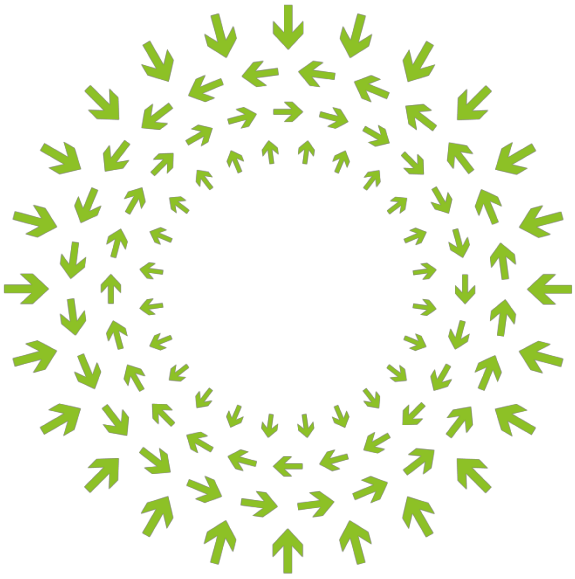

OPEN
Compute Project

Glacier Point V2 Card Design Specification

V0.1

Author:
Hao Shen, Hardware Engineer, Facebook
Pavan Shetty, Hardware Engineer, Facebook
Chenyu Xu, Mechanical engineer, Facebook
Jarrod Clow, Thermal Engineer, Facebook
Yueming Li, Thermal Engineer, Facebook
Ben Wei, Software Engineer, Facebook

Copyrights and Trademarks

1 Scope

This specification describes the hardware design details of the Glacier Point Version 2 (GPv2) card. GPv2 is a carrier card designed for accelerator application in Facebook's Yosemite V2 server system.

2 Contents

3 Overview

This document describes a card named Glacier Point V2 that is plugged in the Yosemite V2 platform. The Yosemite V2 Platform is a next generation multi-node server platform that hosts four Open Compute Platform (OCP) compliant One Socket (1S) server cards, or two sets of 1S server card and device card pairs in a sled that can be plugged into an OCP vCubby chassis, which is a new 4OU form factor design to easily service and accommodate thermal challenges with higher power 1S servers and device cards.

Glacier Point V2 card is one type of the device card that will pair with 1S server card to support accelerator and storage application. it is located on slot 1 and 3 in Yosemite V2 cubby and pair with Twin Lake CPU card which is located on slot 2 and 4. Slot 1 and 2 will form a subsystem; slot 3 and 4 will form another subsystem.

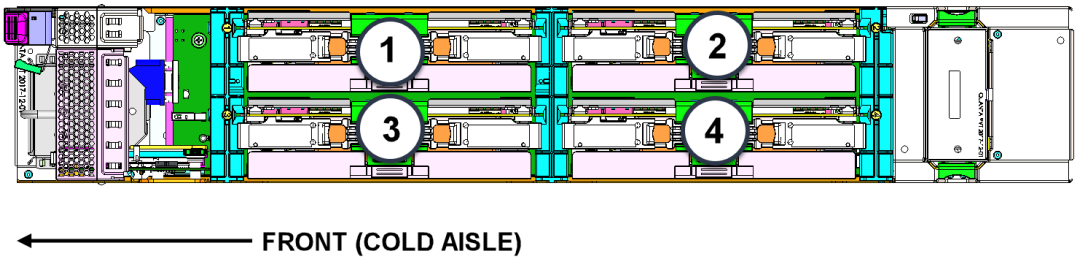

Figure 3-1 Yosemite V2 Cubby top view

The Glacier Point V2 implements primary and extension x16 PCIe edge connectors which is the same as the 1S server specification. The primary x16 PCIe edge connector supports:

- PCIe Gen3 ports x8
- A USB 2.0 port
- A Universal Asynchronous Receiver/Transmitter (UART) port
- An I^2C bus for server management
- A System Management Bus (SMBus) as the sideband of the integrated network controller
- Power and Ground Pins

The extension x16 PCIe edge connector supports:

- PCIe Gen3 ports x16
- Power and Ground Pins

The Glacier Point V2 supports up to 12x M.2 22110 module or up to 6x Dual M.2 modules. M.2 module form factor is defined in PCIe M.2 specification. That will support standard SSD and accelerator modules. Dual M.2 form factor is defined by Facebook. It can be simply treated as the combination of two M.2 form factor with 2mm pitch in between. Dual M.2 form factor provides more surface area and power which is a good extension of M.2 for applications that need bigger ASIC and higher power consumption.

Both M.2 form factor and Dual M.2 form factor has been customized to better support accelerator application. Please refer to M.2 and Dual M.2 accelerator hardware specification for details.

The Glacier Point V2 receives 12.5V from the platform with a maximum current load of 7.7A from the primary edge connector and an additional 7.7A from the extension edge connector. The platform, however, defines and controls the maximum power used by the Glacier Point V2. The Glacier Point V2 uses the INA230 power sensor at the power input to measure the total card power consumption of the whole server with +/-1% accuracy. The power data measured by this sensor can be used by the platform for power management purposes. When the power consumption has exceeded the power limit, the Bridge IC will generate a throttle event to force the all the modules on GPv2 board to throttle to the lowest power state. That throttle capability is defined in accelerator modules only. Storage modules do not support that capability.

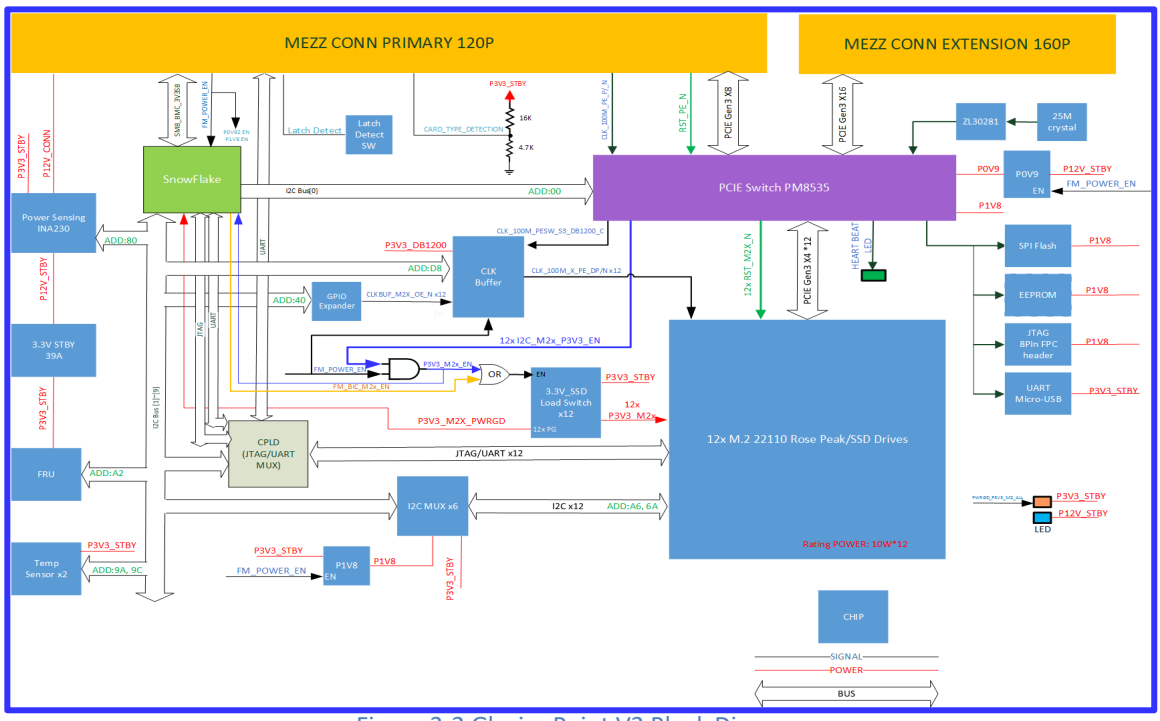

Figure 3-2 Glacier Point V2 Block Diagram

A Bridge IC (Texas Instrument's Tiva microcontroller, show as SnowFlake in Figure 3-2) is used as the management controller on the GPv2 card as the bridge between the BMC and the M.2 modules. The Bridge IC (BIC) manages the M.2 modules on behalf of the BMC. To maximize the communication bandwidth between the BMC and the Bridge IC, a dedicated point-to-point I^2C bus shall be used.

A PCIE Fanout switch (Microchip PM8535) is used to expand 24 PCIE Gen3 lanes to 48 PCIE Gen3 lanes to support up to 12x M.2 modules. The PCIE switch also have the DPC feature that will help to support the failure recovery of end point device.

Glacier Point V2's Field Replaceable Unit (FRU) EEPOM and thermal sensors are connected to the Bridge IC's I^2C buses. Each M.2 module's I2C buses are connected to bridge IC though multiple I2C muxes. We have defined additional debug interface (UART/JTAG) for each M.2 module. Those debug interfaces are connected to the BIC with a CPLD served as a switch in between. Overall, through the interface between BIC, BMC can access the Glacier Point V2's thermal sensors, FRU, M.2 Module's sideband interface and debug interface with standard IPMI commands.

PCIE switch firmware, boot ROM of Bridge IC and other firmwares (CPLD and VRs) are programmable from out-of-band connectivity by the Bridge IC and BMC.

4 License

5 Mechanical Design

5.1 Mechanical Outline

The Glacier Point V2 card uses a riser card adapter with dual PCIe edge connector which is connected to the M.2 PCB by mezzanine connectors. Spacers are used to control the distance between PCBs.

The overall dimensions of the general card are 210 x 148mm (excluding the riser). See Figure 5-1 for the specification drawing including keep-out zones and component restrictions.

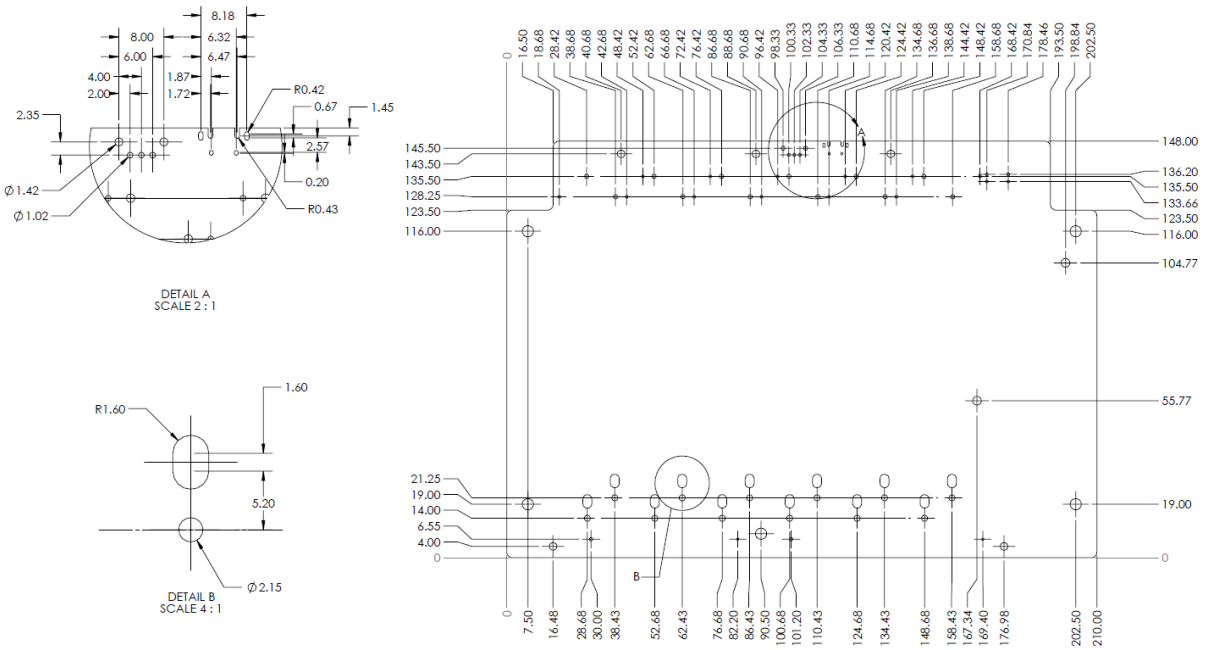

Figure 5-1 Glacier Point V2 PCB Mechanical Drawing

The PCB is mounted to a sheet metal carrier which includes flanges on each perpendicular edge that engage the card guides in the chassis. There are no card guide keep-outs on the PCB edge since all available space is needed for traces and components. The card guide engages the sheet metal flanges instead of the PCB itself.

The carrier assembly includes 2x ejectors which are used for card injection/ejection into the PCIe connectors. These ejectors rotate horizontally before being allowed to swing open, and they include finger access cutouts. One ejector claw engages a limit switch, which allows the BMC to detect if the ejector has been opened. The air duct rotates to open/close, one end fixed as the pivot point, the other end snaps into the retainer as a spring latch.

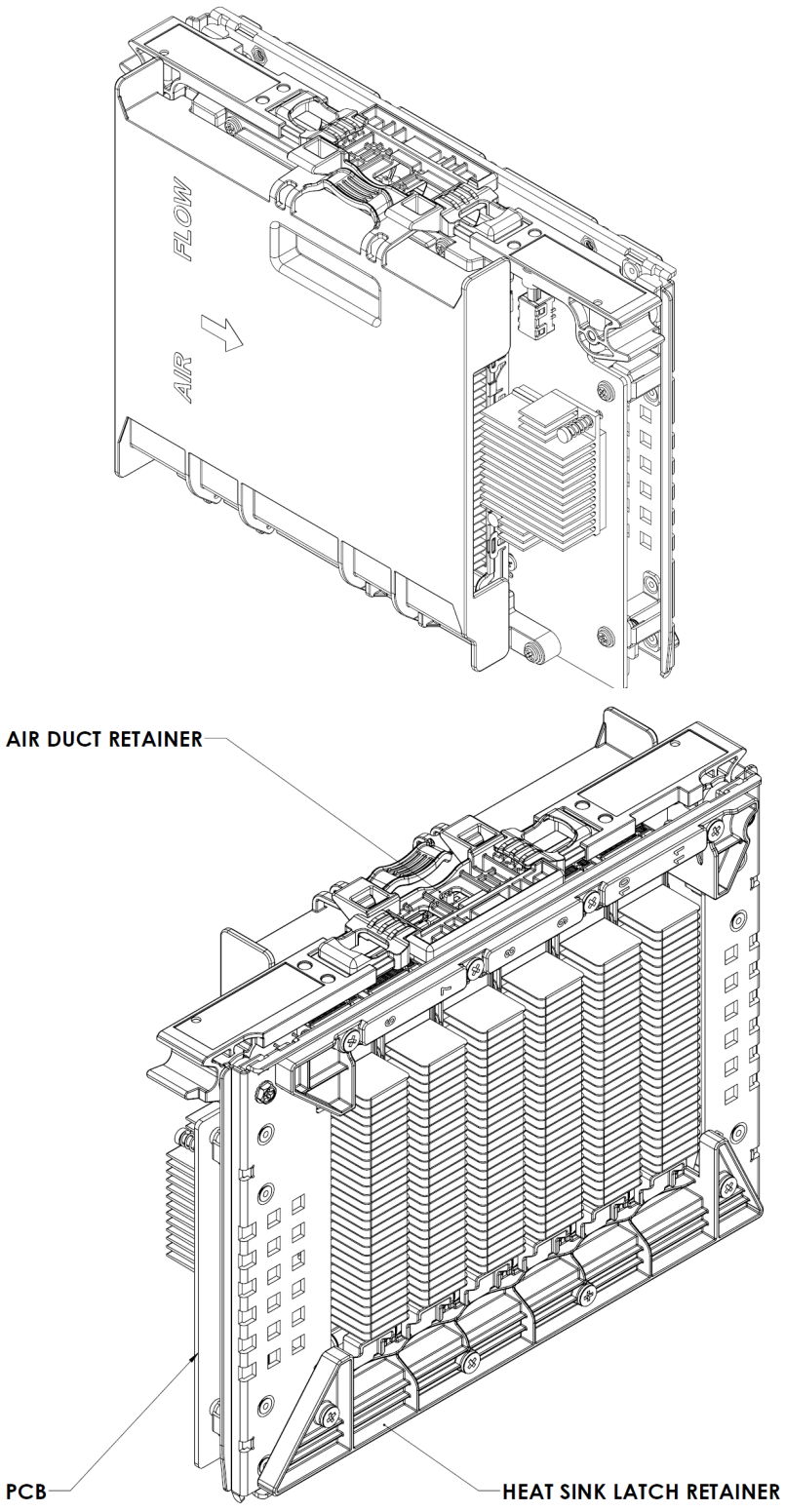

Figure 5-2 Glacier Point V2 Card Mechanical Overview

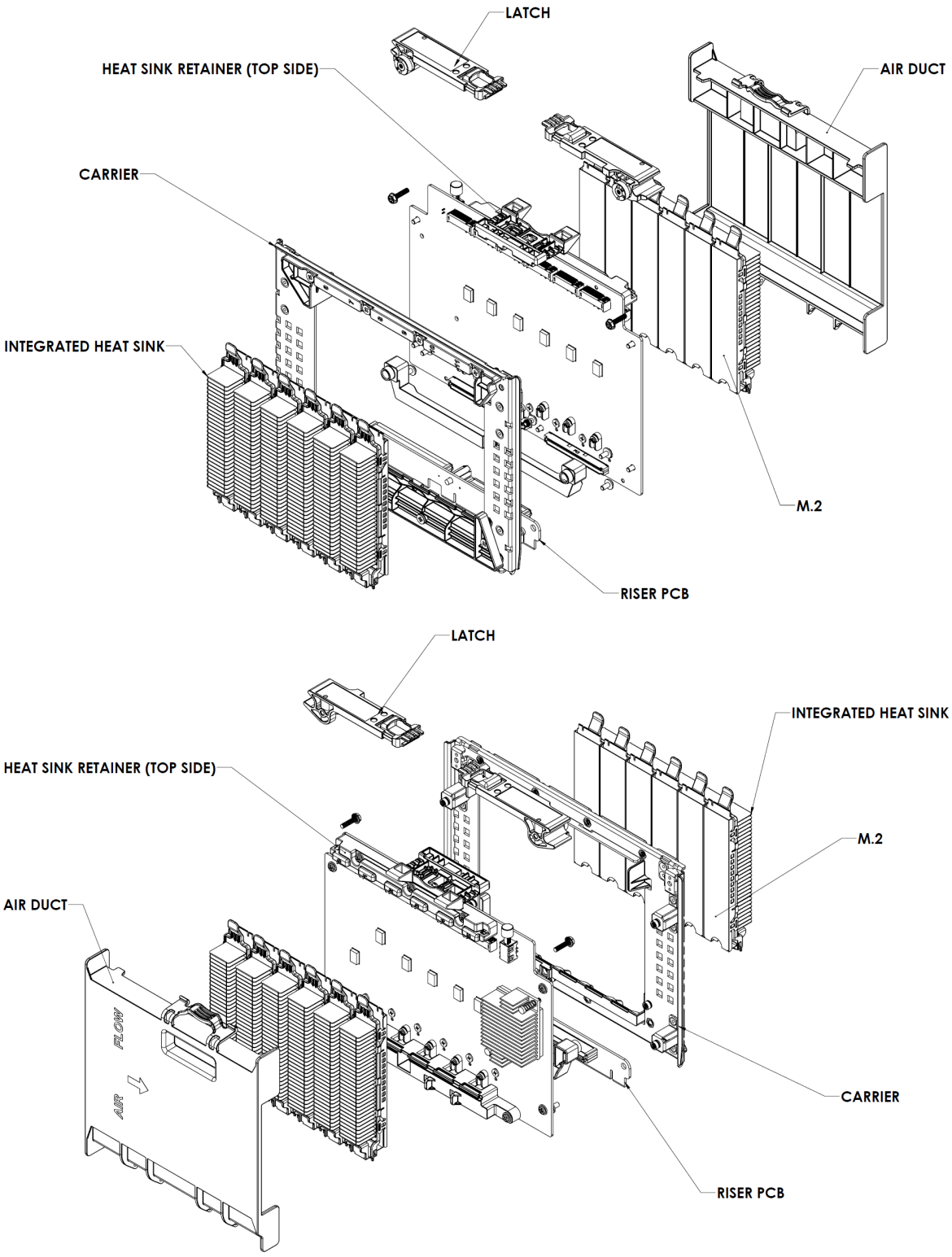

Figure 5-3 Exploded Views of Glacier Point V2 Card

An integrated heat sink was designed for M.2 modules. The M.2 PCBA plus the heat sink forms a whole FRU to replace separate thermal pad plus heatsink design in GPv1 system. This design greatly improved thermal performance and serviceability.

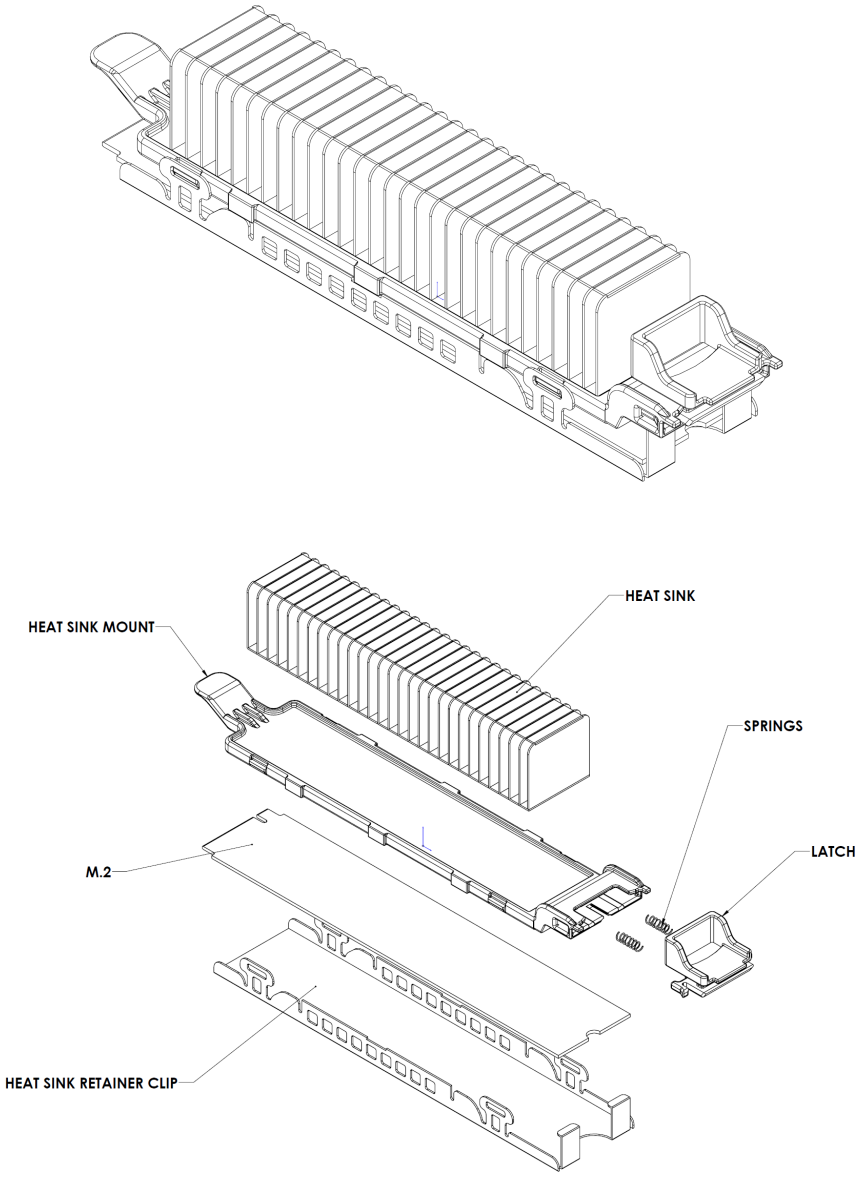

Figure 5-4 Integrated Heat Sink M.2 Module Assembly

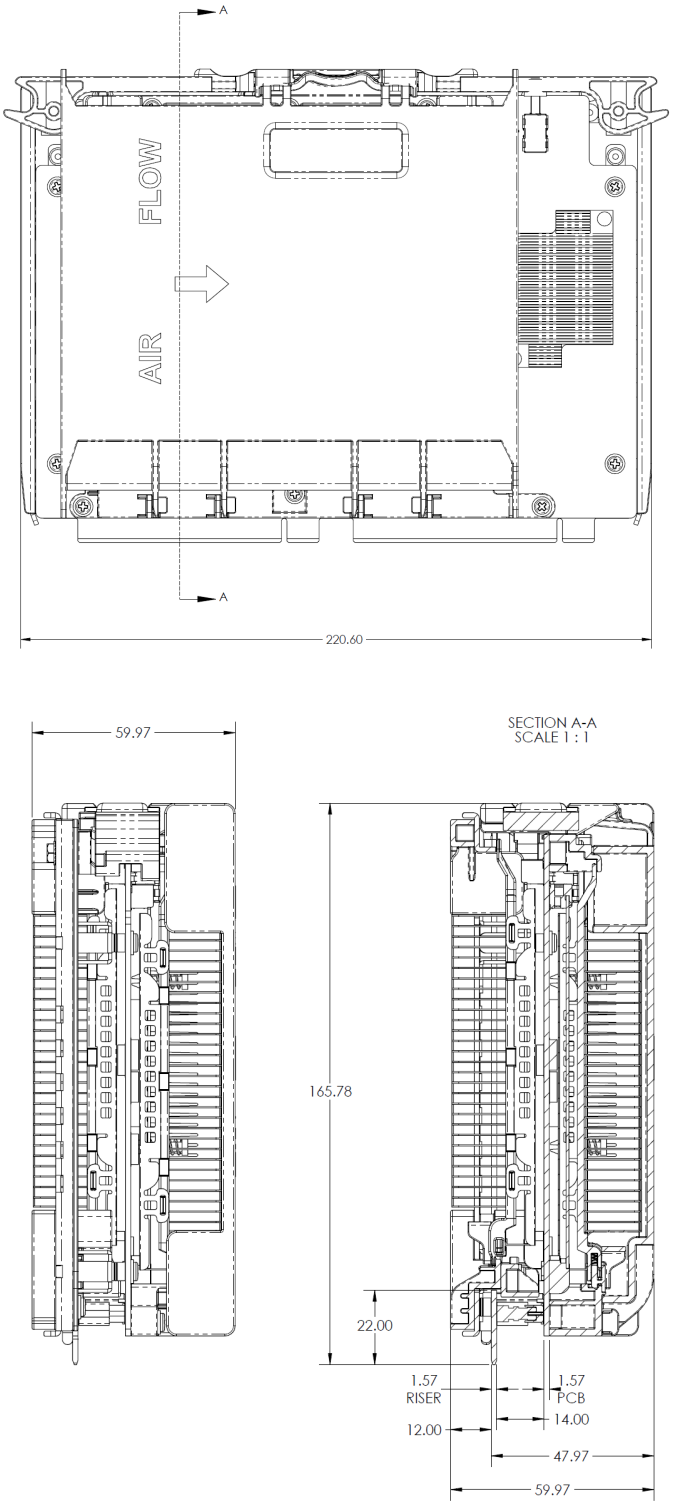

Figure 5-5 Glacier Point V2 Card Dimensions

The thickness of the PCB shall be 1.57±10% mm.

5.2 PCIe Edge Connector

The key dimensions, edge chamfer, pad layout (including a shorter pad for PRSNT# signal), placement, and dimensions of the card edge connector match the PCI Express Card electromechanical specification.

The GND planes underneath the pads for the card's edge connector must be recessed according to the PCI Express Card electromechanical specification to improve signal integrity.

5.3 Platform Design

Platform design details are not discussed in this specification.

6 Thermal Design

6.1 Data Center Environmental Conditions

This section outlines Facebook data center operational conditions.

6.1.1 Location of Data Center/Altitude

Maximum altitude is 6,000 ft above sea level. Any variation of air properties or environmental difference due to the high altitude needs to be deliberated into the thermal design.

6.1.2 Cold-Aisle Temperature

Data centers generally maintain cold aisle temperatures between 18°C and 30°C (65°F to 85°F). The mean temperature in the cold aisle is usually 25°C with 3°C standard deviation. The cold aisle temperature in a data center may fluctuate minutely depending to the outside air temperature. Every component must be cooled and must maintain a temperature below its maximum specification temperature in the cold aisle.

6.1.3 Cold-Aisle Pressurization

Data centers generally maintain cold aisle pressure between 0 inches H2O and 0.005 inches H2O. The thermal solution of the system should consider the worst operational pressurization possible, which generally is 0 inches H2O and 0.005 inches H2O with a single fan (or rotor) failure.

6.1.4 Relative Humidity

Data centers usually maintains a relative humidity between 20% and 90%.

6.2 Server Operational Conditions

6.2.1 System Volumetric Flow

The unit of airflow (or volumetric flow) used for this spec is cubic feet per minute (CFM). The CFM can be used to determine the thermal expenditure or to calculate the

approximate Delta T of the system. The thermal expenditure is quantified by the metric CFM/W, which is calculated by the following formula:

$$\text{Thermal Expenditure} = \frac{\text{System airflow}}{\text{Total system power consumption, including fans}} \ [\text{CFM/W}]$$

The required airflow is 0.115 airflow per watt in the system level at sea level. The desired airflow per watt is 0.1 or lower up to 35°C (113°F) ambient temperature at sea level.

6.2.2 Thermal Margin

The thermal margin is the difference between the maximum theoretical safe temperature and the actual temperature. The board design operates at an inlet temperature of 35°C (95°F) outside of the system with a minimum 5% thermal margin for every component on the card.

6.2.3 Upper Critical Threshold

The upper critical threshold (UCT) setting should allow the detection of abnormal thermal behaviors in the system. The UCT values for the sensors that are not used in Fan Speed Control (FSC) should use a 15% thermal margin from the worst experiment data. The UCT values for the sensors used in FSC, except for CPU, inlet, and outlet sensors, should use a 20% thermal margin from the worst experiment data.

6.2.4 Thermal Testing

Thermal testing must be performed at a low inlet temperature 15°C (59°F) and up to 35°C (95°F) inlet temperature to guarantee the design is free of thermal defect and has high temperature reliability.

6.3 M.2 Cooling Solution

Each card must support up to 12 M.2 or 6 dual M.2. Each M.2 or dual M.2 must have a thermal solution that capable of maintaining all components within their thermal spec with desired thermal margin. The thermal design must be thermally optimized design at the lowest cost. Passive cooling is required with a TIM (Thermal Interface Material) and heat sink. Integrated heat sink solution is desired for improved thermal efficiency and serviceability.

6.3.1 M.2 and Dual M.2 connectors

All M.2s and dual M.2s should use 6.7mm height connectors to allow for airflow for top side and bottom side components.

6.3.2 TIM Requirements

The TIM must be used between top side components and heatsink, and between bottom side components and heatsink retainer clip to fill the gap. The TIM should be soft and

compliant transferring low or no pressure between interfaces. The TIM must have a thermal conductivity of 3W/m-K or higher. A desired TIM bond-line thickness should be less or equal than 0.5mm for components with high power density.

6.3.3 Heat Sink Requirements

Each M.2 or dual M.2 must have a heatsink that is integrated to the M.2 PCB board. The heatsink must be designed to help maintain temperatures of all major components on the M.2 or dual M.2 with desired thermal margin.

6.4 Temperature and Power Sensors

Each card must provide following sensors:

- Temperature sensors for M.2, Dual M.2, PCIe switch and ambient temperature.
- Power sensors for the SOC and the whole card
- Voltage sensors for all voltage rails
- One inlet ambient temperature sensor and one outlet ambient temperature sensor

The BMC on the platform must be able to read all these sensors via the Bridge IC. Additionally, over-temperature thresholds must be configurable and an alert mechanism must be provided to enable thermal shutdown and/or an increase in airflow. The sensors are accurate to +/-2°C and desired to be within 2% tolerance across whole operation temperature range. The goal sensor accuracy is +/-1°C.

Two ambient temperature sensors are placed along the edges of the card on the B-side (the side without PCIe switch). Figure 6-1 depicts the desired areas for ambient temperature sensor placement.

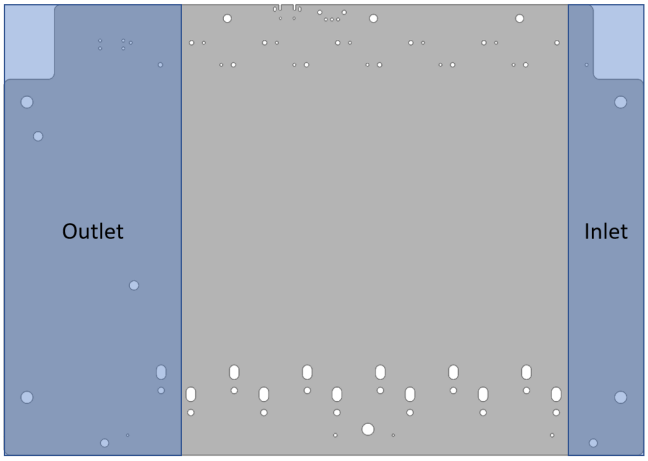

Figure 6-1 Ambient Temperature Sensor Placement Location (Circled Areas in Blue)

6.5 Carrier Card Air Baffle:

The card level air baffle must be designed to help maintain temperatures of all major components on the carrier card by reducing bypass air and increasing airflow through key

components. The air baffle must be easy to service with the goal of requiring no tooling to remove. The air baffle must not have a large adverse effect on system level pressure drop.

7 Electrical Design

7.1 Primary X16 Edge Connector Pinout

The Glacier Point V2 contains two x16 standard PCIe connector, we defined them as primary and second side. The location of as defined in Figure 7-1 below. The PCIe connects sit on the riser card (green). The carrier card (yellow) has the mezzanine connector to riser card.

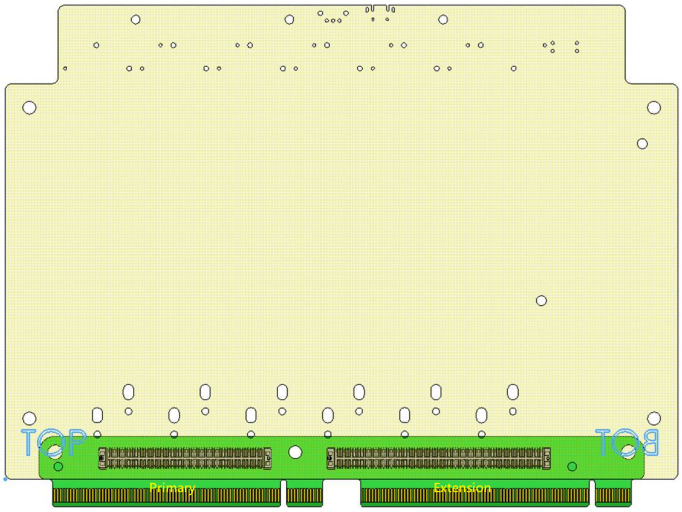

Figure 7-1 location of primary and extension connector of GPv2 Card

The definition of Edge Connector is based on Twin lake server card. Differences are the KR, NCSI and NIC SMBus interface are not connected in GPv2 design.

Table7-1: Glacier Point V2 Primary X16 OCP Edge Connector A Pin-Out

Connector A Default Pin-Out			
Pin Name	**B Side**	**A Side**	**Pin Name**
P12V_CONN	B1	A1	FLASHCARD_PRSNT_N
P12V_CONN	B2	A2	P12V_CONN
P12V_CONN	B3	A3	P12V_CONN
GND	B4	A4	GND
SMB_BMC_3V3SB_CLK_R	B5	A5	CARD_TYPE_DETECTION
SMB_BMC_3V3SB_DAT_R	B6	A6	P3V3_M2_EN_R
GND	B7	A7	COM_GP_TX
NC	B8	A8	COM_GP_RX
NC	B9	A9	LATCH_DETECT

NC	B10	A10	NC
NC	B11	A11	NC
SMB_INA230_ALRT_N	B12	A12	GND
GND	B13	A13	NC
GND	B14	A14	NC
NC	B15	A15	GND
NC	B16	A16	GND
GND	B17	A17	NC
GND	B18	A18	NC
NC	B19	A19	GND
NC	B20	A20	GND
GND	B21	A21	NC
GND	B22	A22	NC
NC	B23	A23	GND
NC	B24	A24	GND
GND	B25	A25	NC
GND	B26	A26	NC
NC	B27	A27	GND
NC	B28	A28	GND
GND	B29	A29	NC
GND	B30	A30	NC
NC	B31	A31	GND
NC	B32	A32	GND
GND	B33	A33	NC
GND	B34	A34	NC
CLK_100M_PE_DP	B35	A35	GND
CLK_100M_PE_DN	B36	A36	GND
GND	B37	A37	NC
GND	B38	A38	NC
RST_PE_SLOT1_1_RESET_N	B39	A39	GND
RST_PE_N_R	B40	A40	GND
GND	B41	A41	NC
GND	B42	A42	NC
NC	B43	A43	GND
NC	B44	A44	GND
GND	B45	A45	NC

GND	B46	A46	NC
NC	B47	A47	GND
NC	B48	A48	GND
GND	B49	A49	P3E_A_TX_DP<0>
GND	B50	A50	P3E_A_TX_DN<0>
P3E_A_RX_DP<0>	B51	A51	GND
P3E_A_RX_DN<0>	B52	A52	GND
GND	B53	A53	P3E_A_TX_DP<1>
GND	B54	A54	P3E_A_TX_DN<1>
P3E_A_RX_DP<1>	B55	A55	GND
P3E_A_RX_DN<1>	B56	A56	GND
GND	B57	A57	P3E_A_TX_DP<2>
GND	B58	A58	P3E_A_TX_DN<2>
P3E_A_RX_DP<2>	B59	A59	GND
P3E_A_RX_DN<2>	B60	A60	GND
GND	B61	A61	P3E_A_TX_DP<3>
GND	B62	A62	P3E_A_TX_DN<3>
P3E_A_RX_DP<3>	B63	A63	GND
P3E_A_RX_DN<3>	B64	A64	GND
GND	B65	A65	P3E_B_TX_DP<0>
GND	B66	A66	P3E_B_TX_DN<0>
P3E_B_RX_DP<0>	B67	A67	GND
P3E_B_RX_DP<1>	B68	A68	GND
GND	B69	A69	P3E_B_TX_DP<1>
GND	B70	A70	P3E_B_TX_DN<1>
P3E_B_RX_DN<0>	B71	A71	GND
P3E_B_RX_DN<1>	B72	A72	GND
GND	B73	A73	P3E_B_TX_DP<2>
GND	B74	A74	P3E_B_TX_DN<2>
P3E_B_RX_DP<2>	B75	A75	GND
P3E_B_RX_DN<2>	B76	A76	GND
GND	B77	A77	P3E_B_TX_DP<3>
GND	B78	A78	P3E_B_TX_DN<3>
P3E_B_RX_DP<3>	B79	A79	GND
P3E_B_RX_DN<3>	B80	A80	GND
GND	B81	A81	P12V_CONN
GND	B82	A82	P12V_CONN

7.2 Extension X16 Edge Connector B

The Glacier Point V2 also implements an extension x16 edge connector to bring out additional x16 PCIe lanes. This extension X16 Edge connector is referred as Connector B.

Table 2: Glacier Point V2 Extension X16 OCP Edge Connector B Pin-Out

Connector B Default Pin-Out			
Pin Name	**B Side**	**A Side**	**Pin Name**
P12V_CONN	B1	A1	FLASHCARD_B_PRSNT_N
P12V_CONN	B2	A2	P12V_CONN
P12V_CONN	B3	A3	P12V_CONN
GND	B4	A4	GND
NC	B5	A5	NC
NC	B6	A6	NC
NC	B7	A7	NC
NC	B8	A8	GND
NC	B9	A9	NC
GND	B10	A10	NC
NC	B11	A11	GND
NC	B12	A12	GND
GND	B13	A13	RESERVED_USB+
GND	B14	A14	RESERVED_USB-
P3E_F_RX_DP<0>	B15	A15	GND
P3E_F_RX_DN<0>	B16	A16	GND
GND	B17	A17	P3E_F_TX_DP<0>
GND	B18	A18	P3E_F_TX_DN<0>
P3E_F_RX_DP<1>	B19	A19	GND
P3E_F_RX_DN<1>	B20	A20	GND
GND	B21	A21	P3E_F_TX_DP<1>
GND	B22	A22	P3E_F_TX_DN<1>
P3E_F_RX_DP<2>	B23	A23	GND
P3E_F_RX_DN<2>	B24	A24	GND
GND	B25	A25	P3E_F_TX_DP<2>
GND	B26	A26	P3E_F_TX_DN<2>
P3E_F_RX_DP<3>	B27	A27	GND
P3E_F_RX_DN<3>	B28	A28	GND
GND	B29	A29	P3E_F_TX_DP<3>
GND	B30	A30	P3E_F_TX_DN<3>
NC	B31	A31	GND

NC	B32	A32	GND
GND	B33	A33	P3E_C_TX_DP<0>
GND	B34	A34	P3E_C_TX_DN<0>
P3E_C_RX_DP<0>	B35	A35	GND
P3E_C_RX_DN<0>	B36	A36	GND
GND	B37	A37	P3E_C_TX_DP<1>
GND	B38	A38	P3E_C_TX_DN<1>
P3E_C_RX_DP<1>	B39	A39	GND
P3E_C_RX_DN<1>	B40	A40	GND
GND	B41	A41	P3E_C_TX_DP<2>
GND	B42	A42	P3E_C_TX_DN<2>
P3E_C_RX_DP<2>	B43	A43	GND
P3E_C_RX_DN<2>	B44	A44	GND
GND	B45	A45	P3E_C_TX_DP<3>
GND	B46	A46	P3E_C_TX_DN<3>
P3E_C_RX_DP<3>	B47	A47	GND
P3E_C_RX_DN<3>	B48	A48	GND
GND	B49	A49	P3E_D_TX_DP<0>
GND	B50	A50	P3E_D_TX_DN<0>
P3E_D_RX_DP<0>	B51	A51	GND
P3E_D_RX_DN<0>	B52	A52	GND
GND	B53	A53	P3E_D_TX_DP<1>
GND	B54	A54	P3E_D_TX_DN<1>
P3E_D_RX_DP<1>	B55	A55	GND
P3E_D_RX_DN<1>	B56	A56	GND
GND	B57	A57	P3E_D_TX_DP<2>
GND	B58	A58	P3E_D_TX_DN<2>
P3E_D_RX_DP<2>	B59	A59	GND
P3E_D_RX_DN<2>	B60	A60	GND
GND	B61	A61	P3E_D_TX_DP<3>
GND	B62	A62	P3E_D_TX_DN<3>
P3E_D_RX_DP<3>	B63	A63	GND
P3E_D_RX_DN<3>	B64	A64	GND
GND	B65	A65	P3E_E_TX_DP<0>
GND	B66	A66	P3E_E_TX_DN<0>
P3E_E_RX_DP<0>	B67	A67	GND
P3E_E_RX_DN<0>	B68	A68	GND
GND	B69	A69	P3E_E_TX_DP<1>
GND	B70	A70	P3E_E_TX_DN<1>

P3E_E_RX_DP<1>	B71	A71	GND
P3E_E_RX_DN<1>	B72	A72	GND
GND	B73	A73	P3E_E_TX_DP<2>
GND	B74	A74	P3E_E_TX_DN<2>
P3E_E_RX_DP<2>	B75	A75	GND
P3E_E_RX_DN<2>	B76	A76	GND
GND	B77	A77	P3E_E_TX_DP<3>
GND	B78	A78	P3E_E_TX_DN<3>
P3E_E_RX_DP<3>	B79	A79	GND
P3E_E_RX_DN<3>	B80	A80	GND
GND	B81	A81	P12V_CONN
NC	B82	A82	P12V_CONN

7.3 Pinout Definitions

Table 7-3 provides a detailed pin defition. In GPv2 card, the pin definition is very similar

Table 7-3: GPv2 Golden Finger Pin Definition

Pin	Direction	Required/ Configurable	Pin Definition
P12V_CONN	Input	Required	12VAUX power from platform
SMB_BMC_3V3SB_CLK_R	Input/Output	Required	I^2C clock signal. I^2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the platform.
SMB_BMC_3V3SB_DAT_R	Input/Output	Required	I^2C data signal. I^2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the platform.
SMB_INA230_ALRT_N	Output	Required	I^2C alert signal. Alerts the BMC that an event has occurred that needs to be processed. 3.3VAUX signal. Pull-up is provided on the platform.
FLASHCARD_PRSNT_N	Output	Required	Present signal. This is pulled low on the card to indicate that a card is installed. 3.3VAUX signal. Pull-up is provided on the platform.

FLASHCARD_B_PRSNT_N	Output	Required	Extension edge connector Present signal. This is pulled low on the card to indicate that a card is installed. 3.3VAUX signal. Pull-up is provided on the platform.
COM_GP_TX	Output	Required	Serial transmit signal. Data is sent from the 1S Server module to the BMC. 3.3VAUX signal.
COM_GP_RX	Input	Required	Serial receive signal. Data is sent from the BMC to the 1S Server module. 3.3VAUX signal.
CARD_TYPE_DETECTION (GPIO0)	Output	Required	CARD_TYPE is an output signal to inform platform that if this card is a server or a device. For a server, this pin should be tied to GND through a 10K resistor on the 1S server card.
P3V3_M2_EN_R (GPIO1)	Input	Required	It is an input signal from a 1S server to enable active power on GPv2 side. It is useful when 1S server is going through DC cycling or AC cycling but the server and devices are not in the same power domain. Active high, 3.3VAUX signal.
LATCH_DETECT (GPIO2)	Output	Required	LATCH_DETECT is an output signal to indicate if the 1S server is fully seated with ejector latch closed and ready for power on. Platform designer can use this signal to control the power to the 1S server and avoid surprise 1S server insertion/removal to/from a hot slot. Active low, 3.3VAUX signal, pull-up should be provided on the platform.
RST_PE_N_R	Input	Required	PCIe reset signal. If a PCIe bus is connected, this signal provides the reset signal indicating the card VRs and clocks are stable when driven high to 3.3V.
P3E_A_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_A_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.

CLK_100M_PE_DP/DN	Input	Required	PCIe reference clock. This signal may or may not be connected on the platform.
RST_PE_SLOT1_1_RESET_N	Output	Required	YV2 UART buffer enable signal. If GPv2 has worked, this signal provides the enable signal for the UART MUX output to YV2 baseboard.
P3E_B_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_B_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_C_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_C_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_D_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_D_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_E_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_E_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.

P3E_F_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_F_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
RESERVED_USB+/-	Input/Output	Required	USB 2.0 differential pair.
GND	Input	Required	GND from platform
Pin	Direction	Required/ Configurable	Pin Definition
P12V_CONN	Input	Required	12VAUX power from platform
SMB_BMC_3V3SB_CLK_R	Input/Output	Required	I^2C clock signal. I^2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the platform.
SMB_BMC_3V3SB_DAT_R	Input/Output	Required	I^2C data signal. I^2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the platform.
SMB_INA230_ALRT_N	Output	Required	I^2C alert signal. Alerts the BMC that an event has occurred that needs to be processed. 3.3VAUX signal. Pull-up is provided on the platform.
FLASHCARD_PRSNT_N	Output	Required	Present signal. This is pulled low on the card to indicate that a card is installed. 3.3VAUX signal. Pull-up is provided on the platform.
FLASHCARD_B_PRSNT_N	Output	Required	Extension edge connector Present signal. This is pulled low on the card to indicate that a card is installed. 3.3VAUX signal. Pull-up is provided on the platform.
COM_GP_TX	Output	Required	Serial transmit signal. Data is sent from the 1S Server module to the BMC. 3.3VAUX signal.
COM_GP_RX	Input	Required	Serial receive signal. Data is sent from the BMC to the 1S Server module. 3.3VAUX signal.
CARD_TYPE_DETECTION (GPIO0)	Output	Required	CARD_TYPE is an output signal to inform platform that if this card is a server or a device. For a server, this pin should be tied to GND through a 10K resistor on the 1S server card.

P3V3_M2_EN_R (GPIO1)	Input	Required	It is an input signal from a 1S server to enable active power on GPv2 side. It is useful when 1S server is going through DC cycling or AC cycling but the server and devices are not in the same power domain. Active high, 3.3VAUX signal.
LATCH_DETECT (GPIO2)	Output	Required	LATCH_DETECT is an output signal to indicate if the 1S server is fully seated with ejector latch closed and ready for power on. Platform designer can use this signal to control the power to the 1S server and avoid surprise 1S server insertion/removal to/from a hot slot. Active low, 3.3VAUX signal, pull-up should be provided on the platform.
RST_PE_N_R	Input	Required	PCIe reset signal. If a PCIe bus is connected, this signal provides the reset signal indicating the card VRs and clocks are stable when driven high to 3.3V.
P3E_A_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_A_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
CLK_100M_PE_DP/DN	Input	Required	PCIe reference clock. This signal may or may not be connected on the platform.
RST_PE_SLOT1_1_RESET_N	Output	Required	YV2 UART buffer enable signal. If GPv2 has worked, this signal provides the enable signal for the UART MUX output to YV2 baseboard.
P3E_B_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_B_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.

P3E_C_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_C_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_D_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_D_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_E_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_E_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
P3E_F_TX _DP/DN<0>/<1>/<2>/<3>	Iutput	Required	PCIe x4 bus transmit signals. Data is sent from the 1S Server module(Twin Lakes) to the platform(GPv2). These signals may or may not be connected on the platform.
P3E_F_RX _DP/DN<0>/<1>/<2>/<3>	Output	Required	PCIe x4 bus receive signals. Data is sent from the platform(GPv2) to the 1S Server module(Twin Lakes). These signals may or may not be connected on the platform.
RESERVED_USB+/-	Input/Output	Required	USB 2.0 differential pair.

7.4 GPv2 Riser Card Mezzanine Connector Pin Definition

To offset the GPv2 card to the middle for better thermal solution, GPv2 card is assembled with GPv2 riser card. The golder finger is on the GPv2 riser card. GPv2 carrier card is

connected with GPv2 riser card with two mezzanine connectors with 14mm stackup height. The location of two mezzanine connector is listed in the drawing below:

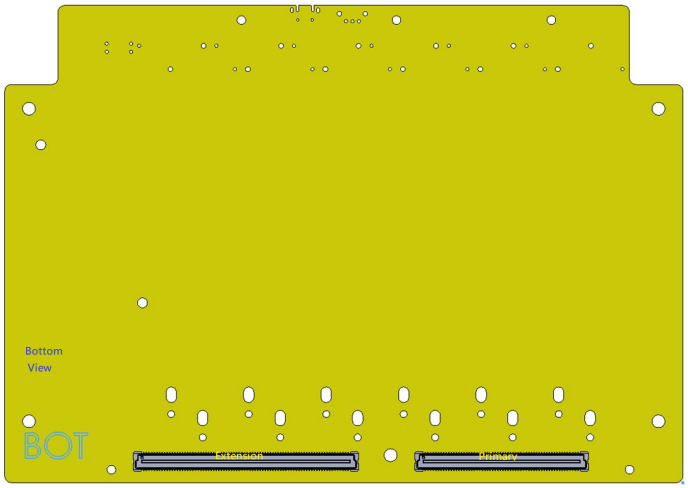

Figure 7-2 Location of mezzanine connector on GPv2 carrier board

Similar to the golden finger idea, we define mezzanine card with primary and extension one. It's pin definition is listed in Table 7-4:

Table 7-4a: GPv2 Carrier Card Mezzanine Connector Pin Definition (Primary)

Primary-120 pin Mezzanine			
Signal	Pin	Pin	Signal
P12V_CONN	2	1	FLASHCARD_PRSNT_N
P12V_CONN	4	3	P12V_CONN
P12V_CONN	6	5	P12V_CONN
P12V_CONN	8	7	P12V_CONN
P12V_CONN	10	9	P12V_CONN
P12V_CONN	12	11	P12V_CONN
P12V_CONN	14	13	P12V_CONN
NC	16	15	P12V_CONN
NC	18	17	NC
NC	20	19	NC
GND	22	21	NC
GND	24	23	GND
GND	26	25	GND
SMB_BMC_3V3SB_CLK	28	27	GND
SMB_BMC_3V3SB_DAT	30	29	GND
SMB_INA230_ALRT_N	32	31	GND
RST_PE_N	34	33	CARD_TYPE_DETECTION
GND	36	35	COM_GP_TX
GND	38	37	COM_GP_RX

SMB_M2_CLK	40	39	GND
SMB_M2_DAT	42	41	GND
GND	44	43	SMB_M2_ALERT_N
GND	46	45	NC
CLK_100M_PE_DP	48	47	GND
CLK_100M_PE_DN	50	49	GND
GND	52	51	P3V3_M2_EN
GND	54	53	LATCH_DETECT
P3E_2_PESW_RX0_C_DP	56	55	GND
P3E_2_PESW_RX0_C_DN	58	57	GND
GND	60	59	P3E_2_PESW_TX0_DP
GND	62	61	P3E_2_PESW_TX0_DN
P3E_2_PESW_RX1_C_DP	64	63	GND
P3E_2_PESW_RX1_C_DN	66	65	GND
GND	68	67	P3E_2_PESW_TX1_DN
GND	70	69	P3E_2_PESW_TX1_DP
P3E_2_PESW_RX2_C_DP	72	71	GND
P3E_2_PESW_RX2_C_DN	74	73	GND
GND	76	75	P3E_2_PESW_TX2_DP
GND	78	77	P3E_2_PESW_TX2_DN
P3E_2_PESW_RX3_C_DP	80	79	GND
P3E_2_PESW_RX3_C_DN	82	81	GND
GND	84	83	P3E_2_PESW_TX3_DP
GND	86	85	P3E_2_PESW_TX3_DN
P3E_1_PESW_RX0_C_DP	88	87	GND
P3E_1_PESW_RX0_C_DN	90	89	GND
GND	92	91	P3E_1_PESW_TX0_DP
GND	94	93	P3E_1_PESW_TX0_DN
P3E_1_PESW_RX1_C_DP	96	95	GND
P3E_1_PESW_RX1_C_DN	98	97	GND
GND	100	99	P3E_1_PESW_TX1_DP
GND	102	101	P3E_1_PESW_TX1_DN
P3E_1_PESW_RX2_C_DP	104	103	GND
P3E_1_PESW_RX2_C_DN	106	105	GND
GND	108	107	P3E_1_PESW_TX2_DP
GND	110	109	P3E_1_PESW_TX2_DN
P3E_1_PESW_RX3_C_DP	112	111	GND
P3E_1_PESW_RX3_C_DN	114	113	GND
GND	116	115	P3E_1_PESW_TX3_DP

GND	118	117	P3E_1_PESW_TX3_DN
NC	120	119	GND

Table 7-4b: GPv2 Carrier Card Mezzanine Connector Pin Definition (Extension)

Extension-160 pin Mezzanine			
Signal	Pin	Pin	Signal
P12V_CONN	2	1	FLASHCARD_B_PRSNT_N
P12V_CONN	4	3	P12V_CONN
P12V_CONN	6	5	P12V_CONN
NC	8	7	P12V_CONN
NC	10	9	NC
NC	12	11	NC
NC	14	13	NC
NC	16	15	NC
NC	18	17	NC
NC	20	19	NC
NC	22	21	NC
NC	24	23	GND
GND	26	25	RESERVED_USB+
GND	28	27	RESERVED_USB-
NC	30	29	GND
GND	32	31	P3E_6_PESW_TX0_DP
GND	34	33	P3E_6_PESW_TX0_DN
P3E_6_PESW_RX0_C_DP	36	35	GND
P3E_6_PESW_RX0_C_DN	38	37	GND
GND	40	39	P3E_6_PESW_TX1_DP
GND	42	41	P3E_6_PESW_TX1_DN
P3E_6_PESW_RX1_C_DP	44	43	GND
P3E_6_PESW_RX1_C_DN	46	45	GND
GND	48	47	P3E_6_PESW_TX2_DP
GND	50	49	P3E_6_PESW_TX2_DN
P3E_6_PESW_RX2_C_DP	52	51	GND
P3E_6_PESW_RX2_C_DN	54	53	GND
GND	56	55	P3E_6_PESW_TX3_DP
GND	58	57	P3E_6_PESW_TX3_DN
P3E_6_PESW_RX3_C_DP	60	59	GND
P3E_6_PESW_RX3_C_DN	62	61	GND
GND	64	63	P3E_3_PESW_TX0_DP
GND	66	65	P3E_3_PESW_TX0_DN

P3E_3_PESW_RX0_C_DP	68	67	GND
P3E_3_PESW_RX0_C_DN	70	69	GND
GND	72	71	P3E_3_PESW_TX1_DP
GND	74	73	P3E_3_PESW_TX1_DN
P3E_3_PESW_RX1_C_DP	76	75	GND
P3E_3_PESW_RX1_C_DN	78	77	GND
GND	80	79	P3E_3_PESW_TX2_DP
GND	82	81	P3E_3_PESW_TX2_DN
P3E_3_PESW_RX2_C_DP	84	83	GND
P3E_3_PESW_RX2_C_DN	86	85	GND
GND	88	87	P3E_3_PESW_TX3_DP
GND	90	89	P3E_3_PESW_TX3_DN
P3E_3_PESW_RX3_C_DP	92	91	GND
P3E_3_PESW_RX3_C_DN	94	93	GND
GND	96	95	P3E_4_PESW_TX0_DP
GND	98	97	P3E_4_PESW_TX0_DN
P3E_4_PESW_RX0_C_DP	100	99	GND
P3E_4_PESW_RX0_C_DN	102	101	GND
GND	104	103	P3E_4_PESW_TX1_DP
GND	106	105	P3E_4_PESW_TX1_DN
P3E_4_PESW_RX1_C_DP	108	107	GND
P3E_4_PESW_RX1_C_DN	110	109	GND
GND	112	111	P3E_4_PESW_TX2_DP
GND	114	113	P3E_4_PESW_TX2_DN
P3E_4_PESW_RX2_C_DP	116	115	GND
P3E_4_PESW_RX2_C_DN	118	117	GND
GND	120	119	P3E_4_PESW_TX3_DP
GND	122	121	P3E_4_PESW_TX3_DN
P3E_4_PESW_RX3_C_DP	124	123	GND
P3E_4_PESW_RX3_C_DN	126	125	GND
GND	128	127	P3E_5_PESW_TX0_DP
GND	130	129	P3E_5_PESW_TX0_DN
P3E_5_PESW_RX0_C_DP	132	131	GND
P3E_5_PESW_RX0_C_DN	134	133	GND
GND	136	135	P3E_5_PESW_TX1_DP
GND	138	137	P3E_5_PESW_TX1_DN
P3E_5_PESW_RX1_C_DP	140	139	GND
P3E_5_PESW_RX1_C_DN	142	141	GND
GND	144	143	P3E_5_PESW_TX2_DP

GND	146	145	P3E_5_PESW_TX2_DN
P3E_5_PESW_RX2_C_DP	148	147	GND
P3E_5_PESW_RX2_C_DN	150	149	GND
GND	152	151	P3E_5_PESW_TX3_DP
GND	154	153	P3E_5_PESW_TX3_DN
P3E_5_PESW_RX3_C_DP	156	155	GND
P3E_5_PESW_RX3_C_DN	158	157	GND
GND	160	159	NC

7.5 M.2 Module Pin Definition

As we mentioned in the overview section, GPv2 card is designed to support accelerator applications to offload some specific workload from host CPU. M.2 form factor is suitable to enable high density deployment in system. It also provides the granuity so system design can deploy different kinds of accelerators in one system and it is easy to change. Last but not least, it is a industrial standard form factor with wide adoption in server system.

Of course, M.2 form factors has some restrictions. One is the debug capability which is critical to accelerators. We have defined the extra debug pins on the golder fingers using NC pins in M.2 standard. The details are listed in Table7-5 and Table 7-6

Table 7-5: M.2 Accelerator Pin Table

M.2 Module Golden Finger			
Signal	Pin	Pin	Signal
3.3V	2	1	GND
3.3V	4	3	GND
PWRDIS	6	5	PETn3
PLN#	8	7	PETp3
LED_1#(0)	10	9	GND
3.3V	12	11	PERn3
3.3V	14	13	PERp3
3.3V	16	15	GND
3.3V	18	17	PETn2
TRST	20	19	PETp2
VIO_1V8	22	21	GND
TDI	24	23	PERn2
TDO	26	25	PERp2
TCK	28	27	GND
PLA_S3#	30	29	PETn1
GND	32	31	PETp1

USB_D+	34	33	GND
USB_D-	36	35	PERn1
GND	38	37	PERp1
SMB CLK (I/O)(0/1.8V)	40	39	GND
SMB DATA (I/O)(0/1.8V)	42	41	PETn0
ALERT# (O)(0/1.8V)	44	43	PETp0
Reserved_UART_Rx	46	45	GND
Reserved_UART_Tx	48	47	PERn0
PERST# (I)(0/3.3V)	50	49	PERp0
CLKREQ# (I/O)(0/3.3V)	52	51	GND
PEWAKE#(I/O)(0/3.3V)	54	53	REFCLKn
Reserved for MFG DATA	56	55	REFCLKp
Reserved for MFG CLOCK	58	57	GND
ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M
NC	68	67	TMS
3.3V	70	69	NC
3.3V	72	71	GND
3.3V	74	73	VIO_CFG_GND
		75	GND

Table 7-6 M.2 Accelerator Pin Definition

Interface	Signal Name	I/O	Description	Voltage	FB Requirement
Power Ground	3.3V(9 pins)	I	3.3V running power source	3.3V	Required
	GND(15 pins)		Ground	0V	Required
	VIO_1V8	I	Reserved 1.8V running power source for future PCI-SIG standard.	1.8V	NC in module, FB platform leave this pin open.
PCIe	PETp0/PETn0	O	PCIe TX/RX Differential signals defined by the PCIe 3.0/4.0 specification. The Tx/Rx are defined on module perspective.		Required
	PETp1/PETn1	O			
	PETp2/PETn2	O			
	PETp3/PETn3	O			
	PERp0/PERn0	I			
	PERp1/PERn1	I			
	PERp2/PERn2	I			
	PERp3/PERn3	I			
	REFCLKp/REFCLKn	I	PCIe Reference Clock signals (100 MHz) defined by the PCIe 3.0/4.0 specification		Required
	PERST#	I	PE-Reset is a functional reset to the card as defined by the PCI Express CEM Rev3.0	3.3V	Required

	CLKREQ#	I/O	Clock Request is a reference clock request signal as defined by the PCIe Mini CEM specification; Open Drain with pull up on Platform; Active Low; Also used by L1 PM Substates.	3.3V	Optional. FB platform leave this pin open.
	PEWAKE#	I/O	Vendor do not need to support this feature	3.3V	Optional. FB platform leave this pin open.
Specific Signals	Reserved for MFG DATA		Manufacturing Data line.		Optional. FB platform leave this pin open.
	Reserved for MFG CLOCK		Manufacturing Clock line.		
	LED1# (O)	O	LED pin	3.3V	Optional. FB platform leave this pin open.
	ALERT#	O	Alert notification to master; Open Drain with pull up on Platform; Active Low.	1.8V	Required. Refer to Sec 5.3 for more details.
	SMB_CLK	I/O	SMBus clock; Open Drain with pull up on Platform, slave on module	1.8V	Required
	SMB_DATA	I/O	SMBus DATA; Open Drain with pull up on Platform, slave on module	1.8V	Required
USB	USB_D+	I/O	USB 2.0 bus reserved for future application.	N/A	NC in module, FB platform leave this pin open.
	USB_D-	I/O	USB 2.0 bus reserved for future application, not required here.	N/A	NC in module, FB platform leave this pin open.
UART	Reserved_UART_RX	I	UART Receive Data connected to TXD on the Platform.	1.8V	Required. Please refer section 5.3 for details.
	Reserved_UART_TX	O	UART Transmit Data connected to RXD on the Platform.	1.8V	Required. Please refer section 5.3 for details.
JTAG	TDI	I	Refer to JTAG Specification (IEEE 1149.1), Test Access Port and Boundary Scan Architecture for definition.	1.8V	Required. Please refer section 5.3 for details.
	TDO	O		1.8V	
	TCK	I		1.8V	
	TMS	I		1.8V	
	TRST	I		1.8V	
Reserved New IOs	PWRDIS	I	Reserved for power disable pin. High: disable power on module. This pin shall be NC on module.	3.3V	FB platform does not support these features.
	PLN#	I	Reserved for Power Loss notification. NC in module.	3.3V	
	PLA_S3#	O	Reserved for Power loss Assert. NC in module.	3.3V	
	VIO_CFG_GND	O	Reserved for IO configure pin. Connected to ground in module.	0V	

7.6 Dual M.2 Module Support

Still M.2 has pretty tight landing footprint and power restrictions which does not match accelerator application. To mitigate those limitations, we proposed Dual M.2 form factor. Dual M.2 form factor is the combination of two sets of M.2 22110 Socket 3 key M Pin Definition. FB defined this form factor to support more power and bigger ASIC chips for specific work loads. We have defined main slot and second slot in this form factor. the main slot has the full set and the 2nd slot just provide additional power and ground pins. In this version each Dual M.2 card shall support 8 PCIE lanes.

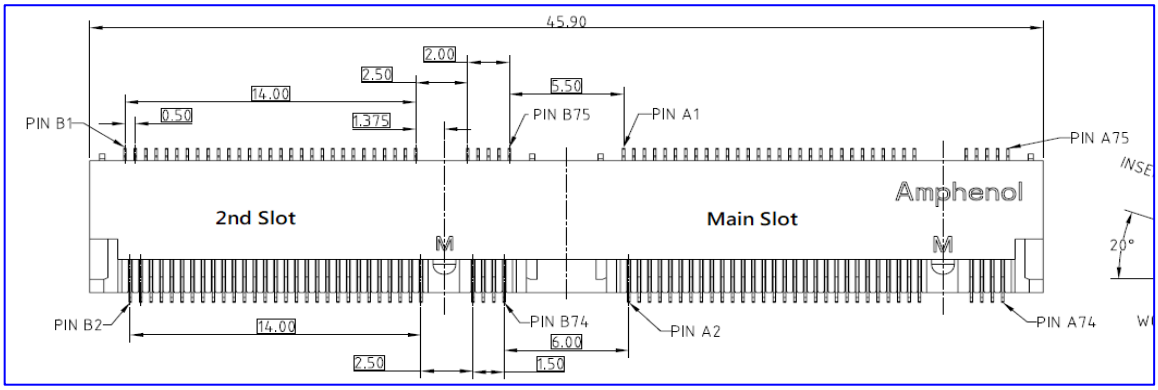

Figure 7-5. Dual M.2 Connector Drawing

The Dual-M.2 module need a specified connector to have better pin position control. The drawing of that connector is shown in Figure 7-5. This connector is backward compatible. You can plug in two standard M.2 module in this connector without any problem.

In Dual M.2 module, Main slot provides PCIe lane0-3, power, ground and all the sideband IOs, second slot provides PCIe lane4-7, power and ground only. The details of pin definition is the same as M.2 module. Dual M.2 module support x8 port as default and can fail over to x4 port Main slot if the host does not provide that many lanes. You can find more details about electrical and mechanical defintion of Dual M.2 form factor in Dual M.2 Accelerator Specification.

Table 7-7 Dual M.2 Accelerator Pin Definition

Dual M.2							
Main Slot (SlotA)					2nd Slot (SlotB)			
Signal	Pin	Pin	Signal		Signal	Pin	Pin	Signal
3.3V	A2	A1	GND		3.3V	B2	B1	GND
3.3V	A4	A3	GND		3.3V	B4	B3	GND
PWRDIS	A6	A5	PETn3		NC	B6	B5	PETn7
PLN#	A8	A7	PETp3		NC	B8	B7	PETp7
LED_1#(0)	A10	A9	GND		NC	B10	B9	GND
3.3V	A12	A11	PERn3		3.3V	B12	B11	PERn7
3.3V	A14	A13	PERp3		3.3V	B14	B13	PERp7
3.3V	A16	A15	GND		3.3V	B16	B15	GND
3.3V	A18	A17	PETn2		3.3V	B18	B17	PETn6
TRST	A20	A19	PETp2		NC	B20	B19	PETp6
VIO_1V8	A22	A21	GND		GND	B22	B21	GND
TDI	A24	A23	PERn2		NC	B24	B23	PERn6
TDO	A26	A25	PERp2		NC	B26	B25	PERp6
TCK	A28	A27	GND		GND	B28	B27	GND
PLA_S3#	A30	A29	PETn1		NC	B30	B29	PETn5
GND	A32	A31	PETp1		NC	B32	B31	PETp5
USB_D+	A34	A33	GND		NC	B34	B33	GND
USB_D-	A36	A35	PERn1		NC	B36	B35	PERn5
GND	A38	A37	PERp1		NC	B38	B37	PERp5
SMB CLK (I/O)(0/1.8V)	A40	A39	GND		NC	B40	B39	GND
SMB DATA (I/O)(0/1.8V)	A42	A41	PETn0		NC	B42	B41	PETn4
ALERT# (O)(0/1.8V)	A44	A43	PETp0		NC	B44	B43	PETp4
Reserved_UART_Rx	A46	A45	GND		NC	B46	B45	GND
Reserved_UART_Tx	A48	A47	PERn0		NC	B48	B47	PERn4
PERST# (I)(0/3.3V)	A50	A49	PERp0		NC	B50	B49	PERp4
CLKREQ# (I/O)(0/3.3V)	A52	A51	GND		NC	B52	B51	GND
PEWAKE#(I/O)(0/3.3V)	A54	A53	REFCLKn		NC	B54	B53	NC
Reserved for MFG DATA	A56	A55	REFCLKp		NC	B56	B55	NC
Reserved for MFG CLOCK	A58	A57	GND		NC	B58	B57	GND
ADD IN CARD KEY M			ADD IN CARD KEY M		ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M		ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M		ADD IN CARD KEY M			ADD IN CARD KEY M
ADD IN CARD KEY M			ADD IN CARD KEY M		ADD IN CARD KEY M			ADD IN CARD KEY M
NC	A68	A67	TMS		NC	B68	B67	NC
3.3V	A70	A69	NC		3.3V	B70	B69	NC
3.3V	A72	A71	GND		3.3V	B72	B71	GND
3.3V	A74	A73	VIO_CFG_GND		3.3V	B74	B73	GND
		A75	GND				B75	GND

7.7 PCIe

GPv2 card supports PCIe Gen3 speed. Comparing to GPv1 design, the biggest change here is GPv2 has a Microchip PM8535 96port PCIE Fanout Switch on the board. That PCIE switch will extend the 24 PCIe lanes from Twin lake to 48 lanes. There are several other benenfits that PCIe switch introduced which we will give a detailed description in this section.

7.7.1 PCIe bifurcation:

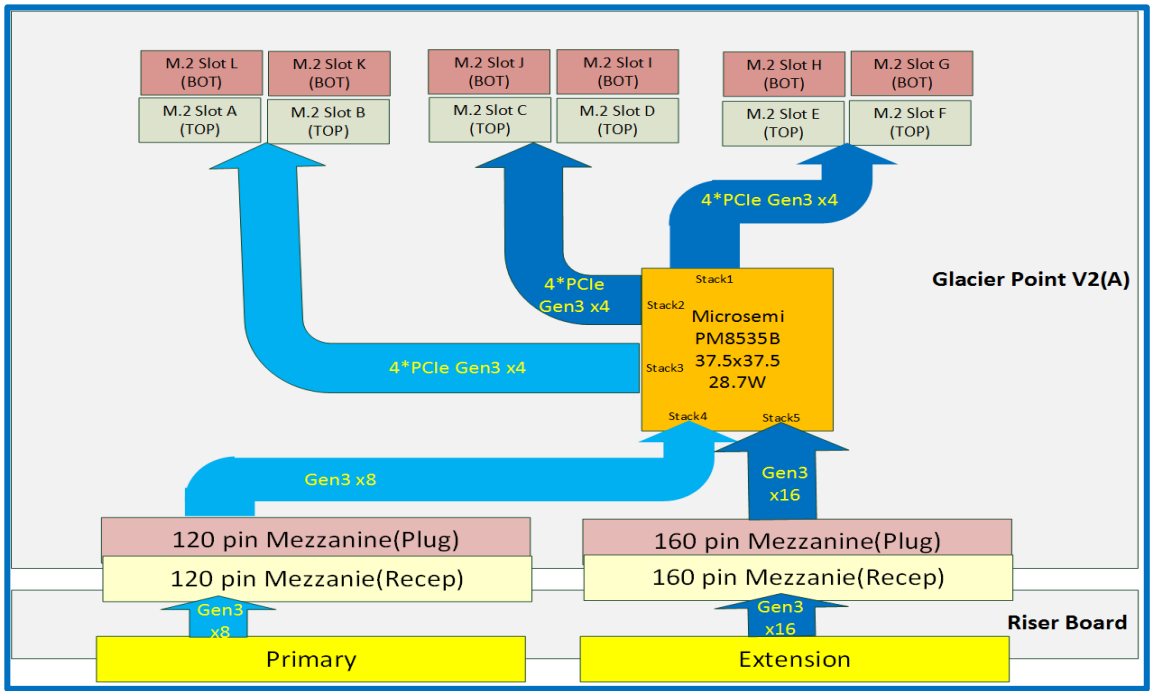

Figure 7-6. PCIe Biffurcation Drawing

Upsteam port to CPU has been separated into two buses. One is bifurcated as Gen3 x8 and another is Gen3 x16. For the downstream port: we have two different bifurcations: once M.2 module is plugged, the down stream port will be bifurcated as 12 x4 ports. Once Dual M.2 module is plugged, the DSP will be configured as Gen3 x8.

Bridge IC will manage the configuration of PM8535 PCIe Switch to select the right bifurcation at power on stage. By reading the pre-defined field in FRU of each accelerator module, we can tell whether the module is storage, M.2 accelerator or Dual M.2 accelerator. PM8535 PCIe switch supports multiple configurations that is pre-stored in its flash. Bright IC will execute the flow in Chapter 9.6 during power on stage to double check the setting based on the real configuration.

Based on this scheme, GPv2 card do have the capability to support the mix of M.2 and Dual M.2 modules. That requests the configuration of modules need to be fixed so that we can limit the types of bifurcations in advance.

7.7.2 PCIe switch

In this design we used PM8535 switch. This chip comes from Microchip (Microsemi is acquired by microchip). It is 37.5x37.5mm with 28.7W TDP corner Tj. It supports up to 80 lanes with 5 stacks. Each stack includes 16 lanes. Figure 7-6 describes the stack setup inside the switches.

PCIe switch serves the following functions:

1. It expand the maximum PCIe lanes from 24 lanes to 48 lanes, which doubles the module density.
2. It provides the flexibility to support different bifurcations. In this case reconfiguration between M.2 and Dual M.2 bifurcation happens at switch level instead of root complex.
3. It provides Downstream Port Containment to better support hot plug features. The support of hot plug is much easier as switch can play the major role to manage the hot plug hardware flow.
4. Swtich firmware is easy to customize so it can provide new control schemes like individual M.2 ac cycle we will describe in section 7.7.3.
5. Switch support near end peer to peer communication between endpoints without compromising the bandwidth. In general root complex could have maximum payload size limitation for peer to peer package which will effectively limit the bandwidth for peer to peer communication. In swtich case, the peer to peer transaction does not need to go up to root complex. So we don't have that restrictions. That requests the peer to peer node under the same USP.

In accelerator use case those are all the features that will bring great benefits.

7.7.3 Module AC Cycle Support

Glacier Point V2 card does not support hot service defined in YV2 system. One we want to service any M.2 card, both GPv2 card and its host Twin lake card need to be turned off. However, we do request a "AC Cycle" action so that host can power cycle the individual M.2 or Dual M.2 card in case the hardware is hang. During this flow, the host should not hang, panic and the other good modules shall not be affected.

This flow includes the following paticipants:

1. A host program that monitor the failure of individual card
2. A load switch hardware for each M.2 connector to turn on/off 3.3V power supply
3. PCIe switch that can execute DPC, monitor channel status register, provide GPIO to control load switch based on the flow defined.
4. OS hotplug driver to manage the managed remove activity (power off) and surprise add activity (power on)
5. Module driver that supports hot plug.

With this feature, we can effectively enhance system robustness and optimize the service requirement.

7.8 I^2C

A single I^2C connection is used to connect the BMC on the platform to the Bridge IC on the GPv2 card as the primary server management interface. It is compatible to SMBus

protocol. 1MHz is the default clock speed in firmware stack. The I²C alert signal is required and is used as an interrupt for both the Bridge IC and the BMC.

Both the BMC and the Bridge IC are I²C master devices on the bus and they communicate with each other via the Intelligent Platform Management Bus (IPMB) protocol. To achieve maximum bandwidth and avoid conflicts, no other devices should use this bus except for the BMC and the Bridge IC.

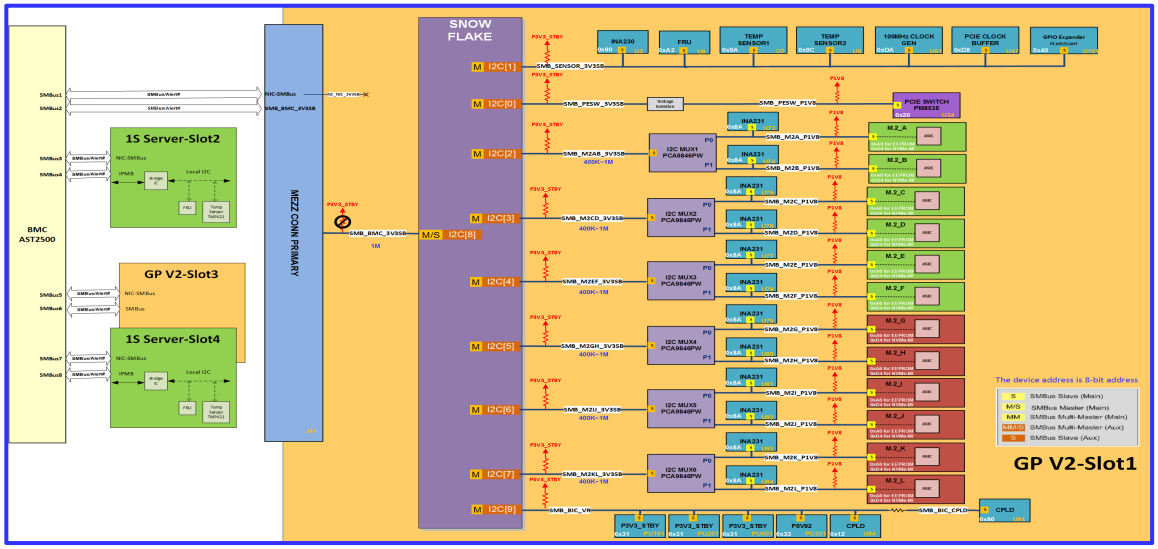

Figure 7-7. GPv2 I²C Topology

As Fig. 7-7, BIC will provide multiple I²C/SMBus connection to M.2 modules and the other components on GPv2 card. Every two adjacent M.2 modules are muxed to one I²C master. In total we designed 6 I²C masters to support 12 modules in M.2 case and 6 modules in Dual M.2 case.

Each M.2 connector has an INA231 voltage and current monitor. BMC chip can use the information from this chip to calculate the real time power consumption of each module.

Each M.2 and Dual M.2 module support NVMe-MI base command interface via I²C/SMbus line that reports the basic status of module.

7.9 Serial port

The serial port shall be routed to the BMC on the platform. Thus, the user can access each module's serial console one at a time through the BMC locally or remotely via Serial Over Lan (SOL). Baud rate is 57600.

Each M.2 connector also have an UART port defined with the same baud rate. In Figure 7-8, a CPLD will mux one of the UART port to BMC. This feature is used mainly for debug purpose. Accelerators could use that port to dump the error logs.

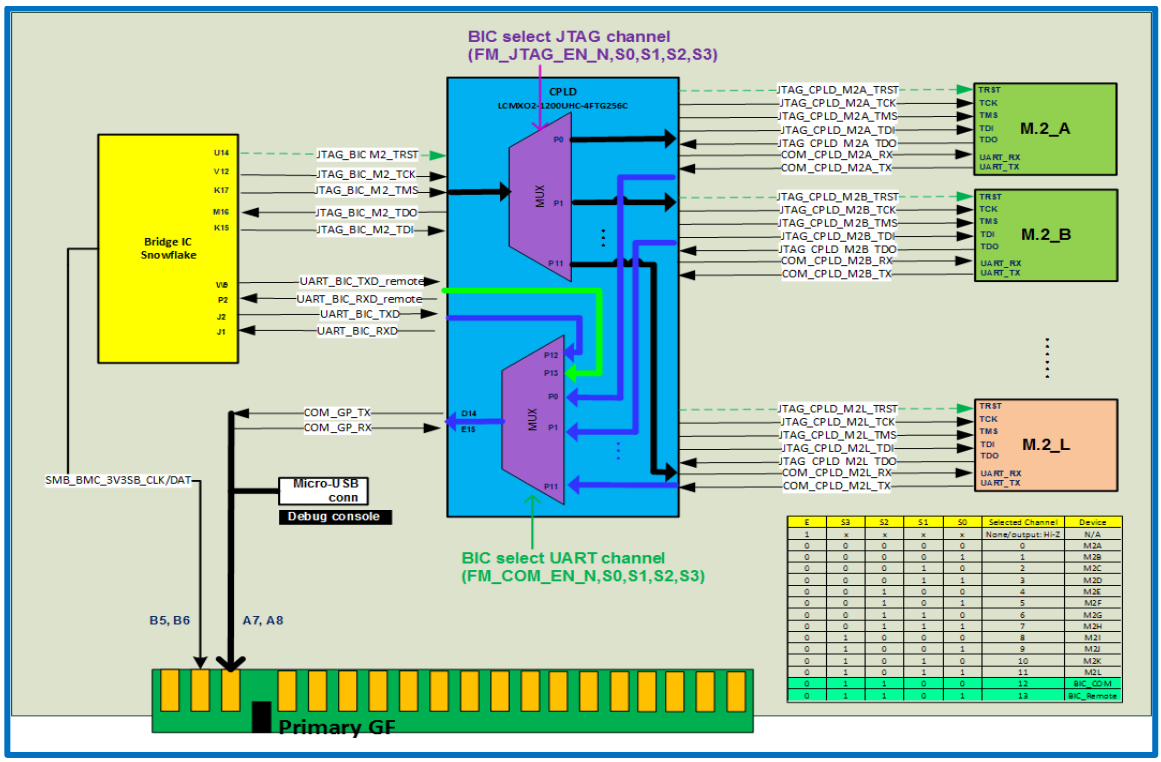

Figure 7-8. GPv2 UART and JTAG Topology

7.10 JTAG Port

In accelerator spec we also defined the JTAG interface from each M.2/Dual M.2 modules. The JTAG interface is also muxed by the CPLD to BIC. BIC could run the JTAG commands to fetch the debug information from M.2 modules. This enables the remote debug capability of each M.2 module which will be very useful in hyper scale data center.

The implementation of remote JTAG interface shall follow the security requirements defined by end user.

7.11 Slot ID

Due to the nature of Yosemite V2 architecture, a BMC is always preferred on the platform side to work with the 1S server. When a BMC is present in the system, the Bridge IC shall always request its slot ID from the BMC but not probe the slot ID by itself.

Glacier Point V2 defined all 4 GPIO pins with specific functions. Table 7-3 already listed the definition of slot ID pins. Let's call them out in Table 7-8 below:

Table 7-8 Slot ID Pin Definition

GPIO	Pin	Min	Active	Max
CARD_TYPE_DETECTION	SVR_ID0	0.55V	0.75V	0.95V
FM_POWER_EN	SVR_ID1	0V	3.3V	3.3V
Latch_Detect_N	SVR_ID2	0V	0V	3.3V
N/A in GPv2	SVR_ID3	0V	0V	3.3V

CARD_TYPE_DETECTION: This pin is defined as an output pin in GPv2. GPv2 shall pull this pin high to 0.75V to identify itself as a “Smart Card” in YV2 system. This signal is a quick way for BMC to identify card types on the Yosemite V2 platform.

FM_POWER_EN: This pin is input pin with pull down on GPv2. It is used to synchronize the power on/off between GPv2 card to the 1S server card. At server card side this pin is a active high output pin to 3.3VAUX power rail.

Latch_Detect_N: This pin is defined as EJECTOR_LATCH_DETECT_N, 3.3VAUX signal, active low. This signal is driven by ejector detection circuits on GPv2 card. The purpose of this signal is to prevent surprise insertion/removal and protect 1S server and device card. This signal is driven low when the ejector is closed, high when the ejector is open. Yosemite V2 platform uses this signal to turn off 12.5V to this particular slot when the ejector is open with a default pull-up.

SVR_ID3 is NCed in GPv2 card.

7.12 Pin Header

Glacier Point V2 defined belwo pin headers with specific functions. Let’s call them out in Table 7-9:

Table 7-9 Pin header Definition

Function	Location	Description
BOOT_RECOVERYB	J46	Holding this pin to '0' during power-on will cause the boot loader to provide a code load option to recover a corrupted flash image.
BIC debug header	J56	It is for Dongle to access GPv2 BIC.
CPLD debug header	J58	It is for Dongle to access GPv2 CPLD.
Disable BIC	J59	Holding this pin to '0' will disable BIC in GPv2.
Remote debug Disable	SW5	0: enable Remote debug (default) 1: disable Remote debug

8 Power

8.1 Input

Power for the card is provided via seven 12V pins on the primary connector and seven more 12V pins on the extension connector. Each pin supports a maximum 1.1A of current. The nominal 12V input voltage is defined as 12.5V, +/-7%.

8.2 Input Capacitance

The capacitance on the input 12V rail of the Glacier Point V2 is optimized to meet load transient filtering requirements and minimize switching noise. However total bus capacitance should be less than 3mF to meet slot level hotswap requirements.

8.3 Glacier Point V2 power capacity

The Glacier Point V2's maximum steady state power draw is 180W with full configuration, with peak transient power draw of up to 235W. However, the individual power rail must be designed to support the sustained and peak power loads specified in Table 8-1. The peak power supported by the power regulators is higher than the overall peak power supported by the platform per slot.

Table 8-1: Glacier point V2 load power specifications

Power rail	Regulator Phase count	Sustained power	Peak power
M.2 3.3V supply	3	145 W	240 W
PCIE switch supply	1	29 W	31 W
Power supplies for Bridge IC, Aux PCIE switch rails and additional rails	1	6 W	6 W

8.4 Power sequence and standby power

Because there is only one 12V power input to the card, there is no power sequence requirement to power on the Glacier Point V2 card from the platform perspective. However, a standby 3.3V_AUX power rail on the card is required to power the Bridge IC at all power states. It is the designer's responsibility to provide proper standby power rails from the main 12V_AUX input with possible specific power sequencing.

For dual M.2 support, load switch enable logic should be designed to ensure both connectors power on at the same time.

Care must be taken in the sequence to avoid any leakage path among the power domains of the PCIE switch during power on.

8.5 VR specifications

The M.2 power supply is designed to support >90% effieciency for 30% to 90% of the power load (Figure 8-1) and is optimized for maximum efficiency at sustained power spec provided in Table 8-1. The PCIE switch power supply should be designed for >82% efficiency at 30% to 100% of specified load. Thermal design should ensure that the regulator junction temperatures are kept within margin for maximum sustained power load.

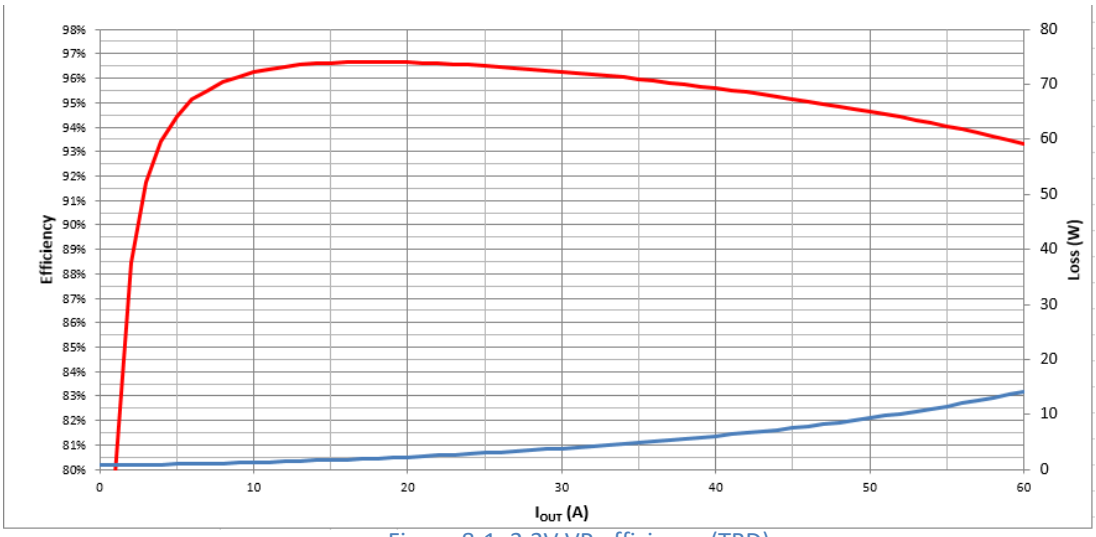

Figure 8-1: 3.3V VR efficiency (TBD)

M.2 power regulator design should also ensure the output regulation is within the +-5% specification required by M.2 loads across load transients. Adequate decoupling shoudn be provided near the M.2 connector to support fast transients with additional MLCC capacitors. Remote sense for regulator should be derived from central M.2 connector to minimize R-drop effects of large 3.3V plane using differential sense lines.

Layout should be validated through DC simulation to ensure via current is less than 2.5A for design reliability, especially in case of the single phase PCIE switch power supply. Analog rails of the PCIE switch can be derived from the core 0.92V with adequate noise filter using an LC filter.

8.6 Power telemtry and power capping

The Glacier Point V2 should implement sophisticated power management features. Each power regulator supporting >10W load power should implement telemetry within +-3% accuracy through PMBUS which is readable by the bridge IC. The Glacier Point V2 should have the power monitoring capability to report a one-second average power reading with 3% accuracy. A power sensor with I^2C interface should be used at the 12.5V power input to the card and which is readable by the Bridge IC.

Each M.2 connector can be power cycled using a load switch. The load switch should also feature soft start to control inrush on the 3.3V supply. The load switch should be designed to support up to 1mF on each M.2 connector. The design should support power telemetry with +-2% accuracy on each M.2 port using a sensor with I2C interface.

System power throttle can be issued though platform BMC command to the bridge IC on the Glacier Point V2 card. The card level throttle is realized by then throttling the M.2 or dual M.2 loads through out of band signaling. The power tree is represented on Figure 8-2.

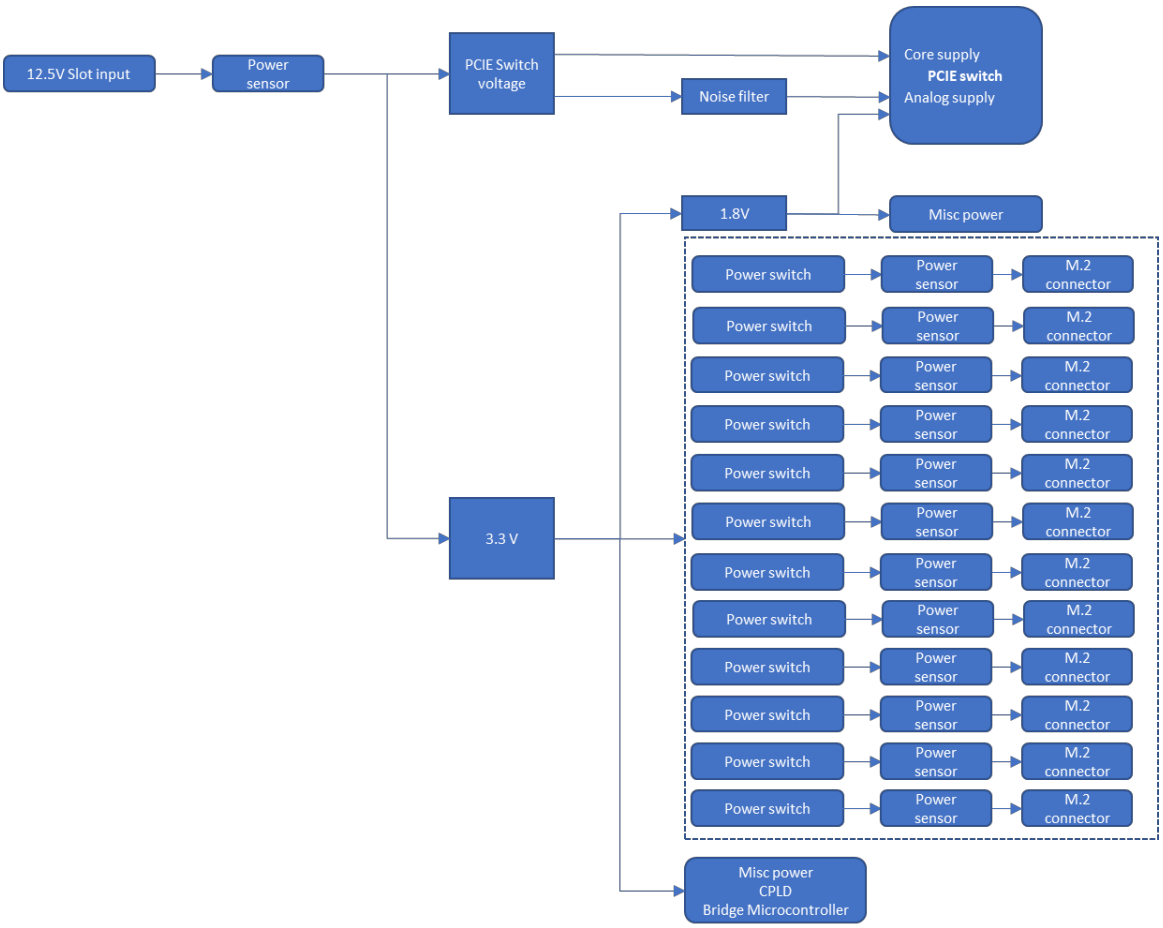

Figure 8-2: Power tree of GPv2 card

9 Function

9.1 Accelerator Carrier card

The Glacier Point V2 card can support up to x12 M.2 modules and up to x6 Dual M.2 modules. The module could be storage or accelerators.

9.2 Debug Interface

The Glacier Point V2 will support UART and JTAG debug interface. A CPLD mux is designed to mux up to 12 JTAG interfaces to BIC and mux up to 13 UART interface (M.2 + BIC) to BMC chip on the baseboard.

On GPv2 board, A micro-usb debug console is attached to the UART interface for bringup and debug purpose.

9.3 Storage

Although Glacier Point V2 card forcus a lot on the accelerator applications, it can support normal 22110 PCIe M.2 SSD without any problem. As mentioned before, GPv2 support mix of different types of accelerator with SSD once it is necessary to certain application.

9.4 EEPROM

The Glacier Point V2 includes a 128Kbits I^2C-accessible Electrically Erasable Programmable Read-Only Memory (EEPROM). The EEPROM is accessible from the platform via the Bridge IC. 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
- Manufacturing Date and Time

9.5 Module type enumeration

Dual M.2 requires PCIe bifurcation to be x8. That adds the complicity of system design. Since M.2 form factor does not have PRNT pin, we need to pay special attention once Dual M.2 module is plugged in. GPv2 platform defines a process for BMC to detect Dual M.2 modules. The information is stored in the FRU of each module to define the type of modules that are inserted (Storage, M.2 Accelerator or Dual M.2 Accelerator). BIC will run a quick scan on each module at power on stage to understand which module is plugged in.

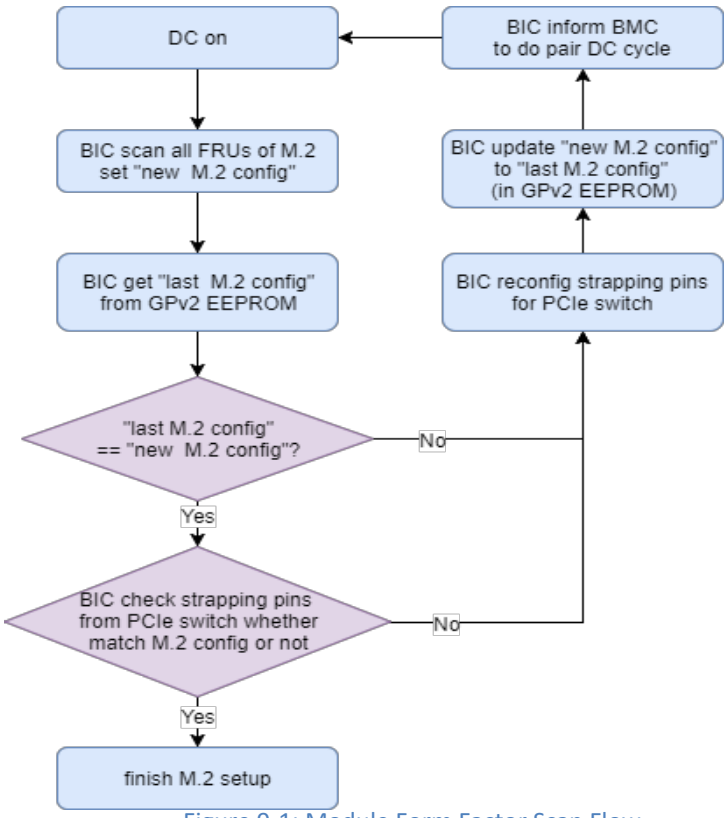

Figure 9-1: Module Form Factor Scan Flow

Currently in GPv2 we only support all M.2 and all Dual M.2 configuration. Although theoretically we can support multiple configurations in hardware, PCIe switch need to have different configuration file to support different bifurcation cases. Limiting the use case can help simply the configuration files we need. For now all M.2 need all DSP to be x4 and all Dual M.2 case need all DSP to be x8. We can add more cases once the application requires so.

9.6 Glacier Point V2 Management

The primary server management functions will be provided using a BMC on the platform. The BMC on the platform will use an I^2C bus as the management interface. This section identifies the required information that must be accessible from the BMC.

9.6.1 NVMe-MI Base command interface

FB defined an NVMe-MI base command interface over SMBus line to communicate with accelerator modules through sideband. The definition of NVMe-MI based command is listed in the Fb accelerator module hardware spec.

9.6.2 Bridge IC

The Glacier Point V2 uses a Bridge IC is defined as the bridging device between the ASIC and the BMC. The Bridge IC is on stand-by power so that it can be accessed by the BMC even when the ASIC is powered down.

On the platform side, the BMC and Bridge IC communicate with each other with IPMI messages over the I^2C bus. To enable prompt communication, this I^2C bus shall be a point-to-point link without any other devices on the same bus. It shall run in high-speed mode with a minimum speed of 400KHz. When possible, a 1MHz or better speed is strongly recommended.

Bridge IC has FRU EEPROM and thermal sensors on a local I^2C bus. The BMC can communicate with the Bridge IC to inquire the FRU and thermal data through the Intelligent Platform Management Bus (IPMB). The FRU EEPROM's data format is defined in Section 8.5. The thermal sensors are mainly used to measure the inlet and outlet temperatures of the device card.

The Bridge IC has a dedicated I^2C bus to ASIC that supports NVME-SI. The Bridge IC now behaves as a transparent bridge to forward NVME-SI messages between the BMC and SoC's Intel® Manageability Engine. With this transparent bridge, the BMC can directly work with SoC's Intel® Manageability Engine to perform most of server management functions.

The Bridge IC monitors the Glacier Point V2's sensors, such as voltage sensors, power sensors, and digital sensors for critical GPIOs. The BMC can inquire about the device status by reading these sensors and taking actions via the Bridge IC.

It is recommended to use a versatile microcontroller as the Bridge IC. The microcontroller has a compact size, uses a low amount of power, and has adequate functions to support all required bridging functions. This microcontroller must have two Serial Peripheral Interfaces (SPI). The first SPI bus is used as its own boot ROM if it does not have an integrated boot ROM. The second SPI can be used to reprogram the SoC's boot ROM when it is corrupted with a multiplexer and BMC's support. The Bridge IC on the Glacier Point V2 is a Texas Instrument's Tiva microcontroller.

9.6.3 I^2C addressing

The Glacier Point V2 and BMC communicates using IPMI 2.0 commands transmitted over the I^2C connection through a Bridge IC on the 1S Server card. The I^2C bus address for the Bridge IC is configured as 0x40. The BMC on the platform is configured as 0x20.

9.6.4 Message Transfer

As the bridge between the ASIC and BMC on the platform, the Bridge IC provides ways to transfer messages between them via NVME-MI base command interfaces. For in-band management, the Bridge IC can forward the ASIC's request to the BMC and then send the received response back to ASIC. When the BMC sends a request on the I^2C bus meant for ASIC, the Bridge IC shall forward the command on the I^2C bus and send the received response back to the BMC.

It is possible to implement an alternative SOL through the Bridge IC. When the alternative SOL feature is enabled, the serial data from the SoC's serial port shall be sent to the BMC via I^2C. When the BMC sends SOL data, it shall be emitted via the serial port. (TBD)

9.6.5 Platform Discovery and Configuration

The Bridge IC provides a way for the BMC to discover platform capabilities such as electrical interface assignment. It also provides a way to discover and configure its own capabilities like enabling or disabling the SOL interface and/or POST code interface.

9.6.6 IPMB Interface

The Bridge IC provides an IPMB interface for the BMC to access various IPMI resources on the Glacier Point V2. To meet this requirement, the Bridge IC shall implement various standard IPMI commands. It shall implement FRUID commands to identify the ASIC type, System Event Log (SEL) commands to store device specific event logs, and Sensor Data Repository (SDR) commands to identify various sensors described for the specific ASIC device

9.6.7 Firmware Update

The Bridge IC can update firmware of the programmable devices on the Glacier Point V2, such as ASIC bootloader, main ASIC firmware, and the Bridge IC's firmware.
Bridge IC provides a way for the BMC to access the version information of various firmware components on the Glacier Point V2, initiate the update process for various firmware components on the 1S Server, and detect and retransmit corrupted firmware image packets during transit from the BMC to the Bridge IC.

9.6.8 GPIO Register

The Bridge IC shall provide a GPIO interface to the BMC through a GPIO register block. In this way, the BMC can control the GPIO behind the Bridge IC (or this hardware abstraction layer) through accessing this register block.

The Bridge IC shall provide a way for BMC to configure GPIO pin direction, interrupt capability, and provide a way to get/set the current status of GPIO signals. It shall send an interrupt message to the BMC when the interrupt enabled GPIO signal changes its state. The GPIO register interface exposed by the Bridge IC shall provide four bytes to represent 32 signals that indicate various conditions as shown in Table 9-1.

Table 9-1: Bridge IC GPIO Table

No.	GPIO offset	Pin name	Description
1	Byte 1 - bit [0]	BMC_HB_LED_N	BIC Heartbeat
2	Byte 1 - bit [1]	I2C_PESW_MULTI_CONFIG_ADDR0_R	PCIE Switch configuration strapping pin
3	Byte 1 - bit [2]	I2C_PESW_MULTI_CONFIG_ADDR1_R	PCIE Switch configuration strapping pin
4	Byte 1 - bit [3]	I2C_PESW_MULTI_CONFIG_ADDR2_R	PCIE Switch configuration strapping pin

5	Byte 1 - bit [4]	I2C_PESW_MULTI_CONFIG_ADDR3_R	PCIE Switch configuration strapping pin
6	Byte 1 - bit [5]	I2C_PESW_MULTI_CONFIG_ADDR4_R	PCIE Switch configuration strapping pin
7	Byte 1 - bit [6]	SMB_BIC_3V3SB_READY_N_R	BIC ready
8	Byte 1 - bit [7]	P3V3_M2A_EN	M.2 Load switch enable signal
9	Byte 2 - bit [0]	P3V3_M2B_EN	M.2 Load switch enable signal
10	Byte 2 - bit [1]	P3V3_M2C_EN	M.2 Load switch enable signal
11	Byte 2 - bit [2]	P3V3_M2D_EN	M.2 Load switch enable signal
12	Byte 2 - bit [3]	P3V3_M2E_EN	M.2 Load switch enable signal
13	Byte 2 - bit [4]	P3V3_M2F_EN	M.2 Load switch enable signal
14	Byte 2 - bit [5]	P3V3_M2G_EN	M.2 Load switch enable signal
15	Byte 2 - bit [6]	P3V3_M2H_EN	M.2 Load switch enable signal
16	Byte 2 - bit [7]	P3V3_M2I_EN	M.2 Load switch enable signal
17	Byte 3 - bit [0]	P3V3_M2J_EN	M.2 Load switch enable signal
18	Byte 3 - bit [1]	P3V3_M2K_EN	M.2 Load switch enable signal
19	Byte 3 - bit [2]	P3V3_M2L_EN	M.2 Load switch enable signal
20	Byte 3 - bit [3]	PWRGD_P3V3_M2A	M.2 3.3V power good
21	Byte 3 - bit [4]	PWRGD_P3V3_M2B	M.2 3.3V power good
22	Byte 3 - bit [5]	PWRGD_P3V3_M2C	M.2 3.3V power good
23	Byte 3 - bit [6]	PWRGD_P3V3_M2D	M.2 3.3V power good
24	Byte 3 - bit [7]	PWRGD_P3V3_M2E	M.2 3.3V power good
25	Byte 4 - bit [0]	PWRGD_P3V3_M2F	M.2 3.3V power good
26	Byte 4 - bit [1]	PWRGD_P3V3_M2G	M.2 3.3V power good
27	Byte 4 - bit [2]	PWRGD_P3V3_M2H	M.2 3.3V power good
28	Byte 4 - bit [3]	PWRGD_P3V3_M2I	M.2 3.3V power good
29	Byte 4 - bit [4]	PWRGD_P3V3_M2J	M.2 3.3V power good
30	Byte 4 - bit [5]	PWRGD_P3V3_M2K	M.2 3.3V power good
31	Byte 4 - bit [6]	PWRGD_P3V3_M2L	M.2 3.3V power good
32	Byte 4 - bit [7]	FM_COM_EN_N_R	UART MUX enable signal
33	Byte 5 – bit [0]	FM_COM_SEL_0_R	UART channel select
34	Byte 5 – bit [1]	FM_COM_SEL_1_R	UART channel select
35	Byte 5 – bit [2]	FM_COM_SEL_2_R	UART channel select
36	Byte 5 – bit [3]	FM_COM_SEL_3_R	UART channel select
37	Byte 5 – bit [4]	FM_JTAG_EN_N_R	JTAG enable signal
38	Byte 5 – bit [5]	FM_JTAG_SEL_0_R	JTAG channel select
39	Byte 5 – bit [6]	FM_JTAG_SEL_1_R	JTAG channel select
40	Byte 5 – bit [7]	FM_JTAG_SEL_2_R	JTAG channel select
41	Byte 6 – bit [0]	FM_JTAG_SEL_3_R	JTAG channel select
42	Byte 6 – bit [1]	BIC_REMOTE_DEBUG_SELECT_N	Disable/Enable remote debug
43	Byte 6 – bit [2]	RST_I2C_MUX1_N_R	I2C mux reset
44	Byte 6 – bit [3]	RST_I2C_MUX2_N_R	I2C mux reset

45	Byte 6 – bit [4]	RST_I2C_MUX3_N_R	I2C mux reset
46	Byte 6 – bit [5]	RST_I2C_MUX4_N_R	I2C mux reset
47	Byte 6 – bit [6]	RST_I2C_MUX5_N_R	I2C mux reset
48	Byte 6 – bit [7]	RST_I2C_MUX6_N_R	I2C mux reset
49	Byte 7 – bit [0]	FM_BIC_DU_M2A_EN_R	BIC to control M.2 3.3V power
50	Byte 7 – bit [1]	FM_BIC_DU_M2B_EN_R	BIC to control M.2 3.3V power
51	Byte 7 – bit [2]	FM_BIC_DU_M2C_EN_R	BIC to control M.2 3.3V power
52	Byte 7 – bit [3]	FM_BIC_DU_M2D_EN_R	BIC to control M.2 3.3V power
53	Byte 7 – bit [4]	FM_BIC_DU_M2E_EN_R	BIC to control M.2 3.3V power
54	Byte 7 – bit [5]	FM_BIC_DU_M2F_EN_R	BIC to control M.2 3.3V power
55	Byte 7 – bit [6]	FM_BIC_DU_M2G_EN_R	BIC to control M.2 3.3V power
56	Byte 7 – bit [7]	FM_BIC_DU_M2H_EN_R	BIC to control M.2 3.3V power
57	Byte 8 – bit [0]	FM_BIC_DU_M2I_EN_R	BIC to control M.2 3.3V power
58	Byte 8 – bit [1]	FM_BIC_DU_M2J_EN_R	BIC to control M.2 3.3V power
59	Byte 8 – bit [2]	FM_BIC_DU_M2K_EN_R	BIC to control M.2 3.3V power
60	Byte 8 – bit [3]	FM_BIC_DU_M2L_EN_R	BIC to control M.2 3.3V power
61	Byte 8 – bit [4]	FM_POWER_EN	Host to enable GPv2 normal power
62	Byte 8 – bit [5]	BIC_BOARD_ID_0	BIC Board ID
63	Byte 8 – bit [6]	BIC_BOARD_ID_1	BIC Board ID
64	Byte 8 – bit [7]	BIC_BOARD_ID_2	BIC Board ID

9.6.9 IPMI commands

The Bridge IC must support the IPMI commands shown in Table 9-2. The command that is striked out here is not supported here since GPv2 is a carrier card without CPU.

Table 9-2: Bridge IC supported IPMI command Table

IPMI Command	Net Function	CMD#
Get Device ID	App	01h
Get Self Test Results	App	04h
Get System GUID	App	37h
Master Write-Read I^2C	App	52h
Get FRU Inventory Area Info	Storage	10h
Read FRU Inventory Data	Storage	11h
Write FRU Inventory Data	Storage	12h
Get SDR Repository Info	Storage	20h
Reserve SDR Repository	Storage	22h
Get SDR	Storage	23h
Get SEL Info	Storage	40h

Get SEL Allocation Info	Storage	41h
Reserve SEL	Storage	42h
Get SEL Entry	Storage	43h
Add SEL Entry	Storage	44h
Clear SEL	Storage	47h
Get Sensor Reading	Sensor/Event	2Dh
Get Sensor Reading	Sensor/Event	2Dh
~~Send request message to BMC~~	~~OEM (0x38)~~	~~01h~~
~~Send request message to Bridge IC~~	~~OEM (0x38)~~	~~02h~~
Get all GPIO status	OEM (0x38)	03h
Set all GPIO status	OEM (0x38)	04h
Get GPIO configuration	OEM (0x38)	05h
Set GPIO configuration	OEM (0x38)	06h
Send interrupt to BMC	OEM (0x38)	07h
~~Send POST Code to BMC~~	~~OEM (0x38)~~	~~08h~~
~~Request POST Code data~~	~~OEM (0x38)~~	~~12h~~
Firmware Update	OEM (0x38)	09h
Firmware Verify	OEM (0x38)	0Ah
Get Firmware version	OEM (0x38)	0Bh
Enable Bridge IC update flag	OEM (0x38)	0Ch
~~Get NIC LED Frequency~~	~~OEM (0x38)~~	~~0Dh~~
~~Bridge IC Discovery~~	~~OEM (0x38)~~	~~0Eh~~
~~Platform Discovery~~	~~OEM (0x38)~~	~~0Fh~~
~~Set Bridge IC Configuration~~	~~OEM (0x38)~~	~~10h~~
Bridge IC Reset Cause	OEM (0x38)	11h
Bridge IC enter update mode	OEM (0x38)	13h
Set VR monitor Enable	OEM (0x38)	14h
Get VR monitor Enable	OEM (0x38)	15h
~~Reset BMC~~	~~OEM (0x38)~~	~~16h~~
~~Read BIOS image~~	~~OEM (0x38)~~	~~18h~~
~~Get Flash size~~	~~OEM (0x38)~~	~~19h~~
~~Set Pump Duty~~	~~OEM (0x38)~~	~~1Bh~~

~~Get Pump Duty~~	~~OEM (0x38)~~	~~1Ch~~
~~Set BIOS Chip Select~~	~~OEM (0x38)~~	~~1D~~
~~Get BIOS Chip Select~~	~~OEM (0x38)~~	~~1F~~
~~Set JTAG Tap State~~	~~OEM (0x38)~~	~~21~~
~~Shift JTAG Data~~	~~OEM (0x38)~~	~~22~~
Set System GUID	OEM (0x38)	EFh

Table 9-3 provides details of the IPMI Original Equipment Manufacturer (OEM) commands details:

Table 9-3: Bridge IC supported IPMI command Table

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
03h	Send request message to BMC	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 – Request interface 01h: Intel® Manageability Engine 02h: SOL 03h: KCS SMS 04h: KCS SMM Byte 5:X – Request data Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5 – Request interface 01h: Intel® Manageability Engine 02h: SOL 03h: KCS Byte 6:X – Response data	This command is used for Bridge IC transfer request to BMC. For example: 1. Bridge IC gets “Get Device ID” command 0x06 0x01 from KCS. 2. Bridge IC will send this command to BMC as below. 0x38 0x01 0x03 0x06 0x01 3. BMC responds Get Device ID data.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
04h	Send request message to Bridge IC	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 –Receive interface 01h: Intel® Manageability Engine 02h: SOL Byte 5:X – Request data from BMC Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5 –Receive interface 01h: Intel® Manageability Engine 02h: SOL Byte 6:X – Response data	This command is used for BMC send request to Bridge IC. For example: 1. When BMC want to send “Get Device ID” command to ME. It can use this command: 0x38 0x02 0x01 0x06 0x01 2. When Bridge IC receive this command, It will send “Get Device ID” command to Intel® Manageability Engine and get the response from Intel® Manageability Engine . 3. Bridge IC responds this command to BMC.
05h	Get all GPIO status	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5:8 – Get all GPIO status 0b: Low 1b: High	This command used by BMC to get GPIO status from Bridge IC. Refer to Table 8 GPIO mapping table.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
06h	Set all GPIO status	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4:7 – GPIO enable mask, refer to GPIO mapping table 0b: Disable 1b: Enable Byte 8:11 – Set all GPIO status 0b: Low 1b: High Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first	This command used by BMC to set GPIO status from Bridge IC. Refer to Table 8 GPIO Mapping Table.
07h	Get GPIO configuration	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4:7 – GPIO enable mask, refer to GPIO mapping table 0b: Disable 1b: Enable Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5:X – GPIO configuration (one byte for one GPIO pin configuration) Bit[0] – Input/output pin 0b : Input pin 1b : Output pin Bit[1] – interrupt disable/enable 0b : Disable 1b : Enable Bit[2] – Edge trigger 0b : Edge trigger (default) Bit[3:4] – Trigger type 00b : Falling edge 01b : Rising edge 10b : Both 11b : Reserved	This command used by BMC to get GPIO configuration from Bridge IC. Refer to Table 8 GPIO Mapping Table.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
09h	Set GPIO configuration	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4:7 – GPIO enable mask, refer to GPIO mapping table 0b: Disable 1b: Enable Byte 8:X – GPIO configuration (one byte for one GPIO pin configuration) Bit[0] – Input/output pin 0b : Input pin 1b : Output pin Bit[1] – interrupt disable/enable 0b : Disable 1b : Enable Bit[2] – Edge trigger 0b : Edge trigger (default) Bit[3:4] – Trigger type 00b : Falling edge 01b : Rising edge 10b : Both 11b : Reserved Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first	This command used by BMC to set GPIO configuration to Bridge IC. Refer to Table 8 GPIO Mapping Table.
0Ah	Send interrupt to BMC	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 – Interrupt GPIO number, refer to GPIO mapping table Byte 5 – Trigger type 00h: Falling edge 01h: Rising edge Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first	This command used for Interrupt notification from Bridge IC sends to BMC. Refer to Table 8 GPIO Mapping Table.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
0Bh	Send POST Code to BMC	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 – Data length Byte 5:X – Port 80 data Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first	Bridge IC support maximum 230 bytes to buffer BIOS POST Code when BMC is not ready. The POST Code data will be in FIFO manner i.e. with the first POST Code as the first byte. But in case BMC is ready, Bridge IC will send one POST Code to BMC at a time.
0Ch	Request POST Code data	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:X – Port 80 data Byte 2:4 – IANA ID – 00A015h, LS byte first	BMC can get all POST Code data by this command. Bridge IC will buffer POST Code data for last boot. The POST Code data will be in LIFO manner i.e. with the latest POST code as the first byte. Bridge IC clear buffer when system power on. The maximum buffer data length is 230 bytes.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
10h	Firmware Update	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 – update target 00h: BIOS 01h: CPLD - bit[7] = 1, last package for image data 02h: Bridge IC boot loader from OOB - bit[7] = 1, last package for image data 03h: Bridge IC boot loader from InBand - bit[7] = 1, last package for image data 04h: VR - bit[7] = 1, last package for image data Byte 5:8 – Offset Byte 9:10 – Data length Byte11:X – Update image data Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) 80h –Write flash error 81h – Power status check fail 82h – Data length error 83h – Flash erase error Byte 2:4 – IANA ID – 00A015h, LS byte first	This command is used to update BIOS and CPLD Firmware from BMC.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
11h	Firmware Verify	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 –update target 00h: BIOS 01h: CPLD 02h: Bridge IC boot loader 03h: VR Byte 5:8 – Offset Byte 9:12 – Data length Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) 80h – Checksum error 82h – Data length error 84h – Read flash error Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5:8 – Checksum	This command is used to verify BIOS and CPLD Firmware from BMC.

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
13h	Get Firmware version	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 –update target 01h: CPLD 02h: Bridge IC 03h: Intel® Manageability Engine version 04h: Bridge IC Bootloader 05h: VCCIO VR 06h: VCCIN VR 07h: VCCSA VR 08h: DDR_AB VR 09h: DDR_DE VR 0Ah: VNNPCH VR 0Bh: 1V05PCH VR Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5:X – CPLD version (Hexadecimal) – CPLD user code 4 bytes. Bridge IC version (Decimal) – 2 bytes length, ex: 1.03. Return data will be 0x01 0x03. Intel® Manageability Engine version (Decimal) – 5 bytes length, ex: Intel® version 03.0.0.010. Return data will be 0x03 0x00 0x00 0x01 0x00. VR version (Hexadecimal) – VR user data, 4 bytes length. Ex: 3d 01 11 00, user data 0 : 0x3d01, user data 1 : 0x1100 Bridge IC bootloader version (Decimal) – 2 bytes length, version 1.08Return data will be 0x01 0x08	

Net Function = OEM (0x38), LUN = 00			
Code	Command	Request, Response Data	Description
14h	Enable Bridge IC update flag	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Byte 4 – Enable update interface flag 00h: UART 01h: I2C 02h: LPC Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first	This command is used to enable Bridge IC update flag from BMC.
15h	Get NIC LED frequency	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5 – LED frequency 00h: No blinking 01h: Solid on 02h: Slow flashing 03h: Fast flashing	
EFh	Get Bridge IC configuration	Request: Byte 1:3 – IANA ID – 00A015h, LS byte first Response: Byte 1 – Completion Code 00h - Success (Remaining standard Completion Codes are shown in IPMI spec v2.0 table 5-2) Byte 2:4 – IANA ID – 00A015h, LS byte first Byte 5 – Bit[0] – SOL interface 0b : Disable 1b : Enable Bit[1] – Port 80 0b : Disable, Bridge IC will not send post code to BMC 1b : Enable	

9.6.10 Thermal Alerts

GPv2 will report the inlet and outlet temperature sensor to BIC and let BMC manage the thermal events through I^2C interface

M.2 module will monitor their own temperature and decide whether they need throttle by themselves. Meanwhile they will report the thermal information via I^2C interface and trigger the alert event through to BIC. However this response loop will be slow since multiple failure events could trigger that alert pin.

9.6.11 Sensors

The following list of analog and discrete sensors are provided and are reported by the Bridge IC to the BMC.

Analog sensors include:

- Outlet Temperature
- Inlet Temperature
- PEIe-switch Temperature
- VR Temperature(s)
- VR Current(s)
- VR Voltage(s)
- VR Power(s)
- Voltage Sensor(s)
- Current Sensor(s)
- Power(s)

Discrete sensors include:

- INA230 Alert

9.7 LEDs

Along the top edge of the card, a blue LED is used to indicate 12V status of the server. There are also two blinking green heartbeat LEDs on the Glacier Point V2 to indicate that the Bridge IC and PCIe switch are in operating mode.

10 Environmental Requirements

The full system with the server card installed meets the following environmental requirements:

- Gaseous contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C (long-term storage) *
- Transportation temperature range: -40°C to +70°C (short-term storage) *
- Operating altitude with no de-rating to 6000 ft

*NOTE: Liquid cooling can meet requirements in pack

10.1 Vibration and Shock (Harsha to revisit)

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

Table 10-1: Vibration and Shock Requirements

	Operating	**Non-Operating**
Vibration	0.5g, 2 to 500 to 2 Hz per sweep, 10 sweeps at 1 octave/minute, test along three axes	1.2g, 2 to 500 to 2 Hz per sweep, 10 sweeps at 1 octave/minute, test along three axes (If shaker can't perform 1.2g with 2Hz, go with 5-500-5 Hz)
Shock	6g, half sine, 11ms, 5 shocks, test along three axes (Test for 7g shock only after system passes 6g shock)	12g, half sine, 11ms, 10 shocks, test along three axes

11 Prescribed Materials

11.1 Disallowed Components

The following components are not used in the design of the motherboard:

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

11.2 Capacitors and Inductors

The following limitations apply to the use of capacitors:

- Only aluminum organic polymer capacitors made by high-quality manufacturers are used and must be rated at 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.
- Surface Mount (SMT) ceramic capacitors with a case size greater than 1206 are forbidden. The 1206 case size is still allowed when installed far from the PCB edge and with a correct orientation that minimizes the risk of cracks.
- Ceramic material for SMT capacitors must be X5R or better (COG or NP0 type are used in critical portions of the design). Only SMT inductors may be used. The use of through-hole inductors is disallowed.

11.3 Component De-rating

For inductors, capacitors, and FETs, derating analysis is based on at least 20% derating.

12 Labels and Markings

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

Description	Type	Barcode Required?
Safety Markings	Silk Screen	No
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
Date Code (Industry Standard: WEEK / YEAR)	Adhesive label	Yes
RoHS compliance	Silk Screen	No
WEEE Symbol. The motherboard will have the crossed out wheeled bin symbol to indicate that it will be taken back by the Manufacturer 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

13 Revision History

Author	Description	Revision	Date
Team	▪ Initial Submission	0.1	02/17/2019