The whole Zion system uses a complex backplane and cable system to connect subsystems together. In this section, we also cover NIC interface and sideband communication schemes to expose how multiple components in Angels Landing work together.

6.3.1 UPI, Intel® Ultra Path Interconnect

Each CPU in Angels Landing has 6 UPI links. The 4 * CPUs are connected through their UPI links with fully connected topology. Figure 8 below illustrates these connections:

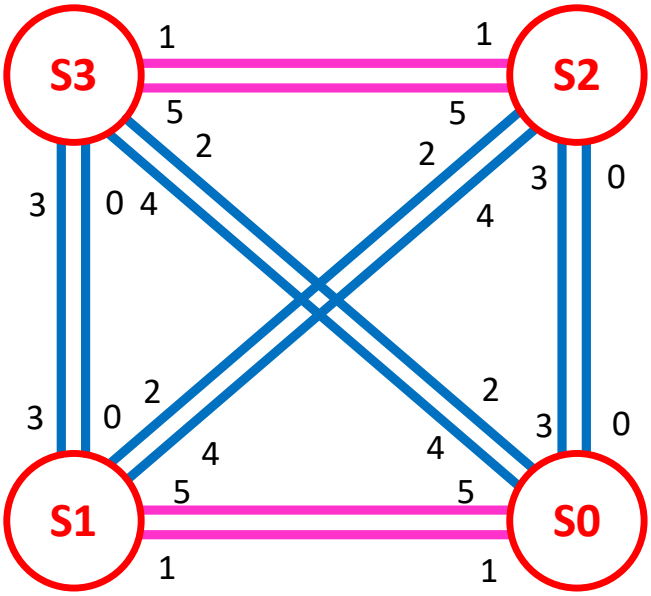


Figure 9: Angels Landing 4 *2S UPI interconnect

Figure 9 illustrates how each CPU UPI port is connected to other ports. There are two motherboards and four UPI backplanes. Each motherboard's two CPUs connect to each other through UPI port 1/5. Each CPU has four UPI ports to UPI backplanes for system interconnect.

Each UPI board has been named as UPI_BP_n where n is from 1 to 4. The design choice here is to keep four CPU MB identical. To support the topology, each UPI board design has to be different.

The motherboard UPI connectors are manufactured by FCI, part number 10131770-101LF

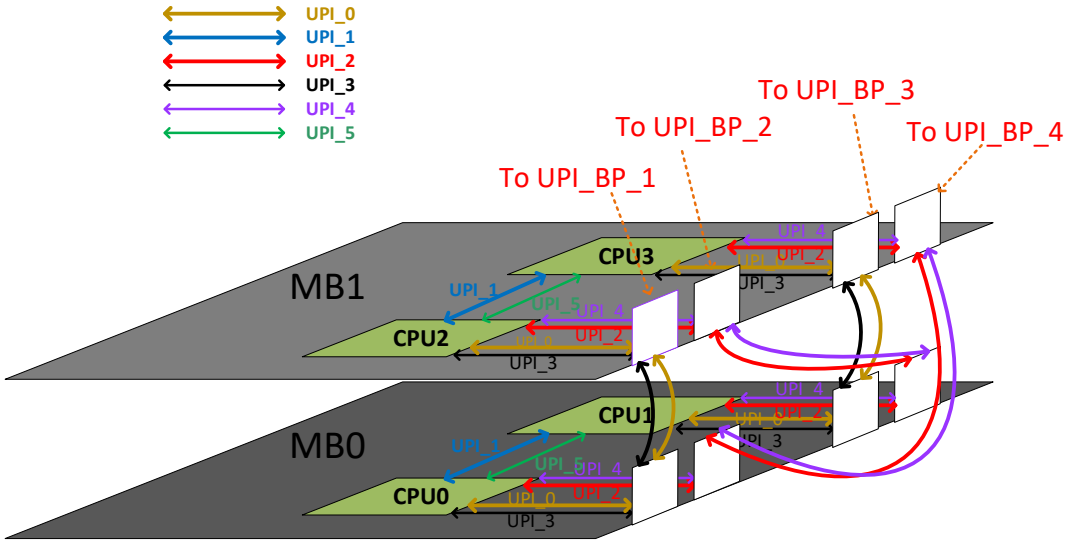


Figure 10: Backplane UPI port connections

4* Amphenol ExaMax 6x16 connectors are used on each motherboard to connect to UPI backplanes. Corresponding ExaMax 6x16 pair connectors are used on the UPI backplane side.

- Motherboard ExaMax connector MPN: 10131770-101LF
- UPI backplane ExaMax connector MPN: 10128316-102LF

There are two additional ExaMax 6x14 connectors and cables on UPI_BP_2 and UPI_BP_4 to create a fully connected UPI topology.

UPI clock is provided from UPI_BP_2 board to two MBs as shown in

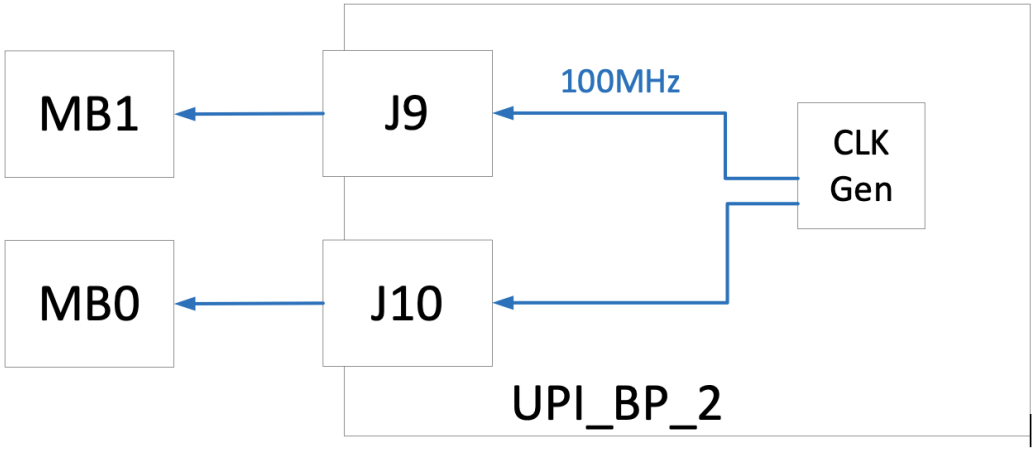


Figure 11: Backplane UPI port connections

6.3.2 PCIe

Intel® CPX-6S processor family provides 48 PCIe Gen3 lanes and Intel® PCH provides up to 16 PCIe Gen3 uplinks and up to 20 PCIe. PCIe lanes are configured as described below:

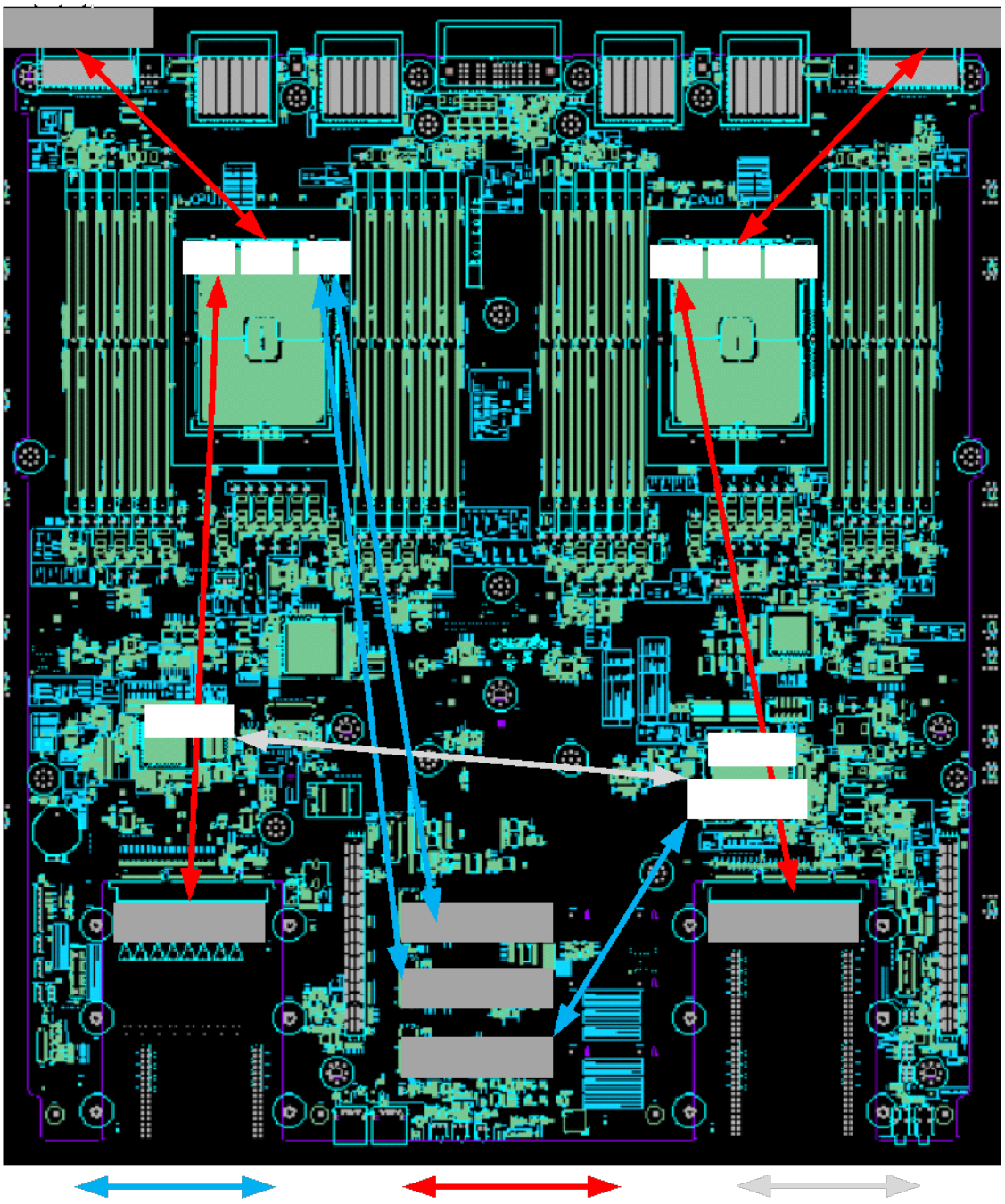


Figure 12: Motherboard CPU and PCH PCIe Block Diagram

The following figure shows the system level PCIe connections, each of the 2S MB has 2 * 4x10 internal Whisper connectors, each of these connectors supports x16.

2* external Y-cables connect the 8 internal Whisper connectors to 4 x32 external Whisper cables

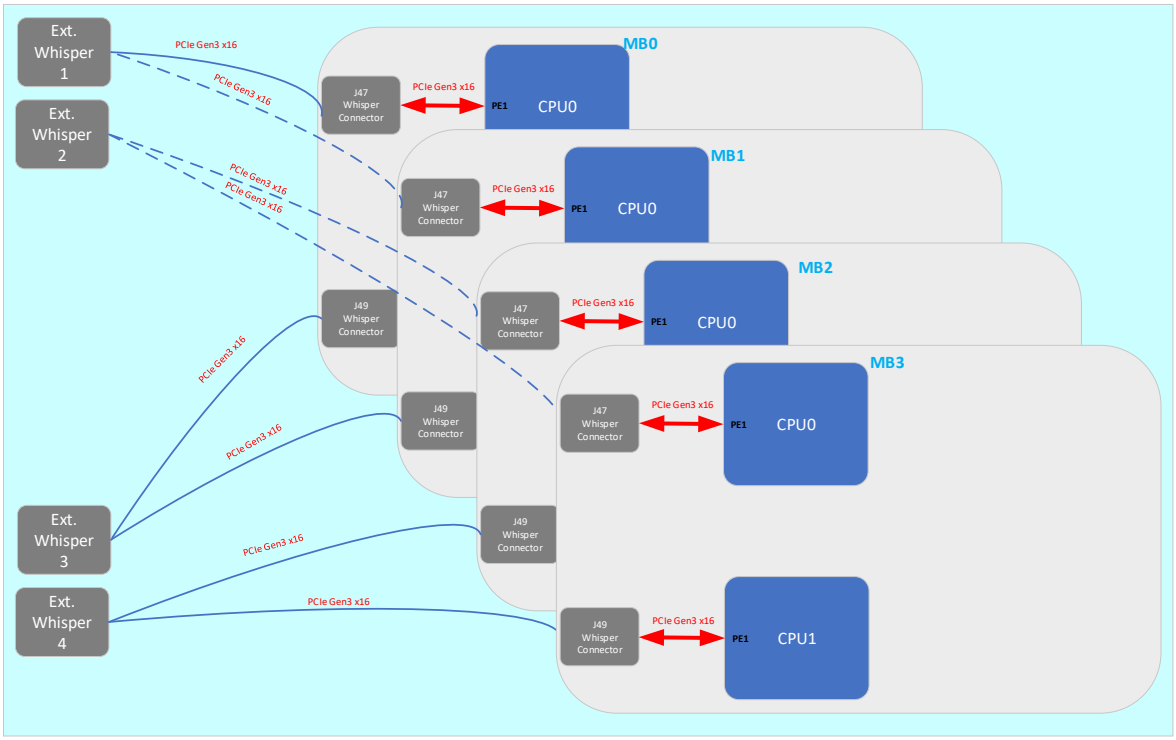


Figure 13: CPU PCIe External Connections

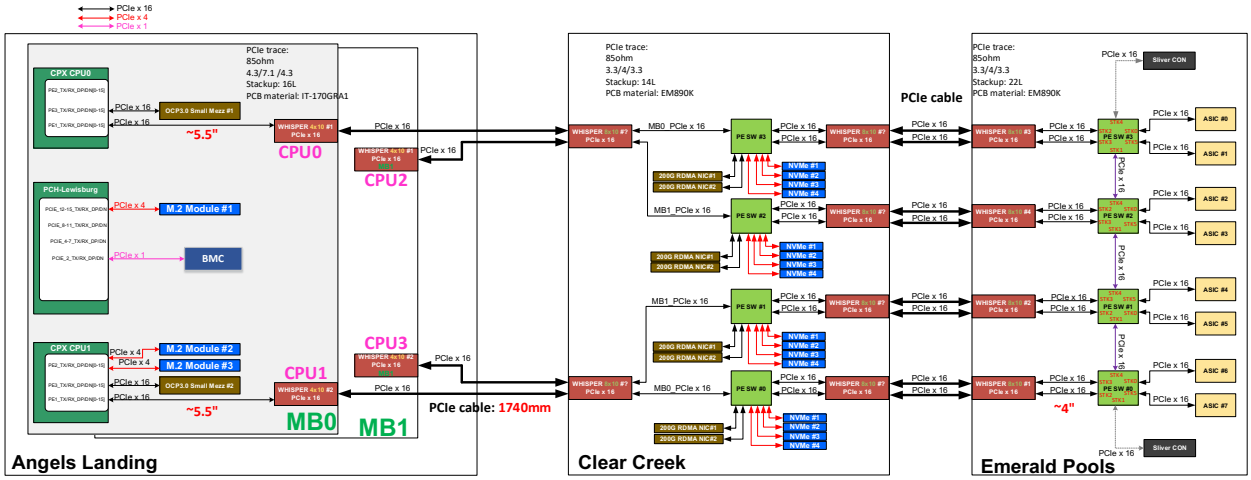


Figure 14: Zion PCIe Connections

2* 4x10 Whisper connectors are used on each motherboard to connect to two Whisper cables.

Table 6.3-1: CPU_0 To Whisper J47 connections

B26,B27	SMB_BMC_ASICBOX1_SCL/SDA	I2C to Emerald Pools
A26,A27	SMB_PCH_RETIMER1_ISO_SCL/SDA	I2C to Retimer Card
A30	RST_PCIE_CABLE_BUF_N	PCIe Reset to Emerald Pools
C26,C27	CLK_100M_RETIMER_S2_DP/N	PCIe Clock to Emerald Pools
D26,D27	USB2_PCH_ASIC_1_BMC_DP/N	PCH USB2 to Emerald Pools
C29,C30	P3E_PCH_PCIE0_RX_C_DP/N<0>	?

B29,B30	P3E_PCH_PCIE0_TX_C_DP/N<0>	?
D29	FM_ASIC_POWER_EN	Emerald Pools power enable
D30	FM_ASIC_PWRGD	Emerald Pools power good
A29	PRSNT_PCIE_CABLE_1_N	Cable presence status
C23,C24	P3E_CPU0_PCIE1_RX_DP/N<0>	High Speed Rx Lane 0
A24,A23	P3E_CPU0_PCIE1_RX_DP/N<1>	High Speed Rx Lane 1
D21,D20	P3E_CPU0_PCIE1_RX_DP/N<2>	High Speed Rx Lane 2
B20,B21	P3E_CPU0_PCIE1_RX_DP/N<3>	High Speed Rx Lane 3
C18,C17	P3E_CPU0_PCIE1_RX_DP/N<4>	High Speed Rx Lane 4
A18,A17	P3E_CPU0_PCIE1_RX_DP/N<5>	High Speed Rx Lane 5
D15,D14	P3E_CPU0_PCIE1_RX_DP/N<6>	High Speed Rx Lane 6
B15,B14	P3E_CPU0_PCIE1_RX_DP/N<7>	High Speed Rx Lane 7
C12,C11	P3E_CPU0_PCIE1_RX_DP/N<8>	High Speed Rx Lane 8
A12,A11	P3E_CPU0_PCIE1_RX_DP/N<9>	High Speed Rx Lane 9
D8,D9	P3E_CPU0_PCIE1_RX_DP/N<10>	High Speed Rx Lane 10
B9,B8	P3E_CPU0_PCIE1_RX_DP/N<11>	High Speed Rx Lane 11
C6,C5	P3E_CPU0_PCIE1_RX_DP/N<12>	High Speed Rx Lane 12
A5,A6	P3E_CPU0_PCIE1_RX_DP/N<13>	High Speed Rx Lane 13
D3,D2	P3E_CPU0_PCIE1_RX_DP/N<14>	High Speed Rx Lane 14
B3,B2	P3E_CPU0_PCIE1_RX_DP/N<15>	High Speed Rx Lane 15
D24,D23	P3E_CPU0_PCIE1_TX_C_DP/N<0>	High Speed Tx Lane 0
B23,B24	P3E_CPU0_PCIE1_TX_C_DP/N<1>	High Speed Tx Lane 1
C21,C20	P3E_CPU0_PCIE1_TX_C_DP/N<2>	High Speed Tx Lane 2
A21,A20	P3E_CPU0_PCIE1_TX_C_DP/N<3>	High Speed Tx Lane 3
D18,D17	P3E_CPU0_PCIE1_TX_C_DP/N<4>	High Speed Tx Lane 4
B18,B17	P3E_CPU0_PCIE1_TX_C_DP/N<5>	High Speed Tx Lane 5
C15,C14	P3E_CPU0_PCIE1_TX_C_DP/N<6>	High Speed Tx Lane 6
A15,A14	P3E_CPU0_PCIE1_TX_C_DP/N<7>	High Speed Tx Lane 7
D12,D11	P3E_CPU0_PCIE1_TX_C_DP/N<8>	High Speed Tx Lane 8
B12,B11	P3E_CPU0_PCIE1_TX_C_DP/N<9>	High Speed Tx Lane 9
C9,C8	P3E_CPU0_PCIE1_TX_C_DP/N<10>	High Speed Tx Lane 10
A8,A9	P3E_CPU0_PCIE1_TX_C_DP/N<11>	High Speed Tx Lane 11
C6,C5	P3E_CPU0_PCIE1_TX_C_DP/N<12>	High Speed Tx Lane 12
B6,B5	P3E_CPU0_PCIE1_TX_C_DP/N<13>	High Speed Tx Lane 13
C2,C3	P3E_CPU0_PCIE1_TX_C_DP/N<14>	High Speed Tx Lane 14
A3,A2	P3E_CPU0_PCIE1_TX_C_DP/N<15>	High Speed Tx Lane 15
A1,A4,A7,A10,A13,A16, A19,A22,A25,A28, B1,B4,B7,B10,B13,B16, B19,B22,B25,B28, C1,C4,C7,C10,C13,C16, C19,C22,C25,C28, D1,D4,D7,D10,D13,D16, D19,D22,D25,D28, GND_REF_A[31:2], GND_REF_B[31:2],	GND	Ground

GND_REF_C[31:2], GND_REF_D[31:2], GND_REF_E[30:2], Via_A31, Via_B31, Via_C31, Via_D31, Via_REF_A1, Via_REF_B1, Via_REF_C1, Via_REF_D1		

Table 6.3-2: CPU_1 To Whisper J49 Connections

C24,C23	P3E_CPU1_PCIE1_RX_DP/N<0>	High Speed Rx Lane 0
A24,A23	P3E_CPU1_PCIE1_RX_DP/N<1>	High Speed Rx Lane 1
D21,D20	P3E_CPU1_PCIE1_RX_DP/N<2>	High Speed Rx Lane 2
B20,B21	P3E_CPU1_PCIE1_RX_DP/N<3>	High Speed Rx Lane 3
C18,C17	P3E_CPU1_PCIE1_RX_DP/N<4>	High Speed Rx Lane 4
A18,A17	P3E_CPU1_PCIE1_RX_DP/N<5>	High Speed Rx Lane 5
D15,D14	P3E_CPU1_PCIE1_RX_DP/N<6>	High Speed Rx Lane 6
B15,B14	P3E_CPU1_PCIE1_RX_DP/N<7>	High Speed Rx Lane 7
C12,C11	P3E_CPU1_PCIE1_RX_DP/N<8>	High Speed Rx Lane 8
A12,A11	P3E_CPU1_PCIE1_RX_DP/N<9>	High Speed Rx Lane 9
D8,D9	P3E_CPU1_PCIE1_RX_DP/N<10>	High Speed Rx Lane 10
B9,B8	P3E_CPU1_PCIE1_RX_DP/N<11>	High Speed Rx Lane 11
C6,C5	P3E_CPU1_PCIE1_RX_DP/N<12>	High Speed Rx Lane 12
A5,A6	P3E_CPU1_PCIE1_RX_DP/N<13>	High Speed Rx Lane 13
D3,D2	P3E_CPU1_PCIE1_RX_DP/N<14>	High Speed Rx Lane 14
B3,B2	P3E_CPU1_PCIE1_RX_DP/N<15>	High Speed Rx Lane 15
D24,D23	P3E_CPU1_PCIE1_TX_C_DP/N<0>	High Speed Tx Lane 0
B23,B24	P3E_CPU1_PCIE1_TX_C_DP/N<1>	High Speed Tx Lane 1
C21,C20	P3E_CPU1_PCIE1_TX_C_DP/N<2>	High Speed Tx Lane 2
A21,A20	P3E_CPU1_PCIE1_TX_C_DP/N<3>	High Speed Tx Lane 3
D18,D17	P3E_CPU1_PCIE1_TX_C_DP/N<4>	High Speed Tx Lane 4
B18,B17	P3E_CPU1_PCIE1_TX_C_DP/N<5>	High Speed Tx Lane 5
C15,C14	P3E_CPU1_PCIE1_TX_C_DP/N<6>	High Speed Tx Lane 6
A15,A14	P3E_CPU1_PCIE1_TX_C_DP/N<7>	High Speed Tx Lane 7
D12,D11	P3E_CPU1_PCIE1_TX_C_DP/N<8>	High Speed Tx Lane 8
B12,B11	P3E_CPU1_PCIE1_TX_C_DP/N<9>	High Speed Tx Lane 9
C9,C8	P3E_CPU1_PCIE1_TX_C_DP/N<10>	High Speed Tx Lane 10
A9,A8	P3E_CPU1_PCIE1_TX_C_DP/N<11>	High Speed Tx Lane 11
D6,D5	P3E_CPU1_PCIE1_TX_C_DP/N<12>	High Speed Tx Lane 12
B6,B5	P3E_CPU1_PCIE1_TX_C_DP/N<13>	High Speed Tx Lane 13
C2,C3	P3E_CPU1_PCIE1_TX_C_DP/N<14>	High Speed Tx Lane 14
A3,A2	P3E_CPU1_PCIE1_TX_C_DP/N<15>	High Speed Tx Lane 15
B26,B27	SMB_BMC_ASICBOX_2_SCL/SDA	I2C to Emerald Pools
A26,A27	SMB_PCH_RETIMER2_ISO_SCL/SDA	I2C (to optional Retimer Card)

A30	RST_PCIE_CABLE_1_BUF_N	PCIe Reset to Retimer and Emerald Pools
C26,C27	CLK_100M_RETIMER_S1_DP/N	PCIe Clock to Emerald Pools
D26,D27	USB2_PCH_ASIC_2_BMC_DP/N	PCH USB2 to Emerald Pools
C29,C30	P3E_PCH_PCIE1_RX_DP/N<0>	PCH to Emerald Pools PCIe x1 Tx
B29,B30	P3E_PCH_PCIE1_TX_C_DP/N<0>	PCH to Emerald Pools PCIe x1 Rx
D29	FM_ASIC_POWER_EN	ASIC power Enable
D30	FM_ASIC_PWRGD	Asic power good
A29	PRSNT_PCIE_CABLE_2_N	Present signal
A1,A4,A7,A10,A13,A16, A19,A22,A25,A28 B1,B4,B7,B10,B13,B16, B19,B22,B25,B28, C1,C4,C7,C10,C13,C16, C19,C22,C25,C28, D1,D4,D7,D10,D13,D16, D19,D22,D25,D28, GND_REF_A[31:2], GND_REF_B[31:2], GND_REF_C[31:2], GND_REF_D[31:2], GND_REF_E[31:2], Via_A31, Via_B31, Via_C31, Via_D31, Via_REF_A1, Via_REF_B1, Via_REF_C1, Via_REF_D1	GND	GND

2 * PCIe Gen3 x16 slot connectors (MFG: TE MFG PN:7-2340321-1) are used for connection to OCP NIC 3.0 cards

Table 6.3-3: CPU to OCP Mezz J74 Connections

OCP_A10,OCP_A13,A1, A2,A3,A4,A5,A6,A13,A16, A19,A22,A25,A28,A29, A32,A35,A38,A41,A43, A46,A49,A52,A55,A58, A61,A64,A67, OCP_B10,OCP_B13,B13, B16,B19,B22,B25,B28, B29,B32,B35,B38,B41, B43,B46,B49,B52,B55, B58,B61,B64,B67	GND	GND
B1,B2,B3,B4,B5,B6	P12V_OCP_V3_1	+12V supply
B11	P3V3_OCP_V3_1	+3.3V supply
B12	OCP_V3_1_A_AUX_PWR_EN	

OCP_B1	OCP_V3_1_NIC_PWR_GOOD	
OCP_B2	OCP_V3_1_MAIN_PWR_R_EN	
OCP_B3	SGPIO_OCP_V3_1_LD_N	
OCP_B4	SGPIO_OCP_V3_1_DATA_IN	
OCP_B5	SGPIO_OCP_V3_1_DATA_OUT	
OCP_B6	SGPIO_OCP_V3_1_CLK	
OCP_A4	OCP_V3_1_ISO_RBT_ARB_IN	
OCP_A5	OCP_V3_1_ISO_RBT_ARB_OUT	
OCP_A6	OCP_V3_1_SLOT_ID1	
OCP_A7	NCSI_OCP_V3_1_ISO_TX_EN	
OCP_A8	NCSI_OCP_V3_1_ISO_TXD1	
OCP_A9	NCSI_OCP_V3_1_ISO_TXD0	
OCP_B7	OCP_V3_1_SLOT_ID0	
OCP_B8	NCSI_OCP_V3_1_ISO_RXD1	
OCP_B9	NCSI_OCP_V3_1_ISO_RXD0	
OCP_B14	NCSI_OCP_V3_1_CRS_R_DV	
OCP_A14	CLK_50M_OCP_V3_1_ISO_RBT_REF	
B10	RST_OCP_V3_1_A_0_N	
A11	RST_OCP_V3_1_A_1_N	
OCP_A1	RST_OCP_V3_1_A_2_N	
OCP_A2	RST_OCP_V3_1_A_3_N	
OCP_A3	OCP_V3_1_WAKE_N	
B42	OCP_V3_1_PRSNTB0_N	
A42	OCP_V3_1_PRSNTB1_N	
A12	OCP_V3_1_PRSNTB2_N	
B70	OCP_V3_1_PRSNTB3_N	
B7	OCP_V3_1_ISO_BIF0_N	
B8	OCP_V3_1_ISO_BIF1_N	
B9	OCP_V3_1_ISO_BIF2_N	
A70	OCP_V3_1_PWRBRK_N	
A7	SMB_BMC_OCP_V3_1_A_SCL	
A8	SMB_BMC_OCP_V3_1_A_SDA	
A9	SMB_BMC_OCP_V3_1_A_RST_N	
OCP_B11,OCP_B12	CLK_100M_DB2001_OCP_V3_1_A_DP/N	100Mhz PCIe reference clock
OCP_A11,OCP_A12	CLK_100M_DB2001_OCP_V3_1_B_DP/N	100Mhz PCIe reference clock
B15,B14	CLK_100M_DB2001_OCP_V3_1_C_DP/N	100Mhz PCIe reference clock
A15,A14	CLK_100M_DB2001_OCP_V3_1_D_DP/N	100Mhz PCIe reference clock
A68,A69	USB2_OCP_1_DP/N	USB 2.0 interface
A17,A18	P3E_CPU0_PCIE3_RX_DP/N<0>	High Speed Rx Lane 0
A20,A21	P3E_CPU0_PCIE3_RX_DP/N<1>	High Speed Rx Lane 1
A23,A24	P3E_CPU0_PCIE3_RX_DP/N<2>	High Speed Rx Lane 2

A26,A27	P3E_CPU0_PCIE3_RX_DP/N<3>	High Speed Rx Lane 3
A30,A31	P3E_CPU0_PCIE3_RX_DP/N<4>	High Speed Rx Lane 4
A33,A34	P3E_CPU0_PCIE3_RX_DP/N<5>	High Speed Rx Lane 5
A36,A37	P3E_CPU0_PCIE3_RX_DP/N<6>	High Speed Rx Lane 6
A39,A40	P3E_CPU0_PCIE3_RX_DP/N<7>	High Speed Rx Lane 7
A44,A45	P3E_CPU0_PCIE3_RX_DP/N<8>	High Speed Rx Lane 8
A47,A48	P3E_CPU0_PCIE3_RX_DP/N<9>	High Speed Rx Lane 9
A50,A51	P3E_CPU0_PCIE3_RX_DP/N<10>	High Speed Rx Lane 10
A53,A54	P3E_CPU0_PCIE3_RX_DP/N<11>	High Speed Rx Lane 11
A56,A57	P3E_CPU0_PCIE3_RX_DP/N<12>	High Speed Rx Lane 12
A59,A60	P3E_CPU0_PCIE3_RX_DP/N<13>	High Speed Rx Lane 13
A62,A63	P3E_CPU0_PCIE3_RX_DP/N<14>	High Speed Rx Lane 14
A65,A66	P3E_CPU0_PCIE3_RX_DP/N<15>	High Speed Rx Lane 15
B17,B18	P3E_CPU0_PCIE3_TX_C_DP/N<0>	High Speed Tx Lane 0
B20,B21	P3E_CPU0_PCIE3_TX_C_DP/N<1>	High Speed Tx Lane 1
B23,B24	P3E_CPU0_PCIE3_TX_C_DP/N<2>	High Speed Tx Lane 2
B26,B27	P3E_CPU0_PCIE3_TX_C_DP/N<3>	High Speed Tx Lane 3
B30,B31	P3E_CPU0_PCIE3_TX_C_DP/N<4>	High Speed Tx Lane 4
B33,B34	P3E_CPU0_PCIE3_TX_C_DP/N<5>	High Speed Tx Lane 5
B36,B37	P3E_CPU0_PCIE3_TX_C_DP/N<6>	High Speed Tx Lane 6
B39,B40	P3E_CPU0_PCIE3_TX_C_DP/N<7>	High Speed Tx Lane 7
B44,B45	P3E_CPU0_PCIE3_TX_C_DP/N<8>	High Speed Tx Lane 8
B47,B48	P3E_CPU0_PCIE3_TX_C_DP/N<9>	High Speed Tx Lane 9
B50,B51	P3E_CPU0_PCIE3_TX_C_DP/N<10>	High Speed Tx Lane 10
B53,B54	P3E_CPU0_PCIE3_TX_C_DP/N<11>	High Speed Tx Lane 11
B56,B57	P3E_CPU0_PCIE3_TX_C_DP/N<12>	High Speed Tx Lane 12
B60,B59	P3E_CPU0_PCIE3_TX_C_DP/N<13>	High Speed Tx Lane 13
B63,B62	P3E_CPU0_PCIE3_TX_C_DP/N<14>	High Speed Tx Lane 14
B66,B65	P3E_CPU0_PCIE3_TX_C_DP/N<15>	High Speed Tx Lane 15

Table 6.3-4: CPU to OCP Mezz J73:

OCP_A10,OCP_A13,A1, A2,A3,A4,A5,A6,A13,A16, A19,A22,A25,A28,A29, A32,A35,A38,A41,A43, A46,A49,A52,A55,A58, A61,A64,A67, OCP_B10,OCP_B13,B13, B16,B19,B22,B25,B28, B29,B32,B35,B38,B41, B43,B46,B49,B52,B55, B58,B61,B64,B67	GND	GND
B1,B2,B3,B4,B5,B6	P12V_OCP_V3_2	+12V supply
B11	P3V3_OCP_V3_2	+3.3V supply
B12	OCP_V3_2_A_AUX_PWR_EN	
OCP_B1	OCP_V3_2_NIC_PWR_GOOD	
OCP_B2	OCP_V3_2_MAIN_PWR_R_EN	

OCP_B3	SGPIO_OCP_V3_2_LD_N	
OCP_B4	SGPIO_OCP_V3_2_DATA_IN	
OCP_B5	SGPIO_OCP_V3_2_DATA_OUT	
OCP_B6	SGPIO_OCP_V3_2_CLK	
OCP_A4	OCP_V3_2_ISO_RBT_ARB_IN	
OCP_A5	OCP_V3_2_ISO_RBT_ARB_OUT	
OCP_A6	OCP_V3_2_SLOT_ID1	
OCP_A7	NCSI_OCP_V3_2_ISO_TX_EN	
OCP_A8	NCSI_OCP_V3_2_ISO_TXD1	
OCP_A9	NCSI_OCP_V3_2_ISO_TXD0	
OCP_B7	OCP_V3_2_SLOT_ID0	
OCP_B8	NCSI_OCP_V3_2_ISO_RXD1	
OCP_B9	NCSI_OCP_V3_2_ISO_RXD0	
OCP_B14	NCSI_OCP_V3_2_CRS_R_DV	
OCP_A14	CLK_50M_OCP_V3_1_ISO_RBT_REF	
B10	RST_OCP_V3_2_A_0_N	
A11	RST_OCP_V3_2_A_1_N	
OCP_A1	RST_OCP_V3_2_A_2_N	
OCP_A2	RST_OCP_V3_2_A_3_N	
OCP_A3	OCP_V3_2_WAKE_N	
B42	OCP_V3_2_PRSNTB0_N	
A42	OCP_V3_2_PRSNTB1_N	
A12	OCP_V3_2_PRSNTB2_N	
B70	OCP_V3_2_PRSNTB3_N	
B7	OCP_V3_2_ISO_BIF0_N	
B8	OCP_V3_2_ISO_BIF1_N	
B9	OCP_V3_2_ISO_BIF2_N	
A70	OCP_V3_2_PWRBRK_N	
A7	SMB_BMC_OCP_V3_2_A_SCL	
A8	SMB_BMC_OCP_V3_2_A_SDA	
A9	SMB_BMC_OCP_V3_2_A_RST_N	
OCP_B11,OCP_B12	CLK_100M_DB2001_OCP_V3_2_A_DP/N	100Mhz PCIe reference clock
OCP_A11,OCP_A12	CLK_100M_DB2001_OCP_V3_2_B_DP/N	100Mhz PCIe reference clock
B15,B14	CLK_100M_DB2001_OCP_V3_2_C_DP/N	100Mhz PCIe reference clock
A15,A14	CLK_100M_DB2001_OCP_V3_2_D_DP/N	100Mhz PCIe reference clock
A68,A69	USB2_OCP_2_DP/N	USB 2.0 interface
A17,A18	P3E_CPU1_PCIE3_RX_DP/N<0>	High Speed Rx Lane 0
A20,A21	P3E_CPU1_PCIE3_RX_DP/N<1>	High Speed Rx Lane 1
A23,A24	P3E_CPU1_PCIE3_RX_DP/N<2>	High Speed Rx Lane 2
A26,A27	P3E_CPU1_PCIE3_RX_DP/N<3>	High Speed Rx Lane 3
A30,A31	P3E_CPU1_PCIE3_RX_DP/N<4>	High Speed Rx Lane 4

A33,A34	P3E_CPU1_PCIE3_RX_DP/N<5>	High Speed Rx Lane 5
A36,A37	P3E_CPU1_PCIE3_RX_DP/N<6>	High Speed Rx Lane 6
A39,A40	P3E_CPU1_PCIE3_RX_DP/N<7>	High Speed Rx Lane 7
A44,A45	P3E_CPU1_PCIE3_RX_DP/N<8>	High Speed Rx Lane 8
A47,A48	P3E_CPU1_PCIE3_RX_DP/N<9>	High Speed Rx Lane 9
A50,A51	P3E_CPU1_PCIE3_RX_DP/N<10>	High Speed Rx Lane 10
A53,A54	P3E_CPU1_PCIE3_RX_DP/N<11>	High Speed Rx Lane 11
A56,A57	P3E_CPU1_PCIE3_RX_DP/N<12>	High Speed Rx Lane 12
A59,A60	P3E_CPU1_PCIE3_RX_DP/N<13>	High Speed Rx Lane 13
A62,A63	P3E_CPU1_PCIE3_RX_DP/N<14>	High Speed Rx Lane 14
A65,A66	P3E_CPU1_PCIE3_RX_DP/N<15>	High Speed Rx Lane 15
B17,B18	P3E_CPU1_PCIE3_TX_C_DP/N<0>	High Speed Tx Lane 0
B20,B21	P3E_CPU1_PCIE3_TX_C_DP/N<1>	High Speed Tx Lane 1
B23,B24	P3E_CPU1_PCIE3_TX_C_DP/N<2>	High Speed Tx Lane 2
B26,B27	P3E_CPU1_PCIE3_TX_C_DP/N<3>	High Speed Tx Lane 3
B30,B31	P3E_CPU1_PCIE3_TX_C_DP/N<4>	High Speed Tx Lane 4
B33,B34	P3E_CPU1_PCIE3_TX_C_DP/N<5>	High Speed Tx Lane 5
B36,B37	P3E_CPU1_PCIE3_TX_C_DP/N<6>	High Speed Tx Lane 6
B39,B40	P3E_CPU1_PCIE3_TX_C_DP/N<7>	High Speed Tx Lane 7
B44,B45	P3E_CPU1_PCIE3_TX_C_DP/N<8>	High Speed Tx Lane 8
B47,B48	P3E_CPU1_PCIE3_TX_C_DP/N<9>	High Speed Tx Lane 9
B50,B51	P3E_CPU1_PCIE3_TX_C_DP/N<10>	High Speed Tx Lane 10
B53,B54	P3E_CPU1_PCIE3_TX_C_DP/N<11>	High Speed Tx Lane 11
B56,B57	P3E_CPU1_PCIE3_TX_C_DP/N<12>	High Speed Tx Lane 12
B59,B60	P3E_CPU1_PCIE3_TX_C_DP/N<13>	High Speed Tx Lane 13
B62,B63	P3E_CPU1_PCIE3_TX_C_DP/N<14>	High Speed Tx Lane 14
B65,B66	P3E_CPU1_PCIE3_TX_C_DP/N<15>	High Speed Tx Lane 15

AL motherboard supports three M.2 connectors. One slot from PCH supports boot drive. Two slots from CPU1 support data drive due to better PCIe bandwidth performance.

There's an additional x1 PCIe channel between PCH and BMC
The table below shows the connectivity:

Table 6.3-5: PCH to BMC

PCH Pin	PCH Pin Name	Net Name	BMC Pin	Description
BL66	USB3_9_PCIE2_TXP	P2E_PCH_BMC_TX_DP	M21	PCIe Gen3 x1 bus TX P
BK65	USB3_9_PCIE2_TXN	P2E_PCH_BMC_TX_DN	M22	PCIe Gen3 x1 bus TX N
AW69	USB3_9_PCIE2_RXP	P2E_PCH_BMC_RX_C_DP	L21	PCIe Gen3 x1 bus RX P
AW71	USB3_9_PCIE2_RXN	P2E_PCH_BMC_RX_C_DN	L22	PCIe Gen3 x1 bus RX N

6.3.3 Network Interface

6.3.3.1 Data

Each motherboard uses 2 * OCP NIC3.0 cards as its primary data network interface at I/O side. There is an option for a single port operation or dual port.

6.3.3.2 Management Network (OOB)

The motherboard has a few options of management network interface for BMC's connection. Management network shares data network's physical interface. Management connection should be independent from data traffic, and OS/driver condition.

- OOB to Angels Landing - Shared-NIC from Mezzanine 100G NIC to BMC through RMII/NC-SI. There are two OOB IP addresses, one on each MB. Each of them is allocated to NIC0 corresponding to the NIC attached to CPU0.
- In-Band to Angels Landing – 100G NIC to CPU through PCIe Gen3 x16
- OOB to Clear Creek - USB 1.1 connection between BMC on Angels Landing (Host) MB1 to BMC on Clear Creek (Device).
- OOB to Emerald pools - USB 1.1 connection between BMC on Angels Landing (Host) MB0 to BMC on Clear Creek (Device).

6.3.4 USB

The following figure shows the system USB connectivity.

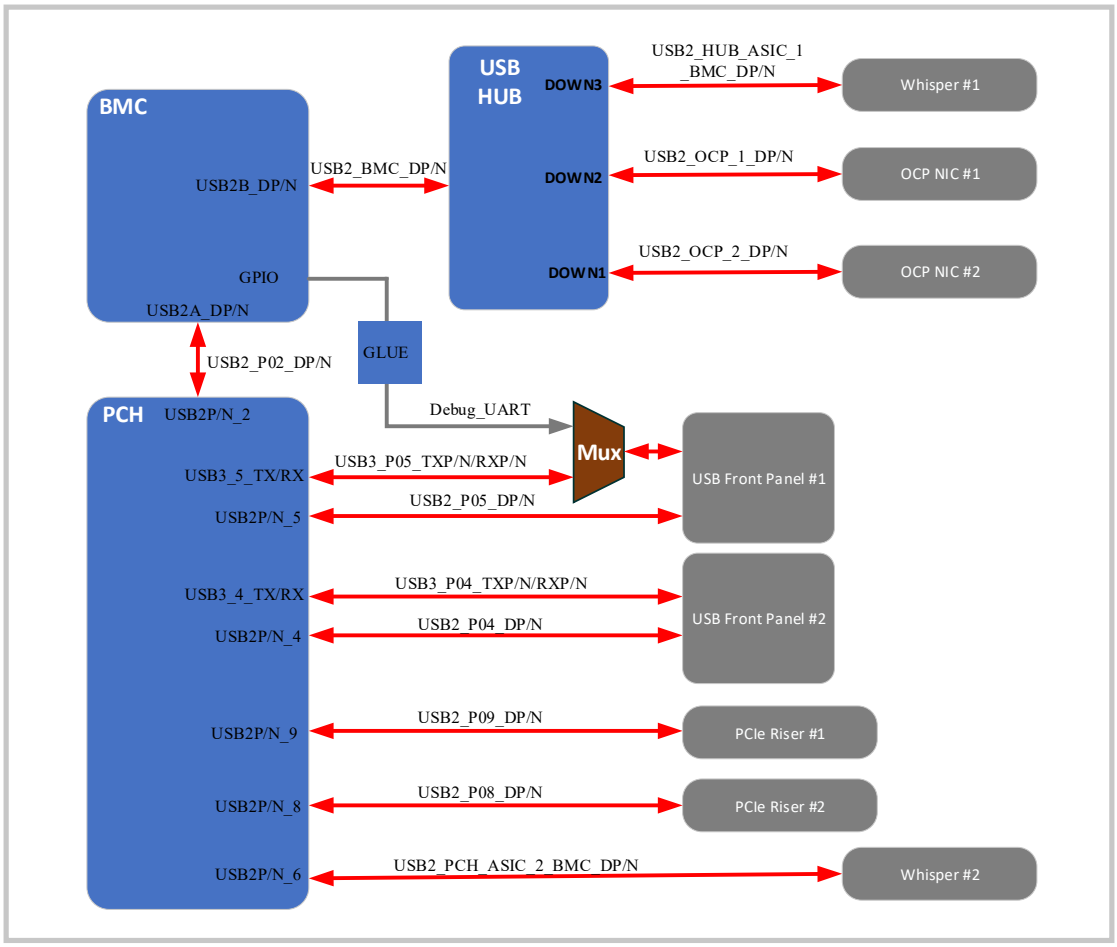


Figure 15: USB Connectivity

The following tables show PCH USB connections to front panel USB connectors:

Table 6.3-6: PCH to Front Panel USB #1

PCH Pin	PCH Pin Name	Net Name	Ext. Connector	Description
BJ59	USB3_5_TXP	USB3_P05_TXP	9	USB3 TXP
BK58	USB3_5_TXN	USB3_P05_TXN	8	USB3 TXN
BE69	USB3_5_RXP	USB3_P05_RXP	6	USB3 RXP
BE71	USB3_5_RXN	USB3_P05_RXN	5	USB3 RXN
BV58	USB2P_5	USB2_P05_DP	3	USB2 DP
BY58	USB2N_5	USB2_P05_DN	2	USB2 DN

Table 6.3-7: PCH to Front Panel USB #2

PCH Pin	PCH Pin Name	Net Name	Ext. Connector	Description
BG56	USB3_4_TXP	USB3_P04_TXP	9	USB3 TXP
BH58	USB3_4_TXN	USB3_P04_TXN	8	USB3 TXN
BF70	USB3_4_RXP	USB3_P04_RXP	6	USB3 RXP
BF72	USB3_4_RXN	USB3_P04_RXN	5	USB3 RXN
BV62	USB2P_4	USB2_P04_DP	3	USB2 DP
BY62	USB2N_4	USB2_P04_DN	2	USB2 DN

The following tables show PCH USB2 connections to PCIe x16 slots

PCH Pin	PCH Pin Name	Net Name	CN35 Connector	Description
BV55	USB2P_9	USB2_P09_DP	A32	USB2 DP
BY55	USB2N_9	USB2_P09_DN	A33	USB2 DN

PCH Pin	PCH Pin Name	Net Name	CN33 Connector	Description
BV64	USB2P_8	USB2_P08_DP	A32	USB2 DP
BY64	USB2N_8	USB2_P08_DN	A33	USB2 DN

The following table shows the USB2 connection between the PCH and the BMC:

PCH Pin	PCH Pin Name	Net Name	BMC	Description
BW61	USB2P_2	USB2_PCH_P02_DP	A7	USB2 DP
CA61	USB2N_2	USB2_PCH_P02_DN	A8	USB2 DN

The following table shows the USB2 connection between the BMC and Emerald Pools through on-board USB hub:

BMC Pin	BMC Pin Name	Net Name	Whisper J47 Pin	Description
B6	USB2B_DP	USB2_BMC_DN	D26	USB2 through Hub to EP (N)
A6	USB2B_DN	USB2_BMC_DP	D27	USB2 through Hub to EP (P)

6.3.5 OOB connectivity

Each AL has its dedicated Baseboard Management Chip (BMC). Only the BMC on the Master MB functions as Master BMC and the other is secondary BMC.

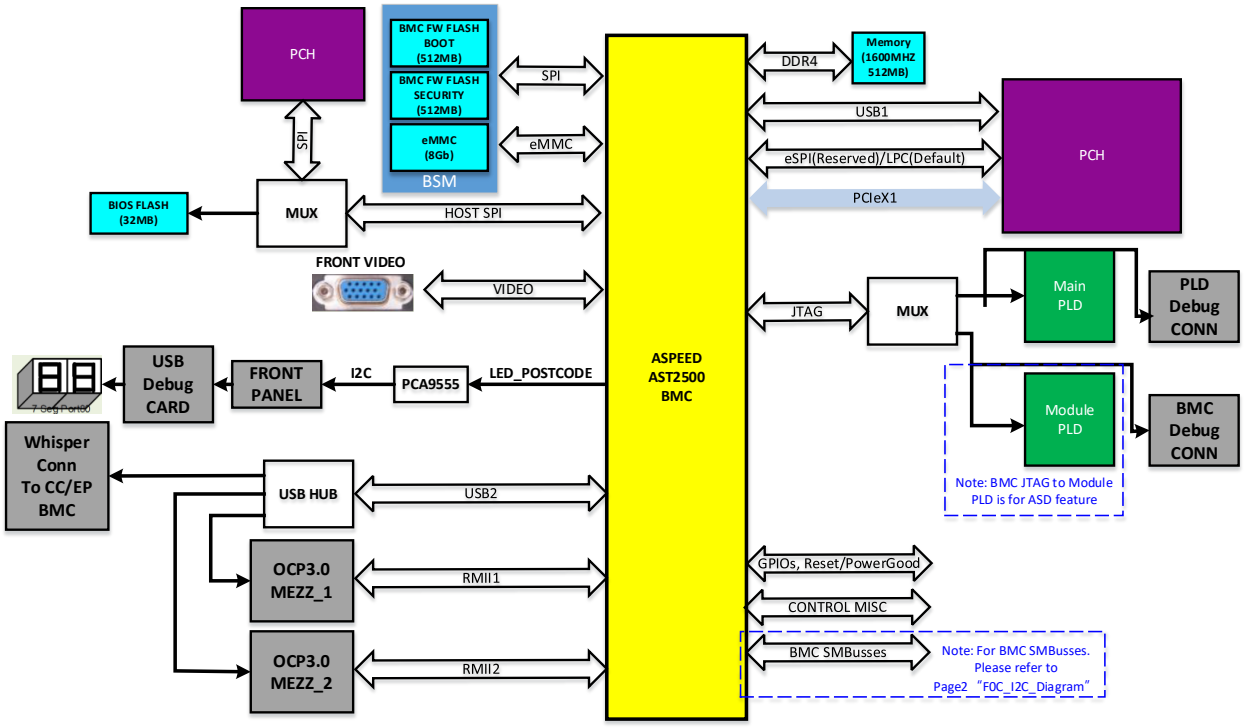


Figure 16: AL BMC connections

As mentioned before, each BMC has an NCSI interface connected to the NIC on the same MB. Each BMC on the board will have its own IP addresses. Emerald Pools' BMC chip has physical USB connections to each BMC chip on AL but only the master BMC has the data flow to EP's BMC. The OOB connection between AL and EP is presented in Figure 17. IP addresses will be allocated to EP BMC through the USB interface. Ethernet-over-USB is the main data path for hosts to access EP's BMC. The I2C interface is reserved to configure the EP BMC. The USB interface connected to PCH is reserved to update the EP BMC image through inband.

We setup Aspeed 2500 series chip's USB interface in USB1.1 speed to provide more stable link between the two BMCs. That port is USB2.0 capable.

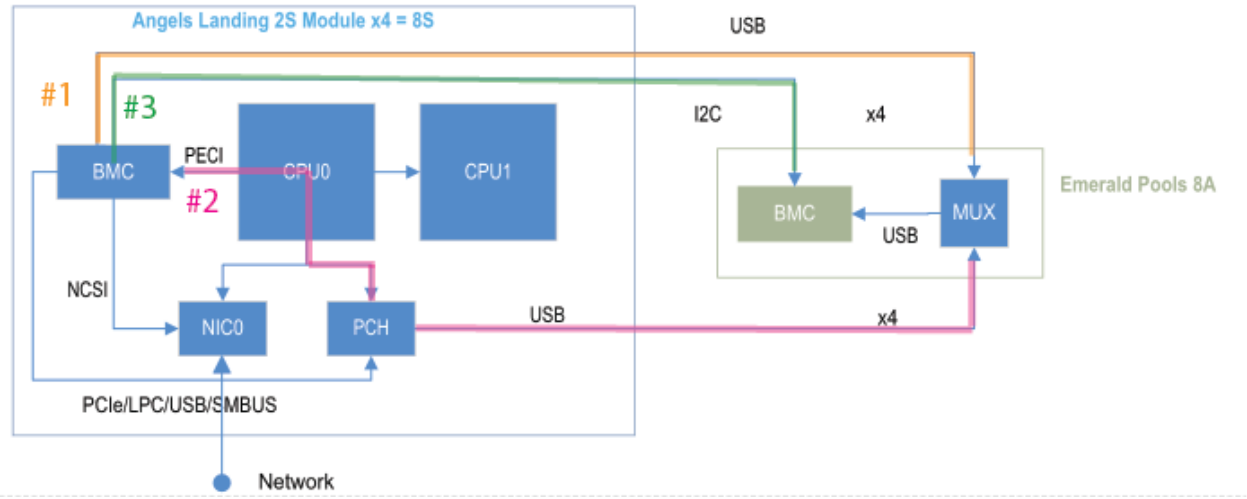


Figure 17: BMC connection between AL and EP

6.4 System Configuration

Angels Landing 4S system has 2*2S MB modules and each module has its own independent BMC. A platform BIC resides on the PDB board and is connected to each 2S module through I2C bus. These 3 microcontrollers are responsible for the system level configuration.

Each 2S module board connects to the CM (snowflake) through I2C bus. CM will be powered on first, then send IDs to configure CPU socket IDs, PCH configuration, BMC IDs and CPLD IDs. CPLD and CPU socket IDs signals will be driven from a GPIO expander PCA9555.

Default configuration sequence:

1. System AC ON CM will read the default configuration from its EEPROM
2. CM enables each 2s-MB HSC
3. P3V3_STBY_CPLD is on while P3V3_STBY is off CPU is in S5
4. CM sets through I2C the SOCKET_ID and BMC_ID to each 2S-MB according to the ID configuration table for each system mode. For default it'll set CASE #1 configuration
5. CM verifies all IDs are okay and issue ID_Okay status to CPLD
6. CPLD starts power up sequence

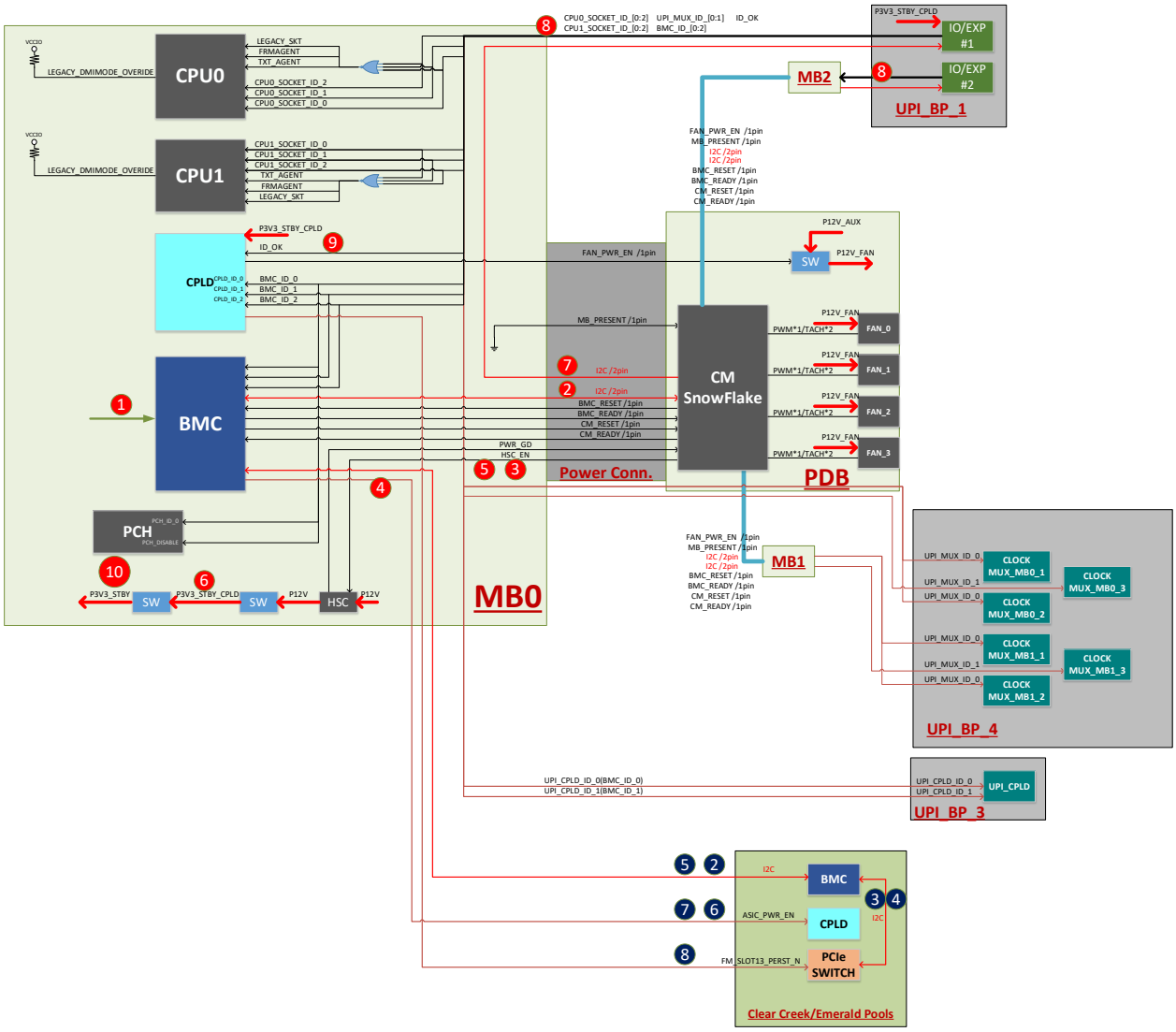


Figure 18: Module and Board Configuration IDs

Figure 18 shows detailed connection diagram which is required for the configuration

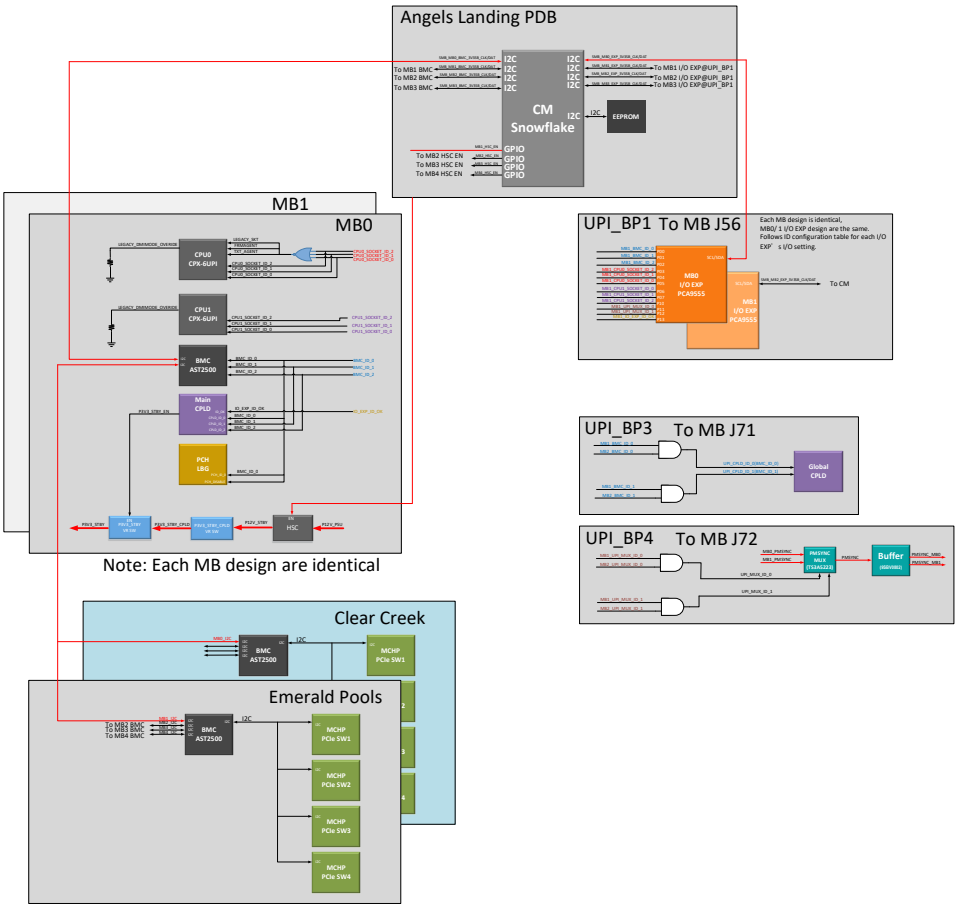


Figure 19: Zion System Level Configuration

6.5 Other configurations

As mentioned before, Angels Landing can be easily configured with our modular design concept. Two CPU motherboards have identical design. There are no hardware dependencies in between so we can configure the system by chassis manager ID pins on PDB.

Figure 20 shows the ID configuration required for each configuration.

Nodes Config	Board#	ID Pins Setting From UPI BP_1's I/O EXPs										
		IO EXP										
		FM_BMC_SKT_ID_0	FM_BMC_SKT_ID_1	FM_BMC_SKT_ID_2	CPU0_SKT_ID_LVC3_0	CPU0_SKT_ID_LVC3_1	CPU0_SKT_ID_LVC3_2	CPU1_SKT_ID_LVC3_0	CPU1_SKT_ID_LVC3_1	CPU1_SKT_ID_LVC3_2	FM_UPI_MUX_ID_0	FM_UPI_MUX_ID_1
2S_M	MB1	0	0	1	0	0	0	1	0	0	0	1
2S_M	MB0	0	0	1	0	0	0	1	0	0	1	1
		IO EXP										
		FM_BMC_SKT_ID_0	FM_BMC_SKT_ID_1	FM_BMC_SKT_ID_2	CPU0_SKT_ID_LVC3_0	CPU0_SKT_ID_LVC3_1	CPU0_SKT_ID_LVC3_2	CPU1_SKT_ID_LVC3_0	CPU1_SKT_ID_LVC3_1	CPU1_SKT_ID_LVC3_2	FM_UPI_MUX_ID_0	FM_UPI_MUX_ID_1
4S_S	MB1	1	1	0	0	1	0	1	1	0	1	1
4S_M	MB0	0	1	0	0	0	0	1	0	0	1	1
		IO EXP										
		FM_BMC_SKT_ID_0	FM_BMC_SKT_ID_1	FM_BMC_SKT_ID_2	CPU0_SKT_ID_LVC3_0	CPU0_SKT_ID_LVC3_1	CPU0_SKT_ID_LVC3_2	CPU1_SKT_ID_LVC3_0	CPU1_SKT_ID_LVC3_1	CPU1_SKT_ID_LVC3_2	FM_UPI_MUX_ID_0	FM_UPI_MUX_ID_1
4S_M	MB1	0	1	0	0	0	0	1	0	0	1	1
4S_S	MB0	1	1	0	0	1	0	1	1	0	1	1

Figure 20: ID Configuration table

The following explains the configurations

1. 4S mode. It is the default Angels Landing setup in the Zion System.

2. 2S x2 mode. Angels landing box can be configured as two individual 2-socket servers. Users can also decide if they just need a 1x 2-socket server by removing the MB1 motherboard. This mode is useful in the production line to update the fw in both MB0 and MB1 even if the final configuration is 4S mode.

It is also possible to swap which MB is MB0 in case there are failures on MB0. However Angels landing system doesn't support degraded mode here. Once CPU or DIMM failures are found in the system, we don't request the system keep running and processing workloads. The system should stop and wait for debugging and service.

7. Angels Landing 4S Sub-Systems Specifications

This section covers the details of each component in the AL system.

7.1 2S Motherboard Configuration

The following figure shows the main blocks of the 2S-MB. The MB is comprised of the following blocks

- CPU
  - Two Intel® Cooper Lake-SP CPU connected through UPI interconnect
  - UPI interconnect to another tray forming up to an 4S socket system
- DDR4 Memory Up to 3200MHz with 24x DDR4 DIMM slots (in a 2DPC config)
- I/O
  - Two OCP NIC3.0 slots
  - Two PCIe Gen3 x16 PCIe external ports for connection to Emerald Pools or other slave server application
  - Two USB3 external connections
  - OOB: two NCSI to BMC connections from NIC3.0 slots
- Storage
  - One 2280 M.2 512GB boot drive
  - Up to two 22110 M.2 2TB data drives
- BMC
  - ASpeed AST2500
  - 512MB DDR4 memory at 1.6GT/S

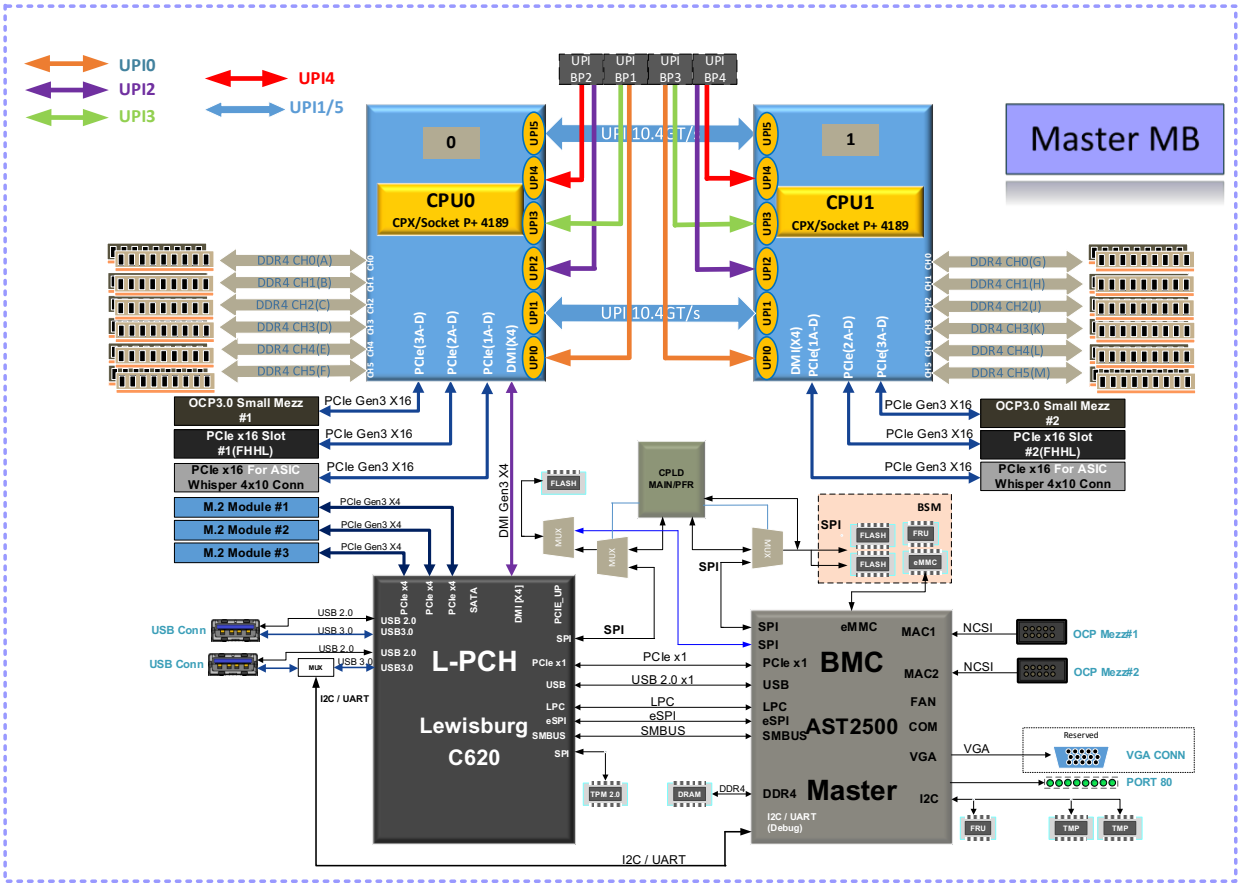


Figure 21: 2S-MB block diagram

In an 4S system both 2S-MBs have identical BOM. Among the two 2S module motherboards, one board is the main MB, its PCH is the only PCH active in this configuration, other MB is secondary, this configuration is done through configuration IDs.

7.2 CPU

Each motherboard supports all Intel® CPX-6S processor family processor SKUs with TDP up to 205W. The motherboard shall provision the support of all future CPUs in Intel® CPX-6S processor Family Platform unless noted otherwise.

The features listed below must be supported by the motherboard:

- Support two Intel® CPX-6S processor family processors up to 205W TDP, and vendors should engage with Intel® to ensure the design is ready for future processors
- Six full-width Intel® UPI links up to 10.4 GT/s/direction for Intel® CPX-6S processor family processor
- Up to 26 cores per CPU (up to 2 threads with Intel® Hyper-Threading Technology).
- In ZionEx configuration we are using Cooperlake-SP,24 Core, 8339HC 1.8GHz 125W TDP.

7.2.1 DIMM

Each motherboard has DIMM subsystem designed as below:

- DDR4 direct-attach memory support on CPU0 and CPU1 on each motherboard.
- SMT DDR4-DIMM slot connectors

[Yellow] Foxconn_ AH58897-Q9Y10-1H

[Black] Foxconn_ AH58897-Q9B10-1H

- 6x channels DDR4 registered memory interface on each CPU
- 2x DDR4 slots per channel, total 24x DIMM per motherboard
- Support DDR4 speeds up to 3200MHz @2DPC
- Support RDIMM. System does not request LRDIMM support.
- Support SR, DR, QR and 8R DIMM
- AL DIMM configuration uses 32GB DDR4 DIMM @3200MT/S @2DPC. Total memory capacity is 1536GB
- Follow updated JEDEC DDR4 specification with 288 pin DIMM socket

7.2.2 Intel® PCH

The motherboard uses Intel® Lewisburg C620 PCH chipset, which supports following features:

- 2x USB 3.0 type A ports for front panel access
- 5x USB2.0 ports: one for BMC in-band firmware update; one each to X16 riser connector; one type A usb2.0 in usb3.0 connector for OCP LCD debug card; one Type A 3.0
- 1x PCIe x4 ports to M.2 connectors
- SPI interface, mux with BMC to enable BMC the capability to perform BIOS upgrade and recovery
- SPI interface for TPM header
- SMBUS interface (master & slave)
- Intel® Server Platform Services 4.0 Firmware (SPS) with Intel® Node Manager
    - PECI access to CPU
    - SMLink0 connect to BMC
    - Intel® Manageability Engine (ME) obtain HSC PMBus related information directly. Intel® Manageability Engine SMLink1 connects to Hotswap controller PMBus interface by default. BMC has connection to HSC PMBus to have flexibility of HSC PMBus related feature support, default is disconnected
    - Power capping capability
    - Temperature sensors reading from BMC
- PCH SKUs
    - Board design supports all PCH SKUs in terms of power delivery and thermal design

Noted that only master PCH is working in this configuration. The PCH on the slave MBs will be disabled.

7.2.3 PCH SPS Firmware

Intel® provides the Server Platform Software (SPS) as a FW stack, there won't be usage of Innovation Engine (IE).

7.2.4 Intel® UPI Connection

Intel® UPI is a low-latency coherent interconnect for scalable multiprocessor systems with a shared address space. It uses a directory-based home snoop coherency protocol with a transfer speed of up to 10.4 GT/s.

7.3 Storage

AL only supports M.2 form factors for storage. MB supports three M.2 solid-state drive in 2280 and 22110 form factors:

- One 2280 M.2 for Boot Drive
- Two 22110 M.2 for Data Drives

Only the master boot driver is functional in 4S mode since the boot driver is connected to the PCIe interface through PCH. Data drivers are connected to the CPU so all data drivers will be accessible in 4S mode.

7.4 EEPROM

Each 2S MB includes a 128Kbits I2C-accessible Electrically Erasable Programmable Read-Only Memory (EEPROM). The EEPROM is accessible from the platform via the BMC. The EEPROM contains the Field Replaceable Unit Identification (FRU ID) information and any additional configuration information that may be required. The FRU ID is formatted in accordance with the IPMI Platform Management FRU Information Storage Definition document.

The EEPROM must contain the following entries:

- Board Manufacturer
- Board Name
- Board Serial Number
- Board Part Number
- Product Manufacturer
- Product Name
- Product Part Number
- Product Serial Number
- Product Asset Tag
- Product Build: e.g. EVT, DVT, PVT, MP
- Product Version: e.g. C1
- Manufacturing Date and Time
- Manufacturing Lot Code: (preferred, but optional)
- Manufacturing Work Order: (preferred, but optional)
- PCB Revision
- CPU Model Name/Number
- CPU Revision
- CPU TjMAX (Maximum Junction Temperature)

8. BIOS

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

8.1 BIOS Chip

The BIOS chip should use PCH's SPI interface through BMC controlled MUX for BMC to perform offline BIOS update or recovery. The vendor is responsible for selecting a specific BIOS chip, which should fit the required functionality in addition to potential additional features that may be required in the future. 64 MB size is recommended considering space needed for both BIOS and Intel® Manageability Engine firmware as well as PFR and other features. Vendors should provide a flash region plan for different code and current used size for each region to justify the sizing of the SPI flash.

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

In 4S configuration, only the bios in the main motherboard is active. The bios in the secondary motherboard is not active. However, we still suggest users to maintain the same bios version in both chips.

8.2 BIOS Source Code

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

8.3 BIOS Feature Requirements

8.3.1 Optimization

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

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

- SPEC power should be used as guidance by ODM to validate BIOS tuning results.

8.3.2 Setup Menu

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

- Settings for adjusting memory speed, UPI speed, Speed-step/Turbo mode and CPU C-state power state: The default follows the CPU and chipset vendor's POR unless otherwise mentioned.
- Settings to enable different Turbo mode tuning settings based on CPU SKU and memory configuration: The default is Turbo enabled with the CPU vendor's POR, unless otherwise mentioned.
- Setting for the power feature after AC failure: The default is set to restore the last power state.
- Setting for the local physical COM port (COM0) and SOL (COM1): The default is enabled console redirection on both ports with baud rate 57600, no flow control, terminal type VT100, 8 data bits, No Parity, 1 Stop Bit.
- Setting for legacy console redirection to be local physical COM port (COM0) and SOL(COM1). The default is SOL(COM1)
- Setting for the altitude of the server deployment location: The default is 300M.
- Setting for the watchdog timer: The default setting for EVT/DVT/PVT is disabled. The default setting for MP is enabled. The timeout value is 15 minutes and reset the system after the timer expires. The watchdog timer is always disabled after POST.
- Setting for ECC error threshold: Available settings are 1, 4, 10 and 1000. The default setting is 1 for EVT, DVT, and PVT and 1000 for MP.
- Setting for ECC error event log threshold: Available settings are disabled, 10, 50, 100. The default setting is 10.
- If a CMOS CRC error happens, the BIOS should load the system default automatically and log the CMOS clear event in SEL.
- The default setting to disable all "wait for keyboard input to continue" types of features is "not to wait for keyboard input to continue".
- Calculate checksum of BIOS setting, display in BIOS setup menu, and output to SMBIOS table.
- Setting to save and load 10 different sets of user default.
- Setting of UEFI and Legacy boot options: The default is UEFI.
- Display SKU and hardware revision in the main page based on BOARD ID and FAB ID.
- Setting of PPIN (Protected Processor Inventory Number) Control: The default setting is unlock/enable.
- Display RC version in main page.
- Display CPU information in the main page including CPU signature, processor cores, and microcode patch version.
- Display memory information in the main page including current memory speed, total memory capacity and type (DDR4).
- Display PCH information in the main page including name and stepping.
- Setting of Setup Prompt timeout: The default is "7 seconds".

8.3.3 Boot Options

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

- 1st: USB device if available
- 2nd: Mezzanine card NIC IPv6
- 3rd: Mezzanine card NIC IPv4
- 4rd: PCIe M.2
- 5th: Reserved
- 6th: Reserved
- 7th: Reserved
- 8th: Reserved

If there is no bootable device found, BIOS should keep searching for a bootable device. BIOS should support UEFI and legacy boot mode options and default is UEFI. UEFI and legacy boot mode have an independent boot loop. Boot mode and boot order can be displayed and changed from BMC with OEM command.

8.3.4 Board SKU ID

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

Board SKU ID0 is to identify MB SKU:
0= Master Board, 1= Slave Board

Board SKU ID2:1 is to identify VR Solution:
0,0 = Texas Instrument , 0,1 = Infineon, 1,0 = RESERVE

SKU ID	SKU_ID5	SKU_ID4	SKU_ID3	SKU_ID[2:1]	SKU_ID0
0	RESERVE	RESERVE	LBF-1G	0 0: Texas Instrument	RESERVE
1	RESERVE	RESERVE	LBG-R	0 1: Infineon	RESERVE

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

8.3.5 FAB Revision ID

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

FAB_ID[2:0]			Description
0	0	0	FAB1 (POC)

0	0	1	FAB2 (EVT)
0	1	0	FAB3 (DVT)
0	1	1	FAB4 (PVT)
1	0	0	FAB5 (MP)

8.3.6 Remote BIOS Update Requirement

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

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

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

- Update from the Operating System through ssh; the current OS based is CentOS 8.1 64-bit with updated Kernel specified by customer
- Require no more than one reset cycle to the system to complete BIOS update or BIOS setup option change
- Require no physical access to system
- BIOS update or BIOS setup option change should not take more than 5 minutes to complete
- BIOS update procedure can be scripted and propagated to multiple machines

Once Angels Landing is in 4S mode. The PCH on MB1 will work in non-legacy mode. The BIOS is not in active mode on MB1. However, we suggest the BIOS version on two MBs are identical for consistency. The update tool shall support OOB update since inband update won't work in MB1.

8.3.7 Event log requirement

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

8.3.8 BIOS Error Code Requirement

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

Table 8.3-1 BIOS Error Code

Fatal Errors	Major Code	Minor Code	Error Description
ERR_NO_MEMORY	0E8h		
ERR_NO_MEMORY_MINOR_NO_MEMORY		01h	1. No memory was detected via SPD read. No warning log entries available. 2. Invalid configuration that causes no operable memory. Refer to warning log entries for details.
ERR_NO_MEMORY_MINOR_ALL_CH_DISABLED		02h	Memory on all channels of all sockets is disabled due to hardware memtest error.
ERR_NO_MEMORY_MINOR_ALL_CH_DISABLED_MIXED		03h	No memory installed. All channels are disabled.
ERR_LT_LOCK	0E9h		Memory is locked by LT, inaccessible.
ERR_DDR_INIT	0EAh		DDR training did complete successfully
ERR_RD_DQ_DQS		01h	Error on read DQ/DQS init
ERR_RC_EN		02h	Error on Receive Enable
ERR_WR_LEVEL		03h	Error on Write Leveling
ERR_WR_DQ_DQS		04h	Error on write DQ/DQS
ERR_MEM_TEST	0EBh		Memory test failure
ERR_MEM_TEST_MINOR_SOFTWARE		01h	Software memtest failure
ERR_MEM_TEST_MINOR_HARDWARE		02h	Hardware memtest failure
ERR_MEM_TEST_MINOR_LOCKSTEP_MODE		03h	Hardware memtest failure in Lockstep channel mode requiring a channel to be disabled. This is a fatal error which requires a reset and calling BIOS with a different RAS mode to retry
ERR_VENDOR_SPECIFIC	0ECh		
ERR_DIMM_COMPAT	0EDh		RDIMMs is present DIMM vendor-specific errors
ERR_MIXED_MEM_TYPE		01h	Different DIMM types are detected installed in the system

ERR_INVALID_POP		02h	Violation of population rules
ERR_INVALID_POP_MINOR_UNSUPPORTED_VOLTAGE		05h	Unsupported DIMM Voltage
Reserved	0EEh		Reserved
ERR_INVALID_BOOT_MODE		01h	Boot mode is unknown
ERR_INVALID_SUB_BOOT_MODE		02h	Sub boot mode is unknown

8.3.9 POST Code for Option ROM Entry and Exit

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

8.3.10 PPIN BIOS Implementation

BIOS shall support PPIN and set default to [Unlock and Enable].
There are two ways for user to access PPIN:

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

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

- Serves SMI# signal from BMC and use Send_PPIN OEM command to communicate PPIN to BMC.
- Use Send_PPIN OEM command to communicate PPIN to BMC when BIOS POST COMPLETE.

8.4 OEM Command

NetFn	Command	Priority	Comments
Chassis	Get Chassis Status (0x01)	P1	BIOS to read the current power restore policy configuration
	Set Power Restore Policy (0x06)	P1	BIOS to set power restore policy configuration
	Get System Restart Cause(0x07)	P1	BIOS to know why the system got restarted e.g. user command vs. power button vs. power policy vs. WDT
	Get Boot Options(0x09)	P4	Boot Order sequence;CMOS settings clearance
App	Get Device ID (0x01)	P0	Generic info purpose

	Cold Reset (0x02)	P0	Reset BMC from Host
	Get Selftest Results (0x04)	P0	BIOS decides weather to install IPMI protocol or not
	Manufacturing Test On (0x05)	P1	Allowed Sled-cycle via KCS
	Get Device GUID (0x08)	P0	To read unique UID for device
	Reset WDT(0x22)	P1	WDT commands are used by FRB2 timer
	Set WDT(0x24)	P1	
	Get WDT(0x25)	P1	
	Set BMC Global Enables (0x2e)	P1	Used by ipmi_si driver
	Get Global Enables (0x2f)	P0	Used by BIOS for deciding to send SEL events to BMC or not
	Clear Message Flags (0x30)	P1	Used by ipmi_si driver
	Get System GUID (0x37)	P0	To read unique UID for System
	Set System Info Params(0x58)	P0	Inform BIOS version information to BMC
	Get System Info Params(0x59)	P0	Read BIOS version info from Host
Storage	Get FRUID Info(0x10)	P0	FRUID info is needed by BIOS to populate SMBIOS tables and used by dmideode command for provisioning
	Get FRUID Data (0x11)	P0	
	Get SEL Info (0x40)	P0	SEL commands are needed by BIOS to log critical events that are found during POST/runtime e.g. PCIe, DIMM ECC etc.
	Reserve SEL (0x42)	P0	
	Get SEL Entry (0x43)	P0	
	Add SEL Entry (0x44)	P0	

	Clear SEL (0x47)	P0	
	Get SEL UTC Offset(0x5C)	P0	
Transport	Get LAN Config (0x02)	P0	Needed by software agent on Host to read IP address of BMC; Supports only three parameters: "Lan Address Enables", "read IPv4", and "read IPv6 address"
	Get SOL Config (0x22)	P0	Needed by BIOS to output console data on both ports during bootup
OEM (0x30)	Set DIMM Info (0x1C)	P0	(OPTIONAL based on BIOS need) Inform BMC about DIMM info Request: Byte 1 – DIMM Index, 1 based Byte 2 – DIMM Type [7:6] Voltage type 00: Normal voltage(1.5V) 01: Ultra Low Voltage DIMM(1.25V) 10: Low voltage(1.35V) 11: DDR4 Normal voltage(1.35V) [5:0] DIMM type 0x00: SDRAM 0x01: DDR-1 RAM 0x02: Rambus 0x03: DDR-2 RAM 0x04: FBDIMM 0x05: DDR-3 RAM 0x06: DDR-4 RAM 0x3F – No DIMM present Byte 3..4 – DIMM speed in MHz, LS-byte first Byte 5..8 – DIMM size in Mbytes, LS-byte first Response: Byte 1 – Completion Code

	Get Board ID(0x37)	P0	(OPTIONAL based on BIOS need) Get Board Information from BMC Request: None Response: Byte 1 – Completion Code Byte2 - Platform ID Byte3 - Board Revision ID Note1:PlatformID SKU[0]:RESERVE SKU[2:1] 00: TI 01: Infineon 10: VR 3RD SKU[3] LBG-R SKU[4] RESERVE SKU[5] RESERVE Note2: Board ID Rev2 Rev1 Rev0 POC: 000 EVT: 001 DVT: 010 PVT: 011 MP: 100
	Set POST Start (0x73)	P0	Inform BMC when POST begins Request: None Response: Byte1 - Completion Code
	Set POST End (0x74)	P0	Inform BMC when POST Ends Request: None Response: Byte1 - Completion Code

	Set PPIN(0x77)	P1	Inform BMC about PPIN data of 8 bytes for each CPU Request: Byte 1:8 – CPU0 PPIN data Optional: Byte 9:16 – CPU1 PPIN data Byte 17:24 – CPU2 PPIN data Byte 25:32 – CPU3 PPIN data Response: Byte 1 – Completion Code
	Set PPR Parameter (0x90)	P2	DDR4 PPR/sPPR support;Need to repair DDR4 memory by using extra row as part of repair flow Request: Byte 1 - PPR Command Selector Byte 2..N - Configuration parameter data per table below Response: Byte 1 - Completion Code Command: PPR Action Selector: 0x01 Byte 1 - Enable or disable PPR function in next host system reboot Bit[7]: 0 for disable;1 for enable Bit[6..0] - 01h for Soft PPR; 02h for Hard PPR If there is no candidate row for repair: When get, response data is zero for PPR function When Set, response completion code: D5h for parameter not support in current state Command: PPR Candidate Row Count Selector: 0x2 Byte1 - PPR Candidate row count, range from 0 to 100 0h - No candidate row Command: PPR Candidate Row Address Selector: 0x3

			Byte1 - Set Selector: PPR candidate row index, from 0 to ("PPR Candidate Row Count" - 1). Byte 2..7 - PPR candidate row address Byte 2: Bit[7:2] - Reserved Bit[1:0] - Logical Rank Byte 3: Bit[7:5] - Socket Number Bit[4:2] - Channel Number Bit[0:1] - DIMM Number Byte 4: Device ID Byte 5: Bit[7:4] - Bank Bit[3:0] - Bank Group Byte 6: Row (LSB) Byte 7: Row (MSB) Command: PPR History Data Selector: 0x04 Byte1: Set Selector: PPR history index, from 0 to ("PPR History Count" - 1) Byte 2..17: PPR history data Byte 2..5 - PPR Table Signature 0xbfeeeefb Byte 6 - Checksum Byte 7..10 - Seconds since Epoch for First PPR Entry Byte 11..14 - Seconds since Epoch for the Most Recent Entry Byte 15 - Number of Successful PPR (LSB) Byte 16 - Number of Successful PPR (MSB) Byte 17 - DIMM ID

	Get PPR Parameter (0x91)	P2	DDR4 PPR/sPPR support;Need to repair DDR4 memory by using extra row as part of repair flow Request: Byte 1 - PPR Command Selector Byte 2 - Set Selector. Selects a given set of parameters under a parameter selector value. 00h if parameter does not require a set selector Response: Byte 1 - Completion Code Byte 2..N - Configuration parameter data per table below Command: PPR Action Selector: 0x01 Byte 1 - Enable or disable PPR function in next host system reboot Bit[7]: 0 for disable;1 for enable Bit[6..0] - 01h for Soft PPR; 02h for Hard PPR If there is no candidate row for repair: When get, response data is zero for PPR function When Set, response completion code: D5h for parameter not support in current state Command: PPR Candidate Row Count Selector: 0x2 Byte1 - PPR Candidate row count, range from 0 to 100 0h - No candidate row Command: PPR Candidate Row Address Selector: 0x3 Byte1 - Set Selector: PPR candidate row index, from 0 to ("PPR Candidate Row Count" - 1). Byte 2..7 - PPR candidate row address Byte 2: Bit[7:2] - Reserved Bit[1:0] - Logical Rank Byte 3: Bit[7:5] - Socket Number Bit[4:2] - Channel Number Bit[0:1] - DIMM Number Byte 4: Device ID

			Byte 5: Bit[7:4] - Bank Bit[3:0] - Bank Group Byte 6: Row (LSB) Byte 7: Row (MSB) Command: PPR History Data Selector: 0x04 Byte1: Set Selector: PPR history index, from 0 to ("PPR History Count" - 1) Byte 2..17: PPR history data Byte 2..5 - PPR Table Signature 0xbfeeeefb Byte 6 - Checksum Byte 7..10 - Seconds since Epoch for First PPR Entry Byte 11..14 - Seconds since Epoch for the Most Recent Entry Byte 15 - Number of Successful PPR (LSB) Byte 16 - Number of Successful PPR (MSB) Byte 17 - DIMM ID

	Write Machine Configuration Information (0x6A)	P1	BIOS to inform BMC about System configuration: #of CPUs, #of DIMMs etc..to help in calculation of Airflow Request: Byte 1 - Chassis Type 00h - ORv1 01h - ORv2 (Tioga pass) FFh - Unknown Byte 2 - Motherboard Type 00h - SS 01h - DS 02h - Type3 03h:FEh - Reserved FFh - Unknown Byte 3 - Processor Count Byte 4 - Memory Count Byte 5 - 3.5" HDD Count (FFh - Unknown) (0 or 1 for Tioga Pass) Byte 6 - 2.5" HDD Count (FFh - Unknown) (0 in tioga pass) Byte 7 - Riser Type 00h - Not Installed 01h - 2-Slot Riser Card 02h - 3-Slot Riser Card FFh - Unknown Byte 8 - PCIe Card Location Bit 0 - SLOT1 (1h - Present; 0h - Absent) Bit 1 - SLOT2 (1h - Present; 0h - Absent) Bit 2 - SLOT3 (1h - Present; 0h - Absent) Bit 3 - SLOT4 (1h -Present; 0h - Absent) Byte 9 - SLOT 1 00h - Absent // Always return 0 since it is an OCP card. Byte 10 – SLOT 2 00h - Absent 01h - AVA (1 x m.2) 02h - AVA (2 x m.2) 03h - AVA (3 x m.2) 04h - AVA (4 x m.2) 05h - Re-timer 06h - HBA 07h - Other flash cards (Intel, HGST) 80h - Unknown

			Byte 11 – SLOT 3 Same as Byte 10 Byte 12 – SLOT 4 Same as Byte 10 Byte 13 – AEP Memory count Note: Bit 0 - presence of C1 [OCP Mezzanine Card] Bit 1 - presence of C2 [bottom card in 2(or 3)-slot riser] Bit 2 - presence of C3 [top card in 2-slot riser or middle slot in 3-slot riser] Bit 3 - presence of C4 [top card in 3-slot riser] Response: Byte 1 - Completion Code 00h - Normal C7h - Request data exceeds beyond data length C9h - Invalid data type of data length specified Bit 3 - Card 4 (1h -Present; 0h - Absent)

	Set BIOS Boot Order (0x52h)	P0	BIOS to set boot order to BMC Request: Byte 1– Boot mode Bit 0 – 0 : Legacy, 1 : UEFI Bit 1 – CMOS clear (Optional, BIOS implementation dependent) Bit 2 - Force boot bios setup (Optional, BIOS implementation dependent) Bit 6:3 – reserved Bit 7 – boot flags valid Byte 2-6– Boot sequence Bit 2:0 – boot device id 000b: USB device 001b: Network 010b: SATA HDD 011b: SATA-CDROM 100b: Other removable Device Bit 7:3 – reserve for boot device special request If Bit 2:0 is 001b (Network), Bit3 is IPv4/IPv6 order Bit3=0b: IPv4 first Bit3=1b: IPv6 first Response: Byte1 - Completion Code

	Get BIOS Boot Order (0x53h)	P0	BIOS to get boot order from BMC Request: None Response: Byte 1 – Completion Code Byte 2– Boot mode Bit 0 – 0 : Legacy, 1 : UEFI Bit 1 – CMOS clear (Optional, BIOS implementation dependent) Bit 6:2 – reserved Bit 7 – boot flags valid Byte 3-7– Boot sequence Bit 2:0 – boot device id 000b: USB device 001b: Network 010b: SATA HDD 011b: SATA-CDROM 100b: Other removable Device Bit 7:3 – reserve for boot device special request If Bit 2:0 is 001b (Network), Bit3 is IPv4/IPv6 order Bit3=0b: IPv4 first Bit3=1b: IPv6 first

	Get FLASH Info (0x55h)	P2	Request: None Response: Byte 1 – Completion Code Byte 2 – Flash Manufacturer ID (*Note1) Byte 3:4 – Flash Device ID (*Note1) Byte 5:6 – Flash Status Register (*Note2) *Note1: MXID mx25l51245GMI-08G Manufacturer id is C2 MXID mx25l51245GMI-08G device id is 0x20 0x1A *Note2: MXID mx25l51245GMI-08G status register: BIT7 BIT6 BIT5 BIT4 BIT3 BIT2 BIT1 BIT0 SRWD QE BP3 BP2 BP1 BP0 WEL WIP

OEM (0x36)	Set Processor Information (0x10)	P2	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZ h, LSB first Byte 4 – Processor Index, 0 base Byte 5 – Parameter Selector Byte 6..N – Configuration parameter data (see below for Parameters of Processor Information) **Response:** Byte 1 – Completion code **Parameter#1: (Processor Product Name)** Byte 1..48 –Product name(ASCII code) Ex. Intel(R) Xeon(R) CPU E5-2685 v3 @ 2.60GHz **Param#2: Processor Basic Information** Byte 1 – Core Number Byte 2 – Thread Number (LSB) Byte 3 – Thread Number (MSB) Byte 4 – Processor frequency in MHz (LSB) Byte 5 – Processor frequency in MHz (MSB) Byte 6..7 – Revision

	Get Processor Information (0x11)	P2	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZ h, LSB first Byte 4 – Processor Index, 0 base Byte 5 – Parameter Selector **Response:** Byte 1 – Completion code Byte 2..N – Configuration Parameter Data (see below for Parameters of Processor Information) **Parameter#1: (Processor Product Name)** Byte 1..48 –Product name(ASCII code) Ex. Intel(R) Xeon(R) CPU E5-2685 v3 @ 2.60GHz **Param#2: Processor Basic Information** Byte 1 – Core Number Byte 2 – Thread Number (LSB) Byte 3 – Thread Number (MSB) Byte 4 – Processor frequency in MHz (LSB) Byte 5 – Processor frequency in MHz (MSB) Byte 6..7 – Revision

	Set DIMM Information (0x12)	P2	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZh, LSB first Byte 4 – DIMM Index, 0 base Byte 5 – Parameter Selector Byte 6..N – Configuration parameter data (see below for Parameters of DIMM Information) **Response:** Byte 1 – Completion code **Param#1 (DIMM Location):** Byte 1 – DIMM Present Byte 1 – DIMM Present 01h – Present FFh – Not Present Byte 2 – Node Number, 0 base Byte 3 – Channel Number , 0 base Byte 4 – DIMM Number , 0 base **Param#2 (DIMM Type):** Byte 1 – DIMM Type Bit [7:6] For DDR3 00 – Normal Voltage (1.5V) 01 – Ultra Low Voltage (1.25V) 10 – Low Voltage (1.35V) 11 – Reserved For DDR4 00 – Reserved 01 – Reserved 10 – Reserved 11 – Normal Voltage (1.2V) Bit [5:0] 0x00 – SDRAM 0x01 – DDR-1 RAM 0x02 – Rambus 0x03 – DDR-2 RAM 0x04 – FBDIMM 0x05 –DDR-3 RAM 0x06 –DDR-4 RAM **Param#3 (DIMM Speed):** Byte 1..2 – DIMM speed in MHz, LSB Byte 3..6 – DIMM size in Mbytes, LSB

			Param#4 (Module Part Number): Byte 1..20 –Module Part Number (JEDEC Standard No. 21-C) **Param#5 (Module Serial Number):** Byte 1..4 –Module Serial Number (JEDEC Standard No. 21-C) **Param#6 (Module Manufacturer ID):** Byte 1 - Module Manufacturer ID, LSB Byte 2 - Module Manufacturer ID, MSB

	Get DIMM Information (0x13)	P2	**Request:** Byte 1:3 – Manufacturer ID – XXYYZZh, LSB first Byte 4 –DIMM Index, 0 base Byte 5 – Parameter Selector **Response:** Byte 1 – Completion code Byte 2..N – Configuration Parameter Data (see Table_1213h Parameters of DIMM Information) **Param#1 (DIMM Location):** Byte 1 – DIMM Present Byte 1 – DIMM Present 01h – Present FFh – Not Present Byte 2 – Node Number, 0 base Byte 3 – Channel Number , 0 base Byte 4 – DIMM Number , 0 base **Param#2 (DIMM Type):** Byte 1 – DIMM Type Bit [7:6] For DDR3 00 – Normal Voltage (1.5V) 01 – Ultra Low Voltage (1.25V) 10 – Low Voltage (1.35V) 11 – Reserved For DDR4 00 – Reserved 01 – Reserved 10 – Reserved 11 – Normal Voltage (1.2V) Bit [5:0] 0x00 – SDRAM 0x01 – DDR-1 RAM 0x02 – Rambus 0x03 – DDR-2 RAM 0x04 – FBDIMM 0x05 –DDR-3 RAM 0x06 –DDR-4 RAM **Param#3 (DIMM Speed):** Byte 1..2 – DIMM speed in MHz, LSB Byte 3..6 – DIMM size in Mbytes, LSB

			Param#4 (Module Part Number): Byte 1..20 –Module Part Number (JEDEC Standard No. 21-C) **Param#5 (Module Serial Number):** Byte 1..4 –Module Serial Number (JEDEC Standard No. 21-C) **Param#6 (Module Manufacturer ID):** Byte 1 - Module Manufacturer ID, LSB Byte 2 - Module Manufacturer ID, MSB

	Set HDD Information (0x14)		BIOS issue this command to provide HDD information to BMC. BIOS just can get information by standard ATA / SMART command for OB SATA controller. BIOS can get 1. Serial Number 2. Model Name 3. HDD FW Version 4. HDD Capacity 5. HDD WWN Use Get HDD info Param #5 to know the MAX HDD info index. Request: Byte 1:3 – Quanta Manufacturer ID – 001C4Ch, LSB first Byte 4 – [7:4] Reserved [3:0] HDD Controller Type 0x00 – BIOS 0x01 – Expander 0x02 – LSI Byte 5 – HDD Info Index, 0 base Byte 6 – Parameter Selector Byte 7..N – Configuration parameter data (see Table_1415h Parameters of HDD Information) Response: Byte 1 – Completion Code Param#0 (HDD Location): Byte 1 – Controller [7:3] Device Number [2:0] Function Number For Intel C610 series (Wellsburg) D31:F2 (0xFA) – SATA control 1 D31:F5 (0xFD) – SATA control 2 D17:F4 (0x8C) – sSata control Byte 2 – Port Number Byte 3 – Location (0xFF: No HDD Present) BIOS default set Byte 3 to 0xFF, if No HDD Present. And then skip send param #1~4, #6, #7 to BMC

			(still send param #5) BIOS default set Byte 3 to 0, if the HDD present. BMC or other people who know the HDD location has responsibility for update Location info Param#1 (Serial Number): Bytes 1..33: HDD Serial Number Param#2 (Model Name): Byte 1..33 – HDD Model Name Param#3 (HDD FW Version): Byte 1..17 –HDD FW version Param#4 (Capacity): Byte 1..4 –HDD Block Size, LSB Byte 5..12 - HDD Block Number, LSB HDD Capacity = HDD Block size * HDD BLock number (Unit Byte) Param#5 (Max HDD Quantity): Byte 1 - Max HDD Quantity Max supported port numbers in this PCH Param#6 (HDD Type) Byte 1 – HDD Type 0h – Reserved 1h – SAS 2h – SATA 3h – PCIE SSD (NVME) Param#7 (HDD WWN) Data 1...8: HDD World Wide Name, LSB

	Get HDD Information (0x15)		BMC needs to check HDD presented or not first. If NOT presented, return completion code 0xD5. Request: Byte 1:3 – Quanta Manufacturer ID – 001C4Ch, LSB first Byte 4 – [7:4] Reserved [3:0] HDD Controller Type 0x00 – BIOS 0x01 – Expander 0x02 – LSI Byte 5 – HDD Index, 0 base Byte 6 – Parameter Selector (See Above Set HDD Information) Response: Byte 1 – Completion Code 0xD5 – Not support in current status (HDD Not Present) Byte 2..N – Configuration parameter data (see Table_1415h Parameters of HDD Information)

9. BMC

Angels Landing uses an ASPEED® AST2500 BMC with one x16 4Gb DDR4 DRAM running at 1.6MT/S for various platform management services and interfaces with hardware, BIOS, Intel® Server Platform Services firmware.

BMC should be a standalone system working side by side with the host. The health status of the host system should not affect the normal operation and network connectivity of BMC. BMC cannot share memory with the host system. BMC management connectivity should work independently from host, and has no NIC driver dependency for Out-of-Band communication if using a shared-NIC.

Angels Landing 4S system total has 2* BMC chips, one on each motherboard. Each motherboard has 2*NICs. BMC's two NCSI ports should connect to the two NIC slots, and request BMC IP from the NIC0 which connects to CPU0 on the motherboard.

Main BMC is using USB and I2C to communicate with other BMCs. The USB port shall be configured at full speed only due to the limitation of Aspeed 2500 chips. High speed mode shall be disabled in BMC driver.

9.1 BMC Connection

9.1.1 Local Serial Console

UART access to BMC (also known as Local Serial Console) is to be implemented with the following requirements

- The local serial console is to be functional in all stages of system operation. During booting, POST codes sent to the 7-segment display should also be available on the Local Serial Console to allow console logging and debugging. For example, codes like “[00]”, “[01]”, “[02]”, “[E0]”, etc, should be displayed on the console upon power up.
- The last POST code received upon entering the normal BMC console is to be displayed in the lower right-hand corner of the console screen.
- The serial console buffer for the last five console screens are to be stored in volatile media; each screen is of standard 80 columns by 24 rows. Screens are to be completely erased within five seconds of standby power removal for security and privacy reasons.
- The OOB raw command is to have the ability to extract and display what is shown in the console screen buffer for remote debugging purposes and can be used for scaling data collection. This is done using a buffer dumping script with OEM commands.

9.1.2 Remote Power Control and Power Policy

The implemented BMC should have the following capabilities:

- Support remote power on/off/cycle and warm reboot either through In-Band or Out-of-Band IPMI command.
- Support power on policy to be last-state, always-on and always-off upon recovery from an AC power loss event. The default setting is to be always-on. Power policy changes can be done through IPMI command and take effect without a cold reset or system reboot.
- Take no more than three seconds upon AC power on for BMC to power up and display system POST on debug support ports and local serial console.
- Take no more than three seconds to be able to process a power button signal upon AC power on.

9.2 Sensors

Various sensors are deployed around in order to support BMC functions like power/thermal monitoring, throttling, fan-speed control. Sensors are classified as Analog, Event-only type sensors.

9.2.1 Analog Sensors

The BMC is to have access to all analog sensors placed in the system and ensure that they are displayed in a sensor data record repository. The tables below list different types of analog sensors and their set limits around the Angels Landing system.

Temperature Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability
MB_PCH_TEMP	0x08	10	N/A	100	DC
NIC_MEZZ0_TEMP	0x10	10	N/A	95	Standby
NIC_MEZZ1_TEMP	0x11	10	N/A	95	Standby
MB_BOOT_DRIVER_TEMP	0x12	10	N/A	70	DC
MB_DATA0_DRIVER_TEMP	0x13	10	N/A	70	DC
MB_DATA1_DRIVER_TEMP	0x14	10	N/A	70	DC
PDB_HSC_TEMP	0x22	10	N/A	85	Standby
MB_CPU0_TJMAX	0x28	10	N/A	95	DC
MB_CPU1_TJMAX	0x29	10	N/A	95	DC
MB_CPU2_TJMAX	0x2A	10	N/A	95	DC
MB_CPU3_TJMAX	0x2B	10	N/A	95	DC
MB_CPU0_THERM_MARGIN	0x38	-81	N/A	-5	DC
MB_CPU1_THERM_MARGIN	0x39	-81	N/A	-5	DC
MB_CPU2_THERM_MARGIN	0x3A	-81	N/A	-5	DC
MB_CPU3_THERM_MARGIN	0x3B	-81	N/A	-5	DC
MB_HSC_TEMP	0x43	10	N/A	85	Standby
MB1_HSC_TEMP	0x47	10	N/A	85	Standby
MB_CPU0_DIMM_A_TEMP	0x70	10	N/A	85	DC
MB_CPU0_DIMM_B_TEMP	0x71	10	N/A	85	DC
MB_CPU0_DIMM_C_TEMP	0x72	10	N/A	85	DC
MB_CPU0_DIMM_D_TEMP	0x73	10	N/A	85	DC
MB_CPU0_DIMM_E_TEMP	0x74	10	N/A	85	DC
MB_CPU0_DIMM_F_TEMP	0x75	10	N/A	85	DC
MB_CPU1_DIMM_A_TEMP	0x76	10	N/A	85	DC
MB_CPU1_DIMM_B_TEMP	0x77	10	N/A	85	DC
MB_CPU1_DIMM_C_TEMP	0x78	10	N/A	85	DC
MB_CPU1_DIMM_D_TEMP	0x79	10	N/A	85	DC
MB_CPU1_DIMM_E_TEMP	0x7A	10	N/A	85	DC
MB_CPU1_DIMM_F_TEMP	0x7B	10	N/A	85	DC
MB_CPU2_DIMM_A_TEMP	0x7C	10	N/A	85	DC
MB_CPU2_DIMM_B_TEMP	0x7D	10	N/A	85	DC

MB_CPU2_DIMM_C_TEMP	0x7E	10	N/A	85	DC
MB_CPU2_DIMM_D_TEMP	0x7F	10	N/A	85	DC
MB_CPU2_DIMM_E_TEMP	0x80	10	N/A	85	DC
MB_CPU2_DIMM_F_TEMP	0x81	10	N/A	85	DC
MB_CPU3_DIMM_A_TEMP	0x82	10	N/A	85	DC
MB_CPU3_DIMM_B_TEMP	0x83	10	N/A	85	DC
MB_CPU3_DIMM_C_TEMP	0x84	10	N/A	85	DC
MB_CPU3_DIMM_D_TEMP	0x85	10	N/A	85	DC
MB_CPU3_DIMM_E_TEMP	0x86	10	N/A	85	DC
MB_CPU3_DIMM_F_TEMP	0x87	10	N/A	85	DC
MB_INLET_TEMP	0xA0	10	N/A	50	Standby
MB_OUTLET_TEMP_R	0xA1	10	N/A	70	Standby
MB_OUTLET_TEMP_L	0xA2	10	N/A	70	Standby
MB_INLET_REMOTE_TEMP	0xA3	10	N/A	55	Standby
MB_OUTLET_REMOTE_TEMP	0xA4	10	N/A	75	Standby
MB_CPU0_TEMP	0xA8	10	N/A	94	DC
MB_CPU1_TEMP	0xA9	10	N/A	94	DC
MB_CPU2_TEMP	0xAA	10	N/A	94	DC
MB_CPU3_TEMP	0xAB	10	N/A	94	DC
MB_VR_CPU0_VCCIN_TEMP	0xB1	10	N/A	115	DC
MB_VR_CPU0_VSA_TEMP	0xB5	10	N/A	115	DC
MB_VR_CPU0_VCCIO_TEMP	0xB9	10	N/A	115	DC
MB_VR_CPU0_VDDQ_ABC_TEMP	0xBD	10	N/A	115	DC
MB_VR_CPU0_VDDQ_DEF_TEMP	0xC1	10	N/A	115	DC
MB_VR_PCH_P1V05_TEMP	0xC5	10	N/A	115	DC
MB_VR_PCH_PVNN_TEMP	0xC9	10	N/A	115	DC
MB_VR_CPU1_VCCIN_TEMP	0xE1	10	N/A	115	DC
MB_VR_CPU1_VSA_TEMP	0xE5	10	N/A	115	DC
MB_VR_CPU1_VCCIO_TEMP	0xE9	10	N/A	115	DC
MB_VR_CPU1_VDDQ_ABC_TEMP	0xED	10	N/A	115	DC
MB_VR_CPU1_VDDQ_DEF_TEMP	0xF1	10	N/A	115	DC

Voltage Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability
PDB_HSC_VIN	0x20	10.8	N/A	13.2	Standby
PDB_P12V	0x26	10.8	N/A	13.2	DC
PDB_P3V	0x27	3.135	N/A	3.465	DC
MB_HSC_VIN	0x40	10.8	N/A	13.2	Standby
MB1_HSC_VIN	0x44	10.8	N/A	13.2	Standby
MB_P12V_P3V3_STBY_VOL	0x5C	10.8	N/A	13.2	Standby
MB_P3V3_M2_1_INA260_VOL	0x5D	3.135	N/A	3.465	DC
MB_P3V3_M2_2_INA260_VOL	0x5E	3.135	N/A	3.465	DC
MB_P3V3_M2_3_INA260_VOL	0x5F	3.135	N/A	3.465	DC
PDB_FAN0_VOLT	0x68	10.8	N/A	13.2	standby
PDB_FAN1_VOLT	0x69	10.8	N/A	13.2	standby
PDB_FAN2_VOLT	0x6A	10.8	N/A	13.2	standby
PDB_FAN3_VOLT	0x6B	10.8	N/A	13.2	standby
MB_VR_CPU0_VCCIN_VOUT	0xB0	1.33	N/A	2.1	DC
MB_VR_CPU0_VSA_VOUT	0xB4	0.437	N/A	1.25	DC
MB_VR_CPU0_VCCIO_VOUT	0xB8	0.852	N/A	1.25	DC
MB_VR_CPU0_VDDQ_ABC_VOUT	0xBC	1.13	N/A	1.4	DC
MB_VR_CPU0_VDDQ_DEF_VOUT	0xC0	1.13	N/A	1.4	DC
MB_VR_PCH_P1V05_VOUT	0xC4	1.0085	N/A	1.2	Standby
MB_VR_PCH_PVNN_VOUT	0xC8	0.80175	N/A	1.2	Standby
MB_P5V	0xD0	4.75	N/A	5.25	DC
MB_P5V_STBY	0xD1	4.75	N/A	5.25	Standby
MB_P3V3_STBY	0xD2	3.135	N/A	3.465	Standby
MB_P3V3	0xD3	3.135	N/A	3.465	DC
MB_P3V3_BAT	0xD4	2.6	N/A	3.4	Standby
MB_CPU_P1V8	0xD5	1.71	1.8	1.89	DC
MB_PCH_P1V8	0xD6	1.71	1.8	1.89	DC
MB_CPU0_PVPP_ABC	0xD7	2.38	2.575	2.75	DC
MB_CPU1_PVPP_ABC	0xD8	2.38	2.575	2.75	DC
MB_CPU0_PVPP_DEF	0xD9	2.38	2.575	2.75	DC
MB_CPU1_PVPP_ABC	0xDA	2.38	2.575	2.75	DC

MB_CPU0_PVTT_ABC	0xDB	0.547	0.62	0.662	DC
MB_CPU1_PVTT_ABC	0xDC	0.547	0.62	0.662	DC
MB_CPU0_PVTT_DEF	0xDD	0.547	0.62	0.662	DC
MB_CPU1_PVTT_ABC	0xDE	0.547	0.62	0.662	DC
MB_VR_CPU1_VCCIN_VOUT	0xE0	1.33	N/A	2.1	DC
MB_VR_CPU1_VSA_VOUT	0xE4	0.437	N/A	1.25	DC
MB_VR_CPU1_VCCIO_VOUT	0xE8	0.852	N/A	1.25	DC
MB_VR_CPU1_VDDQ_ABC_VOUT	0xEC	1.13	N/A	1.4	DC
MB_VR_CPU1_VDDQ_DEF_VOUT	0xF0	1.13	N/A	1.4	DC

Power Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability
PDB_HSC_PIN	0x23	N/A	N/A	N/A	Standby
PDB_HSC_PEAK_IOUT	0x24	N/A	N/A	N/A	Standby
PDB_HSC_PEAK_PIN	0x25	N/A	N/A	N/A	Standby
MB_CPU0_PKG_POWER	0x30	0	N/A	208	DC
MB_CPU1_PKG_POWER	0x31	0	N/A	208	DC
MB_CPU2_PKG_POWER	0x32	0	N/A	208	DC
MB_CPU3_PKG_POWER	0x33	0	N/A	208	DC
MB_HSC_PIN	0x42	N/A	N/A	N/A	Standby
MB1_HSC_PIN	0x46	N/A	N/A	N/A	Standby
MB_VR_CPU0_VCCIN_POUT	0xB3	N/A	N/A	N/A	DC
MB_VR_CPU0_VSA_POUT	0xB7	N/A	N/A	N/A	DC
MB_VR_CPU0_VCCIO_POUT	0xBB	N/A	N/A	N/A	DC
MB_VR_CPU0_VDDQ_ABC_POUT	0xBF	N/A	N/A	N/A	DC
MB_VR_CPU0_VDDQ_DEF_POUT	0xC3	N/A	N/A	N/A	DC
MB_VR_PCH_P1V05_POUT	0xC7	N/A	N/A	N/A	Standby
MB_VR_PCH_PVNN_POUT	0xCB	N/A	N/A	N/A	Standby
MB_VR_CPU1_VCCIN_POUT	0xE3	N/A	N/A	N/A	DC
MB_VR_CPU1_VSA_POUT	0xE7	N/A	N/A	N/A	DC
MB_VR_CPU1_VCCIO_POUT	0xEB	N/A	N/A	N/A	DC

MB_VR_CPU1_VDDQ_ABC_POUT	0xEF	N/A	N/A	N/A	DC
MB_VR_CPU1_VDDQ_DEF_POUT	0xF3	N/A	N/A	N/A	DC

Current Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability
PDB_HSC_IOUT	0x21	N/A	N/A	N/A	NA
PDB_HSC_PEAK_IOUT	0x24	N/A	N/A	N/A	NA
MB_HSC_IOUT	0x41	N/A	N/A	N/A	Standby
MB1_HSC_IOUT	0x45	N/A	N/A	N/A	Standby
PDB_FAN0_CURR	0x6C	N/A	N/A	N/A	Standby
PDB_FAN1_CURR	0x6D	N/A	N/A	N/A	Standby
PDB_FAN2_CURR	0x6E	N/A	N/A	N/A	Standby
PDB_FAN3_CURR	0x6F	N/A	N/A	N/A	Standby
MB_VR_CPU0_VCCIN_IOUT	0xB2	N/A	N/A	N/A	DC
MB_VR_CPU0_VSA_IOUT	0xB6	N/A	N/A	N/A	DC
MB_VR_CPU0_VCCIO_IOUT	0xBA	N/A	N/A	N/A	DC
MB_VR_CPU0_VDDQ_ABC_IOUT	0xBE	N/A	N/A	N/A	DC
MB_VR_CPU0_VDDQ_DEF_IOUT	0xC2	N/A	N/A	N/A	DC
MB_VR_PCH_P1V05_IOUT	0xC6	N/A	N/A	N/A	Standby
MB_VR_PCH_PVNN_IOUT	0xCA	N/A	N/A	N/A	Standby
MB_VR_CPU1_VCCIN_IOUT	0xE2	N/A	N/A	N/A	DC
MB_VR_CPU1_VSA_IOUT	0xE6	N/A	N/A	N/A	DC
MB_VR_CPU1_VCCIO_IOUT	0xEA	N/A	N/A	N/A	DC
MB_VR_CPU1_VDDQ_ABC_IOUT	0xEE	N/A	N/A	N/A	DC
MB_VR_CPU1_VDDQ_DEF_IOUT	0xF2	N/A	N/A	N/A	DC

Fan Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability

PDB_FAN0_INLET_SPEED	0x60	1200	N/A	11990	Standby
PDB_FAN1_INLET_SPEED	0x61	1200	N/A	11990	Standby
PDB_FAN2_INLET_SPEED	0x62	1200	N/A	11990	Standby
PDB_FAN3_INLET_SPEED	0x63	1200	N/A	11990	Standby
PDB_FAN0_OUTLET_SPEED	0x64	1200	N/A	10800	Standby
PDB_FAN1_OUTLET_SPEED	0x65	1200	N/A	10800	Standby
PDB_FAN2_OUTLET_SPEED	0x66	1200	N/A	10800	Standby
PDB_FAN3_OUTLET_SPEED	0x67	1200	N/A	10800	Standby

Miscellaneous Sensors

Miscellaneous Sensor Name	Sensor Number	Lower Critical Threshold	Upper Non-Critical Threshold	Upper Critical Threshold	Availability
MB_PROCESSOR_FAIL	0XDF	N/A	N/A	N/A	NA

9.2.2 Event Sensor

Threshold sensor list

Sensor Descrption	Sensor Name	Sensor#	LCR	UNC	UCR	Reading Available
PCH Temp	AL_MB_PCH_TEMP	0x8	10	NA	100	DC
NIC0 Temp	AL_NIC_MEZZ0_TEMP	0x10	10	NA	95	Standby
NIC0 Temp	AL_NIC_MEZZ1_TEMP	0x11	10	NA	95	Standby
Data0 Driver Temp	AL_MB_DATA0_DRIVE_TEMP	0x13	10	NA	70	DC
Data1 Driver Temp	AL_MB_DATA1_DRIVE_TEMP	0x14	10	NA	70	DC
PDB Hsc Vin	AL_PDB_HSC_VIN	0x20	10.8	NA	13.2	Standby
PDB Hsc Iout	AL_PDB_HSC_IOUT	0x21	NA	NA	36	Standby
PDB Hsc Temp	AL_PDB_HSC_TEMP	0x22	10	NA	85	Standby
PDB Hsc Pin	AL_PDB_HSC_PIN	0x23	NA	NA	NA	Standby
PDB Hsc Peak Iout	AL_PDB_HSC_PEAK_IOUT	0x24	NA	NA	NA	Standby
PDB Hsc Peak Pin	AL_PDB_HSC_PEAK_PIN	0x25	NA	NA	NA	Standby

PDB P12V	AL_PDB_P12V	0x26	10.8	NA	13.2	DC
PDB P3V	AL_PDB_P3V3	0x27	3.14	NA	3.47	DC
P0 Tjmax	AL_MB_CPU0_TJMAX	0x28	NA	NA	NA	DC
P1 Tjmax	AL_MB_CPU1_TJMAX	0x29	NA	NA	NA	DC
P2 Tjmax	AL_MB_CPU2_TJMAX	0x2A	NA	NA	NA	DC
P3 Tjmax	AL_MB_CPU3_TJMAX	0x2B	NA	NA	NA	DC
P0 Package Power	AL_MB_CPU0_PKG_POWER	0x30	NA	NA	NA	DC
P1 Package Power	AL_MB_CPU1_PKG_POWER	0x31	NA	NA	NA	DC
P2 Package Power	AL_MB_CPU2_PKG_POWER	0x32	NA	NA	NA	DC
P3 Package Power	AL_MB_CPU3_PKG_POWER	0x33	NA	NA	NA	DC
P0 Therm Margin	AL_MB_CPU0_THERM_MARGIN	0x38	-81	NA	-5	DC
P1 Therm Margin	AL_MB_CPU1_THERM_MARGIN	0x39	-81	NA	-5	DC
P2 Therm Margin	AL_MB_CPU2_THERM_MARGIN	0x3A	-81	NA	-5	DC
P3 Therm Margin	AL_MB_CPU3_THERM_MARGIN	0x3B	-81	NA	-5	DC
HSC Input Voltage	AL_MB_HSC_VIN	0x40	10.8	NA	13.2	Standby
HSC Output Current	AL_MB_HSC_IOUT	0x41	NA	NA	100	Standby
HSC Input Power	AL_MB_HSC_PIN	0x42	NA	NA	NA	Standby
HSC Temperature	AL_MB_HSC_TEMP	0x43	10	NA	85	Standby
HSC1 Input Voltage	AL_MB1_HSC_VIN	0x44	10.8	NA	13.2	Standby
HSC1 Output Current	AL_MB1_HSC_IOUT	0x45	NA	NA	100	Standby
HSC1 Input Power	AL_MB1_HSC_PIN	0x46	NA	NA	NA	Standby
HSC1 Temperature	AL_MB1_HSC_TEMP	0x47	10	NA	85	Standby
HSC Peak Power IN	AL_MB_HSC_PEAK_PIN	0x50	NA	NA	NA	Standby
HSC1 Peak Power IN	AL_MB1_HSC_PEAK_PIN	0x51	NA	NA	NA	Standby
Fan0 Power		0x54	NA	NA	NA	Standby
Fan1 Power	AL_PDB_FAN1_POWER	0x55	NA	NA	NA	Standby
Fan2 Power	AL_PDB_FAN2_POWER	0x56	NA	NA	NA	Standby
Fan3 Power	AL_PDB_FAN3_POWER	0x57	NA	NA	NA	Standby

Ina260 P12V Stby	AL_MB_P12V_STBY_INA260_VOL	0x5C	10.8	NA	13.2	Standby
Ina260 P3v3 M2_1 Voltage	AL_MB_P3V3_M2_1_INA260_VOL	0x5D	3.14	NA	3.47	DC
Ina260 P3v3 M2_2 Voltage	AL_MB_P3V3_M2_2_INA260_VOL	0x5E	3.14	NA	3.47	DC
Ina260 P3v3 M2_3 Voltage	AL_MB_P3V3_M2_3_INA260_VOL	0x5F	3.14	NA	3.47	DC
Fan0 Inlet Speed	AL_PDB_FAN0_INLET_SPEED	0x60	600	NA	13500	standby
Fan1 Inlet Speed	AL_PDB_FAN1_INLET_SPEED	0x61	600	NA	13500	standby
Fan2 Inlet Speed	AL_PDB_FAN2_INLET_SPEED	0x62	600	NA	13500	standby
Fan3 Inlet Speed	AL_PDB_FAN3_INLET_SPEED	0x63	600	NA	13500	standby
Fan0 Outlet Speed	AL_PDB_FAN0_OUTLET_SPEED	0x64	600	NA	12500	standby
Fan1 Outlet Speed	AL_PDB_FAN1_OUTLET_SPEED	0x65	600	NA	12500	standby
Fan2 Outlet Speed	AL_PDB_FAN2_OUTLET_SPEED	0x66	600	NA	12500	standby
Fan3 Outlet Speed	AL_PDB_FAN3_OUTLET_SPEED	0x67	600	NA	12500	standby
Fan0 Voltage	AL_PDB_FAN0_VOLT	0x68	10.8	NA	13.2	standby
Fan1 Voltage	AL_PDB_FAN1_VOLT	0x69	10.8	NA	13.2	standby
Fan2 Voltage	AL_PDB_FAN2_VOLT	0x6A	10.8	NA	13.2	standby
Fan3 Voltage	AL_PDB_FAN3_VOLT	0x6B	10.8	NA	13.2	standby
Fan0 Current	AL_PDB_FAN0_CURR	0x6C	NA	NA	NA	standby
Fan1 Current	AL_PDB_FAN1_CURR	0x6D	NA	NA	NA	standby
Fan2 Current	AL_PDB_FAN2_CURR	0x6E	NA	NA	NA	standby
Fan3 Current	AL_PDB_FAN3_CURR	0x6F	NA	NA	NA	standby
P0 DIMM A Temp	AL_MB_CPU0_DIMM_A0_C0_TEMP	0x70	10	NA	85	DC
P0 DIMM B Temp	AL_MB_CPU0_DIMM_A1_C1_TEMP	0x71	10	NA	85	DC
P0 DIMM C Temp	AL_MB_CPU0_DIMM_A2_C2_TEMP	0x72	10	NA	85	DC
P0 DIMM D Temp	AL_MB_CPU0_DIMM_A3_C3_TEMP	0x73	10	NA	85	DC

P0 DIMM E Temp	AL_MB_CPU0_DIMM_A4_C4_TEMP	0x74	10	NA	85	DC
P0 DIMM F Temp	AL_MB_CPU0_DIMM_A5_C5_TEMP	0x75	10	NA	85	DC
P1 DIMM A Temp	AL_MB_CPU1_DIMM_B0_D0_TEMP	0x76	10	NA	85	DC
P1 DIMM B Temp	AL_MB_CPU1_DIMM_B1_D1_TEMP	0x77	10	NA	85	DC
P1 DIMM C Temp	AL_MB_CPU1_DIMM_B2_D2_TEMP	0x78	10	NA	85	DC
P1 DIMM D Temp	AL_MB_CPU1_DIMM_B3_D3_TEMP	0x79	10	NA	85	DC
P1 DIMM E Temp	AL_MB_CPU1_DIMM_B4_D4_TEMP	0x7A	10	NA	85	DC
P1 DIMM F Temp	AL_MB_CPU1_DIMM_B5_D5_TEMP	0x7B	10	NA	85	DC
P2 DIMM A Temp	AL_MB_CPU2_DIMM_A0_C0_TEMP	0x7C	10	NA	85	DC
P2 DIMM B Temp	AL_MB_CPU2_DIMM_A1_C1_TEMP	0x7D	10	NA	85	DC
P2 DIMM C Temp	AL_MB_CPU2_DIMM_A2_C2_TEMP	0x7E	10	NA	85	DC
P2 DIMM D Temp	AL_MB_CPU2_DIMM_A3_C3_TEMP	0x7F	10	NA	85	DC
P2 DIMM E Temp	AL_MB_CPU2_DIMM_A4_C4_TEMP	0x80	10	NA	85	DC
P2 DIMM F Temp	AL_MB_CPU2_DIMM_A5_C5_TEMP	0x81	10	NA	85	DC
P3 DIMM A Temp	AL_MB_CPU3_DIMM_B0_D0_TEMP	0x82	10	NA	85	DC
P3 DIMM B Temp	AL_MB_CPU3_DIMM_B1_D1_TEMP	0x83	10	NA	85	DC
P3 DIMM C Temp	AL_MB_CPU3_DIMM_B2_D2_TEMP	0x84	10	NA	85	DC
P3 DIMM D Temp	AL_MB_CPU3_DIMM_B3_D3_TEMP	0x85	10	NA	85	DC
P3 DIMM E Temp	AL_MB_CPU3_DIMM_B4_D4_TEMP	0x86	10	NA	85	DC
P3 DIMM F Temp	AL_MB_CPU3_DIMM_B5_D5_TEMP	0x87	10	NA	85	DC
Inlet Temp	AL_MB_INLET_TEMP	0xA0	10	NA	55	Standby
Outlet Temp Right	AL_MB_OUTLET_TEMP_R	0xA1	10	NA	75	Standby
Outlet Temp Left	AL_MB_OUTLET_TEMP_L	0xA2	10	NA	75	Standby

Inlet Remote Temp	AL_MB_INLET_REMOTE_TEMP	0xA3	10	NA	55	Standby
Outlet Right Remote Temp	AL_MB_OUTLET_R_REMOTE_TEMP	0xA4	10	NA	75	Standby
Outlet Left Remote Temp	AL_MB_OUTLET_L_REMOTE_TEMP	0xA5	10	NA	75	Standby
P0 Temp	AL_MB_CPU0_TEMP	0xA8	10	NA	89	DC
P1 Temp	AL_MB_CPU1_TEMP	0xA9	10	NA	89	DC
P2 Temp	AL_MB_CPU2_TEMP	0xAA	10	NA	89	DC
P3 Temp	AL_MB_CPU3_TEMP	0xAB	10	NA	89	DC
P0 Vccin VR Vol	AL_MB_VR_CPU0_VCCIN_VOUT	0xB0	1.33	NA	2.1	DC
P0 Vccin VR Temp	AL_MB_VR_CPU0_VCCIN_TEMP	0xB1	10	NA	115	DC
P0 Vccin VR Curr	AL_MB_VR_CPU0_VCCIN_IOUT	0xB2	NA	NA	440	DC
P0 Vccin VR Pwr	AL_MB_VR_CPU0_VCCIN_POUT	0xB3	NA	NA	1008	DC
P0 Vsa VR Vol	AL_MB_VR_CPU0_VCCSA_VOUT	0xB4	0.44	NA	1.25	DC
P0 Vsa VR Temp	AL_MB_VR_CPU0_VCCSA_TEMP	0xB5	10	NA	115	DC
P0 Vsa VR Curr	AL_MB_VR_CPU0_VCCSA_IOUT	0xB6	NA	NA	19.6	DC
P0 Vsa VR Pwr	AL_MB_VR_CPU0_VCCSA_POUT	0xB7	NA	NA	28.75	DC
P0 Vccio VR Vol	AL_MB_VR_CPU0_VCCIO_VOUT	0xB8	0.85	NA	1.25	DC
P0 Vccio VR Temp	AL_MB_VR_CPU0_VCCIO_TEMP	0xB9	10	NA	115	DC
P0 Vccio VR Curr	AL_MB_VR_CPU0_VCCIO_IOUT	0xBA	NA	NA	28.8	DC
P0 Vccio VR Pwr	AL_MB_VR_CPU0_VCCIO_POUT	0xBB	NA	NA	71.24	DC
DIMM ABC VR Vol	AL_MB_VR_CPU0_VDDQ_ABC_VOUT	0xBC	1.13	NA	1.4	DC
DIMM ABC VR Temp	AL_MB_VR_CPU0_VDDQ_ABC_TEMP	0xBD	10	NA	115	DC
DIMM ABC VR Curr	AL_MB_VR_CPU0_VDDQ_ABC_IOUT	0xBE	NA	NA	77	DC
DIMM ABC VR Pwr	AL_MB_VR_CPU0_VDDQ_ABC_POUT	0xBF	NA	NA	121.8	DC

DIMM DEF VR Vol	AL_MB_VR_CPU0_VDDQ_DEF_VOUT	0xC0	1.13	NA	1.4	DC
DIMM DEF VR Temp	AL_MB_VR_CPU0_VDDQ_DEF_TEMP	0xC1	10	NA	115	DC
DIMM DEF VR Curr	AL_MB_VR_CPU0_VDDQ_DEF_IOUT	0xC2	NA	NA	77	DC
DIMM DEF VR Pwr	AL_MB_VR_CPU0_VDDQ_DEF_POUT	0xC3	NA	NA	121.8	DC
PCH P1V05 VR Vol	AL_MB_VR_PCH_P1V05_VOUT	0xC4	1.01	NA	1.2	Standby
PCH P1V05 VR Temp	AL_MB_VR_PCH_P1V05_TEMP	0xC5	10	NA	115	Standby
PCH P1V05 VR Curr	AL_MB_VR_PCH_P1V05_IOUT	0xC6	NA	NA	18	Standby
PCH P1V05 VR Pwr	AL_MB_VR_PCH_P1V05_POUT	0xC7	NA	NA	24	Standby
PCH PVNN VR Vol	AL_MB_VR_PCH_PVNN_VOUT	0xC8	0.8	NA	1.2	Standby
PCH PVNN VR Temp	AL_MB_VR_PCH_PVNN_TEMP	0xC9	10	NA	115	Standby
PCH PVNN VR Curr	AL_MB_VR_PCH_PVNN_IOUT	0xCA	NA	NA	26	Standby
PCH PVNN VR Pwr	AL_MB_VR_PCH_PVNN_POUT	0xCB	NA	NA	33.6	Standby
P5V	AL_MB_P5V	0xD0	4.75	NA	5.25	DC
P5V_STBY	AL_MB_P5V_STBY	0xD1	4.75	NA	5.25	Standby
P3V3_STBY	AL_MB_P3V3_STBY	0xD2	3.14	NA	3.47	Standby
P3V3	AL_MB_P3V3	0xD3	3.14	NA	3.47	DC
P3V3_BAT	AL_MB_P3V_BAT	0xD4	2.6	NA	3.4	Standby
CPU_P1V8	AL_MB_CPU_P1V8	0xD5	1.71	NA	1.89	DC
PCH_P1V8	AL_MB_PCH_P1V8	0xD6	1.71	NA	1.89	DC
CPU0_PVPP_ABC	AL_MB_CPU0_PVPP_ABC	0xD7	2.38	NA	2.75	DC
CPU1_PVPP_ABC	AL_MB_CPU1_PVPP_ABC	0xD8	2.38	NA	2.75	DC
CPU0_PVPP_DEF	AL_MB_CPU0_PVPP_DEF	0xD9	2.38	NA	2.75	DC
CPU1_PVPP_DEF	AL_MB_CPU1_PVPP_DEF	0xDA	2.38	NA	2.75	DC
CPU0_PVTT_ABC	AL_MB_CPU0_PVTT_ABC	0xDB	0.55	NA	0.66	DC
CPU1_PVTT_ABC	AL_MB_CPU1_PVTT_ABC	0xDC	0.55	NA	0.66	DC
CPU0_PVTT_DEF	AL_MB_CPU0_PVTT_DEF	0xDD	0.55	NA	0.66	DC
CPU1_PVTT_DEF	AL_MB_CPU1_PVTT_DEF	0xDE	0.55	NA	0.66	DC
PROCESSOR_FAIL	AL_MB_PROCESSOR_FAIL	0XDF	NA	NA	NA	DC

P1 Vccin VR Vol	AL_MB_VR_CPU1_VCCIN_VOUT	0xE0	1.33	NA	2.1	DC
P1 Vccin VR Temp	AL_MB_VR_CPU1_VCCIN_TEMP	0xE1	10	NA	115	DC
P1 Vccin VR Curr	AL_MB_VR_CPU1_VCCIN_IOUT	0xE2	NA	NA	440	DC
P1 Vccin VR Pwr	AL_MB_VR_CPU1_VCCIN_POUT	0xE3	NA	NA	1008	DC
P1 Vsa VR Vol	AL_MB_VR_CPU1_VCCSA_VOUT	0xE4	0.44	NA	1.25	DC
P1 Vsa VR Temp	AL_MB_VR_CPU1_VCCSA_TEMP	0xE5	10	NA	115	DC
P1 Vsa VR Curr	AL_MB_VR_CPU1_VCCSA_IOUT	0xE6	NA	NA	19.6	DC
P1 Vsa VR Pwr	AL_MB_VR_CPU1_VCCSA_POUT	0xE7	NA	NA	28.75	DC
P1 Vccio VR Vol	AL_MB_VR_CPU1_VCCIO_VOUT	0xE8	0.85	NA	1.25	DC
P1 Vccio VR Temp	AL_MB_VR_CPU1_VCCIO_TEMP	0xE9	10	NA	115	DC
P1 Vccio VR Curr	AL_MB_VR_CPU1_VCCIO_IOUT	0xEA	NA	NA	28.8	DC
P1 Vccio VR Pwr	AL_MB_VR_CPU1_VCCIO_POUT	0xEB	NA	NA	71.24	DC
DIMM ABC VR Vol	AL_MB_VR_CPU1_VDDQ_ABC_VOUT	0xEC	1.13	NA	1.4	DC
DIMM ABC VR Temp	AL_MB_VR_CPU1_VDDQ_ABC_TEMP	0xED	10	NA	115	DC
DIMM ABC VR Curr	AL_MB_VR_CPU1_VDDQ_ABC_IOUT	0xEE	NA	NA	77	DC
DIMM ABC VR Pwr	AL_MB_VR_CPU1_VDDQ_ABC_POUT	0xEF	NA	NA	121.8	DC
DIMM DEF VR Vol	AL_MB_VR_CPU1_VDDQ_DEF_VOUT	0xF0	1.13	NA	1.4	DC
DIMM DEF VR Temp	AL_MB_VR_CPU1_VDDQ_DEF_TEMP	0xF1	10	NA	115	DC
DIMM DEF VR Curr	AL_MB_VR_CPU1_VDDQ_DEF_IOUT	0xF2	NA	NA	77	DC
DIMM DEF VR Pwr	AL_MB_VR_CPU1_VDDQ_DEF_POUT	0xF3	NA	NA	121.8	DC

AirFlow	SYSTEM_AIRFLOW	Aggregate sensor	NA	NA	NA	Standby

9.3 System Event Log (SEL)

The BMC needs to support SEL capabilities. The following items are to be logged in the SEL

- Analog sensors exceeding thresholds that are being set.
- Event-only sensors like
  - All types of PCIE errors, status change
  - All types of NVIDIA® NVLink errors, status change
  - All types of POST error
  - All types of Machine Check Error
  - All types of Fan failures (like fan speed out of range), and identify which fan failed
  - I2C/SMA/PMBUS transactions errors or device NACK
  - All PATMA related features
  - Any other system anomalies

9.3.1 Critical SEL Filter

OEM commands are required to set and display two different levels of SEL filtering. The default is to log all errors during EVT/DVT/PVT with the option to log only critical SEL that needs servicing or indicates power cycle state change. Ability to log when a user clears the SEL log or when it has overflowed.

Event Category	Event	Priority	SEL Type	Comments	Event Rename
Time Sync	Log event when BMC time syncs up with System RTC or NTP server when the difference is > 5 seconds	P1			

WDT	WDT Timer expired	P1	Master MB	BIOS or Host OS can use this WDT timer "FRU: X, BIOS FRB2 Watchdog <ACTION>", "FRU: X, BIOS/POST Watchdog <ACTION>", "FRU: X, OS Load Watchdog <ACTION>", "FRU: X, SMS/OS Watchdog <ACTION>", "FRU: X, OEM Watchdog <ACTION>", <ACTION> can be one of: Timer expired Hard Reset Power Down Power Cycle	
	WDT expiry action taken	P1	Master MB	BMC Watchdog reboot event: BMC Reboot detected - unknown kernel flag (0xXXXXXXXX) BMC Reboot detected - caused by reboot command BMC Reboot detected - unknown reboot command flag (0xXXXXXXXX) BMC Reboot detected	
Power State	Power OFF	P0	MBX	SERVER_POWER_OFF successful for FRU: 1	
	Power ON	P0	MBX	SERVER_POWER_ON successful for FRU: 1	
	Power Cycle	P0	MBX	SERVER_POWER_CYCLE successful for FRU: 1	
	Power Reset	P0	MBX	SERVER_POWER_RESET successful for FRU: 1	
	AC-Lost	P1	MBX	SLED Powered OFF at <DATE TIME>	
	AC-Recovery	P1	MBX	SLED Powered ON at <DATE TIME>	
	Warm Reboot	P1	Master MB	BMC Cold Reset BMC Reboot detected	
	Graceful Shutdown	P0	MBX	SERVER_GRACEFUL_SHUTDOWN successful for FRU: 1	
	SLED Cycle	P0	MBX	power-util SLED_CYCLE starting…	
	Power Fail	P1	MBX	ASSERT: XXXXXX power rail fails discrete - raised - FRU: 1, num: 0x9C, snr: MB_POWER_FAIL val:0xX	

BIOS Generated	Memory ECC Error	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: MEMORY_ECC_ERR(0xXX), DIMM Slot Location: Sled XX/Socket XX, Channel XX, Slot XX, DIMM Failure Event: <DESCRIPTION> 0000b = Memory training error 0001b = Memory correctable error 0010b = Memory uncorrectable error 0011b = Memory correctable error (Patrol scrub) 0100b = Memory uncorrectable error (Patrol scrub) 0101b = Memory Parity Error event 11b = Reserved See BIOS_SEL Sheet for more details	
	PCIe Error	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: x86/PCIeErr(0xXX), Bus XX/Dev XX/Fun XX, TotalErrID1Cnt: 0xXXXX, ErrID2: 0xXX, ErrID1: 0xXX See BIOS_SEL Sheet for more details	
	POST Error	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: POST(0xXX), POST Failure Event: <DESCRIPTION> 0000b = System PXE boot fail 0001b = CMOS/NVRAM configuration cleared 0010b = TPM Self-Test Fail See BIOS_SEL Sheet for more details	
	Other IIO Error	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: IIOErr(0xXX), IIO Port Location: Sled XX/Socket %XX, Stack 0xXX, Error ID: 0xXX See BIOS_SEL Sheet for more details	
	Machine Check Error	P0	Master MB	SEL Entry: FRU: X, Record: Standard (0x02), Time: xxxx-xx-xx xx:xx:xx, Sensor: MACHINE_CHK_ERR (0x40), Event Data: (XXXXXX) Correctable, Machine Check bank Number xx , CPU x, Core x Assertion See BIOS_SEL Sheet for more details	
	UPI Error Event	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: UPIErr(0xXX), UPI Port Location: Sled xx/Socket xx, Port xx, UPI Failure Event: <DESCRIPTION> See BIOS_SEL Sheet for more details	

	PPR or sPPR Event	P0	Master MB	SEL Entry: FRU: X, Record: Facebook Unified SEL (0xFB), GeneralInfo: MemEvent(0xXX), PPR fail/success, DIMM Failure Event: <DESCRIPTION> See BIOS_SEL Sheet for more details	
	Software NMI	P1	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> See BIOS_SEL Sheet for more details	
ME Generated	SPS FW Health	P1	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> Please refer to NM4.0 #550710 ch2.6.5 "ME Firmware Health Event"	
	NM Exception	P1	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> Please refer to NM4.0 #550710 ch3.5.15	
	Pwr Thresh Evt	P1	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> Please refer to NM4.0 #550710 Table 3-11 "Power Threshold Event"	
	CPU0/CPU1 Therm Status	P0	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> Please refer to NM4.0 #550710 Table A-2 "CPU Thermal Status" and ch2.10.1 for more detail	
	ME Status events	P2	Master MB	ASSERT: ME Status - Controller Unavailable on the mb DEASSERT: ME Status - Controller Unavailable on the mb ASSERT: ME Status - Controller Access Degraded or Unavailable on the mb DEASSERT: ME Status - Controller Access Degraded or Unavailable on the mb	

	Host Partition Reset Warning	P2	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), OEM data: (<6 BYTE OEM data>) <DESCRIPTION> Please refer to NM4.0 #550710 Table 2-10 "Event Messages Definition"	
	DC Cycle	P2	Master MB	SEL Entry: FRU: X, Record: Standard (0x02), Time: YYYY-MM-DD HH:MM:SS, Sensor: ME_POWER_STATE (0x16), Event Data: (020000) POWER_OFF Assertion	
	AC Cycle	P2	Master MB	SEL Entry: FRU: X, Record: Standard (0x02), Time: YYYY-MM-DD HH:MM:SS, Sensor: ME_POWER_STATE (0x16), Event Data: (000000) RUNNING Assertion	
Log	Log Clear	P1	MBX	Time: YYYY-MM-DD HH:MM:SS log-util: User cleared all logs Time: YYYY-MM-DD HH:MM:SS log-util: User cleared sys logs	
		P1	MBX	Time: YYYY-MM-DD HH:MM:SS log-util: User cleared FRU: x logs	
	Log Over Flow	P1			
Anolog Sensors	Temperature Sensors: - Inlet, Outlet, PCH, VR, CPU core, DIMM, HSC, LAN card	P0	MBX	ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> _x000D_ DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	
	Power Sensors:- VRs for CPU/DIMM, HSC, Package Power, INA230	P0	MBX	ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> _x000D_ DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	
	Current Sensors:- VRs for CPU/DIMM, HSC, INA230	P0	MBX	ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> _x000D_ DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	

	Voltage Sensors:- VRs for CPU/DIMM, HSC, INA230	P0	MBX	ASSERT: Lower Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>_x000D_ DEASSERT: Lower Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>_x000D_ ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> _x000D_ DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	
	Voltage Rail Monitor:- P3V3, P5V, P12V, P1V05_STBY, P1V8_AUX, P3V3_AUX, P5V_AUX, P3V_BAT	P0	MBX	ASSERT: Lower Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>_x000D_ DEASSERT: Lower Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>_x000D_ ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> _x000D_ DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	
	Fan Speed	P0	Master MB	ASSERT: Lower Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> DEASSERT: Lower Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> ASSERT: Upper Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> DEASSERT: Upper Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> ASSERT: Upper Non Critical threshold - raised - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME> DEASSERT: Upper Non Critical threshold - settled - FRU: X, num: 0xXX curr_val: XX, thresh_val: XX, snr: <SNR_NAME>	

	Switch Inlet, AirFlow- Need to findout how to monitor these sensors and report errors	P1			
Discrete Sensors					
	CPU Error#0	P0	MBX	FRU: X ASSERT: GPIOF0 - FM_CPU_ERR0_LVT3_N FRU: X DEASSERT: GPIOF0 - FM_CPU_ERR0_LVT3_N	
	CPU Error#1	P0	MBX	FRU: X ASSERT: GPIOF1 - FM_CPU_ERR1_LVT3_N FRU: X DEASSERT: GPIOF1 - FM_CPU_ERR1_LVT3_N	
	CPU Error#2	P0	MBX	FRU: X ASSERT: GPIOF2 - FM_CPU_ERR2_LVT3_N FRU: X DEASSERT: GPIOF2 - FM_CPU_ERR2_LVT3_N	
	CPU#0 ProcHOT	P0	MBX	FRU: X ASSERT: GPIOB6 - FM_CPU0_PROCHOT_LVT3_ BMC_N FRU: X DEASSERT: GPIOB6 - FM_CPU0_PROCHOT_LVT3_ BMC_N	
	CPU#1 ProcHOT	P0	MBX	FRU: X ASSERT: GPIOB7 - FM_CPU1_PROCHOT_LVT3_ BMC_N FRU: X DEASSERT: GPIOB7 - FM_CPU1_PROCHOT_LVT3_ BMC_N	
	VR PVCCIN HOT for CPU#0	P0	MBX	FRU: X ASSERT: ExIO_0 - IRQ_PVCCIN_CPU0_VRHOT_LVC3_N (reason: XXXX) FRU: X DEASSERT: ExIO_0 - IRQ_PVCCIN_CPU0_VRHOT_LVC3_N	CPU0 PVCCIN VR HOT Warning Assertion CPU0 PVCCIN VR HOT Warning Dessertion
	VR PVCCIN HOT for CPU#1	P0	MBX	FRU: X ASSERT: ExIO_1 - IRQ_PVCCIN_CPU1_VRHOT_LVC3_N (reason: XXXX) FRU: X DEASSERT: ExIO_1 - IRQ_PVCCIN_CPU1_VRHOT_LVC3_N	CPU1 PVCCIN VR HOT Warning Assertion CPU1 PVCCIN VR HOT

					Warning Dessertion
	CPU0 VR PVDDQ HOT for DIMM ABC	P0	MBX	FRU: X ASSERT: ExIO_2 - IRQ_PVDDQ_ABC_CPU0_VRHOT_LVC3_N FRU: X DEASSERT: ExIO_2 - IRQ_PVDDQ_ABC_CPU0_VRHOT_LVC3_N	CPU0 PVDDQ ABC VR HOT Warning Assertion CPU0 PVDDQ ABC VR HOT Warning Dessertion
	CPU0 VR PVDDQ HOT for DIMM DEF	P0	MBX	FRU: X ASSERT: ExIO_3 - IRQ_PVDDQ_DEF_CPU0_VRHOT_LVC3_N FRU: X DEASSERT: ExIO_3 - IRQ_PVDDQ_DEF_CPU0_VRHOT_LVC3_N	CPU0 PVDDQ DEF VR HOT Warning Assertion CPU0 PVDDQ DEF VR HOT Warning Dessertion
	CPU1 VR PVDDQ HOT for DIMM ABC	P0	MBX	FRU: X ASSERT: ExIO_4 - IRQ_PVDDQ_ABC_CPU1_VRHOT_LVC3_N FRU: X DEASSERT: ExIO_4 - IRQ_PVDDQ_ABC_CPU1_VRHOT_LVC3_N	CPU1 PVDDQ ABC VR HOT Warning Assertion CPU1 PVDDQ ABC VR HOT Warning Dessertion
	CPU1 VR PVDDQ HOT for DIMM DEF	P0	MBX	FRU: X ASSERT: ExIO_5 - IRQ_PVDDQ_DEF_CPU1_VRHOT_LVC3_N FRU: X DEASSERT: ExIO_5 - IRQ_PVDDQ_DEF_CPU1_VRHOT_LVC3_N	CPU1 PVDDQ DEF VR HOT Warning Assertion CPU1 PVDDQ

					DEF VR HOT Warning Dessertion
	CPU0 Memory Hot for DIMM	P0	MBX	FRU: X ASSERT: GPIOU5 - FM_CPU0_MEMHOT_OUT_N FRU: X DEASSERT: GPIOU5 - FM_CPU0_MEMHOT_OUT_N	
	CPU1 Memory Hot for DIMM	P0	MBX	FRU: X ASSERT: GPIOL3 - FM_CPU1_MEMHOT_OUT_N FRU: X DEASSERT: GPIOL3 - FM_CPU1_MEMHOT_OUT_N	
	HSC Fault Signal	P0	MBX	FRU: X ASSERT: GPIOL2 - IRQ_HSC_FAULT_N FRU: X DEASSERT: GPIOL2 - IRQ_HSC_FAULT_N	
	CPU#0 socket Occupied State	P0	MBX	FRU: X ASSERT: ExIO_6 - FM_CPU0_SKTOCC_LVT3_PLD_N FRU: X DEASSERT: ExIO_6 - FM_CPU0_SKTOCC_LVT3_PLD_N	
	CPU#1 socket Occupied State	P0	MBX	FRU: X ASSERT: ExIO_7 - FM_CPU1_SKTOCC_LVT3_PLD_N FRU: X DEASSERT: ExIO_7 - FM_CPU1_SKTOCC_LVT3_PLD_N	
	Memory Thermal Event	P0	MBX	FRU: X ASSERT: GPIOB0 - FM_MEM_THERM_EVENT_CPU0_LVT3_N	
	Memory Thermal Event	P0	MBX	FRU: X ASSERT: GPIOB1 - FM_MEM_THERM_EVENT_CPU1_LVT3_N	
	Thermal Trip for CPU#0	P0	MBX	FRU: X ASSERT: GPIOA1 - FM_CPU0_THERMTRIP_LVT3_PLD_N	
	Thermal Trip for CPU#1	P0	MBX	FRU: X ASSERT: GPIOD0 - FM_CPU1_THERMTRIP_LVT3_PLD_N	
	PCH Thermal Trip	P0	MBX	FRU: X ASSERT: GPIOG2 - FM_PCH_BMC_THERMTRIP_N FRU: X DEASSERT: GPIOG2 - FM_PCH_BMC_THERMTRIP_N	
	FIVR Fault for CPU#0	P0	MBX	FRU: X ASSERT: GPIOB2 - FM_CPU0_FIVR_FAULT_LVT3_PLD FRU: X DEASSERT: GPIOB2 - FM_CPU0_FIVR_FAULT_LVT3_PLD	
	FIVR Fault for CPU#1	P0	MBX	FRU: X ASSERT: GPIOB3 - FM_CPU1_FIVR_FAULT_LVT3_PLD FRU: X DEASSERT: GPIOB3 - FM_CPU1_FIVR_FAULT_LVT3_PLD	
	CATERR	P0	MBX	ASSERT: CPU MCERR/CATERR	
	MSMI	P0	MBX	ASSERT: CPU MCERR/MSMI	

	System Throttle	P0	MBX	FRU: X ASSERT: GPIOR7 - FM_SYS_THROTTLE_LVC3 FRU: X DEASSERT: GPIOR7 - FM_SYS_THROTTLE_LVC3	
	SYS Power OK	P0	MBX	FRU: X ASSERT: GPIOZ1 - PWRGD_CPU0_LVC3 FRU: X DEASSERT: GPIOZ1 - PWRGD_CPU0_LVC3	
	Under Voltage Detect Signal	P0	MBX	FRU: X ASSERT: GPIOM0 - IRQ_UV_DETECT_N FRU: X DEASSERT: GPIOM0 - IRQ_UV_DETECT_N	HSC Under Voltage Warning Assertion HSC Under Voltage Warning Dessertion
	Over Current Detect Signal	P0	MBX	FRU: X ASSERT: GPIOM1 - IRQ_OC_DETECT_N FRU: X DEASSERT: GPIOM1 - IRQ_OC_DETECT_N	HSC Surge Current Warning Assertion HSC Surge Current Warning Dessertion
	PMBus Alert Signal	P0	MBX	FRU: X ASSERT: GPIOAA1 - IRQ_SML1_PMBUS_BMC_ALERT_N FRU: X DEASSERT: GPIOAA1 - IRQ_SML1_PMBUS_BMC_ALERT_N	HSC OC Warning Assertion HSC OC Warning Dessertion
	HSC Timer Expiry Signal	P0	MBX	FRU: X ASSERT: GPIOM2 - FM_HSC_TIMER_EXP_N FRU: X DEASSERT: GPIOM2 - FM_HSC_TIMER_EXP_N	HSC OCP Fault Warning Assertion HSC OCP Fault Warning Dessertion
	Platform Reset Event	P1	MBX	FRU: X DEASSERT: GPIOF6 - RST_PLTRST_BMC_N FRU: X ASSERT: GPIOF6 - RST_PLTRST_BMC_N	
	PVCC CPU0 power in	P1	MBX	FRU: X DEASSERT: GPIOAA2 - FM_PVCCIN_CPU0_PWR_IN_ALERT_N FRU: X ASSERT: GPIOAA2 - FM_PVCCIN_CPU0_PWR_IN_ALERT_N	CPU0 PVCCIN POWER FAULT Assertion CPU0 PVCCIN POWER

					FAULT Dessertion
	PVCC CPU1 power in	P1	MBX	FRU: X DEASSERT: GPIOAA2 - FM_PVCCIN_CPU1_PWR_IN_ALERT_N FRU: X ASSERT: GPIOAA2 - FM_PVCCIN_CPU1_PWR_IN_ALERT_N	CPU1 PVCCIN POWER FAULT Assertion CPU1 PVCCIN POWER FAULT Dessertion
	FRB3 Failure	P1	MBX	ASSERT: FRB3 failure discrete - raised - FRU: 1	
User Actions	Power Button	P0	MBX	FRU: X ASSERT: GPIOE2 - FM_BMC_PWR_BTN_R_N FRU: X DEASSERT: GPIOE2 - FM_BMC_PWR_BTN_R_N	
	Reset Button	P0	MBX	FRU: X ASSERT: GPIOE0 - FP_BMC_RST_BTN_N FRU: X DEASSERT: GPIOE0 - FP_BMC_RST_BTN_N Reset Button pressed for FRU: 1	
	NMI Signal from Button	P0	MBX	FRU: X ASSERT: GPIOE4 - FP_NMI_BTN_N FRU: X DEASSERT: GPIOE4 - FP_NMI_BTN_N	
	Debug Card Insert/Removal	P0	MBX	FRU: X ASSERT: GPIOQ6 - FM_POST_CARD_PRES_BMC_N FRU: X DEASSERT: GPIOQ6 - FM_POST_CARD_PRES_BMC_N	
	Base OS Installation started	P2	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), Event data: (<6 BYTE OEM data>) Base OS/Hypervisor Installation started	
	Base OS Installation completed	P2	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), Event data: (<6 BYTE OEM data>) Base OS/Hypervisor Installation completed	
	Base OS Installation aborted	P2	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX), Event data: (<6 BYTE OEM data>) Base OS/Hypervisor Installation aborted	
	Base OS Installation failed	P2	Master MB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: YYYY-MM-DD HH:MM:SS, Sensor: <NAME> (0xXX),	

				Event data: (<6 BYTE OEM data>) Base OS/Hypervisor Installation failed	
BMC Generated	BMC Memory utilization	P2	MBX	ASSERT: BMC Memory utilization (XX%) exceeds the threshold (95%). DEASSERT: BMC Memory utilization (XX%) is under the threshold (95%).	
	BMC CPU utilization	P2	MBX	ASSERT: BMC CPU utilization (XX%) exceeds the threshold (85%). DEASSERT: BMC CPU utilization (XX%) is under the threshold (85%).	
	Verified boot status	P2	MBX	ASSERT: Verified boot failure (XX%:XX%) ASSERT: Verification of BMC image failed Notable errors: (7,76) - TPM is not present (3.35) - CS1 image is unsigned (9,91) - Rollback was prevented (4,43) - U-Boot was not verified using the intermediate keys (3,30) - Invalid FDT magic number for U-Boot FIT at 0x28080000	
CM	CM cold reset		PDB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: Time: YYYY-MM-DD HH:MM:SS, Sensor: <name> (0xXX), Event Data: (0xXX) PDB_Event_CM_Reset	
	Power cycle		PDB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: Time: YYYY-MM-DD HH:MM:SS, Sensor: <name> (0xXX), Event Data: (0xXX) PDB_Event_Power_Cycle, MBX	
	SLED cycle		PDB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: Time: YYYY-MM-DD HH:MM:SS, Sensor: <name> (0xXX), Event Data: (0xXX) PDB_Event_SLED_Cycle, MBX	
	Reconfig System		PDB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: Time: YYYY-MM-DD HH:MM:SS, Sensor: <name> (0xXX), Event Data: (0xXX) PDB_Event_Reconfig_System, MBX, Mode=X	
	CM external reset		PDB	SEL Entry: FRU: X, Record: <RECORD TYPE> (0xXX), Time: Time: YYYY-MM-DD HH:MM:SS, Sensor: <name> (0xXX), Event Data: (0xXX) PDB_Event_CM_External_Reset	

	Crash dump		MBX	Crashdump for FRU: X is generated at /mnt/data/autodump.tar.gz	
FSCD	No sane fan		MBX	ASSERT: ZoneX No sane fan speed could be calculated! Using transitional speed. DEASSERT: ZoneX No sane fan speed could be calculated! Using transitional speed.	
	Debug Card UART_SEL button pressed		SYS	FRU: X ASSERT: GPIOL0 - FM_UARTSW_LSB_N FRU: X DEASSERT: GPIOL0 - FM_UARTSW_LSB_N FRU: X ASSERT: GPIOL1 - FM_UARTSW_MSB_N FRU: X DEASSERT: GPIOL1 - FM_UARTSW_MSB_N	
FW update	BMC		SYS	BMC fw upgrade initiated BMC fw upgrade completed. Version: XXXX	
	BIOS & CPLD		SYS	Component XXXX upgrade initiated Component XXXX upgrade completed	
	ME		SYS	SEL Entry: FRU: 1, Record: Standard (0x02), Time: YYYY-MM-DD HH:MM:SS, Sensor: SPS_FW_HEALTH (0x17), Event Data: (A00600) Direct Flash update SEL Entry: FRU: 1, Record: Standard (0x02), Time: YYYY-MM-DD HH:MM:SS, Sensor: SPS_FW_HEALTH (0x17), Event Data: (A00600) Direct Flash update SEL Entry: FRU: 1, Record: Standard (0x02), Time: YYYY-MM-DD HH:MM:SS, Sensor: SPS_FW_HEALTH (0x17), Event Data: (A00004) Recovery GPIO forced	
NIC	ncsid event		MBX	FRU: x NIC AEN Supported: 0x7, AEN Enable Mask=0x7 FRU: x PLDM type supported = 0x35 FRU: x PLDM type 0 version = x.x.x.x FRU: x PLDM type 2 version = x.x.x.x FRU: x PLDM type 4 version = x.x.x.x FRU: x PLDM type 5 version = x.x.x.x FRU: x PLDM sensor monitoring enabled	

9.4 FSC in BMC

BMC firmware has an FSC lookup table to control fan speed.

9.5 BMC FW chip and Firmware Update

The BMC FW flash chip sits on an BMC Storage Module (BSM). This module can be treated as a FRU and it is easily swappable. During service we can easily change the flash or wipe out the data during the decommission process. Below is the drawing of the BSM module. FB has a BSM specification that you can refer to.

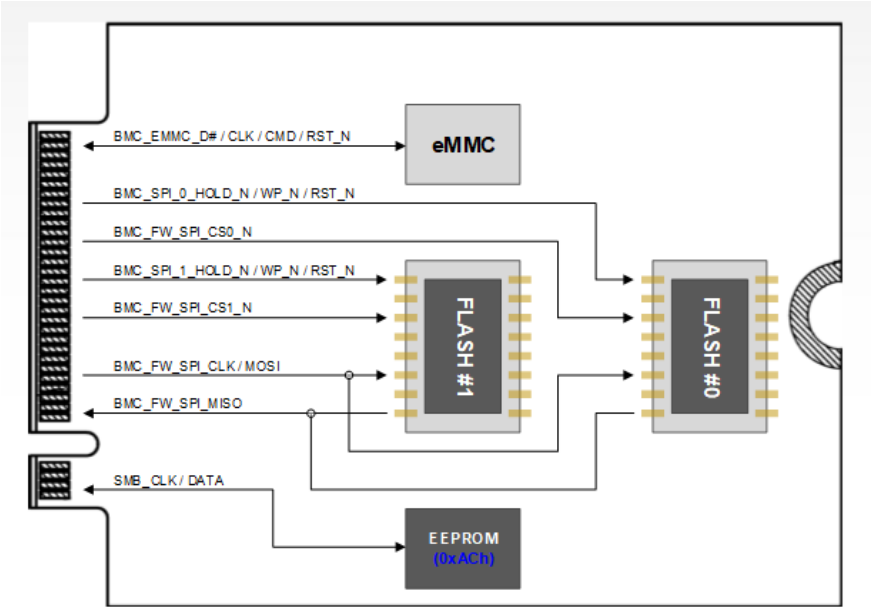


Figure 22. BSM block diagram

The vendor is responsible for selecting a specific flash chip, which should fit the required functionality in addition to potential additional features that may be required in the future. 128 Mbyte size is recommended to support additional features such as bigger storage.

Vendors should provide tool(s) to implement remote BMC firmware update, which will not require any physical access to the system. Remote update means either through the management network or through In-Band by logging into the local OS (CentOS) with the data network. Tool(s) shall support CentOS 8.1 64-bit with updated Kernel specified by the customer.

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

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

The main BMC should update 2nd BMC through NCSI by default, then over the USB interface as in Figure 17.

9.6 BMC MSR Dump

Vendor shall implement BMC to dump MSR from both CPUs through Intel® Manageability Engine and PECI. Vendor shall provide utility that supports CentOS 7.x 64-bit with updated Kernel specified by customer.
This is a debug feature to allow the user to access critical debug information from faulty SUT on a server rack, without removing the system from a failure status and risking the loss of critical debug information.
BMC firmware shall apply MSR dump automatically when there is IERR or MCERR. The BMC shall store the MSR dump in BMC flash. During the dump, the BMC shall reject chassis power related commands to avoid interrupting the dump.

9.7 BMC Verified boot

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

9.8 I2C/PMBus Circuits

The following figure shows the I2C/PMBus connections between the BMC, PCH, CPU, sensors and other hardware components.

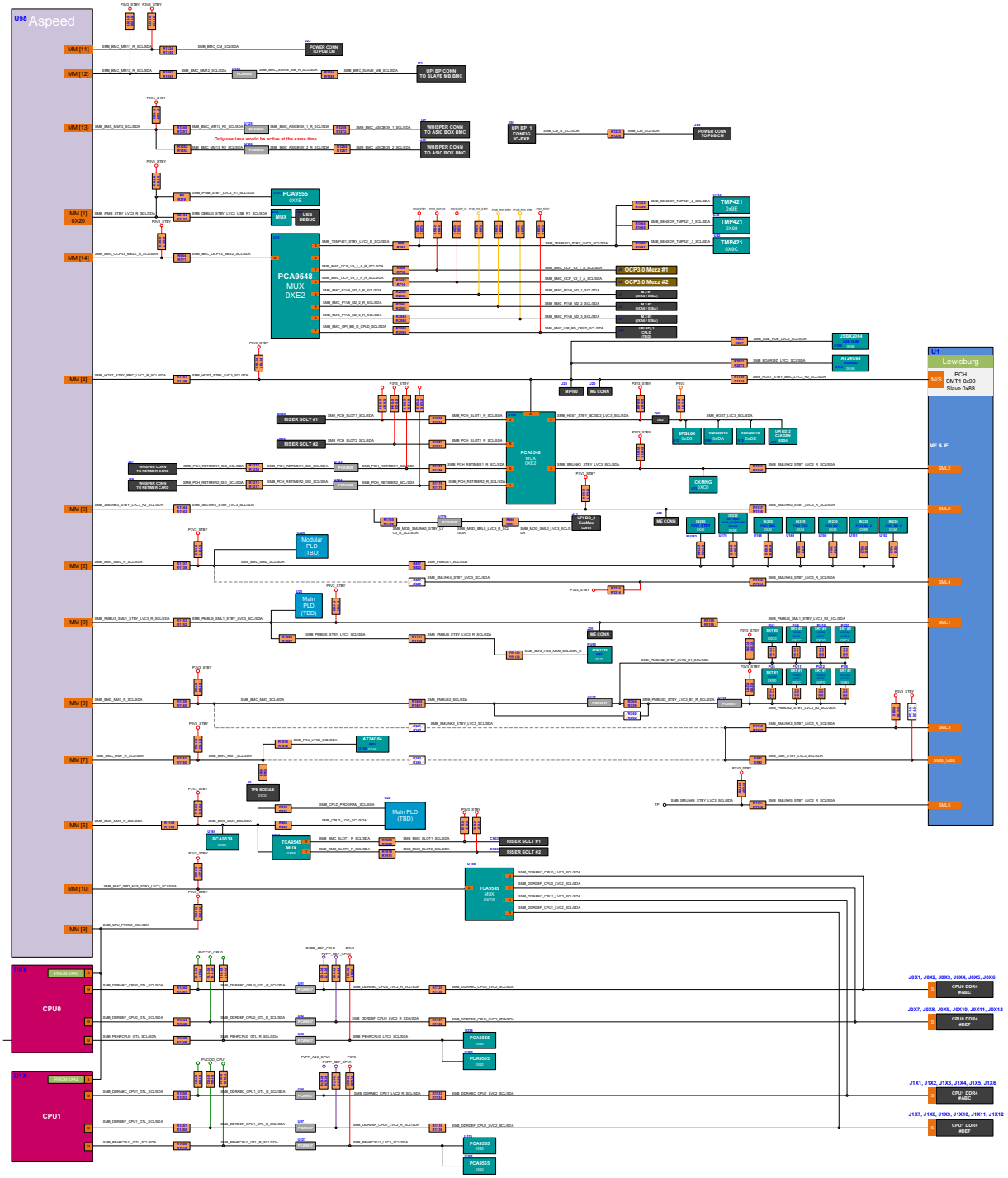


Figure 23: I2C/PMBUS board circuits

The following set of figures show each I2C/PMBUS bus separately:

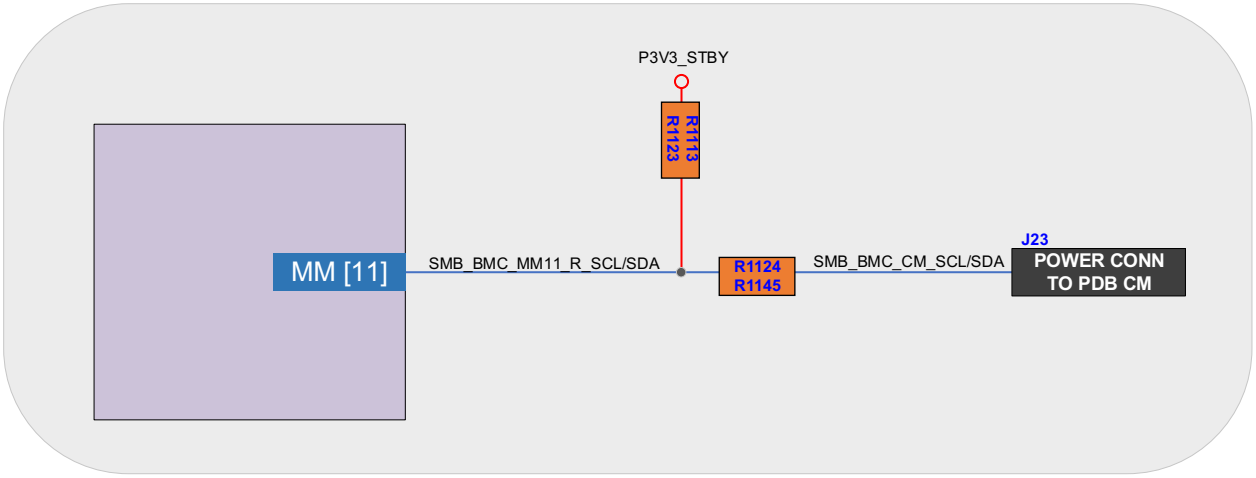


Figure 24: BMC MM[11] bus

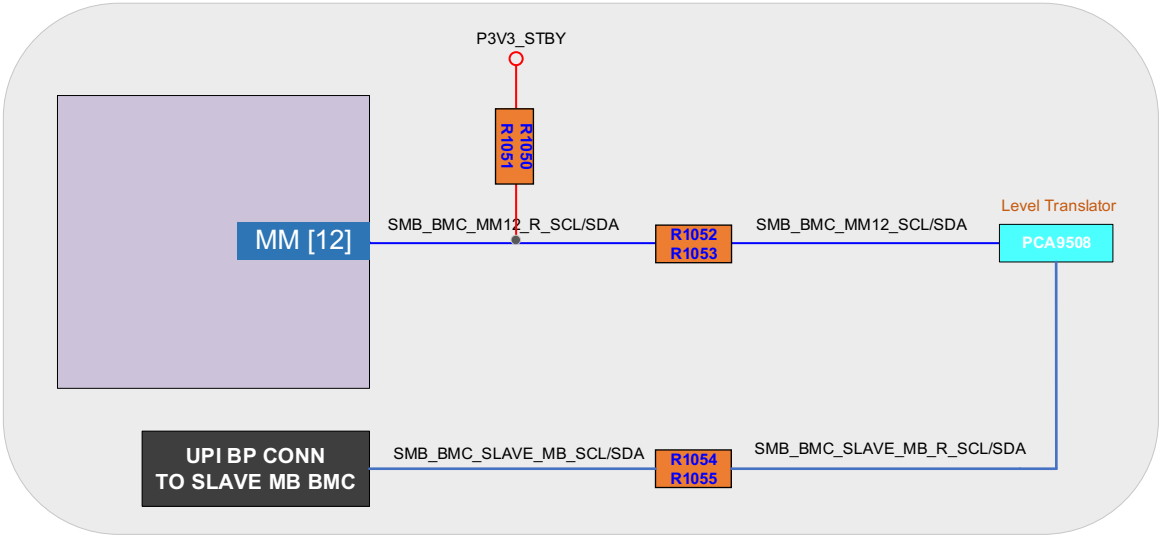


Figure 25: BMC MM[12] Bus

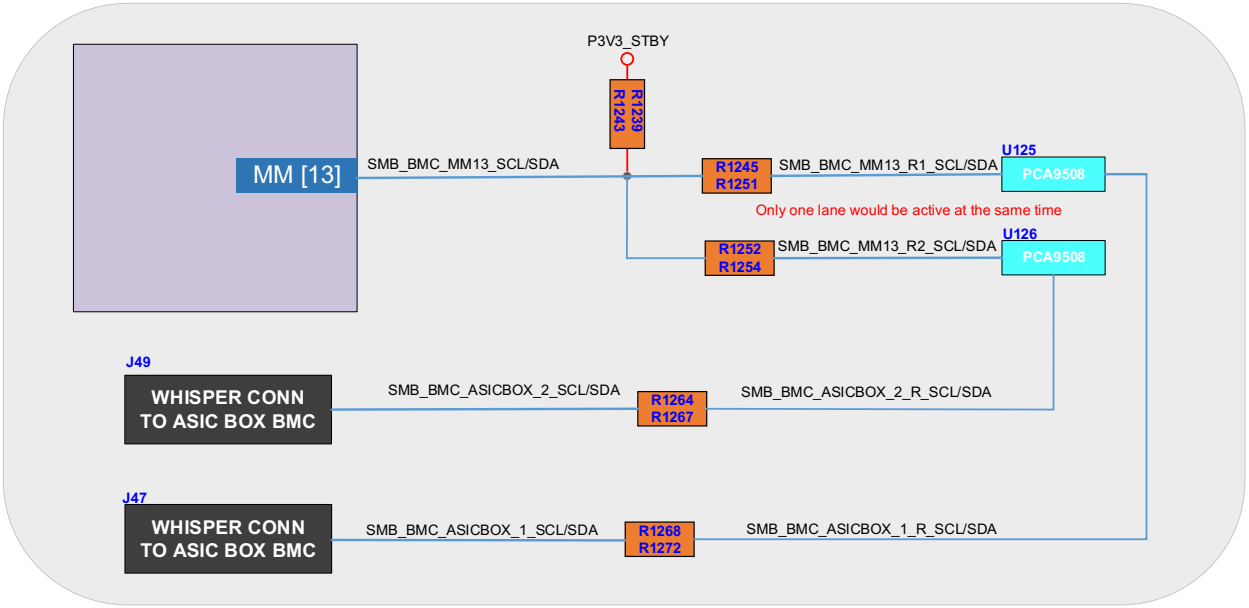


Figure 26: BMC MM[13] bus

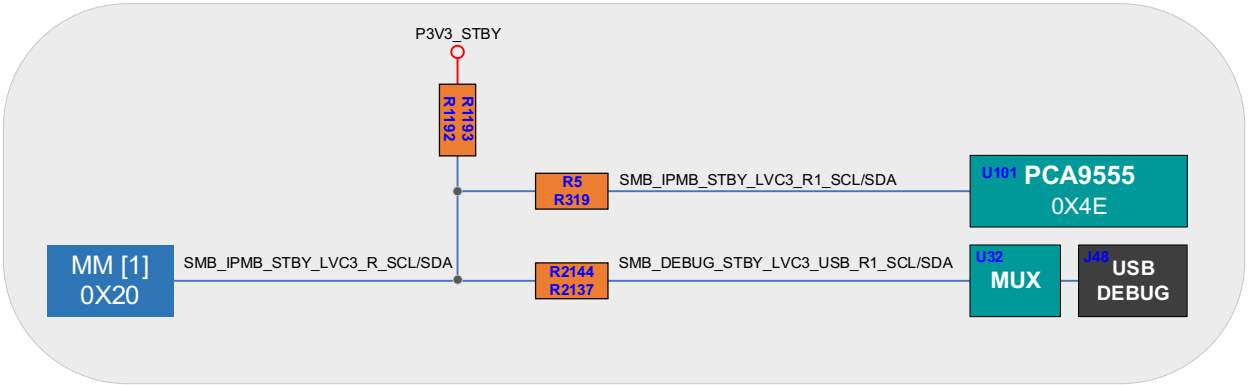


Figure 27: BMC MM[1] bus

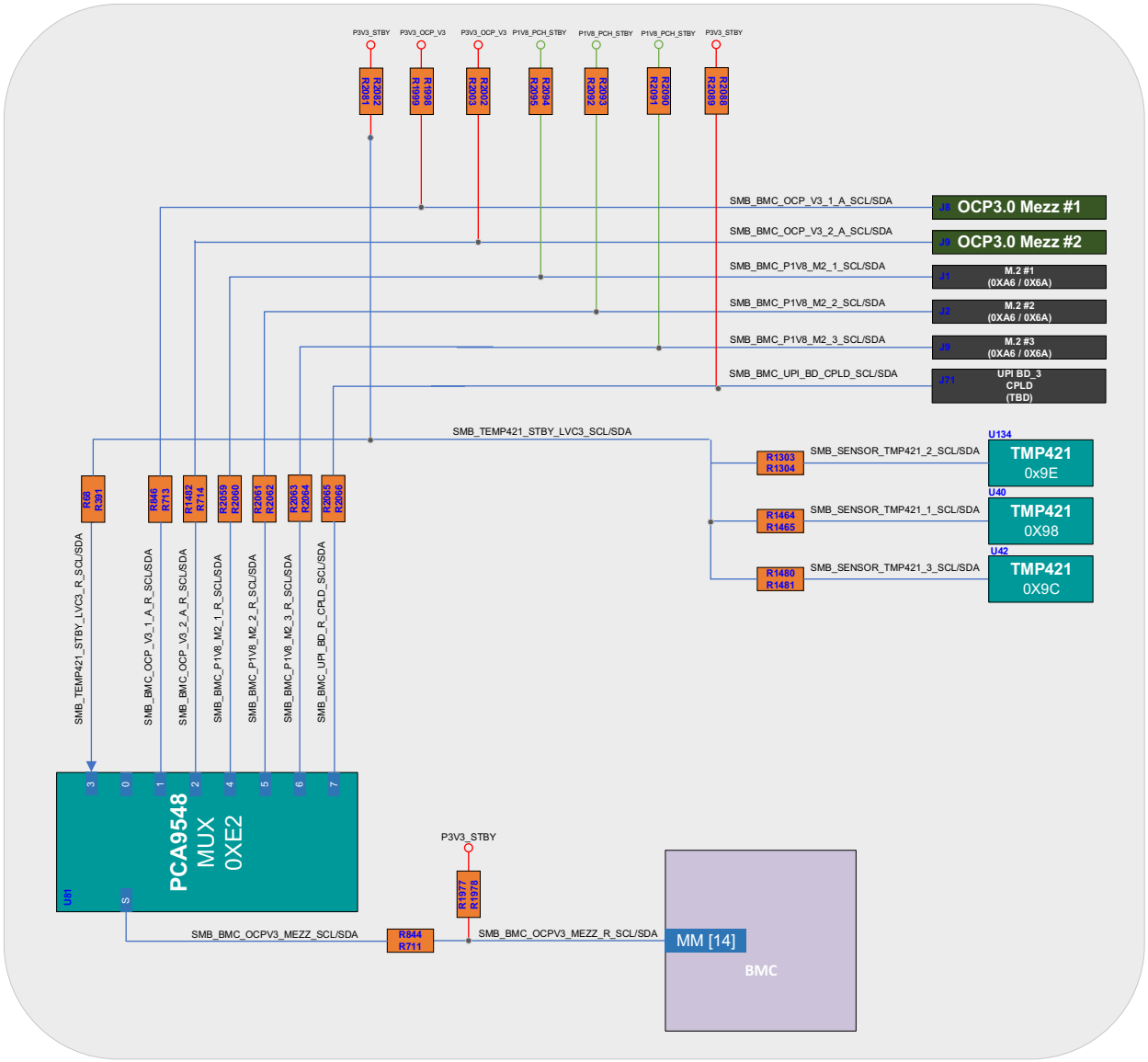


Figure 28: BMC MM[14] bus

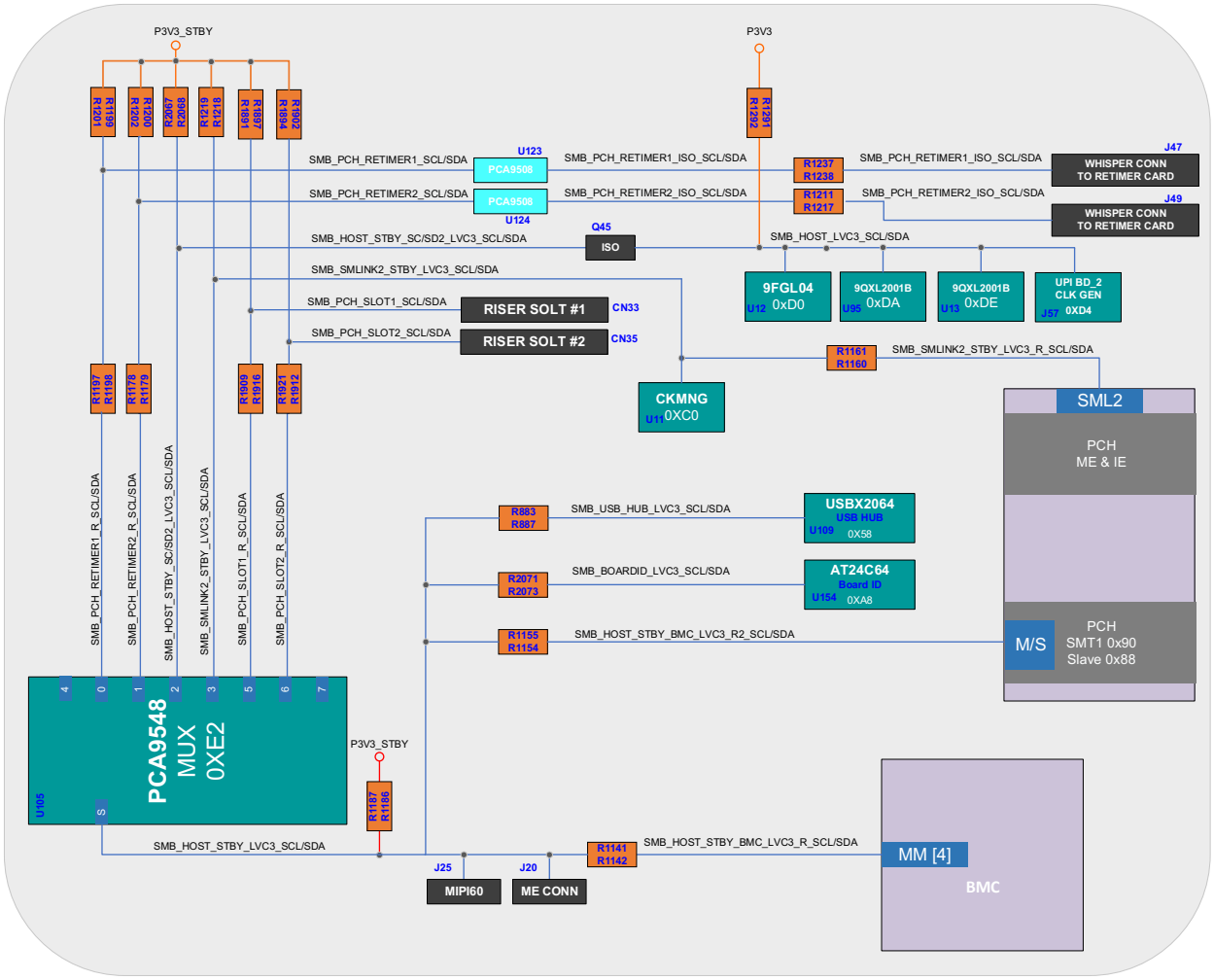


Figure 29: BMC MM[4] bus

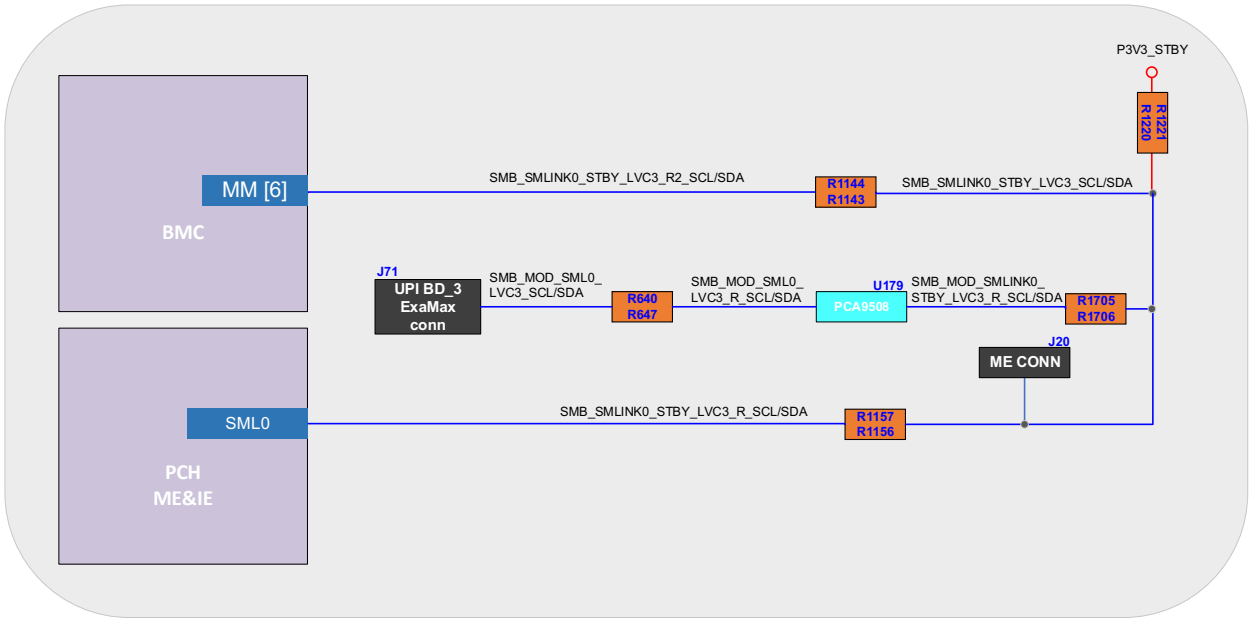


Figure 30: BMC MM[6] bus

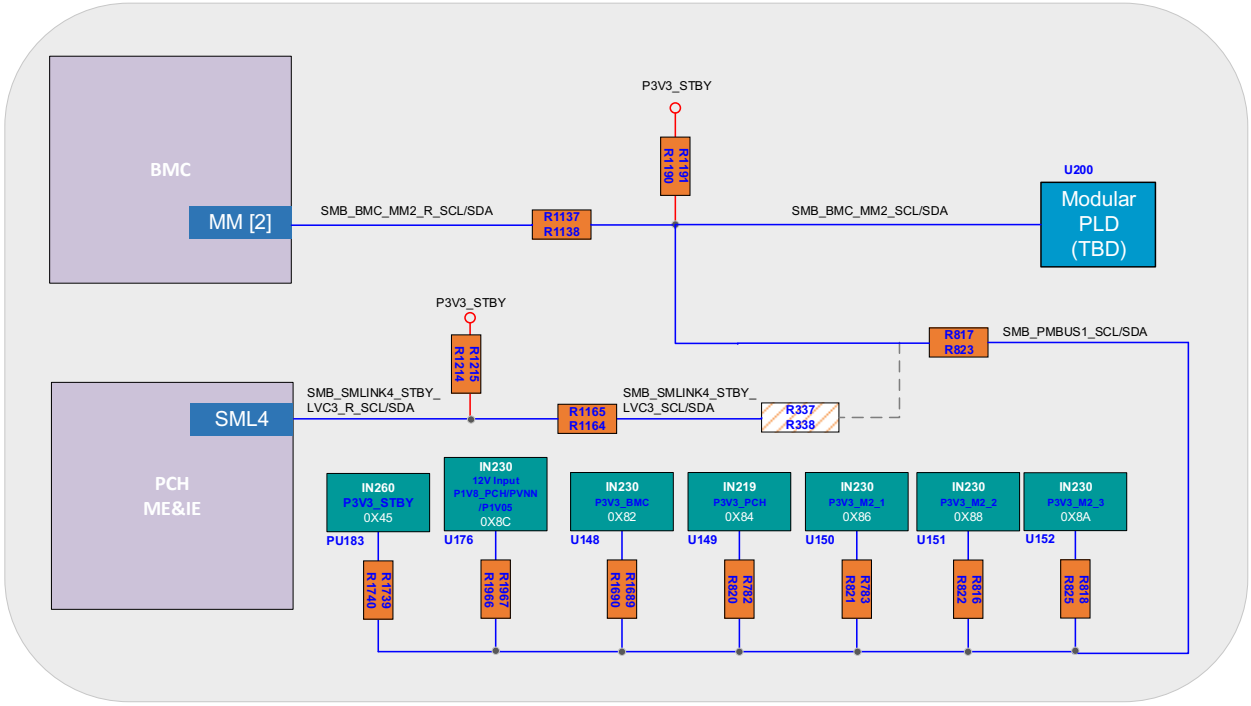


Figure 31: BMC MM[2] bus

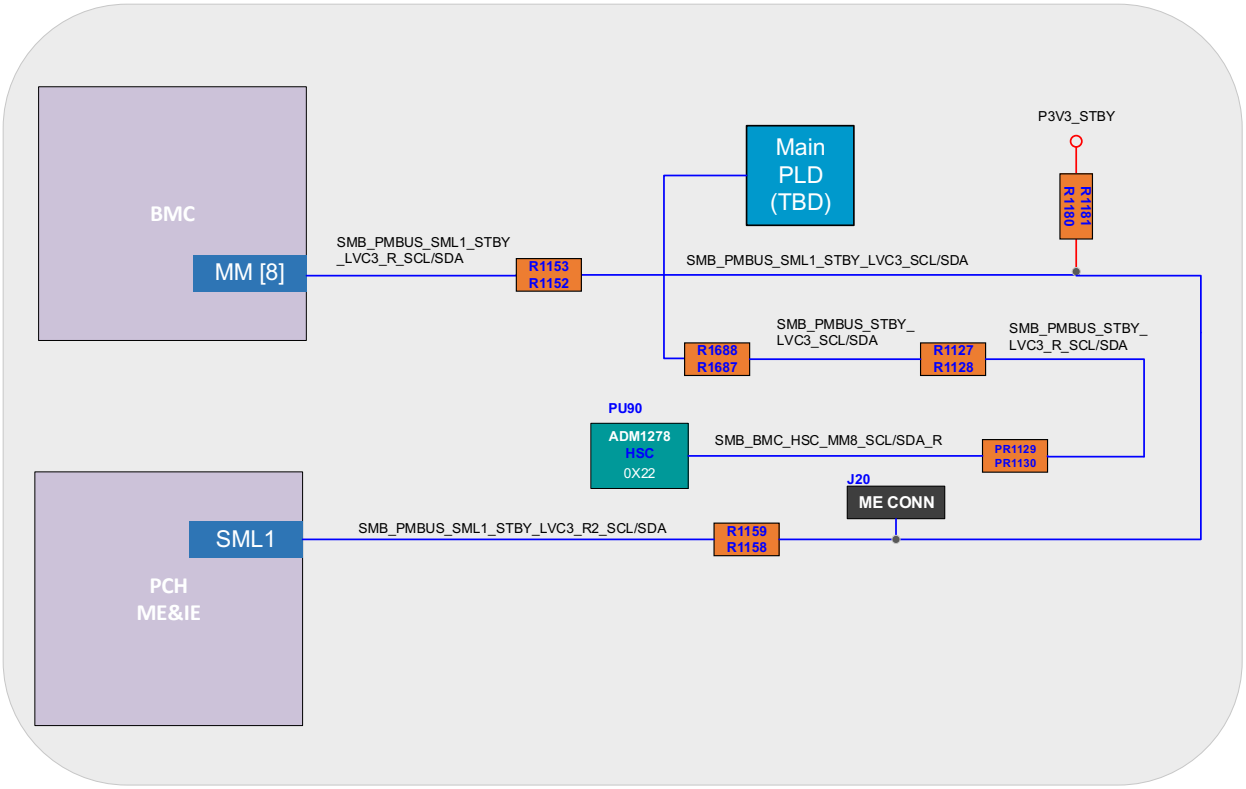


Figure 32: BMC MM[8] bus

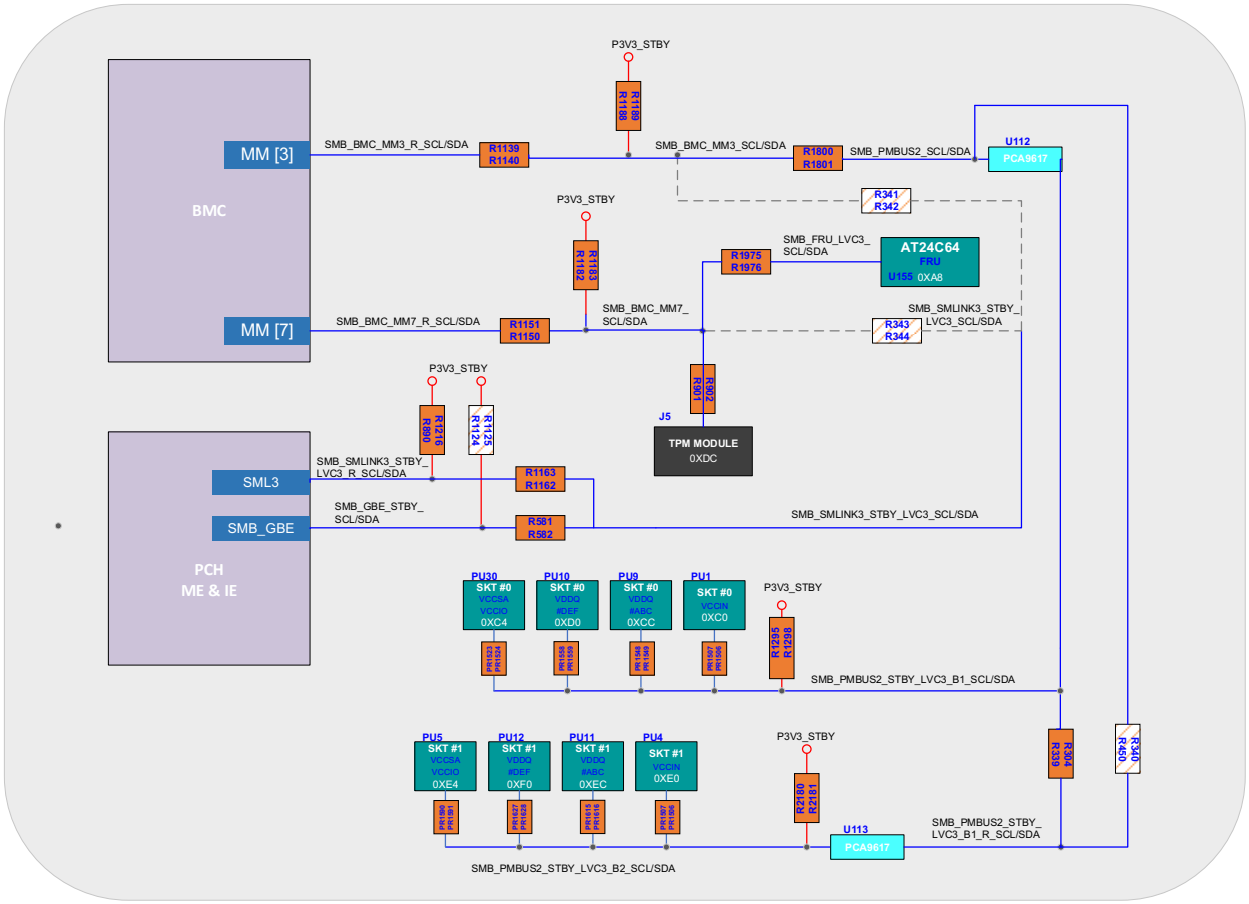


Figure 33: BMC MM[3] and MM[7] bus

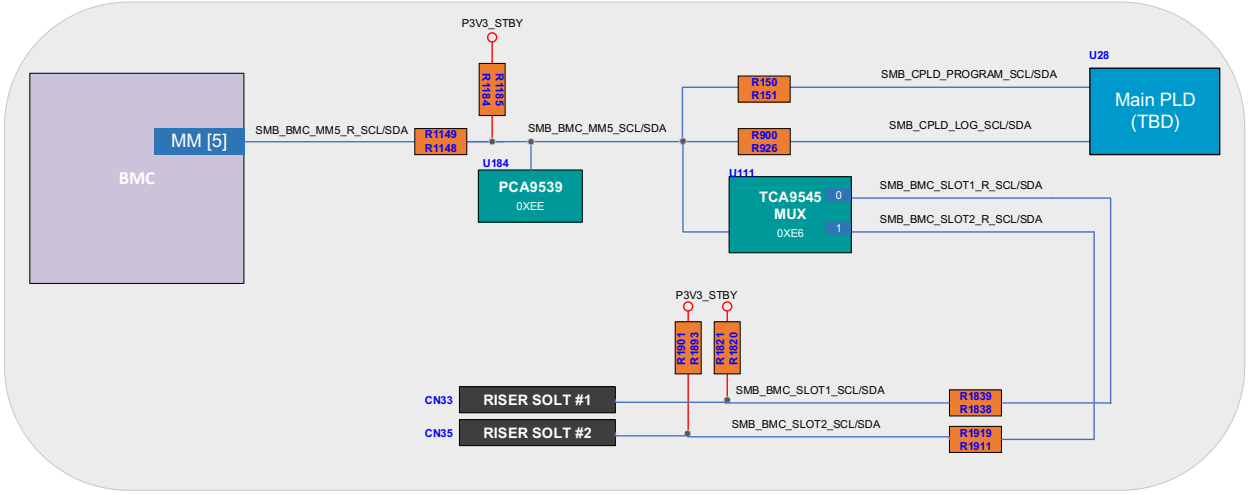


Figure 34: BMC MM[5] bus

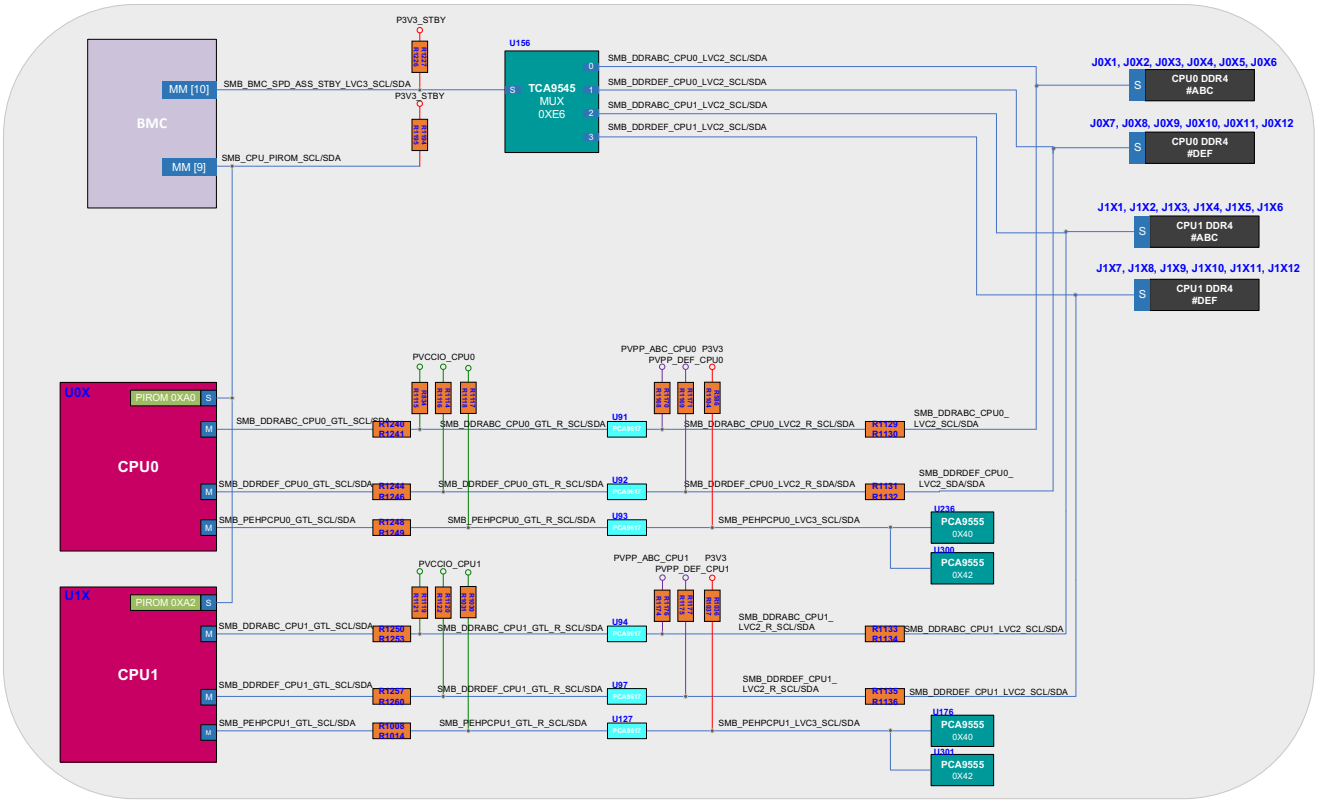


Figure 35: BMC MM[9] and MM[10] bus

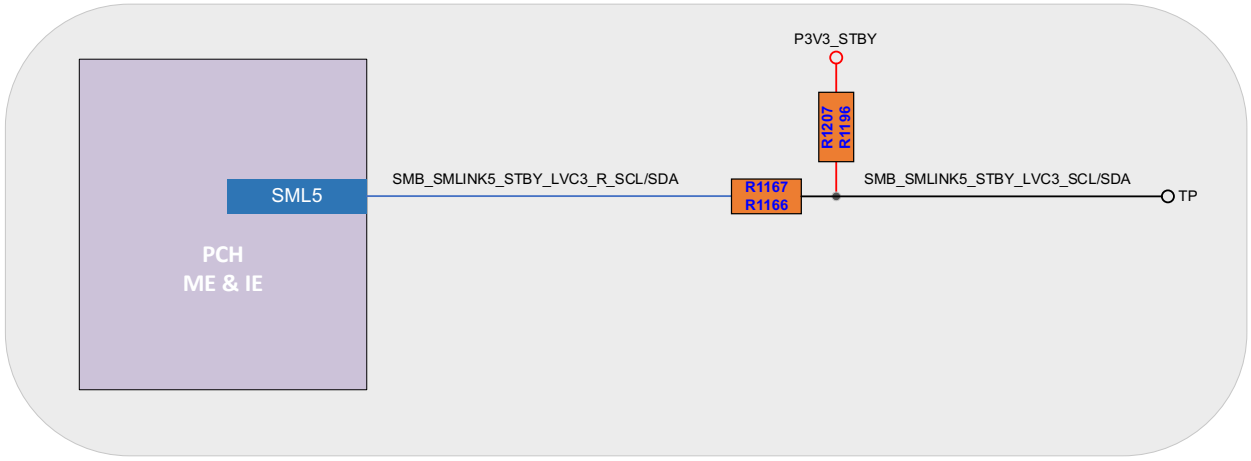


Figure 36: PCH SML5 bus

9.9 Platform Agnostic Throttle Management Architecture (PATMA)

Angels Landing shall support the PATMA implementation to enable:

1. Monitoring of individual throttle sources and log an appropriate SEL event for each occurrence.

2. Provide hardware abstraction through BMC utility (e.g. throttle_util) to individually toggle each source on/off.

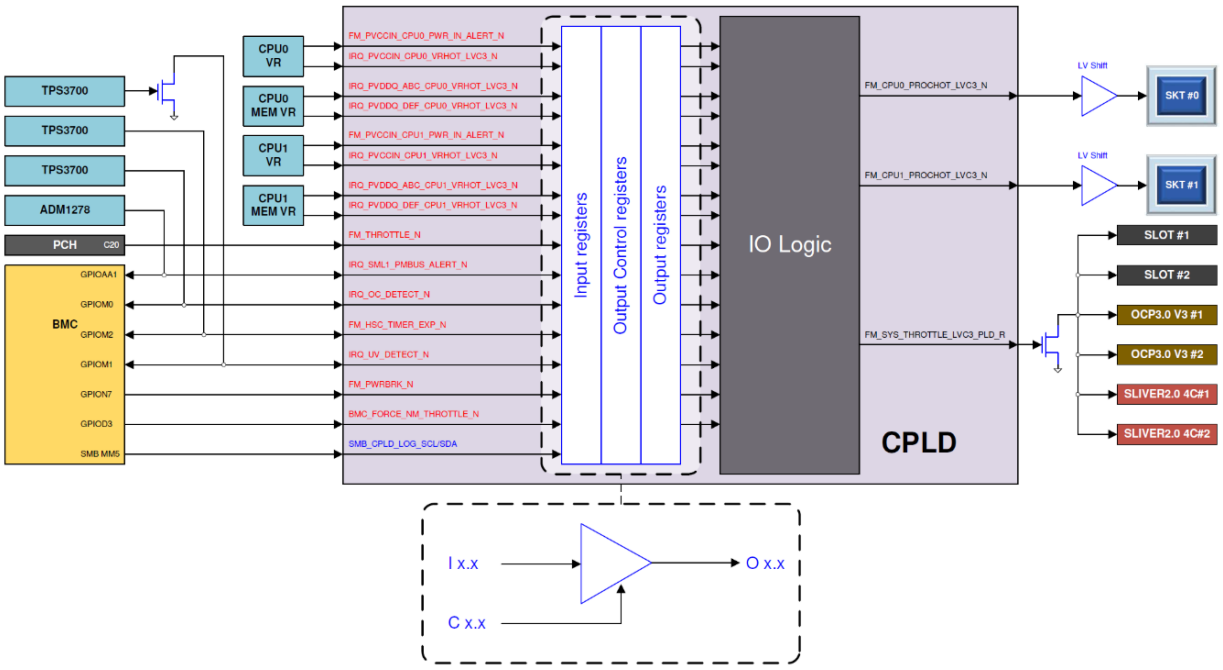


Figure 37: Platform System Throttle Sources and Description

10. Electrical Design

10.1 PCB Design

10.1.1 PCB Stackup

Following PCB stackup should be followed for motherboard/UPI backplane/retimer card design. The vendor needs to check with PCB fab vendors to fine tune the impedance based on the impedance control table below before starting PCB design. PCB material will be low loss, targeting insertion loss is around -0.5db/inch at 4Ghz.

Table 10.1 Motherboard PCB Stackup

STACKUP							
Layer#	**Material**	**Description**	**Copper Weight (oz)**	**Thickness (mil)**	**Tolerance (mil)**	**Glass Fabric**	**Er**
		Soldermask		0.60			3.8
1		TOP	0.5+plating	1.95			
	IT-958GQ	PP		2.70	±0.709	1080x1	3.7
2		GND	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
3		IN1	1 (RG311)	1.30			
	IT-958GQ	PP		5.65	±0.984	2116x1	3.8
4		GND1	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
5		IN2	1 (RG311)	1.30			
	IT-958GQ	PP		5.65	±0.984	2116x1	3.8
6		GND2	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
7		IN3	1 (RG311)	1.30			
	IT-958GQ	PP		9.00	±1.5	1086x1 + 2116x1 + 1086x1	3.8
8		VCC	2 (RTF)	2.60			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
9		VCC1	2 (RTF)	2.60			
	IT-958GQ	PP		9.00	±1.5	1086x1 + 2116x1 + 1086x1	3.8
10		IN4	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
11		GND3	1 (RG311)	1.30			
	IT-958GQ	PP		5.65	±0.984	2116x1	3.8
12		IN5	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
13		GND4	1 (RG311)	1.30			
	IT-958GQ	PP		5.65	±0.984	2116x1	3.8
14		IN6	1 (RG311)	1.30			
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7
15		GND5	1 (RG311)	1.30			
	IT-958GQ	PP		2.70	±0.709	1080x1	3.7
16		BOTTOM	0.5+plating	1.95			
		Soldermask		0.60			3.8
			Total	**92.90**	**±10%**		

Table 10.2 UPI Board Stackup

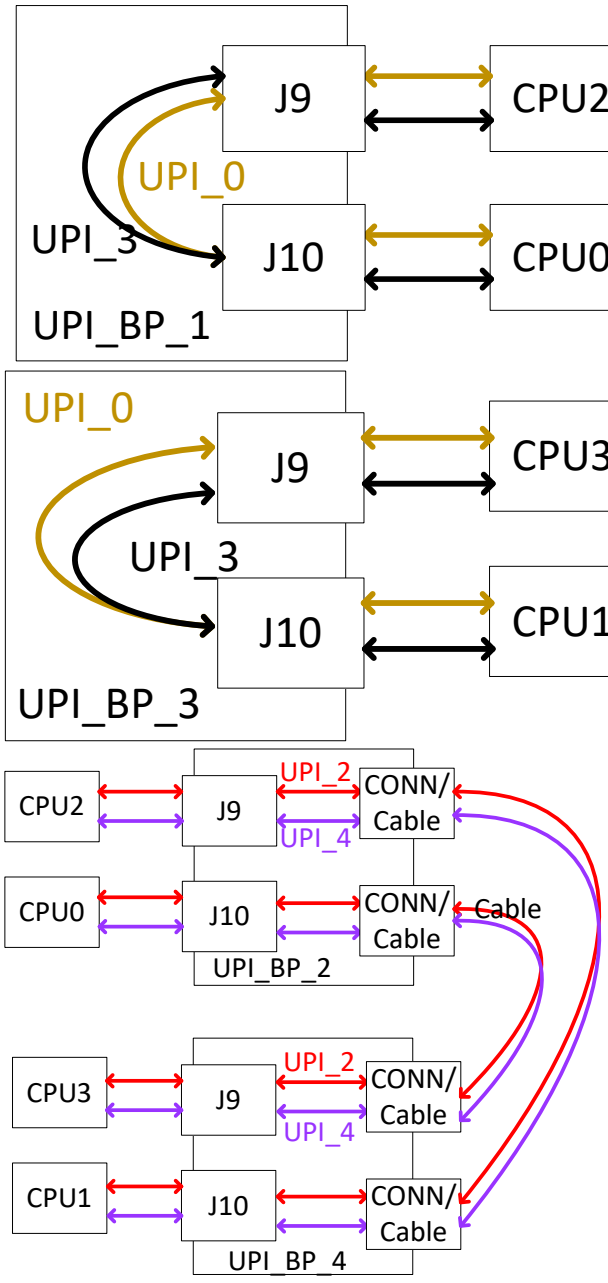
STACKUP													
						Target Z (ohms) - MicroStrip	50	Breakout	85		Breakout		
						Target Z (ohms) - StripLine	50		85				
						Z tolerance	±10%		±10%				
						Z Type	Single	Single	Differential		Differential		
Layer#	**Material**	**Description**	**Copper Weight (oz)**	**Thickness (mil)**	**Tolerance (mil)**	**Glass Fabric**	**Er**	**Width**	**Width**	**Width**	**Space**	**Width**	**Space**
		Soldermask		0.60			3.8						
1		TOP	0.5+plating	1.95				4.4	3.6	5	7	3.5	4
	IT-958GQ	PP		2.70	±0.709	1080x1	3.7						
2		GND	1 (RG311)	1.30									
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
3		IN1	1 (RG311)	1.30				3.7	3.6	4.6	6	3.5	4
	IT-958GQ	PP		7.50	±1.5	3313x2	3.8						
4		GND1	1 (RG311)	1.30									
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
5		IN2	1 (RG311)	1.30				3.7	3.6	4.6	6	3.5	4
	IT-958GQ	PP		7.50	±1.5	3313x2	3.8						
6		GND2	1 (RG311)	1.30									
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
7		IN3	1 (RG311)	1.30				3.55	3.55	4.5	6.1	3.5	4
	IT-958GQ	PP		6.60	±1.5	1080x2	3.8						
8		GND3	1 (RG311)	1.30									
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
9		GND4	1 (RG311)	1.30									
	IT-958GQ	PP		6.60	±1.5	1080x2	3.8						
10		IN4	1 (RG311)	1.30				3.55	3.55	4.5	6.1	3.5	4
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
11		GND5	1 (RG311)	1.30									
	IT-958GQ	PP		7.50	±1.5	3313x2	3.8						
12		IN5	1 (RG311)	1.30				3.7	3.6	4.6	6	3.5	4
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
13		GND6	1 (RG311)	1.30									
	IT-958GQ	PP		7.50	±1.5	3313x2	3.8						
14		IN6	1 (RG311)	1.30				3.7	3.6	4.6	6	3.5	4
	IT-958GQ	CORE		3.00	±0.709	1086x1	3.7						
15		GND7	1 (RG311)	1.30									
	IT-958GQ	PP		2.70	±0.709	1080x1	3.7						
16		BOTTOM	0.5+plating	1.95				4.4	3.6	5	7	3.5	4
		Soldermask		0.60			3.8						
			Total	**92.90**	**±10%**								

Table 10.3 PDB stackup

STACKUP									
						Target Z (ohms) - MicroStrip	50	Breakout	
						Target Z (ohms) - StripLine	50		
						Z tolerance	±10%		
						Z Type	Single	Single	
Layer#	**Material**	**Description**	**Copper Weight (oz)**	**Thickness (mil)**	**Tolerance (mil)**	**Glass Fabric**	**Er**	**Width**	**Width**
		Soldermask		0.60			3.8		
1		TOP	0.5+plating	1.95				4.3	4
	IT-170GRA1	PP		2.70	±0.709	1080x1	3.8		
2		GND	1 (RTF)	1.30					
	IT-170GRA1	CORE		10.00	±1.5	2116x2	3.9		
3		VCC	2 (RTF)	2.60					
	IT-170GRA1	PP		54.60	±5.12	Vendor Define	4.2		
4		VCC1	2 (RTF)	2.60					
	IT-170GRA1	CORE		10.00	±1.5	2116x2	3.9		
5		GND1	1 (RTF)	1.30					
	IT-170GRA1	PP		2.70	±0.709	1080x1	3.8		
6		BOTTOM	0.5+plating	1.95				4.3	4
		Soldermask		0.60			3.8		
			Total	**92.90**	**±10%**				

10.2 UPI Backplane Boards

UPI_BP_1 backplane board connects the 2 * 2S MB through 4x ExaMax connectors, Figure 38 illustrates the connections diagram:



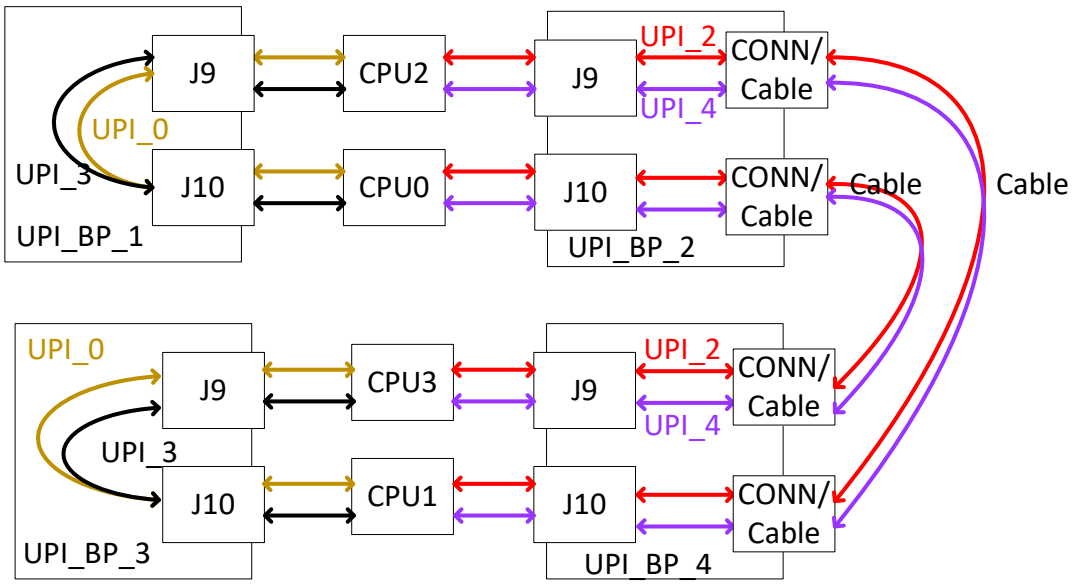


Figure 38: UPI_BP_1 Block Diagram

There are 4x UPI boards with 2 cables that support the fully connected topology between 4 CPUs in 4S configuration.

In 2S configuration once only one MB is present, the UPI boards are not needed.

10.3 LEDs and Buttons

Each MB has 2 buttons for reset and power controls (MFG: Diptronics MFG_PN: TA3-4W2-Q-T/R) these buttons are located in the front of the board.

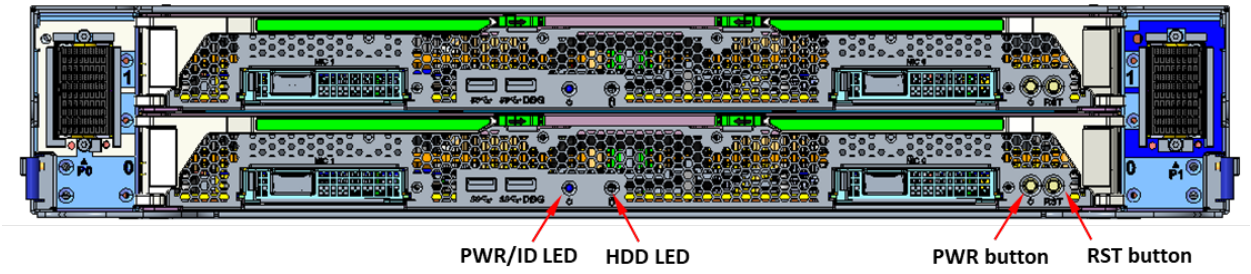


Figure 39: Angels Landing Front Panel LEDs and Buttons Drawing

LED Table

Location	Function	LED Color	Behavior	Behavior Condition
LED57		LED_BLUE	BLINK	M.2 drive active

	HDD ACTIVITY LED		OFF	No M.2 drive
LED58	PWR/ID LED	LED_BLUE/AMBER	BLUE ON, AMBER OFF	Power ON
			BLUE OFF, AMBER OFF	Power OFF
			AMBER BLINK	Node Identified
			AMBER ON	System fault

10.4 Clock Circuit

The following figure shows the clocking scheme for Zion. The 100MHz reference clock can be fed by either individual 2S_MB

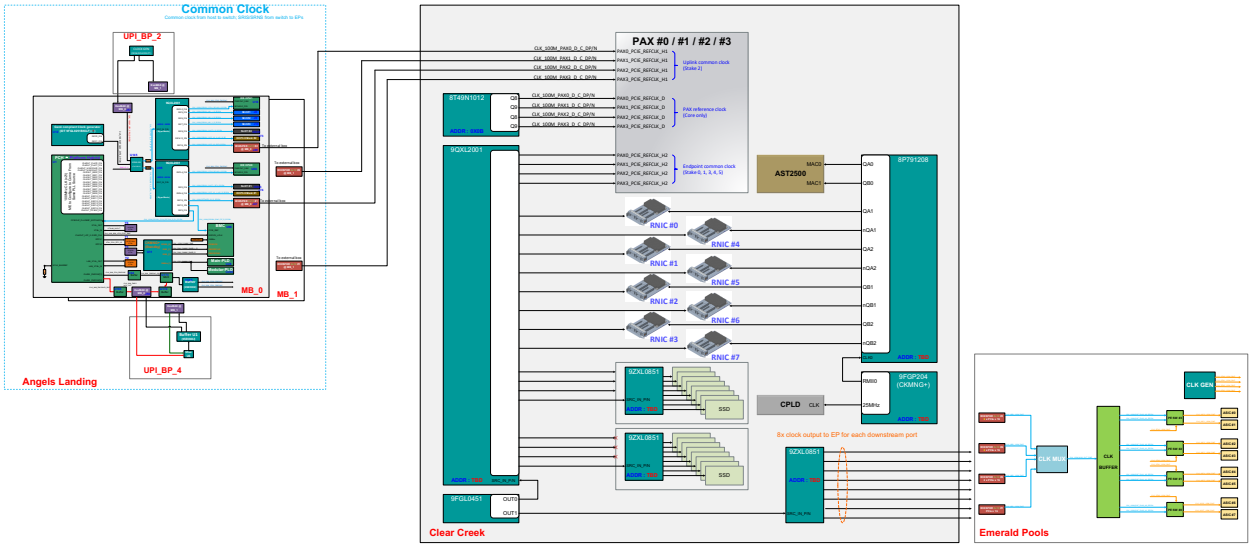


Figure 40: Clock Distribution circuit

10.5 JTAG Circuit

The following figure is a high-level block diagram which describes the JTAG chains on Angels Landing

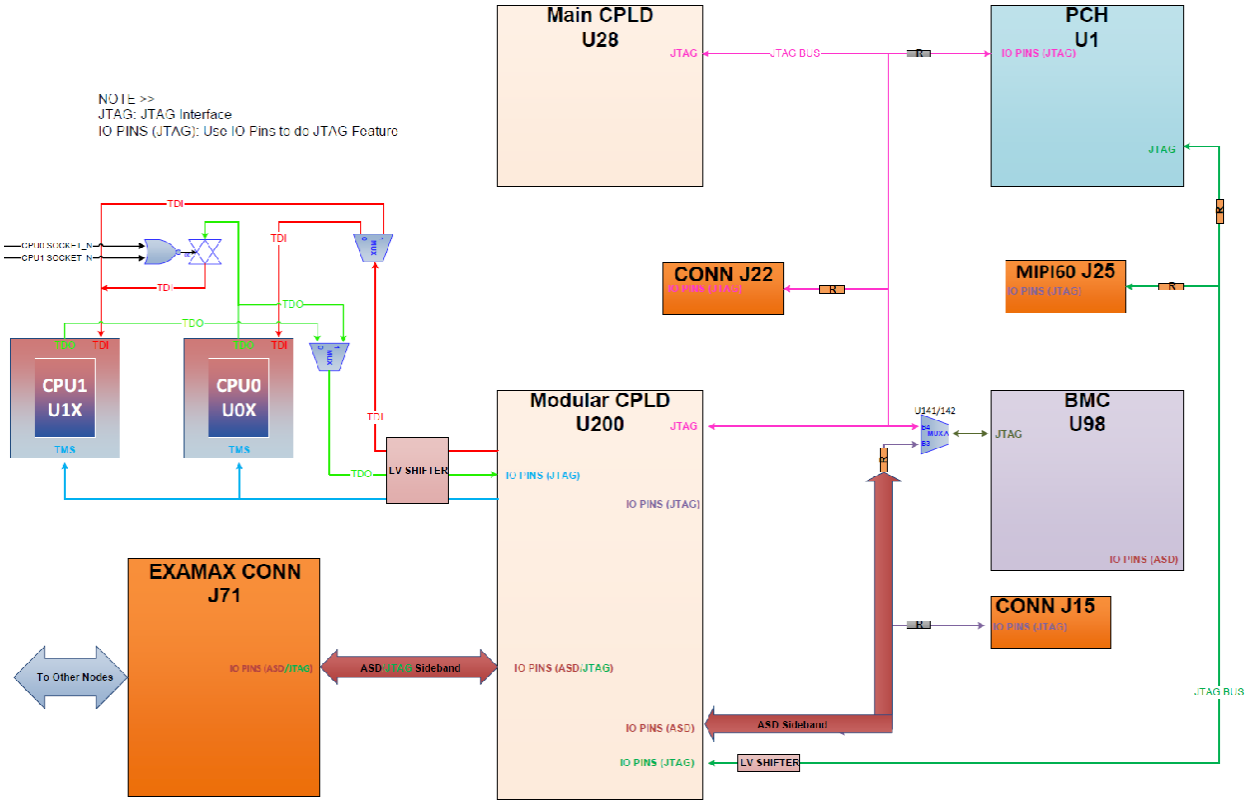


Figure 41: High Level JTAG Chains Block Diagram

Figure 42 below describes all physical connections of the Angels Landing JTAG chains, to see details of each connection you can request a high definition picture that can zoom out.

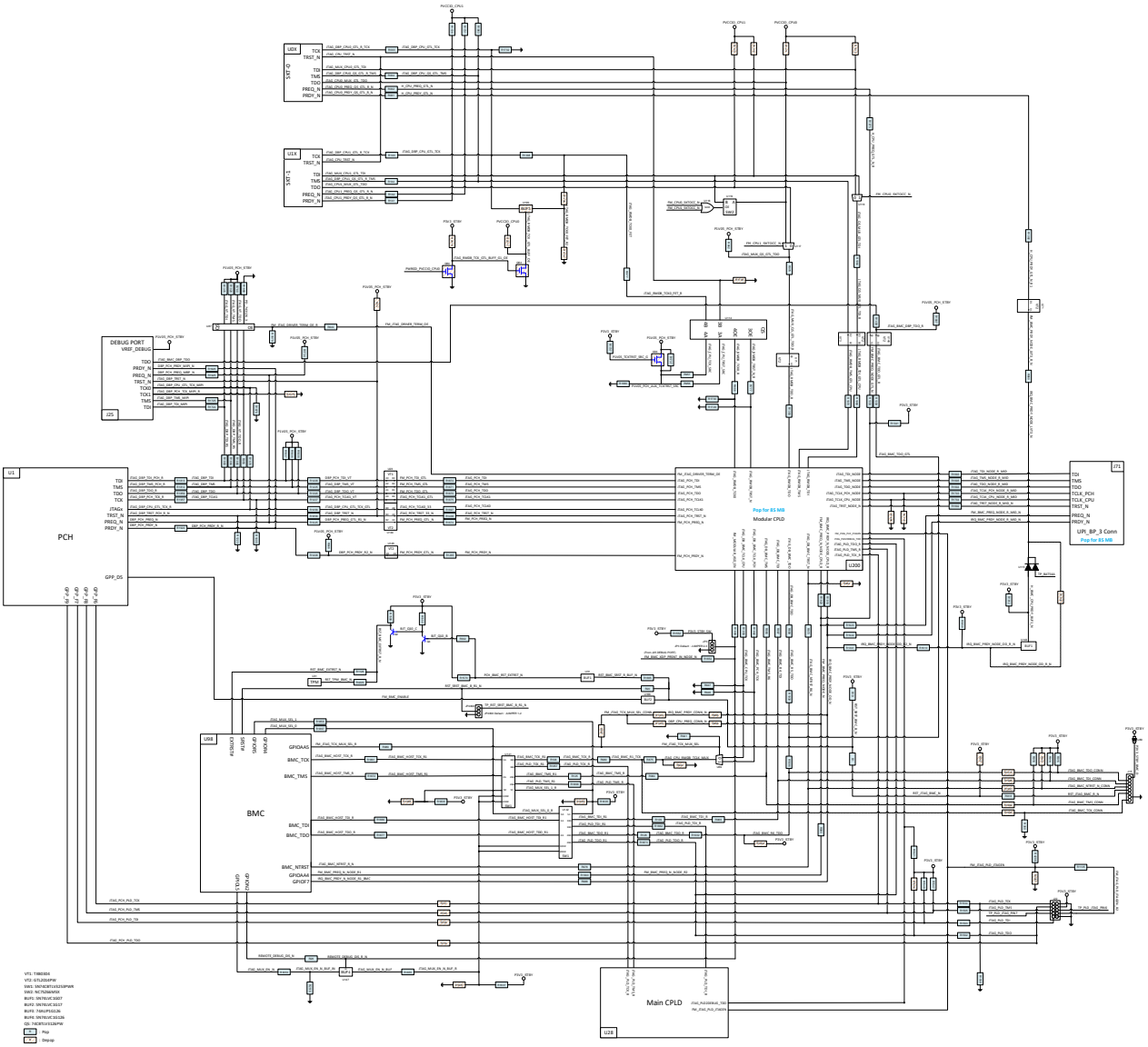


Figure 42: Detailed JTAG Diagram

10.6 TPM circuit

An 11-pin vertical receptacle connector is defined on MB for SPI and I2C TPM module. FCI/91931-31111LF Receptacle or equivalent should be used on the motherboard.

The TPM module is a 32.3mm (L) x 13mm (W) x 0.8mm (T) PCB with FCI/91911-31511LF header or equivalent located at the center of the module.

Angels Landing 4S platform does support intel PFR (Platform Firmware Resilience). Users can choose which one they want to support. The firmware package we released is still using TPM for security.

10.7 Debug

Angels landing system supports OCP debug cards for general debug purposes.

11. Power Distribution Network

11.1 PDB Board

Angels landing's landing has a Power Distribution Board that supplies power to the motherboard. The main features that PDB supports are

- Power Distribution to Angels Landing Hardware
- Low level platform chassis management through CM

11.1.1. PDB Block Diagrams

The following diagram shows the power distribution block diagram in PDB

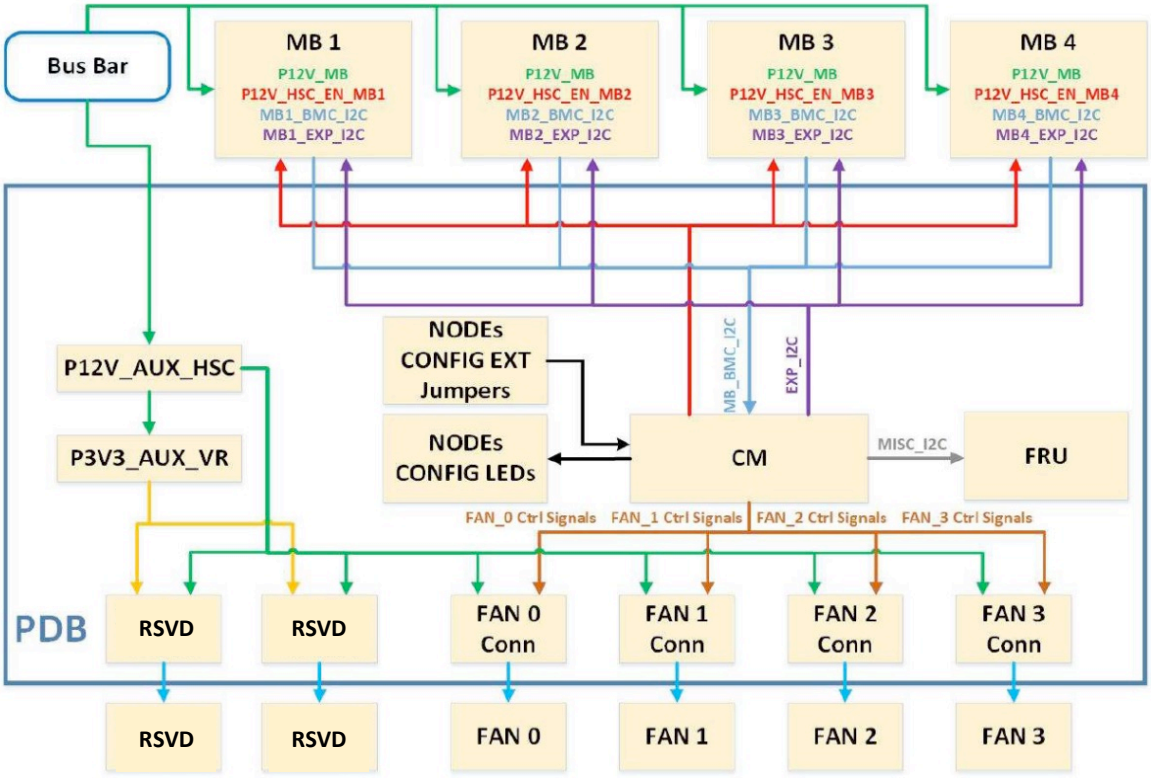


Figure 43: PDB High Level Block Diagram

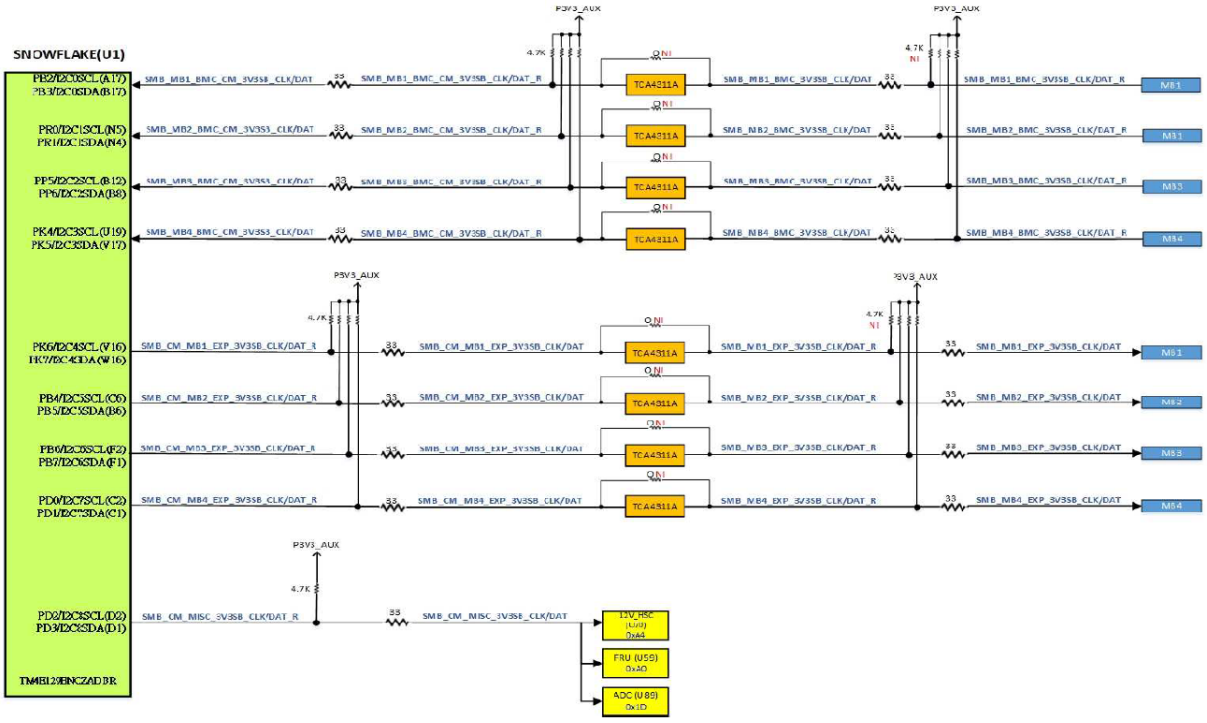


Figure 44: PDB I2C Block Diagram

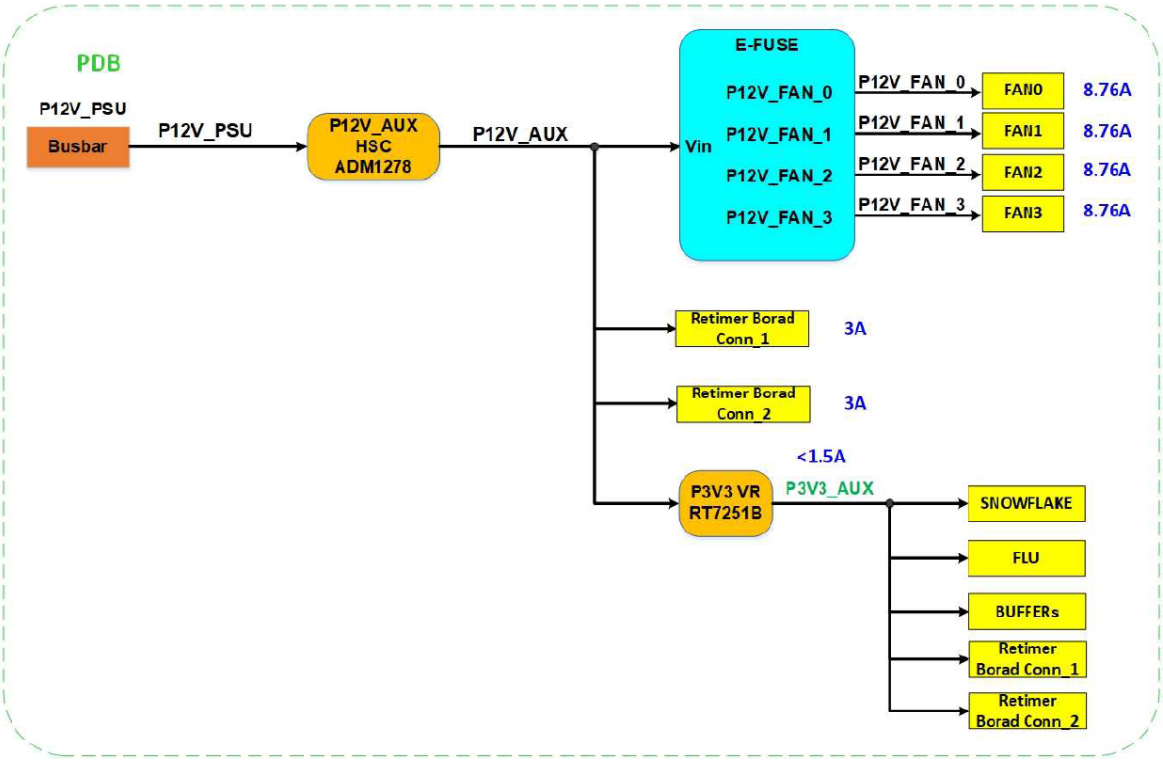


Figure 45: PDB Power Block Diagram

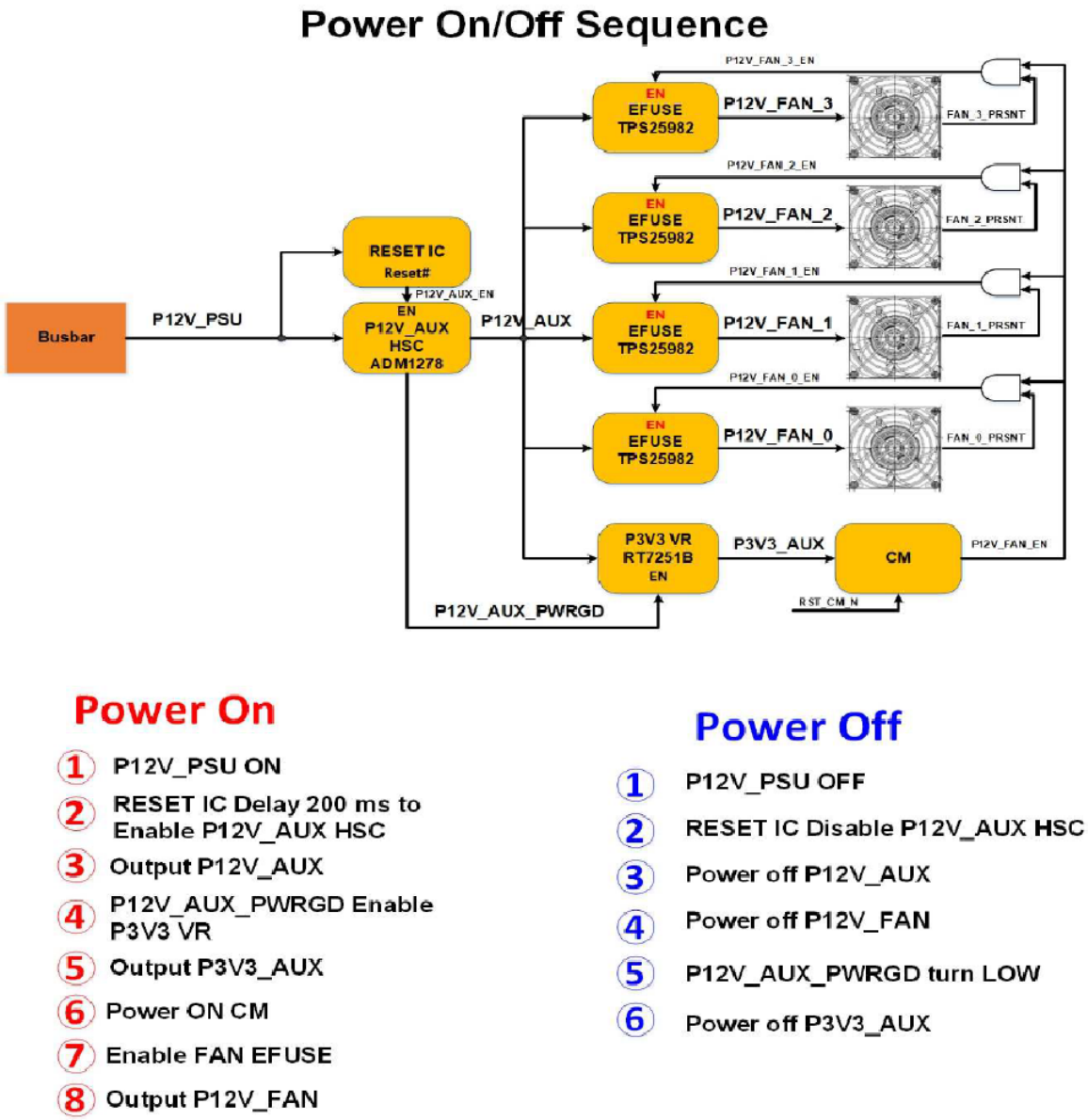


Figure 46: Power Sequence

11.2 Motherboard Power Circuit

The 2S MB gets its 12V Dc supply from the PDB through a power connector (MFG: Amphenol MFG_PN: 10106265-4003001LF). The following figure shows the power distribution on the motherboard:

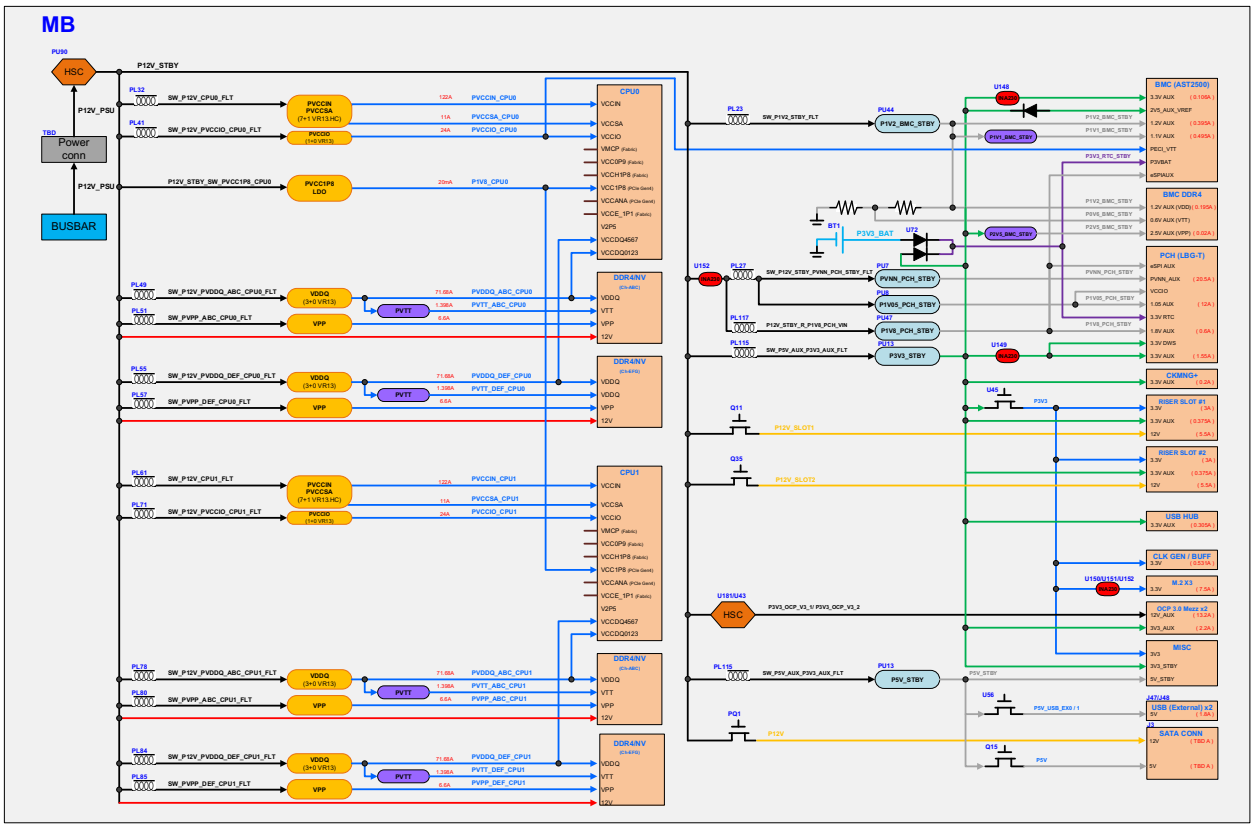


Figure 46: Angels Landing Motherboard Power Distribution scheme

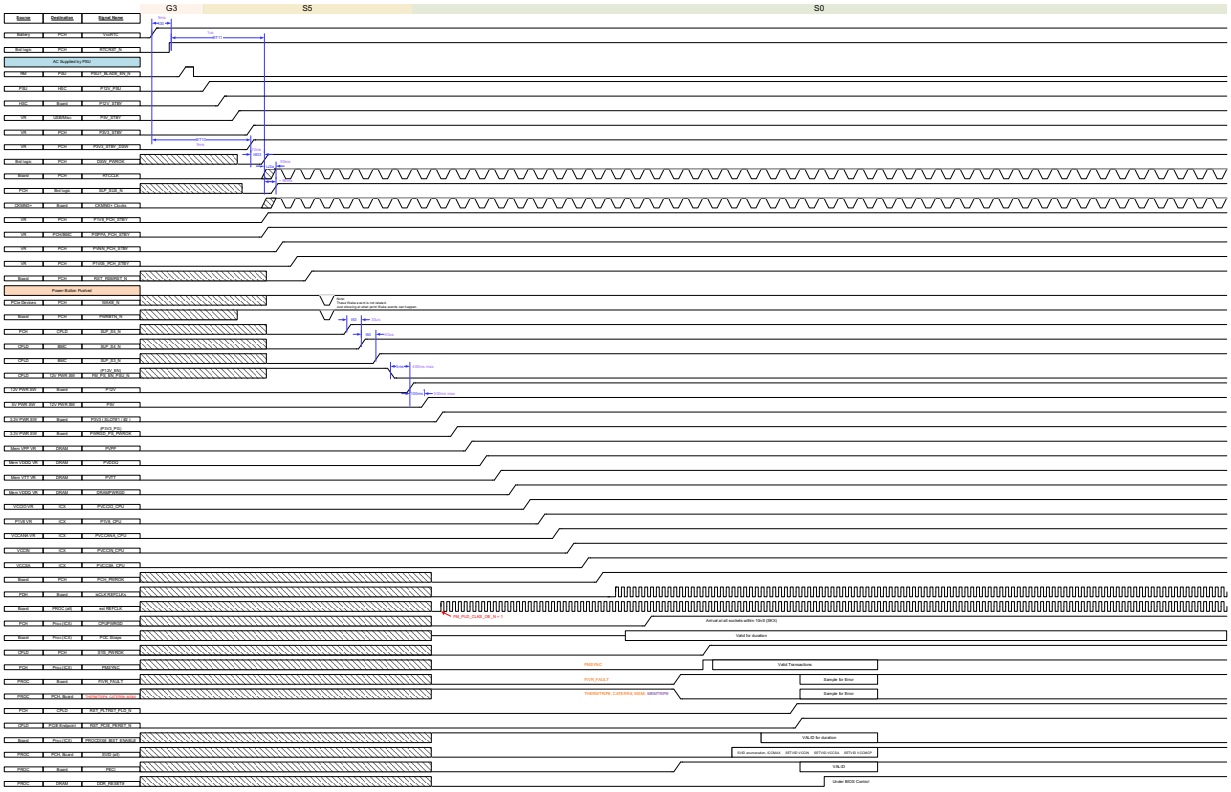


Figure 47: Angels Landing Motherboard Power on sequence

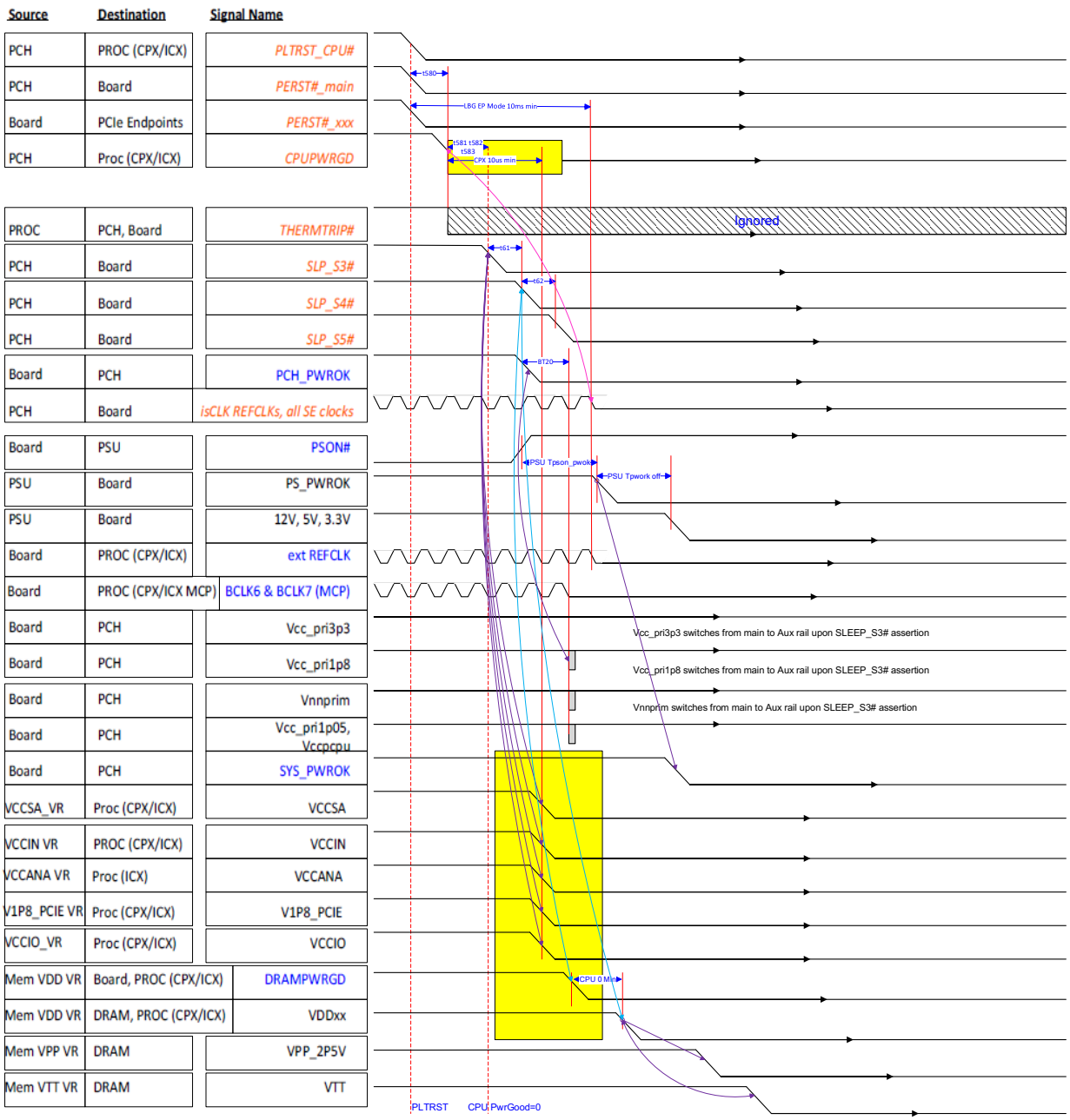


Figure 48: Angels Landing Motherboard Power off sequence

11.3 Input Voltage level

The expected nominal input voltage delivered by the power supply is 12.5 VDC; however, has a varying range of 11.5V to 13.5V. The motherboard shall accept and operate normally with an input voltage tolerance range between 11.25V and 13.75V.

11.4 Platform Power Budget

Each Angels Landing motherboard shall be designed to support a maximum sustained 1500W (12.5V nominal) of distributed power among its subsystems. The power budget has been defined in a way such that it allows for greater platform configuration flexibility, but design consideration be taken to ensure all aspects of the platform can be sufficiently cooled and without overstress of devices. Depending on the configuration of each motherboard, the total power budget is governed by the current carrying capability of various connectors as shown below:

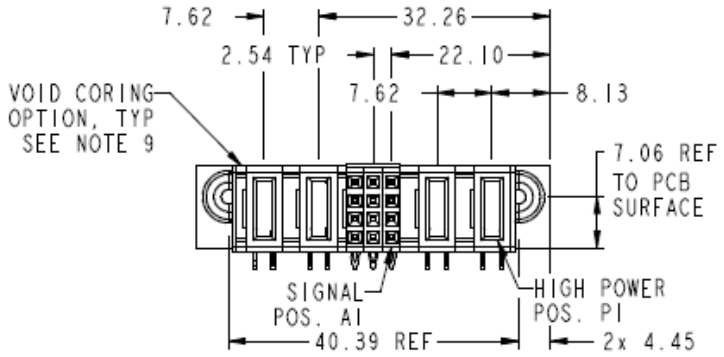


Table 1A: Current Rating (Amperes)

Configuration of PwrBlade+ Cable Assembly mated with PwrBlade+ Board Connector

Wire Size	High Power Contact					Low Power Contact			Signal Contact		
	1P	2P	4P	8P	12P	1LP	4LP	8LP	1S	24S	48S
6AWG	70	57	52	48	45	N/A			N/A		
10AWG	58	N/A	36	33	33						
12AWG	N/A					42	26	21			
14AWG						36	26	19			
16AWG						32	21	17			
22AWG	N/A								6A	3A	1A

Figure 49: Angels Landing Motherboard Power Connector

Note: Limits defined in table above are based on still air measurements taken at 25C ambient and current is increased until temperature rise exceeds 30C.

A dedicated HSC shall be implemented on each Angels Landing motherboard which will actively monitor the total board power consumption and report this sensor information to upstream devices such as BMC and/or PCH whereby an appropriate throttle or power policy may be applied if necessary (both aspects are further explained in subsequent sections).

11.5 Capacitive Load

To minimize the inrush current applied to the Open Rack V2 Power Shelf during initial rack power on and assertion of Angels Landing motherboards into a live bus bar, special design consideration around input capacitance must be considered:

1. Between the bus bar and input of any load switch or HSC, the placement of input capacitors is disallowed (if truly necessary, this must be thoroughly investigated and validated)
2. The placement of input capacitors at the output of any load switch or HSC is allowed, but the total sum among all server systems within the same power zone shall not exceed 226mF

11.6 Hot Swap Controller Circuit

Each Angels Landing motherboard shall implement a dedicated HSC which is expected to support the following features as mentioned below:

1. In-rush current control when motherboard is inserted and powered up
2. MOSFETs must be kept within safe operating area during all operational conditions such as power on/off and fault conditions
3. Signals that indicate power status, alerts, interrupts are expected to allow rapid response upon impeding fault conditions and/or warnings
4. Current limit protection for over current and short circuit whereby overcurrent threshold should be configured to 126A for the motherboard
5. Undervoltage and overvoltage protection shall be configured to 10.09V and 14.33V respectively
6. Default HSC response for fault conditions shall be latch off with retry as a stuff option
7. PMBus interface that supports the following features:
    a. Report voltage, current, and power (VIP) telemetry with accuracy of +/- 2.0% or better when operating above 10% of the maximum range
    b. Status registers that allow the definition of upper and lower critical thresholds for VIP of which are logged upon being triggered
    c. Implements a fast (<20us) overcurrent monitoring scheme that generates an alert based on a remotely programmable threshold that triggers system throttling (Fast PROCHOT#) either using the HSC itself or external circuits. The recommended threshold for Fast PROCHOT# shall be slightly lower than the overcurrent limit such that there is no tolerance overlap. Based on design tolerances, the final value for Fast PROCHOT# was determined to be 123A.

The voltage drop on the HSC current-sense resistor should be less than or equal to 25mV at full loading.

The power reporting of the HSC must be better than 2%, from 50W to full loading at room temperature. Further optimizations to power telemetry accuracy shall be performed through firmware based on characterized results from multiple boards based on entire load range and operating temperature requirements.

11.7 Power Reading and Capping

Since each Angels Landing motherboard implements its own dedicated HSC, it is responsible for monitoring its own power consumption from the single input power source. Like previous generation of 2S server designs, the PCH is responsible for the calculation of a one second

average based on the telemetry readings from the HSC which is periodically polled by the BMC for platform power management.

11.8 VR Efficiency

High efficiency Voltage Regulators (VRs) shall be used on the Angels Landing platform with at least 91% efficiency over the 30% to 90% load range. If higher efficiency VRs are available at additional cost and/or design complexity, then the vendor is encouraged to present the tradeoffs prior to implementation.

11.9 Power Policy

The power policy of each Angels Landing motherboard shall be configurable by the BMC to either Always On or Last Power State. When the power policy is Always On, the motherboard will be powered on automatically regardless of their last power state (default option). When the power policy is Last Power State, the server modules will restore the last power state after AC cycling.

11.10 P12V_PSU to GND Clearance

Design consideration must be taken when routing unprotected power planes such as P12V_PSU which is responsible for carrying current from the bus bar to input of HSC. Below are layout recommendations that should be followed when possible:

1. On same and adjacent layers, P12V_PSU shape to all other nets, including GND ≥ 40 mil
2. On different layers, from P12V_PSU shape to all other nets, including GDN ≥ 2 layers of dielectrics if overlapping
3. P12V_PSU traces are typically needed to provide biasing for HSC and related circuitry. Such traces must be ≤ 20 mil and have 40mil clearance to other signals on the same layer. On the adjacent layer, it is preferred to generate void in plane to provide clearance to P12V_PSU where there is no other tradeoff.

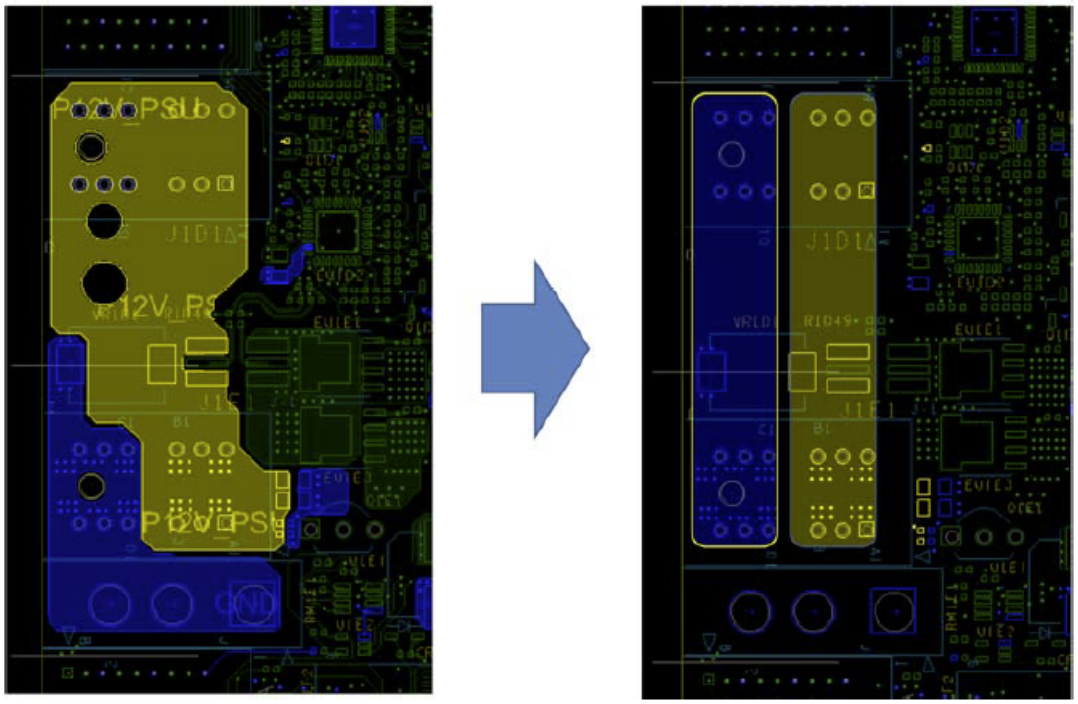


Figure 50: Example P12V_PSU and GND Layout

12. Mechanical

The Angels Landing system is a 2OU, Open Rack V2 compatible 4 CPU server. The system consists of two front accessible CPU trays each containing 2 CPUs. The CPU Trays communicate with one another using 4 orthogonal interposer cards. Two Interposers are also connected through high speed cable. Medusa power cables bring power directly from the rack busbar to the back of the CPU Trays and Power Distribution Board. The Power Distribution Board delivers power to the 4 rear serviceable, hot swappable Fan Modules.

The Angels Landing system communicates to other elements in the Zion Platform with PCIE cables attached to the rear edge of the CPU tray. The cables are routed around the sides of the CPU tray so that the cables are not removed when servicing the components on the CPU Tray. All commodities, PCBAs, and cables are field serviceable. Most field serviceable units can be toolessly accessed and swapped. The high speed PCIE cable swap requires a #2 Philips and a T10 screwdriver. All field serviceable units can be serviced without removing the system from the rack, except the Power Distribution Board, the Interposer Boards and the high speed cable in between the interposer boards.

The Angels Landing system can be configured in 4 socket or 2 socket configurations. In the 2 socket configuration, One CPU tray is removed and a blanking panel is populated. Also, in the 2 socket configuration, the Interposer boards and Interposer cables are removed.

For more details on the mechanical design, see the latest revision of the Angels Landing CAD available on the OCP Server Wiki.

13. Thermal Design Requirement

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. The thermal solution should be found by setting a high power target for initial design in order to avoid redesign of cooling solution; however, the final thermal solution of the system should be most optimized and energy efficient under data center environmental conditions with the lowest capital and operating costs. Thermal solution should not allow any overheating issue for any components in the system. Components with throttling features (such as CPU, Memory, OAM) shall not throttle due to any thermal issue under the following conditions.

- Inlet temperature lower than or equal to 35°C, and +/- 0.005 InH_2O datacenter pressurization with all FANs in each thermal zone running properly
- Inlet temperature lower than or equal to 35°C, and +/- 0.005 InH_2O datacenter pressurization with one FAN (or one rotor) in each thermal zone failed

13.1 Data Center Environmental Conditions

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

13.1.1 Location of Data Center/Altitude

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

13.1.2 Cold-Aisle temperature

In nominal conditions, data centers will maintain cold aisle temperatures between 18°C and 30°C (65°F to 85°F). The mean temperature in the cold aisle is 24°C with 3°C standard deviation. The cold aisle temperature in a data center may fluctuate minutely depending on the outside air temperature of the data center. Every component in the system must be cooled and maintained below its maximum spec temperature in any cold aisle temperature in a data center.

13.1.3 Cold/Hot Aisle Pressurization

In nominal conditions, data centers will maintain the cold/hot aisle pressure difference to be between -0.005 InH_2O and 0.005 InH_2O. The thermal solution of the system accommodates the worst-case operational pressurization in the data centers.

13.1.4. R.H

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

13.2. Server validation condition

13.2.1. System loading

The power consumption of individual components on the system motherboard will vary by application or by motherboard SKU. Plan of record worst case configuration for thermal and power delivery design is 2 x 125W TDP CPU with 48x 32G RDIMM for the whole system. The max supportable CPU TDP needs to be at least 250W.

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

13.2.2 DIMM Operation

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

13.2.3 Inlet temperature

The inlet air temperature will vary. The cooling system should cover inlet temperatures as 20ºC, 25ºC, 30ºC, 35ºC and 40ºC. Ambient temperatures above 30ºC are beyond Data Center general environment conditions but used during validation to demonstrate design margin. CPU throttling is not allowed over the range 20ºC ~ 35ºC.

13.2.4. Pressurization

In data center environment, the system may operate in under/over pressurized conditions, ranging between +0.005 InH2O and -0.005 InH2O, defined as pressure difference between cold aisle and hot aisle. The thermal design and fan speed control shall enable the system to operate in such environment without significant penalty on thermal margin or extra airflow consumption. In validation, pressure up to +/- 0.05InH2O were adopted to test robustness of the design and collect airflow variation data.

13.2.5. Fan Redundancy

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

memory throttling. System shall also allow a 30 seconds minimum (5minutes desired) time window for single fan service before throttling happens, when inlet temperature is at 30°C or below.

13.2.6. System airflow Budget

CFM/W - The unit of airflow (or volumetric flow) used for this spec is CFM (cubic feet per minute). The maximum allowable airflow-to-power ratio for the system must be 0.145 CFM/W at sea level when the inlet temperature is equal or lower than 30°C. The desired ratio is 0.1 CFM/W at such conditions.

Delta T - Another correlated metric is the air temperature difference across the system. Instead of direct measurement, it is derived based on the airflow rate, heat dissipation, and heat capacity of air (at specified temperature and altitude). The allowable server-level delta T is 12.2°C or higher when the inlet air temperature is equal to or lower than 30°C, which is equivalent to airflow-to-power ratio of 0.145 CFM/W or lower at sea level. Desired delta T is 17.8°C or higher, which is equivalent to 0.1 CFM/W or lower at sea level.

13.2.7 Thermal margin

The thermal margin is the difference between the maximum operational component temperature without performance penalty and the current temperature reading. The system operating at an inlet temperature between 30°C and 35°C shall have a minimum 2% thermal margin for every component in the system. The system operating at an inlet temperature up to 30°C shall have a minimum thermal margin of 7% for every component in the system.

The basis of 100% refers to the absolute temperature spec of the component at the unit of Celsius degree. For instance, with a component spec of 85°C, a margin of 3°C is considered as 3.5%.

13.3. Thermal kit requirements

Thermal testing must be performed up to 40°C inlet temperature to guarantee high temperature reliability.

13.3.1 Heat Sink

The heat sink design should choose to be the most optimized design with lowest cost. The heat sink design should be reliable and the most energy efficient design that satisfies all the conditions described above. The ODM can always propose alternative designs if there's performance or cost benefit. The heat sink should be without complex installation guidance, such as air-flow direction.

13.3.2 System Fan

The system fan must be highly power-efficient with enough airflow delivery capability to support worst case operation conditions. The propagation of vibration caused by fan rotation should be

minimized and limited. This system has 4 fan slots with frame size of 80x80x56mm, and connectors are designed for dual rotor models.
The operational fan power consumption for inlet temperature below 30ºC shall not exceed 5% of system power consumption excluding fan power.

System fans should not have back rush current in all conditions. System fan should have an inrush current of less than 11.13A on 12V per fan. When there is a step change on fan PWM signal from low PWM to high PWM, there should be less than 10% of overshoot or no overshoot for fan input current. System should stay within its power envelope in all conditions of fan operation.

13.3.3. Air-Duct

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

13.3.4. Thermal sensor

The maximum allowable tolerance of thermal sensors in the motherboard is ±5°C. For inlet temperature sensors this request is ±2°C.

14. Environmental Regulations

14.1 Environmental Requirements

The system shall meet the following environmental requirements:

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

The full system shall meet the following environmental requirements:

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

System would be deployed into datacenter with following environment.
Site 1:

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

Site 2:

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

Site 3:

- Temperature: 65F to 85F
- Humidity: 30% to 85%
- Altitude: 1800m (6000 feet)

14.2. Vibration & Shock

The system shall meet shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) & IEC721-3-(*) Standard & Levels; the testing requirements are listed in Table 9.2-1. Angels Landing server shall exhibit fully compliance to the specification without any electrical discontinuities during the vibration and shock tests. No physical damage or limitation of functional capabilities (as defined in this specification) shall occur to the Angels Landing server during the non-operational vibration and shock tests.

Table 14.2-1 Vibration and Shock Requirements

	Operating	Non-Operating
Vibration	0.5g acceleration, 1.5mm amplitude, 5 to 500 Hz, 10 sweeps at 1 octave / minute per each of the three axes (one sweep is 5 to 500 to 5 Hz)	1g acceleration, 3mm amplitude, 5 to 500 Hz, 10 sweeps at 1 octave / minute per each of the three axes (one sweep is 5 to 500 to 5 Hz)
Shock	6g, half-sine 11mS, 5 shocks per each of the three axes	12g, half-sine 11mS, 10 shocks per each of the three axes

14.3. Regulations

The vendor must provide CB reports of the Angels Landing and tray in component level. These documents are needed to have rack level CE. The sled should be compliant with RoHS and WEEE. Angels Landing's PCBs should have a UL 94V-0 certificate. The vendor should design an EMI panel kit and pass FCC Class A.

15. Labels and Markings

Angels Landing boards shall include labels on the component side, the labels shall not be placed in a way that may cause them to disrupt the system functionality or the air flow path. The following table lists the required labels:

Description	Type	Barcode Required?
Vendor P/N, S/N, REV (Revision would increment for any approved changes)	Adhesive label	Yes
Vendor Logo, Name & Country of Origin	Silk Screen	No
PCB Vendor Logo, Name	Silk Screen	No

Purchaser P/N	Adhesive label	Yes
Date Code (Industry Standard: WEEK / YEAR)	Adhesive label	Yes
RoHS Compliance	Silk Screen	No
WEEE symbol:  Angels Landing will have the crossed out wheeled bin symbol to indicate that it will be taken back by the Manufacturer for recycle at the end of its useful life. This is defined in the European Union Directive 2002/96/EC of January 27, 2003 on Waste Electrical and Electronic Equipment (WEEE) and any subsequent amendments.	Silk Screen	No
CE Marking	Silkscreen	No
UL Marking	Silkscreen	No
Vendor Asset Tag[1]	Adhesive label	Yes

16 Prescribed Materials

16.1. Disallowed Components

The following components shall not be used in the design of the Zion.

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

16.2. Capacitors & Inductors

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

- Only aluminum organic polymer capacitors made by high-quality manufacturers are used; they must be rated 105°C.
- All capacitors have a predicted life of at least 50,000 hours at 45°C inlet air temperature, under 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 cracking).
- X7R ceramic material for SMT capacitors should be used by default and at minimum X6S for portions of design subject to thermal hotspots such as CPU and/or DIMM cavities. COG or NP0 type should be used for tolerance sensitive portions of design Conditional usage of X5R ceramic material must be based on evaluation of worst-case thermal conditions and upon approval from Facebook The following limitations apply to the use of inductors:
- Only SMT inductors may be used as the use of through-hole inductors is disallowed.

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

17. System level RAS Requirement

There is a need for improving RAS (Reliability, Availability and Servicibility) solution for this system. The goals of the RAS solution are:

- To improve the Services Reliability or Uptime. Key metric to measure the reliability is Annual Crash Rate (ACR) where the goal is to lower ACR. Target is to reduce hardware/firmware related ACR per year.
- To improve the Serviceability or Root Cause Analysis (RCA) capabilities. Key metric to measure the serviceability is RCA time. The goal is to reduce Root Cause Analysis (RCA) time. Target will be evaluated based on the field data.

In order to meet the target ACR reduction goal, Facebook infrastructure needs to consider the following two key areas:

- Deploy Predictive Failure Analysis (PFA) and Proactive Action. A robust telemetry infrastructure is required to deploy effective PFA algorithms. Telemetry solution would assist in detecting any type of service degradation (e.g., persistent memory corrected error due to stuck bit in a DRAM) including monitoring various types of HW corrected errors, monitoring any BW congestion or performance degradation, and monitoring any thermal events. Proactive actions are generally taken at cluster management level and could be:
    a. Node Health State Update (e.g., VM drain)
    b. Service migration away from the impacted machine and moving the impacted machine to “Maintenance” state for RCA.
- Improve System Reliability. The objective is to reduce the probability of crash and its impact on the service due to hardware failure. This is achieved by implementing various Reliability features as part of RAS technology. This allows minimizing the probability of a crash due to machine faults (e.g., memory errors, PCIe and Intel QPI/UPI link errors, PCIe Device errors, or silicon errors). RAS actions could be taken at HW level, FW level, or OS level. Key RAS action could be:
    a. HW managed (i.e., no FW or OS involvement)
        i. Intel® UPI Link Level Retry
        ii. Intel® UPI Dynamic Link width reduction
        iii. PCIe Link Retraining and Recovery
    b. FW+HW managed (i.e., no OS involvement)
        i. Map-out memory device failures at various granularity e.g., cell, bank, or rank.
        ii. Partial Memory Mirror Failover
    c. OS+FW+HW managed
        i. MCA Recovery
        ii. Memory page off-line

iii. PCIe Live Error Recovery or PCIe Enhanced Downstream Port Containment and Recovery

17.1 Error Reporting Requirements

Error Reporting is a key RAS feature and foundational building block to implement Telemetry, PFA and various “standard” and “advanced” RAS features. At the HW platform level, Error Reporting consists of two key components:

1. Error Logging: HW based error logging, e.g., MCA logs, AER logs, Memory CE logs, Intel® UPI logs, and IIO Error logs.
2. Error Signaling: HW based signaling method
    a. Corrected Error Signaling, e.g., CMCI, MSI, CSMI, ERROR_PIN [0], GPE, or SCI
    b. Uncorrected Error Signaling, e.g., MCERR, MSI, MSMI, ERROR_PIN [1/2], GPE, SCI, or NMI.

In case of a Firmware First (FWF) model of error reporting, HW first notifies to Firmware via SMI/MSMI/CSMI or ERROR_PIN [0/1/2] event. Here Firmware could involve run-time SMM handler, BMC FW, or SPS FW. Then FW handler optionally signals to OS using SCI, CMCI, MCERR, or NMI.

Once Firmware notifies the OS, then OS uses in-band methods to collect various error logs by either directly reading HW based log registers or FW created logs, e.g., ELOG, CPER logs. In case of corrected errors, OS can continue to monitor errors as part of telemetry infrastructure and once certain criteria is met, it can trigger appropriate action, e.g., OS page off-line.

An alternative is defined as Operating System First (OSF) model of error reporting, where HW first notifies to OS via CMCI, MCERR, and SCI/GPE event. OS could optionally make an ACPI call to Firmware for additional platform specific action.

Key error reporting requirements are described in a table below:

Table 17.1-1 Vibration and Shock Requirements

#	Category	Requirement	Comments
1	Compatibility	OS level Error Reporting should be compatible with industry standard specifications	Error reporting Specifications, e.g., PCIe AER, ACPI/APEI, UEFI CPER, MCA/CMCI
2	Performance Impact	OS level Error reporting should not cause system hang.	For example: Frequent intermittent HW errors resulting in triggering back-to-back SMI event. Node level implementation shall incorporate appropriate back-off methods such as disabling the SMI interrupt and configuring periodic SMI.
3		OS level Error reporting should not cause severe system performance loss.	For example: persistent memory cell fault resulting excessive error correction and triggering threshold based SMI/CMCI too frequently.

			Node level implementation shall incorporate methods to detect such persistent faults using leaky bucket and Moving time-window based error rate methods. Additionally shall trigger appropriate PFA action, e.g., page off-line, or RAS action, e.g., ADDDC/ADC.
4		Platform Firmware should try best effort to continue error report.	During normal operation it is expected that errors occur infrequently and intermittently, therefore, FW shall keep sufficiently low threshold to maximize the observability of even a small perturbation.
5	Operational	Platform must provide sufficient information for FRU isolation	Unified SEL Record and ACD Format for following types of errors: 1. Boot time Error - Ex. MRC, TPM, SSD security etc. 2. Run time Error - PCIe CE/UCE/Fatal, Memory CE/UCE etc. 3. Crash time Error – Autonomous Crash Dump (ACD)
6	Reporting Interface	Platform must implement appropriate HW, FW, and OS interfaced for maximum observability of the error events.	For example: HW based signaling method, e.g., CMCI, SMI, SCI, and ERROR_PIN [0]. In case of a Firmware First (FWF) model of error reporting, HW first triggers SMI or ERROR_PIN [0] to the SMM or BMC FW handler and then FW handler optionally signals to OS using SCI or CMCI. Once OS is notified then OS implements in-band method to monitor corrected errors.
7	Fault Coverage	Boot time, Run time, and Crash time error reporting must cover various types of errors detected by the HW and FW.	Boot time Errors: ● Memory Training/CE/UCE ● TPM initialization failure ● PCH/ME CE/UCE ● SSD initialization failure (security check) Run time Errors: ● Memory CE/UCE-fatal/UCE-recoverable ● Intel® UPI CE/UCE ● PCIe CE/UCE/Fatal ● Intel DMI CE/UCE ● Network Device CE/UCE ● CPU Cache CE/UCE/Catastrophic (CMCI/MCERR/IERR) ● Other Component specific
8	Modes of Error Reporting	Platform shall provide provisions for various modes of error reporting, e.g., Native OS, FWF, OSF	Operating Modes: ● Native OS mode aka IA-32 Legacy MCA mode ● Firmware First (FWF) mode aka EMCA gen2 mode ● OS First (OSF) mode where error event is first notified to OS and then OS/FW interface extracts additional error information for various telemetry and RAS actions Error Types: ● Corrected error ● Uncorrected non-fatal (aka Uncorrected Recoverable) ● Uncorrected fatal ● Uncorrected Catastrophic Signaling mode: ● Corrected errors can be signaled via CMCI, CSMI followed by CMCI, MSI, or SMI followed by SCI. ● Uncorrected errors can be signaled via MCERR, MSMI followed by MCERR, or SMI followed by NMI

			Error detection and logging logic: ● Memory sub-system ● Intel UPI/QPI ● IIO ● CPU core and Uncore Note: MCA logs reported via CPU are not necessary point to CPU as a faulty component. Often the source of fault is outside of the CPU.
9	Stage control of Error Reporting	NPI: Single error exposure with collateral information MP: Event based failure handling	In NPI, every single hardware component error should be reported individually, and follow Facebook error reporting format. In MP, platform shall provide the capability to manage thresholding of error rate and reporting of error as per Facebook error reporting format.

In addition to Error Reporting feature and various base-line RAS features, next section described additional RAS features required to enable in Facebook's 8-socket platforms designed based on Intel's Cedar Island (CPX6-SP) platforms.

17.2. RAS Feature List

RAS Feature	Settings
Memory	
Memory disable/map-out for FRB (S_MEM01)	Enable
Memory Single Device Data Correction (SDDC) (R_MEM01)	Disable
MEMHOT Pin support for reporting (S_MEM06)	Enable
Memory corrected error reporting (S_MEM07)	Disable(by user)
Memory Data Scrambling with Command and Address (R_MEM07)	Enable
Address range/partial mirroring (R_MEM14)	Disable
Memory Demand and Patrol Scrubbing (R_MEM08)	Enable
Memory Thermal Throttling (R_MEM12)	Enable
ADC(SR) (R_MEM22)	Disable
ADDDC(MR),+1 (R_MEM23)	Disable
Power-up DDR4 Post Package Repair (PPR) (R_MEM24)	Disable (by user)
Mem SMBus hang recovery (R_MEM25)	Enable
CPU	
Error detection and correction (coverage at the socket level) (R_CPU01)	Enable
PCIe Corrected Error Information (Error counters and Leaky-bucket logic) (S_CPU10)	Disable
Corrupt Data containment core (R_CPU03)	Enable
Corrupt Data containment uncore (R_CPU02)	Enable

Processor BIST(S_CPU05)	Enable
Thresholding for Corrected Errors (Intel UPI, PCIe) (S_CPU11)	Disable
UPI	
Intel UPI Protocol Protection via CRC (16bit) (R_QPI01)	Enable
Intel UPI Link Level Retry (R_QPI03)	Enable
Intel UPI Dynamic Link Width Reduction (R_QP04)	Enable
PCIE	
PCIe Link Retraining and Recovery (R_PCI01)	Enable
PCIe Link CRC Error Check and Retry (R_PCI02)	Enable
PCIe Card Hot Plug (Add/Remove/Swap) (S_PCI02)	Disable
PCIe Card Hot Plug Surprise (S_PCI03)	Disable
PCIe Corrupt Data Containment (Data Poisoning) (R_PCI03)	Enable
PCI Express ECRC (R_PCI05)	Enable
System	
Failed DIMM Isolation (S_SYS01)	Enable
Autonomous Crash Dump	Enable
Core Disable for FRB (S_SYS05)	Enable

18. Security

Angels Landing default security solution is still using TPM2.0 and verified boot. The system design has the hardware that could support Intel's PFR. It is possible to enable PFR in angels landing platform with additional software development work done.

All products seeking OCP Inspired™ or OCP Accepted™ Product Recognition shall have a completed Security Profile in the 2021 Supplier Requirements Checklist. Whether the answer is a yes or no, the profile must be completed. For Additional Security Badges (Bronze/Silver/Gold), please fill out the Security Profile in accordance with the requirements for that level. Security Badges will be reassessed on an annual basis as requirements are subject to change.

19. Reliability and Quality

19.1 Specification Compliance

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

19.2 Change Orders

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

19.3 Failure Analysis

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

19.4. Warranty

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

19.5. MTBF Requirements

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

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

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

19.6. Quality Control

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

- Incoming product must have less than 0.1% rejections
- Cpk values will exceed 1.33 (Pilot Build & Production)
- Vendors will implement a quality control procedure during Production, by sampling Angels Landings at random from the production line and running full test to prove ongoing compliance to the requirements. This process shall be documented and submitted prior to Production. The relative reports shall be submitted on an ongoing basis.
- Vendors will conduct an ongoing burn-in procedure for use in Production (Production will not start without an agreement on some sort of burn-in procedure). Vendors shall submit documentation detailing the burn in procedure.

19.7. Change Authorization and Revision Control

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

Any request for changes must be submitted to the customer with proper documentation showing details of the changes, and reason for the changes. This includes changes affecting form, fit, function, safety, or serviceability of the product. Major changes in the product (or in the manufacturing process) will require requalification and/or re-certification to the product. A new set of First Article Samples may be required to complete the Engineering Change Order (ECO) process. Any modifications after approval shall phase-in during production without causing any delays or shift of the current production schedule. Vendors shall provide enough advance notice to the customer to prevent any discontinuation of production.

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

19.8. PCB Tests

The vendor should arrange an independent third party lab testing on Delta-L, IST, and IPC-6012D for each baseboard and middle plane PCB from every PCB vendor. The server ODM cannot use the PCB vendor for these tests. The server ODM should submit reports for review and approval before a PCB vendor can be used in mass production. The testing lots should be manufactured at the same facility of a PCB vendor with the same process that is planned to be used by mass production.

SET2DIL/Delta-L requires 5x different PCB fabrication lots from the PCB vendor. Environmental shipping, packaging, and handling of this board is vital to test success; overnight shipping direct from the PCB vendor to SET2DIL/Delta-L independent lab is recommended.

IST is done once. It is required to be tested on a board manufactured at the same time as a board that completely passes SET2DIL/Delta-L. (Run SET2DIL/Delta-L, and if it passes then ask the IST lab to run IST on the board they receive.) The IST test profile is 3x cycles to 250°C and up to 1000x cycle to 150°C. Passing criteria is 150x cycles average, and 100x cycles minimum for 35x coupons.
IPC-6012D is done when 2x of the 5x SET2DIL tests pass from a PCB vendor. (Passing at the independent test lab)
The ODM should work with then PCB house to implement IST and SET2DIL/Delta-L coupon to break off panel without increasing unit cost of PCB.

19.9. Secondary Component

Secondary component planning should start from EVT and reach 80% of total number of BOM items in PCBA BOM in EVT. The rest of the secondary component should be included in DVT.
It is recommended that PCB is planned with three vendors at EVT. The EVT and DVT build plan should cover all possible combinations of key components of DC to DC VR including output inductor, MOSFETs and driver.

The vendor should provide second source plan and specification comparison before each build stage.

20. Deliverables

20.1. OS Support

Angels Landing system should support 64-bit CentOS version 8.1 and above with an updated Kernel specified by customers, and should pass Red Hat certification tests.

20.2 Accessories

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

20.3. Documentation

The vendor shall supply the following documentation to the customer:

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

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

20.4. Mass Production First Article Samples

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

- All the pertinent documentation described in section 15.3 and any other documents and reports, necessary for customer to release the product to mass production
- Pilot samples that are built in the allocated facility for mass production
- A full Specification Compliance Matrix
- A full Test/Validation Report
- Production line final Test ‘PASS’ tickets
- Samples that have passed the production burn-in process
- Samples shipped using the approved-for-production shipping box described in section 0.

21. Shipping

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

22. References (recommended)

[1] Facebook Zion System Specification
[2] Facebook Clear Creek System Specification
[3] Facebook Emerald Pools System Specification

Appendix A - Requirements for IC Approval (to be completed Contributor(s) of this Spec)

List all the requirements in one summary table with links from the sections.

Requirements	Details	Link to which Section in Spec
Contribution License Agreement	OCP-CLA	Link to Sec 1
Are All Contributors listed in Sec 1: License?	Yes	
Did All the Contributors sign the appropriate license for this spec? Final Spec Agreement/HW License?	Yes	
Which 3 of the 4 OCP Tenets are supported by this Spec?	Openness Efficiency Impact Scale	List reasons here. Link to presentation if separate.
Is there a Supplier(s) that is building a product based on this Spec? (Supplier must be an OCP Solution Provider)	Yes	List Supplier Name(s)
Will Supplier(s) have the product available for GENERAL AVAILABILITY within 120 days?	Yes	Please have each Supplier fill out Appendix B.

Appendix B-____ - OCP Supplier Information (to be provided by each Supplier of Product)

Company:
Contact Info:

Product Name:
Product SKU#:
Link to Product Landing Page:

Please complete the following 2021 Supplier Requirements. This link will allow you to create a copy for your product-specific requirements.

For OCP Inspired™,

- All Suppliers must be a OCP Solution Provider.
- All Suppliers must run the Hardware Management Conformance Checks and all products must meet the OCP Hardware Baseline Profile v1.0.0.
- All Suppliers must fill out a Security Profile (No Badge Level) for their product.

For OCP Accepted™, Supplier details are required.

- All Suppliers must be a OCP Solution Provider.
- All Suppliers must run the Hardware Management Conformance Checks and all products must meet the OCP Hardware Baseline Profile v1.0.0.
- All Suppliers must fill out a Security Profile (No Badge Level) for their product.
- All Products must meet the Open System Firmware requirements.
- All Products must have source code for BMC, if applicable. This must be in the OCP Github repository.

List all the requirements in one summary table with links from the sections.

Requirements	Details	Links
Which Product recognition?	OCP Accepted™ or OCP Inspired™	Provide Marketplace Link
If OCP Accepted™, who provided the Design Package?		Link
2021 Supplier Requirements for your product(s)		Link