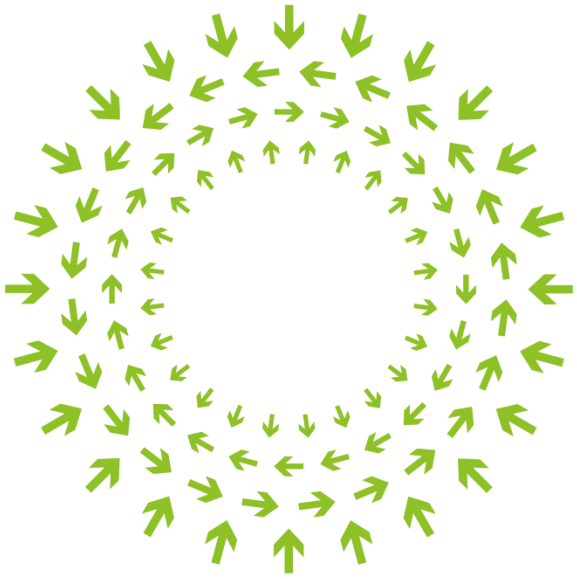


OPEN

Compute Project

Project Olympus Hyperscale GPU Accelerator (HGX-1) Specification

Author:

Siamak Tavallaei, Principal Architect, Microsoft

Revision History

Date	Description
11/1/2017	Version 1.0

© 2017 Microsoft Corporation.

As of November 1, 2017, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0

Microsoft Corporation.

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at Project Olympus License Agreements, which may also include additional parties to those listed above.

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

CONTRIBUTORS AND LICENSORS OF THIS SPECIFICATION MAY HAVE MENTIONED CERTAIN TECHNOLOGIES THAT ARE MERELY REFERENCED WITHIN THIS SPECIFICATION AND NOT LICENSED UNDER THE OWF CLA OR OWFa. THE FOLLOWING IS A LIST OF MERELY REFERENCED TECHNOLOGY: INTELLIGENT PLATFORM MANAGEMENT INTERFACE (IPMI), I^2C TRADEMARK OF PHILLIPS SEMICONDUCTOR. IMPLEMENTATION OF THESE TECHNOLOGIES MAY BE SUBJECT TO THEIR OWN LEGAL TERMS.

Contents

Table of Figures

Table of Tables

1 Project Olympus Specifications List

Project Olympus base specification defines a modular architecture with clear internal and external interfaces. These modules include:

Hardware

- Rack
- Universal PDU
- Rack Manager
- 1U/2U Server and mechanical Enclosure
- Power Supply (PSU)
- Universal Motherboard
- PCIe Riser Boards
- PCIe and other Cables
- Expansion Modules (for storage etc.)

Software/Firmware

- Software RESTful API
- Rack Manager Software/Firmware/Interface
- BMC Firmware
- System Firmware (BIOS/UEFI code)
- ResetAPI Spec
- Software APIs

Project Olympus comprises the above Modules to realize a holistic rack architecture. As specification, design, or product, individual Modules are also applicable to Racks, Chassis, Rack Managers, PDUs, PSUs, Blades, and Motherboards from other architectures such as OCP Open Rack, Open Rack 19, Scorpio, 19" EIA Racks, Rack-mount Servers, and Tower Servers.

Table 1 lists Project Olympus system specifications.

Table 1. List of Specifications

Specification title	Description
Project Olympus HGX-1 Specification	Describes the modules that comprise a flexible PCIe expansion box for Artificial Intelligence Workloads
Project Olympus HGX-1 Mechanical Specification	Describes the mechanical structure for an expansion box which may be cabled to Project Olympus Server or other Servers
Project Olympus JBOD Specification	Describes the modules that comprise a HDD expansion box for Project Olympus Server
Project Olympus JBOF Specification	Describes the modules that comprise an SSD expansion box for Project Olympus Server

Project Olympus Server Rack Specification	Describes the mechanical rack hardware used in Project Olympus system
Project Olympus Server Mechanical Specification	Describes the mechanical structure for a server which may be used as a Host Node to couple to HGX-1
Project Olympus Universal Server Motherboard Specification	Describes general requirements of the Host server motherboard
Project Olympus PSU Specification	Describes the custom Power Supply Unit (PSU) used in the Host server
Project Olympus Universal PMDU Specification	Describes the Universal Power and Management Distribution Unit (PMDU)
Project Olympus Rack Manager Specification	Describes the Rack Manager used to couple to the Project Olympus Server
Project Olympus Standalone Rack Manager Specification	Describes the Standalone Rack Manager used to couple to the HGX-1

This document is intended for designers and engineers who will build PCIe expansion boxes for Project Olympus systems.

2 Overview

This specification focuses on modules that comprise Project Olympus Hyperscale GPU Accelerator (HGX-1) which provides acceleration capabilities for various frameworks and workloads such as HPC, Video Transcoding, Search, and Artificial Intelligence (AI), Deep Neural Networks (DNN) including Training and Inference.

Refer to respective specifications for other elements of the Project Olympus system such as Power Supply Unit (PSU), Standalone Rack Manager (RM), Power and Management Distribution Unit (Universal PDU, Universal PMDU), and Server Rack.

3 Background

Project Olympus Architecture defines a modular system which includes a Rack, power distribution within the Rack, a Blade Server Chassis, and a Rack Manager.

The Server Chassis consists of a Universal Motherboard, CPU/Memory/Storage, power supply, cooling components, PCIe Risers, and IO interfaces.

Project Olympus Universal Motherboard (MB) provides an efficient CPU/Memory compute node in a 1U Chassis. Project Olympus specification allows for this MB to include single- and multi-socket CPUs and

SoCs of various types to meet different compute needs of cloud-scale datacenters. Mechanical enclosures other than Project Olympus may deploy such Universal Motherboards.

To increase the capabilities of the Server, Project Olympus includes a set of expansion boxes to augment bulk storage (Project Olympus JBOD), high-performing storage (Project Olympus JBOF), and accelerated computing via an HGX-1. These expansion boxes appropriately couple to the Server via PCIe or SAS cables.

At Cloud-scale, the accompanying Project Olympus Hyperscale GPU Accelerator (HGX-1 Expansion Box) described here augments I/O and accelerator capabilities for various frameworks and workloads such as HPC, Artificial Intelligence (AI), Deep Machine Learning (DL), Training, Inference, Graphics, Transcoding, and Search.

The Host Interface and the PCIe Switch Fabric flexibility of Project Olympus HGX-1 provide efficient Host-to-IO binding for single- and multi-CPUs to PCIe device sets of various sizes.

Since its baseline is Project Olympus Server, this expansion box benefits from the streamlined datacenter management required of Cloud-Scale.

4 Architectural Requirements

The architecture of Project Olympus Hyperscale Graphics Expansion Chassis (HGX-1) is based on the following tenets:

- 4U Chassis (HGX-1 Expansion Chassis)
- Power: (over 4800W DC available)
    - Six highly efficient, 12V Power Supply Units (PSU) conform to MSFT PSU specification
        - 1800W and 2000W options
        - Two sets of 3-phase balanced power feeds (Primary Feed A and Secondary Feed B)
        - Each of A and B inputs feeds three single-phase units
    - The power supplies current-share the outputs of each feed
    - The power supply sub-system may choose to select feed A, feed B, or both to produce 12V output.
- Cooling: streamlined airflow for efficient cooling of up to 5000W thermal load
    - e.g., 12 Fans with N+2 redundant rotors
    - Cooling 5000W thermal load with less than 790 CFM of airflow (with the choice of different power supplies, the Expansion Chassis may receive 5kW of sustained power.)
- Management:
    - The HGX-1 is a member of the Project Olympus Rack Management on a 1GE Fabric.
    - The Expansion Chassis interfaces to a Project Olympus Standalone Rack Manager via an RJ45 connector/cable for sideband signals Presence#, PowerEn#, and Throttle
        - PowerEn# controls Chassis and BMC power.
    - AST2500 BMC family
        - Uses an RJ45 for 1GE BMC interface (at the rear of the chassis and optionally both at front of the Chassis)

        - Based on OpenBMC firmware
        - BMC controls Power Supplies, Fans, JTAG, Firmware Updates, etc.
- PCIe Switch Fabric:
    - The HGX-1 interconnects four 96-lane PCIe Switches capable of efficient peer-to-peer traffic, multi-Hosting, and Shared IO.
    - Each PCIe Switch has one x1 PCIe port for Fabric Management.
    - Each PCIe Switch provides six x16 Ports:
        - 1 x16 Ports used for inter-Switch Link (ISL)
        - 3 x 16 Ports connect to three PCIe Devices within HGX-1
        - 2 x 16 Ports are available for external connection. (The physical connectors are x8; each x16 Port may bifurcate to two x8 or four x4 Ports.)
- External PCIe Links:
    - The HGX-1 provides sixteen x8 PCIe Ports (via eight x16 cable/connectors).
    - These PCIe Ports may interconnect to external PCIe Hosts, external HGX-1, PCIe Target Devices, or to internal Switches for better peer-to-peer PCIe traffic for Devices within the HGX-1.
    - The pinout of the x8 or x16 Connectors and the circuitry at the end of the of the Cable at the Head Node interface provide Host-presence, Host-power-presence, cable-presence, and correct-seating indications.
- PCIe Fabric Management Network
- PCIe Fabric Manager (FM)
- Internal Head Node:
    - The same SoC PCIe Card that performs as a FM may plug into a UBB PCIe Slot and perform as the Head Node for the Expansion box.
    - This Head Node couples to the HGX-1 BMC via a cable to control Power Supplies, Fans, etc.
- Interface to an External Head Node(s):
    - Project Olympus Server (Head Node) is optimized for front-IO access
        - Using one x16 PCIe Link, up to eight Head Nodes may interface HGX-1
        - The HGX-1 may link to up to sixteen Hosts or External PCIe Ports using x8 Links (via x8 or x16 Cables)
        - Each x8 Link may bifurcate to two x4 Links (at the destination of the x8 Cable).
    - Each external Head Node resides underneath an Expansion Chassis and provides x16 PCIe Links to HGX-1. (HGX-1 provides eight x16 PCIe Links.)
    - To facilitate cable management, serviceability, and EMI containment, a portion of the front bezel of the HGX-1 is offset from the front plane of the Rack (and that of the Head Node) to allow front-cabling behind the Chassis Door of the HGX-1. This Chassis Door along with the rest of the Chassis forms a Faraday Cage for electromagnetic field containment to reduce EMI.
- Up to four HGX-1 Expansion Chassis may interface via a vertical Chimney behind the Chassis Door(s).
- Head Node:
    - An external Head Node such as Project Olympus Server is optimized for front-IO.

  - A Head Node may couple to up to four Expansion Chassis. (Each Expansion Chassis may include up to 36 PCIe devices.)
  - For various balance of CPU-to-IO ratio, up to sixteen Head Nodes may bind to up to 36 PCIe devices in an HGX-1.
- Upper Baseboard (UBB):
  - The Expansion Box consists of a UBB which provides four front-accessible, x16, single-width, ¾LFH PCIe Slots.
  - Within HGX-1, behind the row of PCIe Slots in the base configuration, there are four PCIe Switches toward the front of the Chassis for ease of interconnection to the Host Node which resides underneath the HGX-1. (no PCIe Re-timer is needed for this configuration.)
  - Using 400-pin MegArray® connectors, eight embedded PCIe Devices within the Chassis reside behind the PCIe Switches on UBB.
  - Using a second 400-pin MegArray® connector and PCB traces, UBB interconnects these 8 embedded PCIe Devices to each other to form a private peer-to-peer network such as nVidia's NVLINK in a hyper-cube + double-ring topology.
  - UBB can be flipped around by 180 degrees in a different Chassis so that it may receive PCIe Links from the rear and may present PCIe Cards and Devices with front-accessible connectors.
- Mezzanine:
  - Using Mezzanine boards plugged into the MegArray® connectors, UBB offers up to eight double-width, 300W, x16 PCIe add-in Cards.
  - Several variants of Mezzanine boards provide flexible PCIe Slot or Device configurations.
  - MEZZ1x16 includes one x16 PCIe Slot connected to a PCIe Switch for Host Interface.
  - MEZZ2x16 includes two x16 PCIe Connectors to provide for two x8 PCIe Slots connected to the Host via the PCIe Switch Fabric. The upper x8 port of the two x16 PCIe connectors are interconnected to form a private Card-to-Card Link.
  - MEZZ4x16 includes four x16 PCIe Connectors to provide for four x4 PCIe Slots connected to the Host via the PCIe Switch Fabric. The three upper x4 port of each x16 PCIe connector are interconnected to form a fully-connected, private Card-to-Card network.
- Double-device Cards (e.g., FPGA):
  - Via a PCIe Riser board and a x8 OCuLink Cable, one of the PCIe Slots provides the support for an add-in Card implementation with two PCIe Devices.
  - One PCIe device receives its PCIe Link via the normal PCIe Slot Connector; while, the other PCIe Device receives its x8 PCIe Link via a x8 OCuLink Cable.
- JTAG Tree:
  - The HGX-1 provides a JTAG Tree to program all programable devices such as CPLDs or PCIe Devices/Slots. JTAG signals connected to PCIe Slots are isolated from each other. The BMC is the primary JTAG Root with a 6-bit Device Select.
- I^2C:
  - I^2C GPIO Device: JTAG Device Select
  - Temp Monitors
  - Clock Synthesizer Control
  - PCIe Slots and Devices
  - Power Supplies power monitoring

- GPO for various Reset signals

Panel-mount connector:

- Provisions for 32 x QSFP+ at the IO-plate of the HGX-1 (two Rows of 16 each)
- Panel-mount QSFP+ connectors cable to internal dual-SERDES 50GE or quad-SERDES 100GE Ports.

Internal 100GE Switch in the top U:

- As an alternative, a 48-ported 100GE Switch Board interfaces to the 32 QSFP+ Ports of the HGX-1 IO-Plate in the upper U. Up to 32 internal devices connect to this Switch using internal cables (32 x2, 16 x4, or a mix without over-subscription). With 2:1 oversubscription, 32 internal 100GE ports connect to an External ToR via this Switch.
- Notes:
  - Internal 100GE cables are not as thick as the ones for external cabling.
  - With smaller bend-radius of internal cables, up to 32 internal PCIe Cards may cable to this Switch using the upper U space above PCIe Cards.
  - This Ethernet Switch may be designed so that it does not block the inter-HGX-1 Chimney.

5 High-level Specification

The following sections describe a high-level specification for realizing the architectural goals outlined in the previous section.

Table 2 shows the high-level mechanical specification of Microsoft Project Olympus HGX-1.

Table 2 High-level Mechanical Specification

Feature/ Function	Project Olympus Server HGX-1 Specification
Mechanical Specification	
Form Factor Height	4U (172mm)
Chassis Dimensions: Width (W) x Length (L)	441mm (W) x 947mm (L)
Rack Width	482.6 mm (19-inch EIA)
Upper Base Board Dimensions – Width (W) x Depth (L)	UBB Dimensions: 425mm (W) x 576mm (L)
Total Mass	Less than 45 kg
Power blind mating	6 x C14 Power Plugs
I/O cards support	Four FHHL x16 Gen3 PCIe Cards Eight SXM2 Modules Or Eight double-width FH¾L PCIe Cards via MEZZ1x16
Cabling access	Cold aisle cabling for IO and inter-chassis Links Rear-access Power and Management
Servicing	Cold aisle servicing PCIe Cards and Accelerators Hot aisle servicing PSUs

Power Supply Unit	Two 3-pahse feeds: 6 x 1600W (86mm x 40mm x 197mm)
SATA devices support	N/A
Fans	12 (N+2) non-hot-swap (60mm x 60mm x 38mm)
Optional remote heat sink	N/A
Front Panel	Four FHHL x 16 PCIe Cards RJ45 for 1GE Chassis Management Port RJ45 for sideband power control and presence
Rear Panel	RJ45 for 1GE Chassis Management Port RJ45 for sideband power control and presence Three x4 PCIe Ports for Fabric Management Six C14 Power Plugs
Electromagnetic Interference (EMI) Mitigation	For EMI containment, EMI shielding, and grounding are considered at the Chassis assembly level. The Chassis supports a top cover that fits within the 4U envelope to prevent leakage of electromagnetic fields and airflow
Grounding and Return	The chassis grounding/return is provided to the UBB from the tray assembly through the alignment and mounting holes that secure the UBB to the tray. The UBB is also tied to the PSU ground through the 12 V connector.
Environmental	Inlet Temperatures: Operating 50°F to 95°F (10°C to 35°C) Inlet Temperature: Non-Operating -40°F to 140°F (-40°C to 60°C) Rate of change less than 36°F (20°C)/hour Acoustics Less than 7.5 bells at system idle mode Non-operational vibration Random 1.09Grms, 30 minutes per axis, each of 3 mutually orthogonal axes. Non-operational shock Half sine 20G/11ms, 1 shock per direction of 3 mutually perpendicular axes, total 6 shocks.
Regulatory and Safety Requirements	UL, BSMI
Thermals	The required airflow to maintain thermal integrity shall not be more than 158CFM per kW of consumed power (even with failed fans). The Expansion Chassis must be able to operate at its maximum power configuration without performance degradation while meeting reliability requirements with two failed fans for an N+2 configuration (12 fans running). Failed fan testing should be conducted on all fans with the worst case failed fan locations as the minimum requirement. Variable fan speed capability shall be implemented to enable the rack to minimize energy consumption of the air movers and facilities in conditions that permit it. The speed of airflow is based on component temperature requirements within the chassis.

Table 3 shows a high-level specification of Project Olympus Server and the associated HGX-1.

Table 3 High-level Specification

Feature/ Function	Description
	Project Olympus Server HGX-1 Specification
Processors	
SoC	Fabric Manager SoC
Number of Sockets	(1) Optional
I/O Devices/Functions and Slots	
PCI Express Connectors	Four x16 Eight x16 using MEZZ1x16
SXM2	Eight nVidia SXM2 Modules (Volta 100 + six NVLINKs)
SoC Ethernet interfaces	One 10GE Fabric Manger
System Management	
BMC device	ASPEED AST2400 or AST2500 family, on Chassis Management Board
BMC Boot Flash	BMC SPI Flash – Minimum 32 MB device (Windbond W25Q256FVFIG)
1 GbE PHY	Ethernet Interface – Dedicated BMC 1GE (RJ45) at front-panel and rear-panel.
CPU Thermal Monitoring	N/A
Telemetry through BMC I^2C	The following functions support I^2C access from the BMC: ▪ Local Hot Swap Controllers ▪ Power Supplies ▪ Key Voltage Regulators (processors and memory) ▪ Temperature sensors ▪ FRUID PROM
BMC UART	N/A
JTAG Master	▪ The BMC provides JTAG controller for all PCIe Slots and Devices. ▪ Debug Header overrides BMC.
Voltage Regulators	▪ Voltage Regulators that support I^2C or PMBus should be available to the BMC. ▪ Clock Circuitry that supports I^2C should be available to the BMC.
FRUID PROM	Accessible via local BMC: ▪ The system includes a 256Kb serial EEPROM (AT24C256) for storing manufacturing data and the serial number of MAC. Accessible via local External Hosts: ▪ One FRUID PROM for each one of the eight External x16 PCIe Links ▪ Up to eight external Hosts may connect to this Expansion Chassis and see the FRUID PROMs. ▪ These FRUID PROMs may be programmed the same or as Host-specific configuration
Temperature sensors	▪ The system supports I^2C temperature sensors (TMP75) for monitoring the inlet and outlet air temperatures. ▪ For accurate temperature reading, care shall be taken to not place these temperature sensors close to component or board heat sources.
Hot Swap Controllers	▪ The system includes I^2C support for nine hot-swap controllers for power monitoring and power capping. ▪ All the controllers are located on PDB

Feature/ Function	Description
	Project Olympus Server HGX-1 Specification
LEDs	Each PSU comprises two individual status indications using one LED ▪ Green: Output On and OK ▪ Amber: Fault
Fan Control	The system shall support control of twelve 60 mm fans located on the middle wall of the chassis. Separate six fan zones for six fan modules (two fans per module). Support 12V power, six PWM, and twelve TACH signals for fan control
Security	
TPM 2.0	N/A
FPGA Card Support	
FPGA Card	The HGX-1 supports a dual-device FHHL x16 PCIe form factor Card for applications such as an FPGA+NIC. Such a Card installs in a x16 Riser in a x16 PCIe Slot. This Riser provides a x8 Link to the Card via Gold Fingers and a x8 Link via an OCuLink cable.
Power Input and Voltage Regulators	
Power Supply Unit interface to Motherboard	The Upper Baseboard isolates power into 9 power zones. Using HSCs, the power distribution board protects each power plane against over-current.
Rack Management	The Chassis Manager (BMC) communicates with Project Olympus Standalone Rack Manager via an 1GE link (RJ45) and a sideband control link (RJ45) which provides presence, power enable, and throttle functions.
PSU Management	The Chassis Manger manages the six PSUs
Hot Swap Controller	The Chassis Manger controls Hot-swap Controllers (HSC) via I^2C bus. The system shall include I^2C support for nine hot swap controllers for power monitoring and power capping. All the controllers are located on PDB.
Power Capping	N/A
Overcurrent Protection	HSCs are responsible for detecting current levels that indicate catastrophic failure. In such events, the HSC that detects a failure disables 12V rail to the section it monitors (typically by disabling the HSC's input FETs).
M.2 Support	The HGX-1 supports PCIe Cards as M.2 Carriers
Service Requirements	
USB Service Port	N/A
LED visibility	The Expansion Box includes the following LEDs: ▪ UID LED ▪ Power Status LED ▪ Attention LED

6 Mechanical Features

Figure 1 shows the dimensions of a 4U chassis suitable for Project Olympus HGX-1 Expansion Chassis.

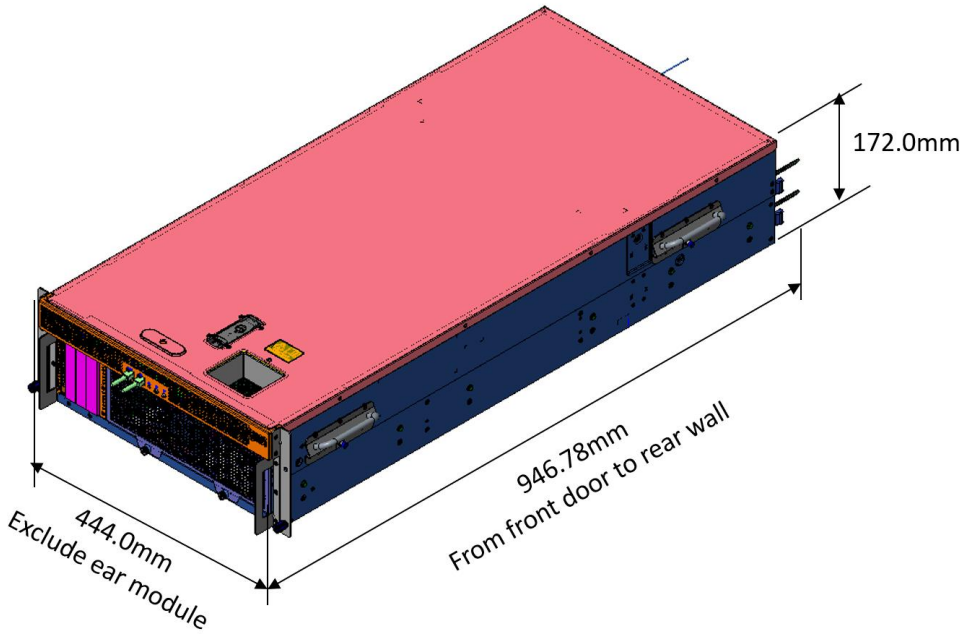


Figure 1 4U HGX-1

The HGX-1 may house various PCIe-connected Accelerators and IO peripherals. Figure 2 shows an isometric view of the HGX-1 with eight SXM2 Modules such as those for nVidia Volta V100 family of GPGPUs.

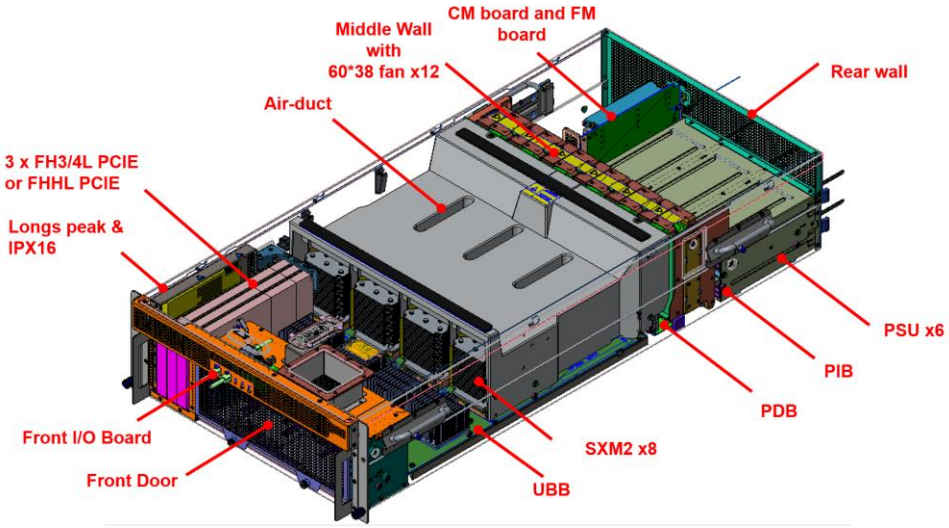


Figure 2 Eight SXM2 Modules

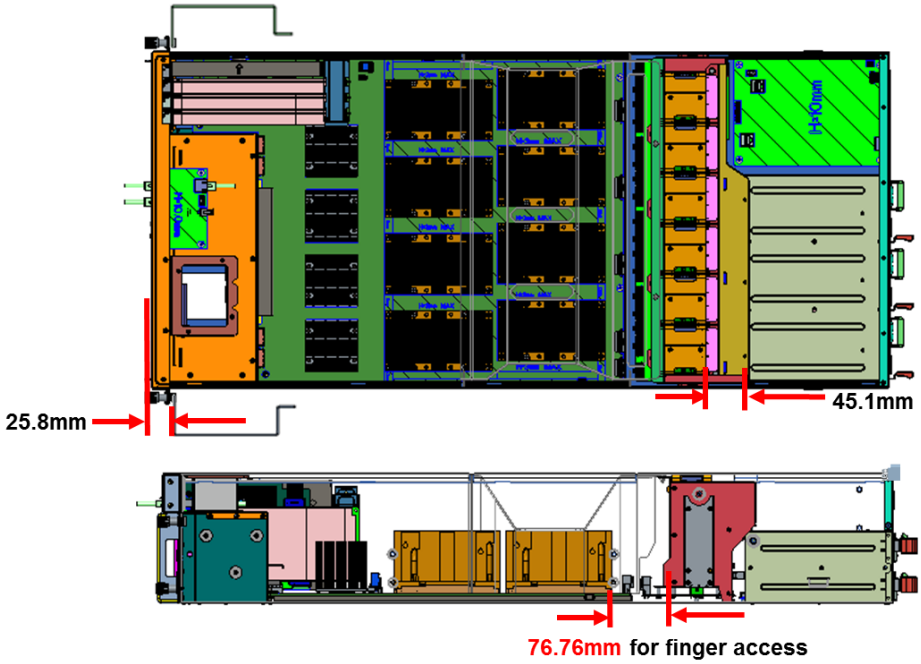


Figure 3 Top-view and Side-view of HGX-1 with SXM2 Modules

The HGX-1 provides an Upper Baseboard (UBB) with four PCIe Switches to couple to an external Motherboard (MB) via a set of cables. Figure 4 shows the UBB and the placement of eight SXM2 Modules.

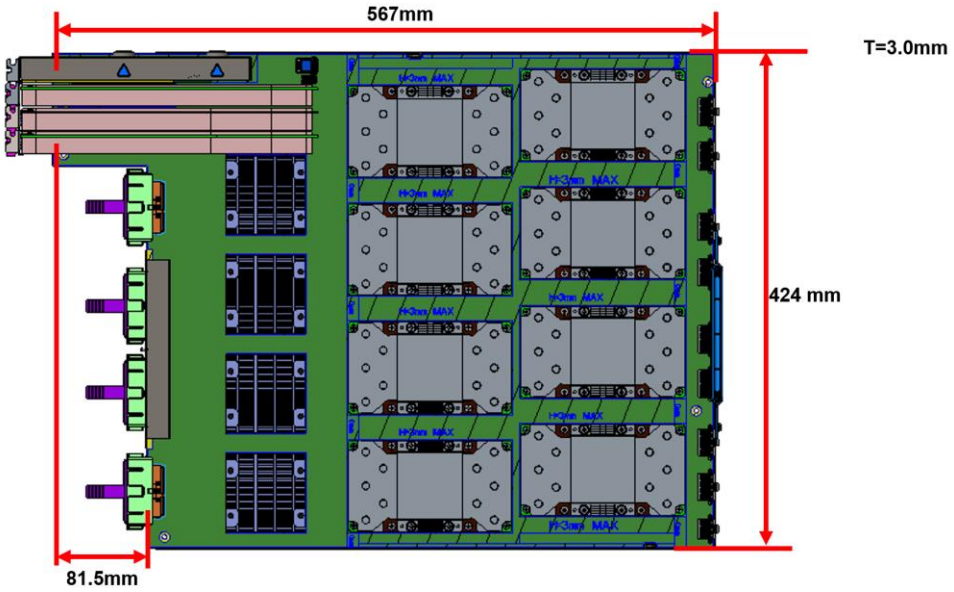


Figure 4 UBB with eight SXM2 Modules

Figure 5 shows component placement on UBB such as PCIe Switches, eight pairs of 400-pin Meg-Array® Connectors, front-accessible x16 PCIe Slots, and eight x16 external PCIe connectors (four of which are under the PCB).

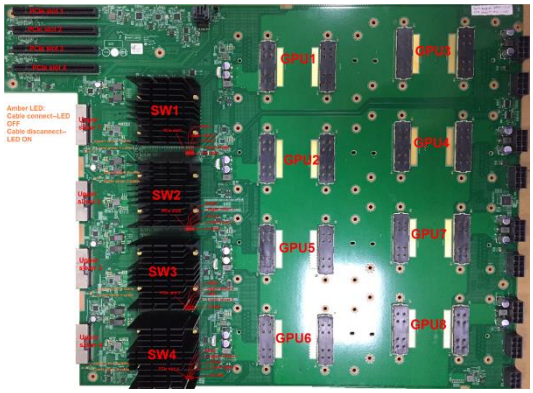


Figure 5 UBB Component Placement

Figure 6 shows an isometric view of the HGX-1 with eight double-width FH¾L 300W PCIe Cards.

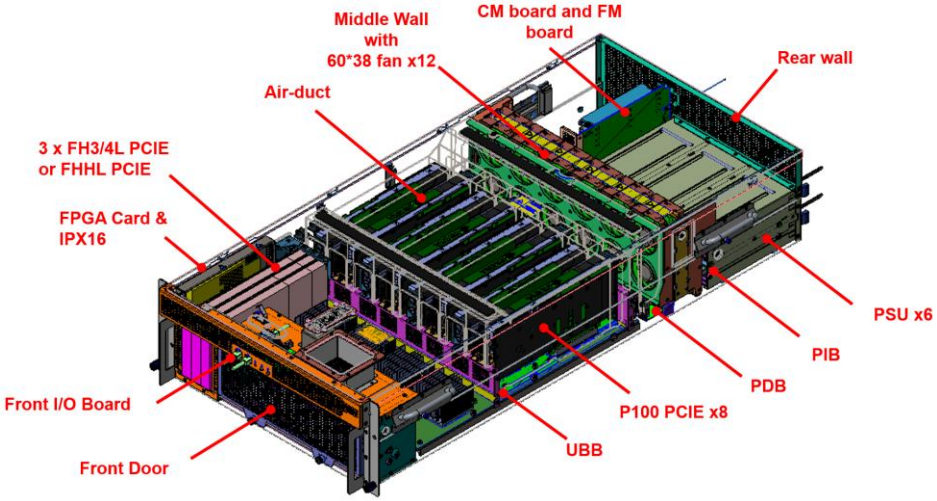


Figure 6 HGX-1 with eight double-width PCIe Cards

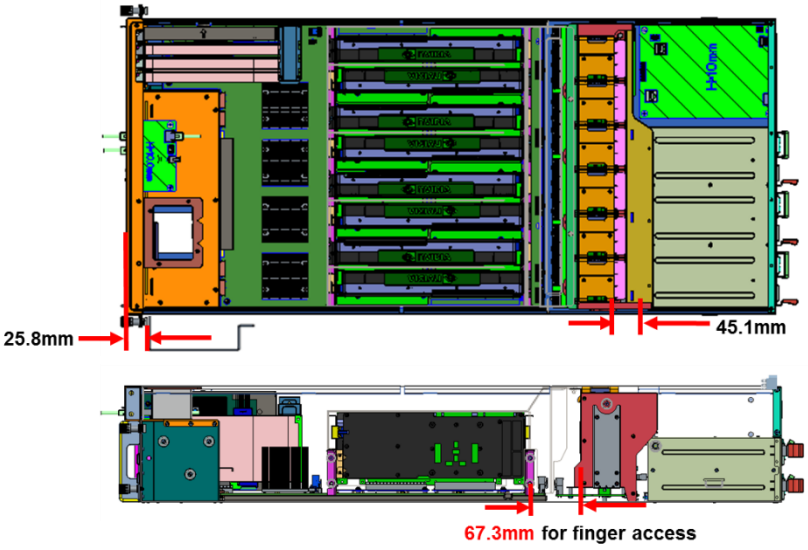


Figure 7 Eight MEZZ1x16 mezzanines support eight x16 PCIe Cards on UBB

Figure 7 and Figure 8 show MEZZ1x16 which plugs into Meg-Array® connectors on UBB and provides eight x16 PCIe Slots.

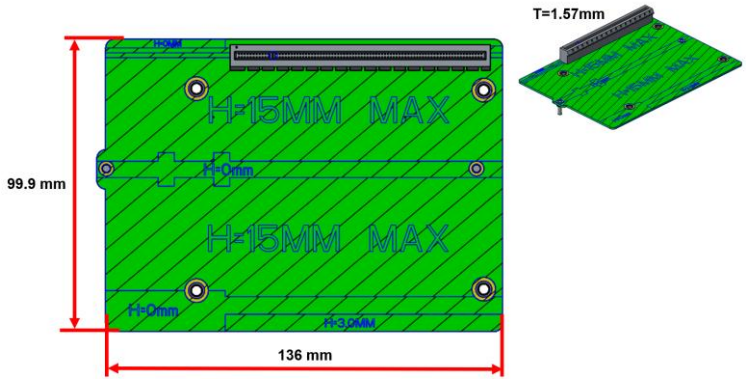


Figure 8 MEZZ1x16

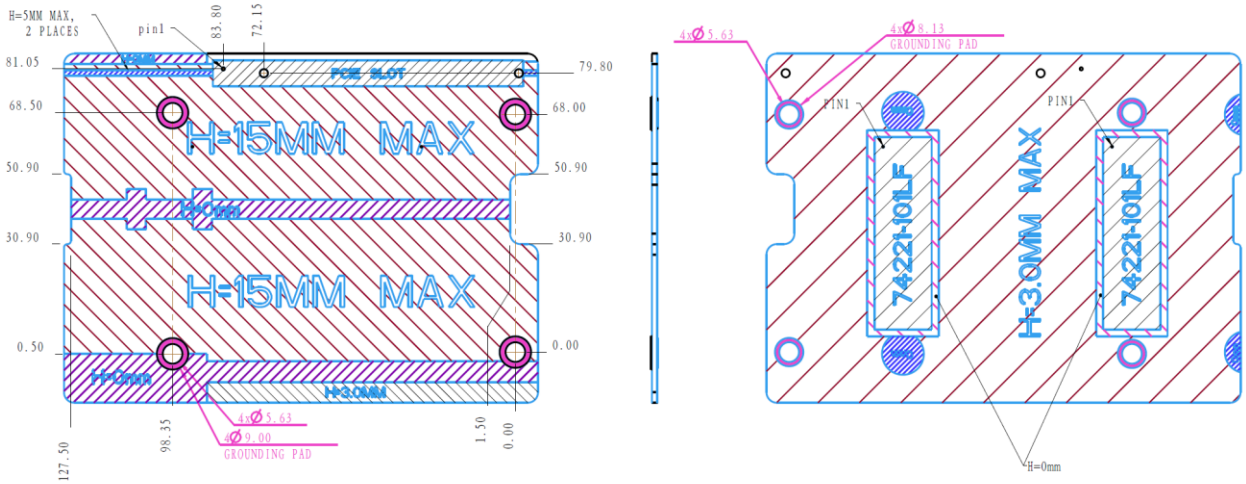


Figure 9 MEZZ1x16 Mechanical Dimensions

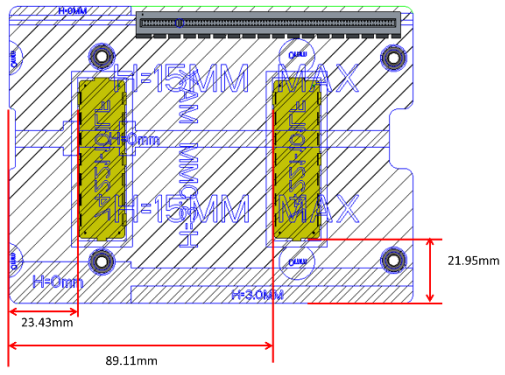


Figure 10 MEZZ1x16 Connector Placement

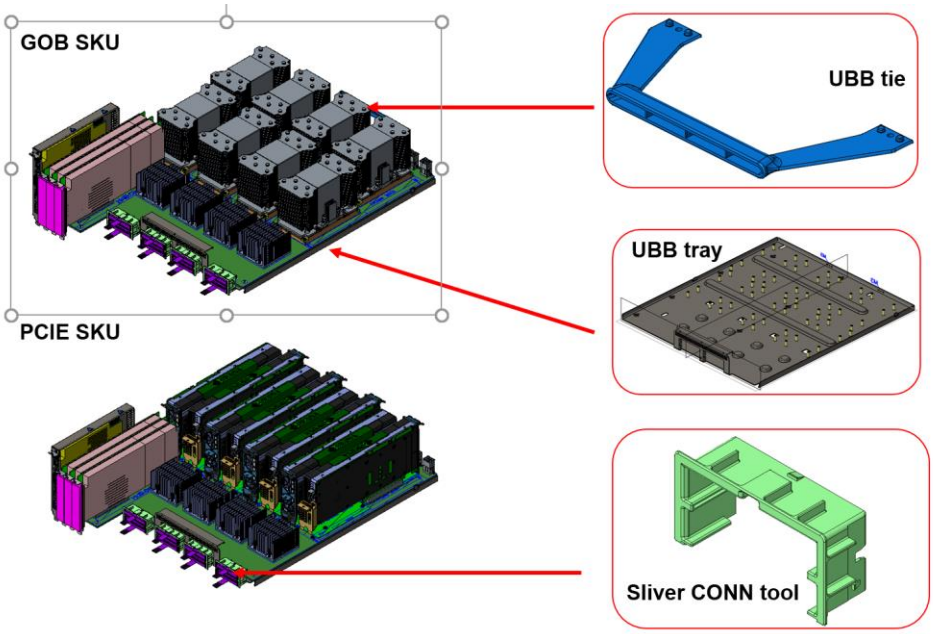


Figure 11 UBB Tray with PCIe Card and SXM2 options

Figure 12 shows the front and rear views of the HGX-1.

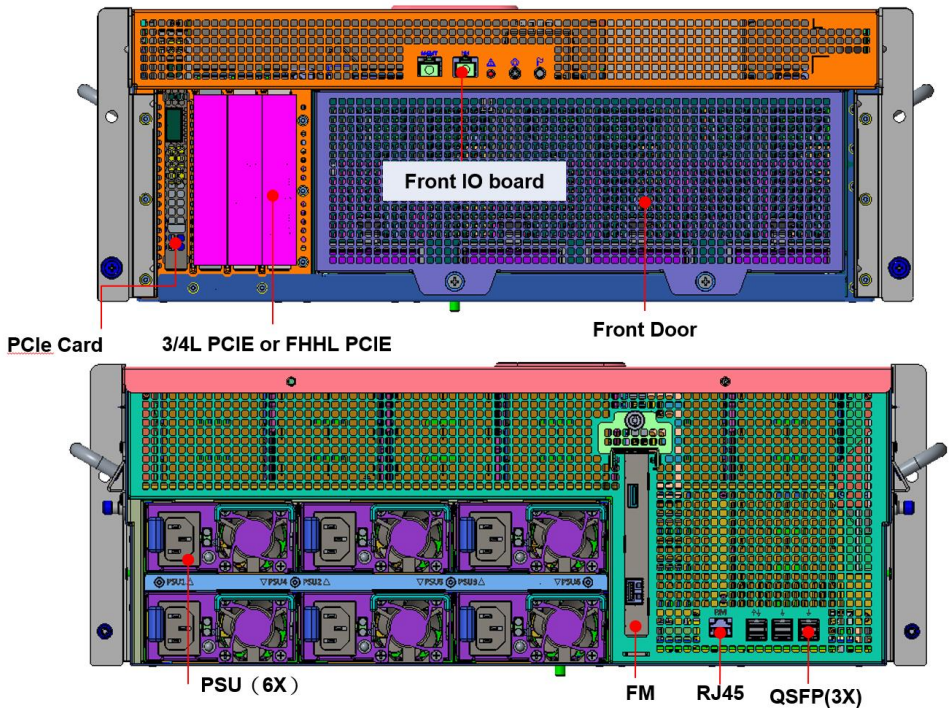


Figure 12 Front and Rear Views of the HGX-1

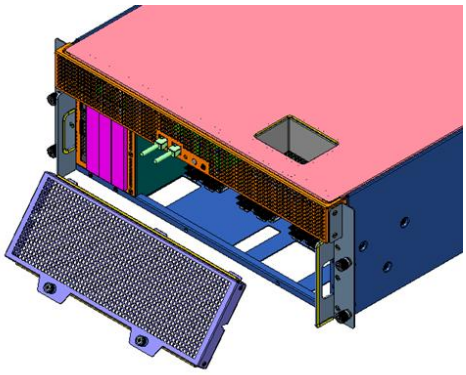


Figure 13 Front Door completes the Faraday Cage

7 Power

Consistent with Project Olympus Rack Architecture, Project Olympus HGX-1 provides balanced 3-phase power with two redundant feeds (A and B). The HGX-1 includes 6 Power Supplies in an N+N redundant configuration and provides load-balancing over three phases. (two feeds of 3-phase each.) These Power Supplies provide 1600W each.

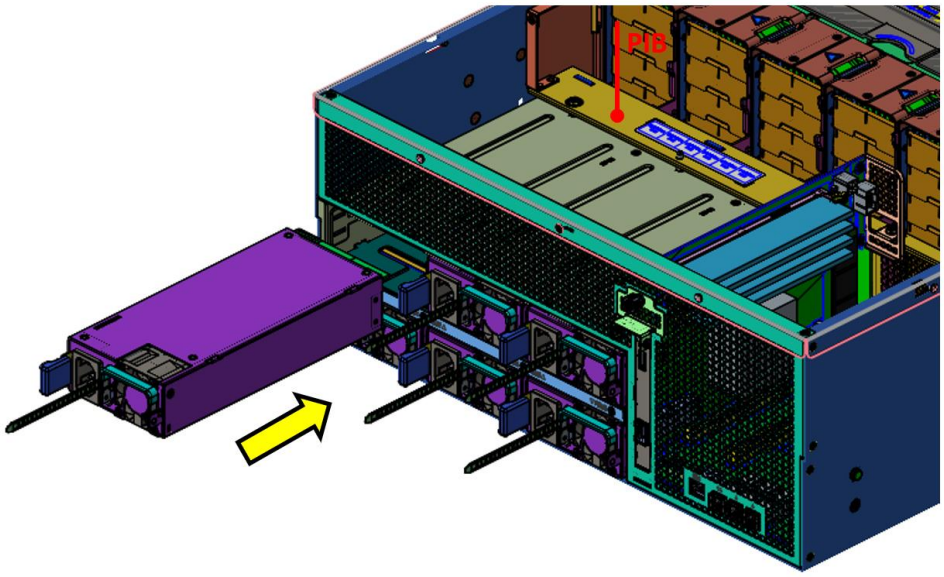


Figure 14 Six 1600W Power Supplies in the rear of the HGX-1 (PSU dimensions: 86mm x 40mm x 197mm)

Figure 21 shows UBB Power Islands. To enable the e-fuses (via Hot-swap Controller) detect over-current, the UBB has several power islands. Separate e-fuses (HSC) monitor and control each power island.

7.1 Power Adapter Module (PAM)

While HGX-1 Expansion Chassis includes six PSUs with C14 power receptacles, it couples to Project Olympus-compatible Server Blades. Project Olympus Rack architecture defines a Power and Management Distribution Unit (PMDU) for power and rack management interfaces.

To enable compatible power cabling, the Power Adapter Module (PAM) shown in Figure 15 through Figure 19 is suitable for interfacing the FCI power connector of a Project Olympus Blade to two C14 power receptacles and one RJ45 connector for power control signals.

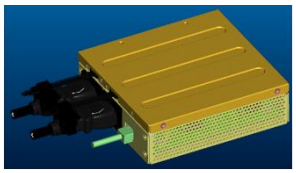


Figure 15 Power Adapter Module (PAM)

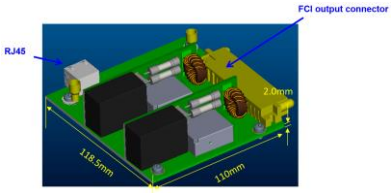


Figure 16 An inside view of PAM

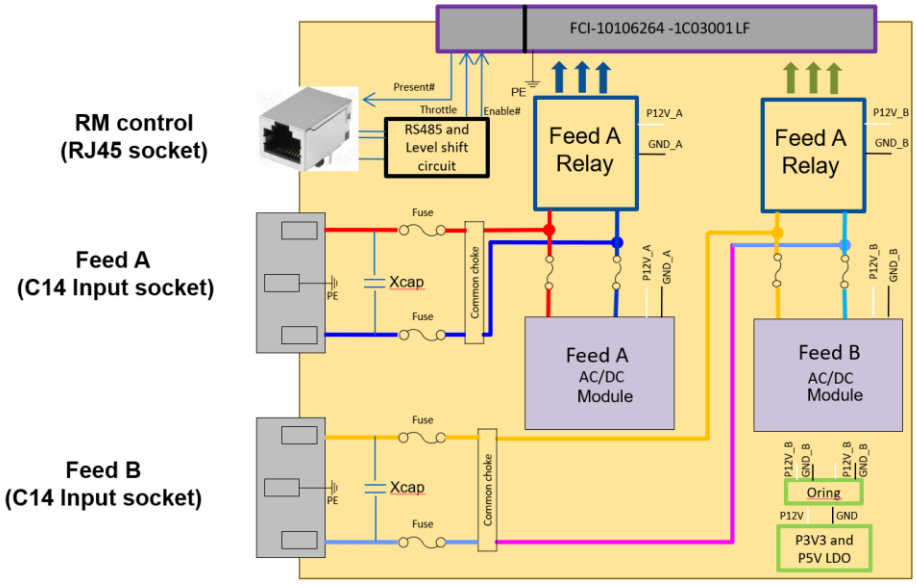


Figure 17 Power Adapter Module Block Diagram

Pin #	Signal	I/O	Voltage	Description
1	RM_THROTTLE+	I	RS485	Set server to throttle mode
2	RM_THROTTLE-	I	RS485	Set server to throttle mode
3	RM_BOOTSTRAP+	I	RS485	DEPOP
4	/A			
5	BLADE_PRESENT#	O	3.3V	Indicates Blade is present
6	RM_BOOTSTRAP-	I	RS485	DEPOP
7	GND		0V	Ground
8	RM_OFF	I	5V	

Figure 18 PAM RJ45 pinout (interface to Project Olympus Rack Manager)

Pin #	Signal	I/O	Voltage	Description
P1	Safety Ground	-		
P2	Not Used	-		
P3	Feed B Neutral	O		
P4	Feed B Phase C	O		
P5	Feed A Neutral	O		
P6	Feed A Phase C	O		
P7	Feed B Neutral	O		
P8	Feed B Phase B	O		
P9	Feed A Neutral	O		
P10	Feed A Phase B	O		
P11	Feed B Neutral	O		
P12	Feed B Phase A	O		
P13	Feed A Neutral	O		
P14	Feed A Phase A	O		
A1	LR_SELECT	O	3.3V	DEPOP
A2	-			
A3	Analog Return		0V	
B1	-			
B2	-			
B3	-			
C1	BLADE_THROTTLE	O	3.3V	Default Disable(Low)
C2	BLADE_ENABLE#	O	3.3V	Default Enable(Low)
C3	BLADE_PRESENT#	I	3.3V	
D1	-			
D2	-			
D3	-			

Figure 19 PAM interface pinout for interfacing to Project Olympus Blade's FCI (PMDU) connector

8 Cooling

Twelve 60mm Fans provide up to 5000W of cooling for the HGX-1 in an N+2 redundant configuration with less than 790 CFM of airflow to meet airflow specification of no more than 158 CFM per kW.

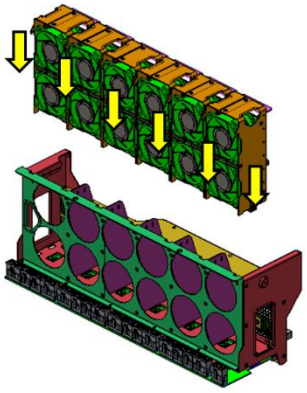


Figure 20 Twelve N+2 redundant Fans

9 Accelerators and Compute Elements

Project Olympus HGX-1 supports various compute accelerator (GPGPUs, FPGAs, ASICs, SoCs) or PCIe peripheral devices. For example, Figure 2 shows a configuration with eight nVidia Pascal P100 GPGPU family in SXM2 form factor suitable for Deep Machine Learning (DL, Training); Figure 6 shows eight accelerators in double-width, 300W PCIe form factor for visualization or HPC; and Figure 38 shows thirty-two 75W accelerators in single-width FHHL PCIe form factor for DL/Inference or Transcoding.

10 PCIe Switch Fabric

The Upper Baseboard of the HGX-1 implements a PCIe Gen-3 Switch Fabric. The HGX-1 interconnects four 96-lane PCIe Switches (PEX9797 from Broadcom) capable of efficient peer-to-peer traffic, multi-Hosting, and Shared IO. Each PCIe Switch has one x1 PCIe port for Fabric Management. Each PCIe Switch provides six x16 Ports. One x16 Port is dedicated for inter-Switch Link (ISL). Three x16 Ports connect to three PCIe Devices on UBB. Two x16 Ports are available for external connection.

The physical connectors for the external Links are x16 TE Sliver Cable/Connectors. Each x16 Port may bifurcate to two x8 or four x4 Ports. Using x8 cables, the HGX-1 can export up to sixteen x8 PCIe Links.

These sixteen PCIe Ports may interconnect to external PCIe Hosts, external HGX-1, PCIe Target Devices, or to internal Switches for better peer-to-peer PCIe traffic for Devices within the HGX-1.

While, this PCIe Switch Fabric is cable of multi-Hosting and Shared IO which require enabling the Fabric Mode and a Fabric Management Controller, the simplest implementation includes PCIe traffic multiplexing between one Host and PCIe end Devices and peer-to-peer traffic between PCIe Devices.

Other configurations include one Host connecting to up to four interconnecting HGX-1 or up to sixteen Hosts to one HGX-1.

11 Interconnects

Figure 21 depicts various HGX-1 Modules and PCIe Cabling to an External Chassis.

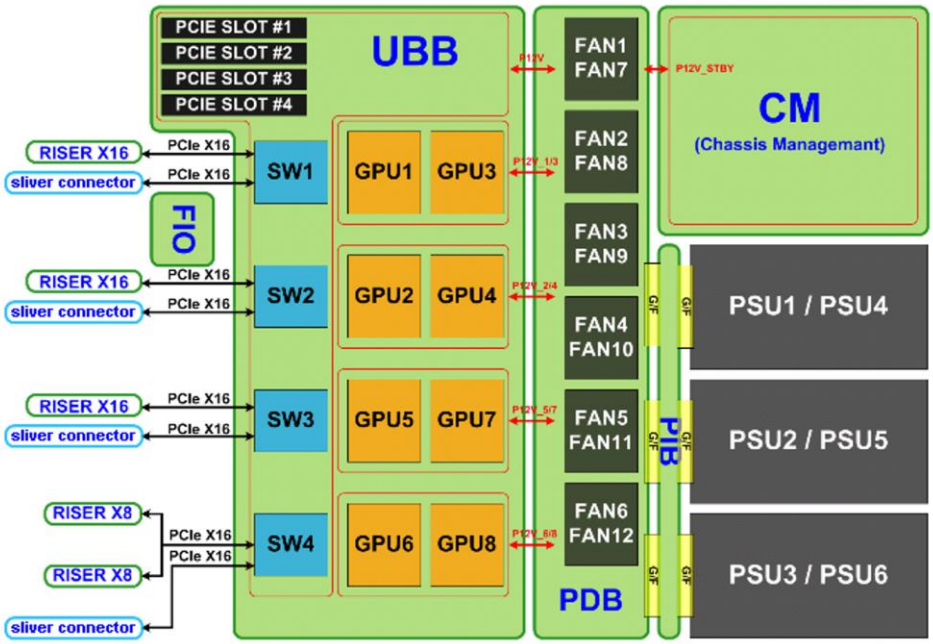


Figure 21 Various PCB (Modules) and External PCIe Cabling

11.1 Inter-Chassis PCIe Cables

Figure 21 shows eight x16 connectors that are available to connect the HGX-1 PCIe Ports to either external Hosts, to itself for peer-to-peer interconnect, or to other expansion boxes.

The pinout of the x16 connector and the circuitry at the end of the of the Cable at the Head Node interface provide Host-presence, Host-power-presence, cable-presence, and correct-seating indications.

11.2 Riser Boards

Figure 21 showed that HGX-1 provides eight x16 PCIe Links to connect to Project Olympus Server.

Figure 22 and Figure 23 show various Riser Boards suitable for extracting PCIe Links from standard PCIe Slots. Three x16 and two x8 PCIe Risers plug into Project Olympus Universal Motherboard to provide up to four x16 PCIe Links (plus miscellaneous sideband signals) to the HGX-1.

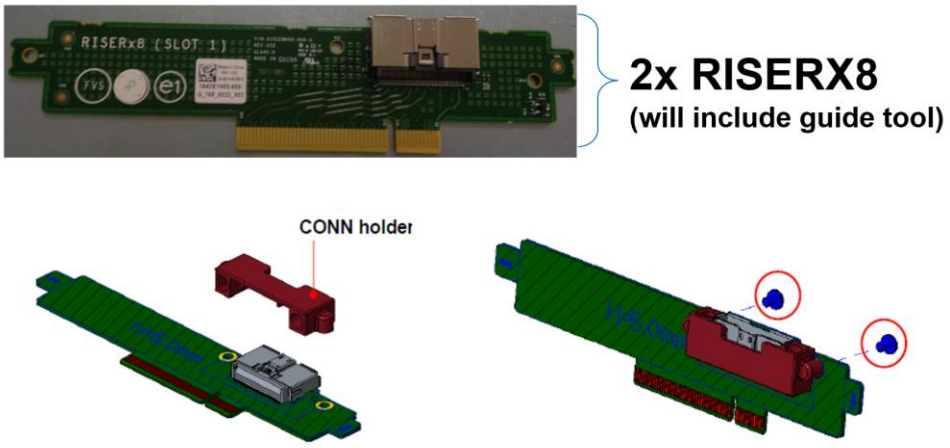


Figure 22 PCIe Riser for x8 PCIe Slots (x8 PCIe Cable + Misc. sideband signals)

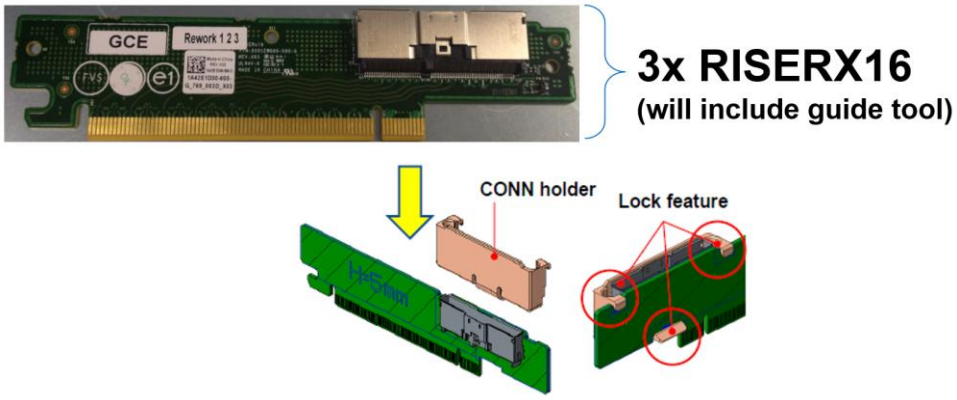


Figure 23 PCIe Riser for x16 PCIe Slots (x16 PCIe Cable + Misc. sideband signals)

Various connector/cable assemblies may accommodate Host Nodes of different capabilities.

HGX-1 Expansion Chassis incorporates TE Sliver connector/cable assemblies for the eight x16 PCIe Ports.

Figure 24 shows the pinout of UBB x16 PCIe Connector.

UBB (TE-2294190-3)							
A SIDE				B SIDE			
Pin	Signal Name	Direction	Description	Pin	Signal Name	Direction	Description
1	GND	-		124	GND	-	
2	PRSNT_A_N	I , PU	Cable Present Signal on A side	123	GND	O	Present GND for UBB Cross Link
3	ANY_HOST_PRSNT_N	I , PU	Present Signal for Power control	122	SSC_ISO_EN	I , PU	Present Signal for PEX9797 Configation
4	GND	-		121	GND	-	
5	PE_RST_N	I , PU	PCIe Reset signal from HOST	120	PE_RST_N	I , PU	PCIe Reset signal to next stage UBB
6	P12V_HOST_GD	I , PU	12V Power on control form HOST	119	P12V_HOST_GD	I , PU	12V Power on control to next stage UBB
7	GND	-		118	GND	-	
8	REFCLK+	I	PCIe Reference Clock Input	117	SMCLK	I	SMbus from HOST
9	REFCLK-	I		116	SMDAT	I/O	
10	GND	-		115	GND	-	
11	P3V3_HOST	I	3.3V Power input form HOST	114	NC	-	
12	GND	O	Present GND for UBB Cross Link	113	PRSNT_B_N	I , PU	Cable Present Signal on B side
13	GND	-		112	GND	-	
14	PETp0	O	PEX9797 PCIe Transmit Path	111	PERp0	I	PEX9797 PCIe Receive Path
15	PETn0	O		110	PERn0	I	
16	GND	-		109	GND	-	
17	PETp1	O	PEX9797 PCIe Transmit Path	108	PERp1	I	PEX9797 PCIe Receive Path
18	PETn1	O		107	PERn1	I	
19	GND	-		106	GND	-	
20	PETp2	O	PEX9797 PCIe Transmit Path	105	PERp2	I	PEX9797 PCIe Receive Path
21	PETn2	O		104	PERn2	I	
22	GND	-		103	GND	-	
23	PETp3	O	PEX9797 PCIe Transmit Path	102	PERp3	I	PEX9797 PCIe Receive Path
24	PETn3	O		101	PERn3	I	
25	GND	-		100	GND	-	
26	PETp4	O	PEX9797 PCIe Transmit Path	99	PERp4	I	PEX9797 PCIe Receive Path
27	PETn4	O		98	PERn4	I	
28	GND	-		97	GND	-	
29	PETp5	O	PEX9797 PCIe Transmit Path	96	PERp5	I	PEX9797 PCIe Receive Path
30	PETn5	O		95	PERn5	I	
31	GND	-		94	GND	-	
32	PETp6	O	PEX9797 PCIe Transmit Path	93	PERp6	I	PEX9797 PCIe Receive Path
33	PETn6	O		92	PERn6	I	
34	GND	-		91	GND	-	
35	PETp7	O	PEX9797 PCIe Transmit Path	90	PERp7	I	PEX9797 PCIe Receive Path
36	PETn7	O		89	PERn7	I	
37	GND	-		88	GND	-	
38	GND	-		87	GND	-	
39	PETp8	O	PEX9797 PCIe Transmit Path	86	PERp8	I	PEX9797 PCIe Receive Path
40	PETn8	O		85	PERn8	I	
41	GND	-		84	GND	-	
42	PETp9	O	PEX9797 PCIe Transmit Path	83	PERp9	I	PEX9797 PCIe Receive Path
43	PETn9	O		82	PERn9	I	
44	GND	-		81	GND	-	
45	PETp10	O	PEX9797 PCIe Transmit Path	80	PERp10	I	PEX9797 PCIe Receive Path
46	PETn10	O		79	PERn10	I	
47	GND	-		78	GND	-	
48	PETp11	O	PEX9797 PCIe Transmit Path	77	PERp11	I	PEX9797 PCIe Receive Path
49	PETn11	O		76	PERn11	I	
50	GND	-		75	GND	-	
51	PETp12	O	PEX9797 PCIe Transmit Path	74	PERp12	I	PEX9797 PCIe Receive Path
52	PETn12	O		73	PERn12	I	
53	GND	-		72	GND	-	
54	PETp13	O	PEX9797 PCIe Transmit Path	71	PERp13	I	PEX9797 PCIe Receive Path
55	PETn13	O		70	PERn13	I	
56	GND	-		69	GND	-	
57	PETp14	O	PEX9797 PCIe Transmit Path	68	PERp14	I	PEX9797 PCIe Receive Path
58	PETn14	O		67	PERn14	I	
59	GND	-		66	GND	-	
60	PETp15	O	PEX9797 PCIe Transmit Path	65	PERp15	I	PEX9797 PCIe Receive Path
61	PETn15	O		64	PERn15	I	
62	GND	-		63	GND	-	

Figure 24 UBB x16 PCIe Connector Pinout

Figure 25 shows the pinout of the TE Sliver connector for a PCIe Riser board with a single x16 cable connector.

RISERX16 (TE-2294190-3)							
A SIDE				B SIDE			
Pin	Signal Name	Direction	Description	Pin	Signal Name	Direction	Description
1	GND	-		124	GND	-	
2	NC	-		123	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
3	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	122	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
4	GND	-		121	GND	-	
5	NC	-		120	PE_RST_N	O	PCIe Reset signal to UBB
6	NC	-		119	P12V_HOST_GD	O	12V Power on control to UBB
7	GND	-		118	GND	-	
8	SMCLK	O	SMbus to UBB	117	REFCLK+	O	PCIe Reference Clock Output
9	SMDAT	I/O		116	REFCLK-	O	
10	GND	-		115	GND	-	
11	NC	-		114	P3V3_HOST	I	3.3V Power output to UBB
12	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	113	NC	-	
13	GND	-		112	GND	-	
14	PETp0	O	CPU PCIe Transmit Path	111	PERp0	I	CPU PCIe Receive Path
15	PETn0	O		110	PERn0	I	
16	GND	-		109	GND	-	
17	PETp1	O	CPU PCIe Transmit Path	108	PERp1	I	CPU PCIe Receive Path
18	PETn1	O		107	PERn1	I	
19	GND	-		106	GND	-	
20	PETp2	O	CPU PCIe Transmit Path	105	PERp2	I	CPU PCIe Receive Path
21	PETn2	O		104	PERn2	I	
22	GND	-		103	GND	-	
23	PETp3	O	CPU PCIe Transmit Path	102	PERp3	I	CPU PCIe Receive Path
24	PETn3	O		101	PERn3	I	
25	GND	-		100	GND	-	
26	PETp4	O	CPU PCIe Transmit Path	99	PERp4	I	CPU PCIe Receive Path
27	PETn4	O		98	PERn4	I	
28	GND	-		97	GND	-	
29	PETp5	O	CPU PCIe Transmit Path	96	PERp5	I	CPU PCIe Receive Path
30	PETn5	O		95	PERn5	I	
31	GND	-		94	GND	-	
32	PETp6	O	CPU PCIe Transmit Path	93	PERp6	I	CPU PCIe Receive Path
33	PETn6	O		92	PERn6	I	
34	GND	-		91	GND	-	
35	PETp7	O	CPU PCIe Transmit Path	90	PERp7	I	CPU PCIe Receive Path
36	PETn7	O		89	PERn7	I	
37	GND	-		88	GND	-	
38	GND	-		87	GND	-	
39	PETp8	O	CPU PCIe Transmit Path	86	PERp8	I	CPU PCIe Receive Path
40	PETn8	O		85	PERn8	I	
41	GND	-		84	GND	-	
42	PETp9	O	CPU PCIe Transmit Path	83	PERp9	I	CPU PCIe Receive Path
43	PETn9	O		82	PERn9	I	
44	GND	-		81	GND	-	
45	PETp10	O	CPU PCIe Transmit Path	80	PERp10	I	CPU PCIe Receive Path
46	PETn10	O		79	PERn10	I	
47	GND	-		78	GND	-	
48	PETp11	O	CPU PCIe Transmit Path	77	PERp11	I	CPU PCIe Receive Path
49	PETn11	O		76	PERn11	I	
50	GND	-		75	GND	-	
51	PETp12	O	CPU PCIe Transmit Path	74	PERp12	I	CPU PCIe Receive Path
52	PETn12	O		73	PERn12	I	
53	GND	-		72	GND	-	
54	PETp13	O	CPU PCIe Transmit Path	71	PERp13	I	CPU PCIe Receive Path
55	PETn13	O		70	PERn13	I	
56	GND	-		69	GND	-	
57	PETp14	O	CPU PCIe Transmit Path	68	PERp14	I	CPU PCIe Receive Path
58	PETn14	O		67	PERn14	I	
59	GND	-		66	GND	-	
60	PETp15	O	CPU PCIe Transmit Path	65	PERp15	I	CPU PCIe Receive Path
61	PETn15	O		64	PERn15	I	
62	GND	-		63	GND	-	

Figure 25 Pinout of the TE Sliver Connector of the single x16 Riser (RISERx16)

Figure 26 shows the pinout of the TE Sliver connector for the PCIe Riser with a single x8 PCIe Link.

RISERX8 (TE-2292069-1)							
A SIDE				B SIDE			
Pin	Signal Name	Direction	Description	Pin	Signal Name	Direction	Description
1	GND	-		74	GND	-	
2	NC	-		73	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
3	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	72	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
4	GND	-		71	GND	-	
5	NC	-		70	PE_RST_N	O	PCIe Reset signal to UBB
6	NC	-		69	P12V_HOST_GD	O	12V Power on control to UBB
7	GND	-		68	GND	-	
8	SMCLK	O	SMbus to UBB	67	REFCLK+	O	PCIe Reference Clock Output
9	SMDAT	I/O		66	REFCLK-	O	
10	GND	-		65	GND	-	
11	NC	-		64	P3V3_HOST	I	3.3V Power output to UBB
12	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	63	NC	-	
13	GND	-		62	GND	-	
14	PETp0	O	CPU PCIe Transmit Path	61	PERp0	I	CPU PCIe Receive Path
15	PETn0	O		60	PERn0	I	
16	GND	-		59	GND	-	
17	PETp1	O	CPU PCIe Transmit Path	58	PERp1	I	CPU PCIe Receive Path
18	PETn1	O		57	PERn1	I	
19	GND	-		56	GND	-	
20	PETp2	O	CPU PCIe Transmit Path	55	PERp2	I	CPU PCIe Receive Path
21	PETn2	O		54	PERn2	I	
22	GND	-		53	GND	-	
23	PETp3	O	CPU PCIe Transmit Path	52	PERp3	I	CPU PCIe Receive Path
24	PETn3	O		51	PERn3	I	
25	GND	-		50	GND	-	
26	PETp4	O	CPU PCIe Transmit Path	49	PERp4	I	CPU PCIe Receive Path
27	PETn4	O		48	PERn4	I	
28	GND	-		47	GND	-	
29	PETp5	O	CPU PCIe Transmit Path	46	PERp5	I	CPU PCIe Receive Path
30	PETn5	O		45	PERn5	I	
31	GND	-		44	GND	-	
32	PETp6	O	CPU PCIe Transmit Path	43	PERp6	I	CPU PCIe Receive Path
33	PETn6	O		42	PERn6	I	
34	GND	-		41	GND	-	
35	PETp7	O	CPU PCIe Transmit Path	40	PERp7	I	CPU PCIe Receive Path
36	PETn7	O		39	PERn7	I	
37	GND	-		38	GND	-	

Figure 26 Pinout of the connector for the single x8 PCIe Riser (RISERx8)

Figure 27 shows the pinout for one of the connectors (Port A) of a dual-x8 PCIe Riser (RISERx16 Port A).

RISERX16 PORT-A (TE-2292069-1)							
A SIDE				B SIDE			
Pin	Signal Name	Direction	Description	Pin	Signal Name	Direction	Description
1	GND	-		74	GND	-	
2	NC	-		73	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
3	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	72	PEX_PRSNT_GND	O	Present GND form C2010 to UBB
4	GND	-		71	GND	-	
5	NC	-		70	PE_RST_N	O	PCIe Reset signal to UBB
6	NC	-		69	P12V_HOST_GD	O	12V Power on control to UBB
7	GND	-		68	GND	-	
8	SMCLK	O	SMbus to UBB	67	REFCLK+	O	PCIe Reference Clock Output
9	SMDAT	I/O		66	REFCLK-	O	
10	GND	-		65	GND	-	
11	NC	-		64	P3V3_HOST	I	3.3V Power output to UBB
12	PEX_PRSNT_GND	O	Present GND form C2010 to UBB	63	NC	-	
13	GND	-		62	GND	-	
14	PETp0	O	CPU PCIe Transmit Path	61	PERp0	I	CPU PCIe Receive Path
15	PETn0	O		60	PERn0	I	
16	GND	-		59	GND	-	
17	PETp1	O	CPU PCIe Transmit Path	58	PERp1	I	CPU PCIe Receive Path
18	PETn1	O		57	PERn1	I	
19	GND	-		56	GND	-	
20	PETp2	O	CPU PCIe Transmit Path	55	PERp2	I	CPU PCIe Receive Path
21	PETn2	O		54	PERn2	I	
22	GND	-		53	GND	-	
23	PETp3	O	CPU PCIe Transmit Path	52	PERp3	I	CPU PCIe Receive Path
24	PETn3	O		51	PERn3	I	
25	GND	-		50	GND	-	
26	PETp4	O	CPU PCIe Transmit Path	49	PERp4	I	CPU PCIe Receive Path
27	PETn4	O		48	PERn4	I	
28	GND	-		47	GND	-	
29	PETp5	O	CPU PCIe Transmit Path	46	PERp5	I	CPU PCIe Receive Path
30	PETn5	O		45	PERn5	I	
31	GND	-		44	GND	-	
32	PETp6	O	CPU PCIe Transmit Path	43	PERp6	I	CPU PCIe Receive Path
33	PETn6	O		42	PERn6	I	
34	GND	-		41	GND	-	
35	PETp7	O	CPU PCIe Transmit Path	40	PERp7	I	CPU PCIe Receive Path
36	PETn7	O		39	PERn7	I	
37	GND	-		38	GND	-	

Figure 27 Connector pinout for Port A of the dual-x8 Riser (RISERx16 Port A)

Figure 28 shows the pinout for Port B of the dualx8 Riser (RISERx16 Port B).

RISERX16 PORT-B (TE-2292069-1)							
A SIDE				B SIDE			
Pin	Signal Name	Direction	Description	Pin	Signal Name	Direction	Description
1	GND	-		74	GND	-	
2	NC	-		73	NC	-	
3	NC	-		72	NC	-	
4	GND	-		71	GND	-	
5	NC	-		70	NC	-	
6	NC	-		69	NC	-	
7	GND	-		68	GND	-	
8	NC	-		67	NC	-	
9	NC	-		66	NC	-	
10	GND	-		65	GND	-	
11	NC	-		64	NC	-	
12	NC	-		63	NC	-	
13	GND	-		62	GND	-	
14	PETp8	O	CPU PCIe Transmit Path	61	PERp8	I	CPU PCIe Receive Path
15	PETn8	O		60	PERn8	I	
16	GND	-		59	GND	-	
17	PETp9	O	CPU PCIe Transmit Path	58	PERp9	I	CPU PCIe Receive Path
18	PETn9	O		57	PERn9	I	
19	GND	-		56	GND	-	
20	PETp10	O	CPU PCIe Transmit Path	55	PERp10	I	CPU PCIe Receive Path
21	PETn10	O		54	PERn10	I	
22	GND	-		53	GND	-	
23	PETp11	O	CPU PCIe Transmit Path	52	PERp11	I	CPU PCIe Receive Path
24	PETn11	O		51	PERn11	I	
25	GND	-		50	GND	-	
26	PETp12	O	CPU PCIe Transmit Path	49	PERp12	I	CPU PCIe Receive Path
27	PETn12	O		48	PERn12	I	
28	GND	-		47	GND	-	
29	PETp13	O	CPU PCIe Transmit Path	46	PERp13	I	CPU PCIe Receive Path
30	PETn13	O		45	PERn13	I	
31	GND	-		44	GND	-	
32	PETp14	O	CPU PCIe Transmit Path	43	PERp14	I	CPU PCIe Receive Path
33	PETn14	O		42	PERn14	I	
34	GND	-		41	GND	-	
35	PETp15	O	CPU PCIe Transmit Path	40	PERp15	I	CPU PCIe Receive Path
36	PETn15	O		39	PERn15	I	
37	GND	-		38	GND	-	

Figure 28 Connector pinout for Port B of the dual-x8 Riser (RISERx16 Port B)

11.3 Host Presence Detection

Figure 29 depicts the Host Presence Detect logic implemented via Riser Boards Cables for each Port. This logic facilitates Reset and Power Sequencing Handshake between the Host and the Expansion Chassis.

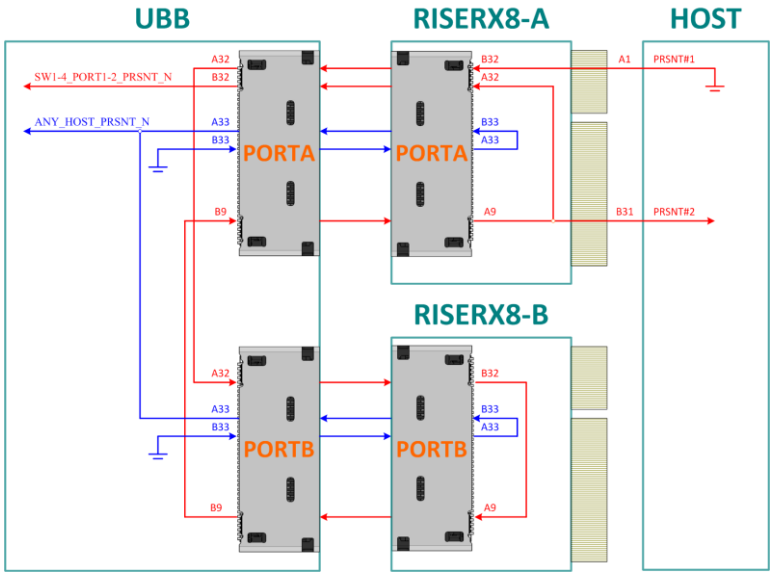


Figure 29 Host and Cable Presence Detect (Riser x8 A & B)

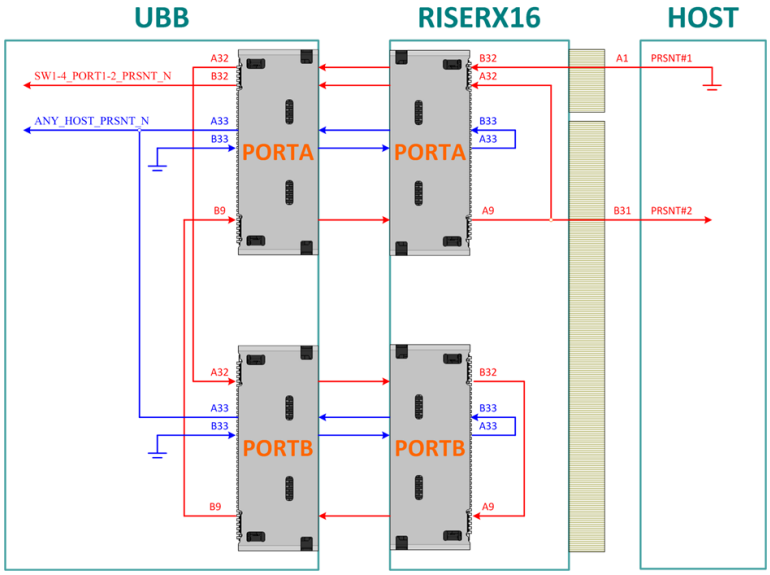


Figure 30 Host and Cable Presence Detect (Riser x16)

11.4 Connecting to External Chassis

The HGX-1 Expansion Chassis provides eight x16 PCIe Ports (bifurcatable to sixteen x8 or thirty-two x4 Links). These PCIe Ports may interconnect to external PCIe Hosts, external HGX-1, PCIe Target Devices, or to internal Switches for better peer-to-peer PCIe traffic for Devices within the HGX-1 Chassis.

As Figure 31 shows, Project Olympus Blade Chassis couples to Project Olympus HGX-1 via Cables.

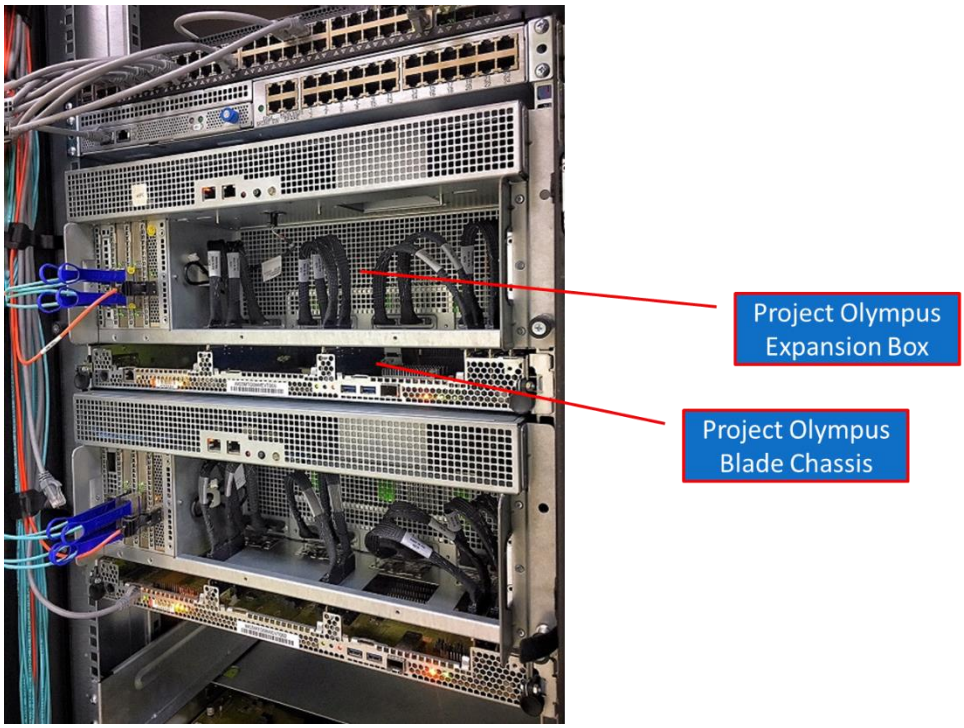


Figure 31 External PCIe Links

11.5 Power Sequencing and PCIe Port Enable Control

These Riser Boards emulate each of the Expansion Box PCIe Ports as a PCIe Card from the viewpoint of the Host CPU of Project Olympus Motherboard.

From the point of view of the Expansion Box, each x8 PCIe Cable may be plugged into a different Host. The Expansion Box BMC and CPLD logic monitor “Host Presence” and “Host 12V Presence” for power sequencing and to determine if a Host Port is to be enabled.

The pinout of the PCIe Cable/Connectors and the circuitry at the end of the of the Cable at the Head Node interface (Riser Boards) provide Host-presence, Host-power-presence, cable-presence, and correct-seating indications.

11.6 PCIe, NVLINK, and PCIe Fabric Manager Network Topologies

Figure 32 shows PCIe, NVLINK, and PCIe Fabric Manager Network Topologies. NVLINK is nVidia's GPGPU-to-GPGPU interconnect for Pascal SXM2 family of GPGPUs.

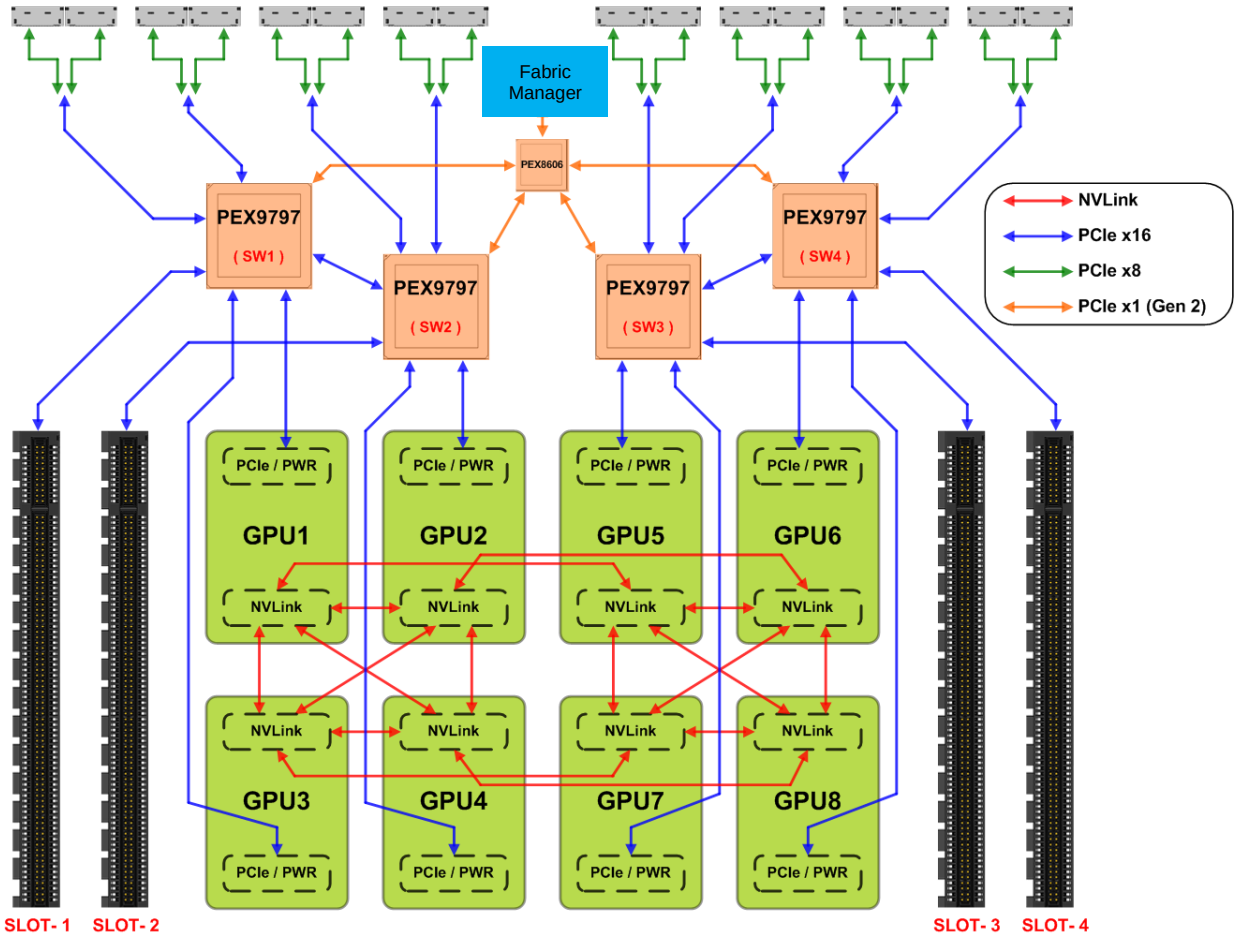


Figure 32 PCIe, NVLINK, and PCIe Fabric Management Topologies (not all NVLINK connections are shown)

Figure 33 shows PCIe Clock distribution map.

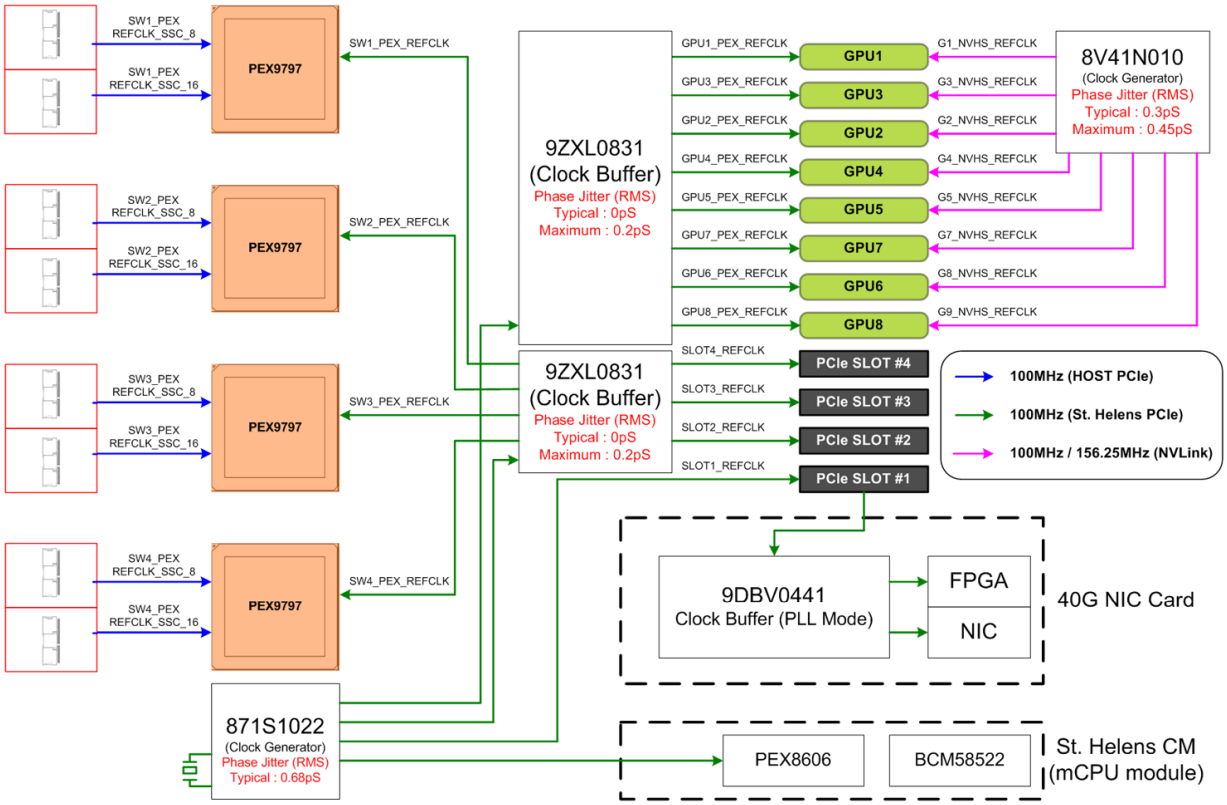


Figure 33 PCIe Clock Distribution Map

Figure 34 shows the block diagram of the Chassis Management (CM). It includes a 32-Lane, 8-ported PCIe Switch as part of PCIe Switch Management Network.

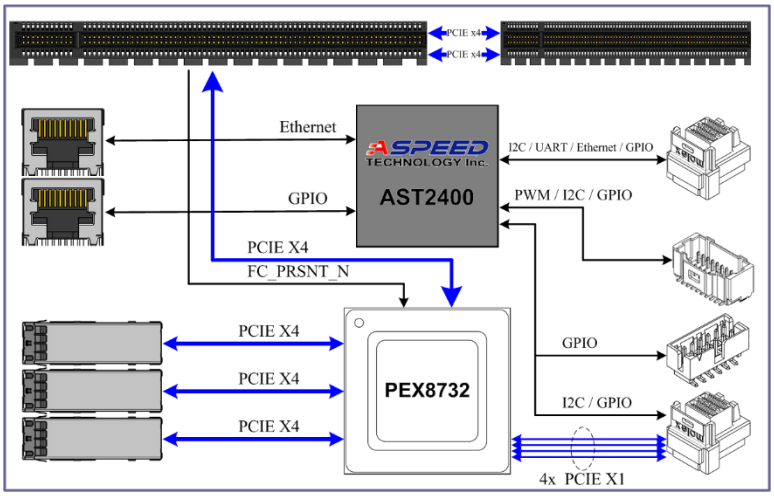


Figure 34 Chassis Management (CM) Module Block Diagram

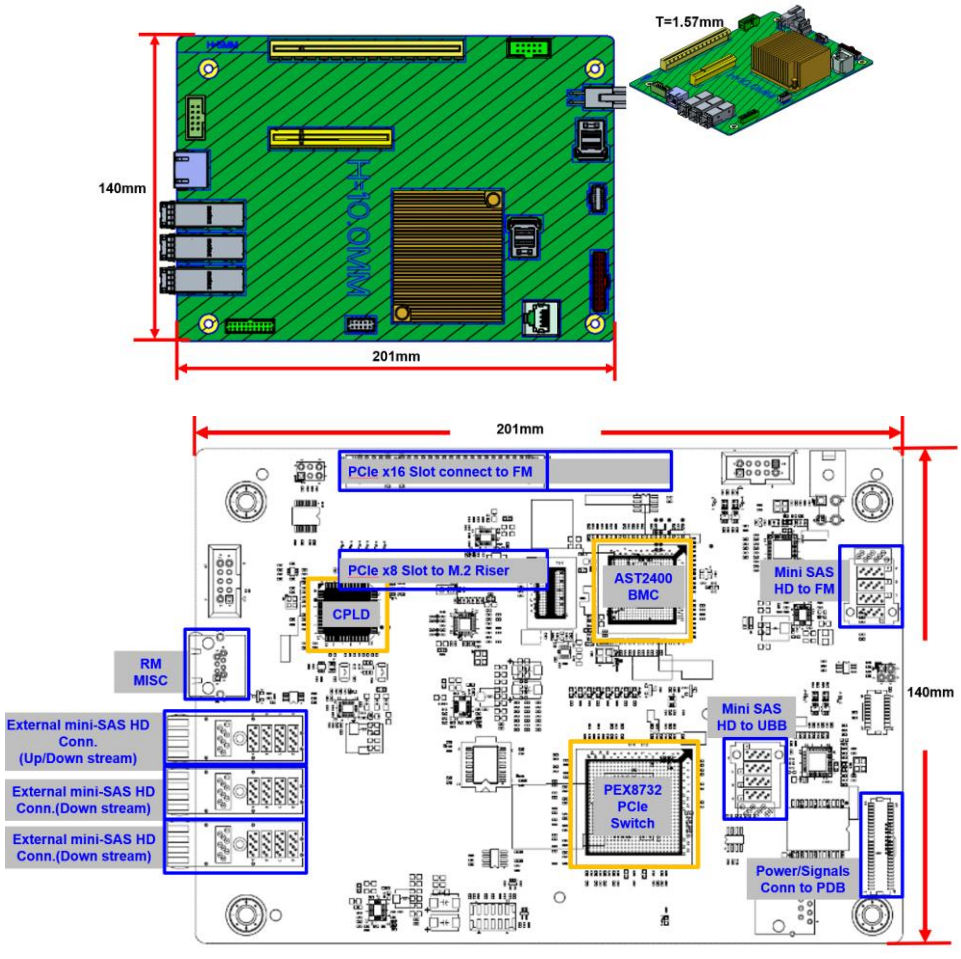


Figure 35 Chassis Management (CM) Board

11.7 Mezzanine Variants

In addition to MEZZ1x16 shown in Figure 8 which provides for eight double-width, x16 PCIe Cards in HGX-1, other mezzanine variants provide flexible choice of PCIe Slot configurations. For example, Figure 36 shows MEZZ2x16 which provides two x8 PCIe Slots using two x16 PCIe connectors. With eight MEZZ2x16s, the HGX-1 provides for four front-accessible x16 Slots and sixteen other x8 PCIe Slots.

While, MEZZ2x16 provides a standard x8 Port to each PCIe Slot, it also interconnects the upper x8 Port of the x16 PCIe connector as a private, Card-to-Card interconnect.

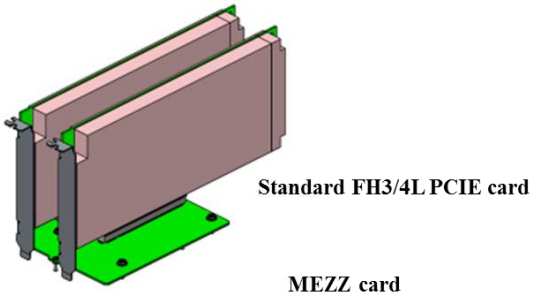


Figure 36 MEZZ2x16

As another example, Figure 37 shows the MEZZ4x16 which provides four interconnecting PCIe Slots. Each x16 PCIe connector provides one x4 Port to the Host via PCIe Switch and three x4 Ports to the other peers on the Mezzanine.

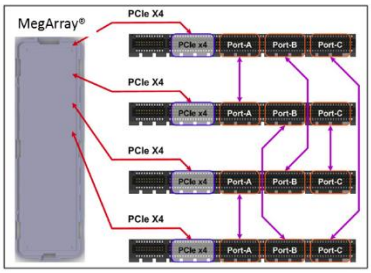


Figure 37 MEZZ4x16

With this MEZZ4x16 Mezzanine board, the Expansion Box supports four front-accessible x16 PCIe Slots for FH¾L Add-in Cards and thirty-two x4 PCIe Slots for FHHL add-in Cards.

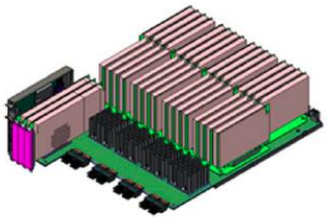


Figure 38 HGX-1 with 4x16 + 32x4 PCIe Slots

12 Management Subsystems

The Expansion Box management consist of a Chassis Management (CM) for BMC and a PCIe Fabric Management (FM) subsystem.

12.1 Fabric Manager (FM)

PCIe Fabric Management Network interconnects up to four HGX-1.

As Figure 39 depicts, the PCIe Fabric Management Network consists of one PCIe Switch in each HGX-1 which is connected to four 96-lane Switches within each HGX-1, 3 other HGX-1, and optionally to a Fabric Manager SoC.

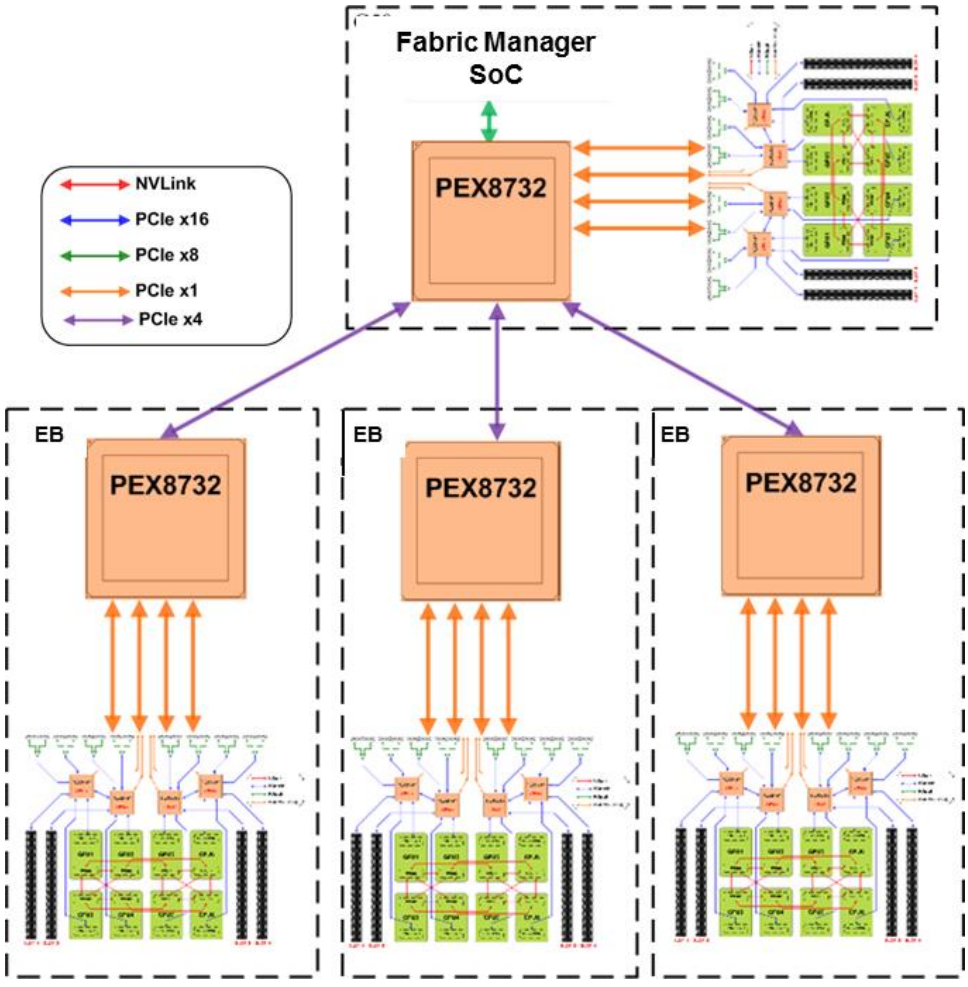


Figure 39 Fabric Management Network

Figure 40 shows the block diagram of the PCIe Fabric Manager Controller (FM).

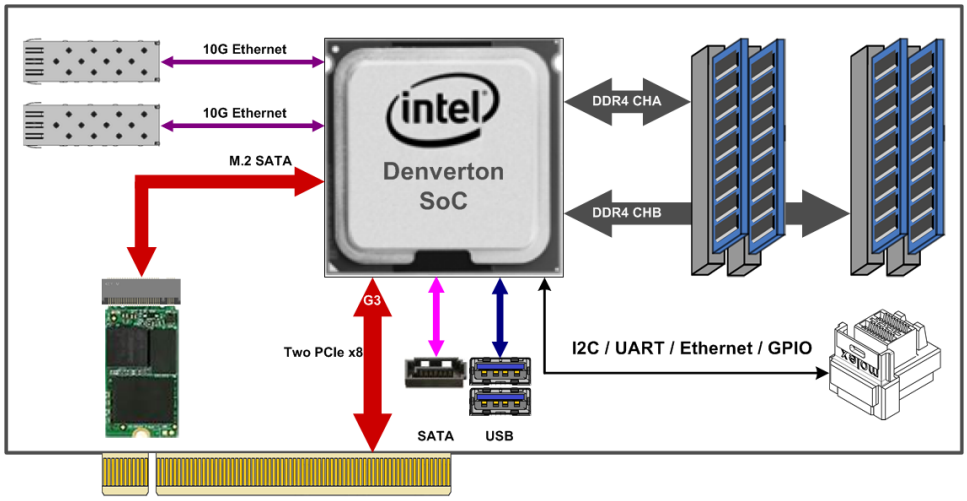


Figure 40 Fabric Manager Controller (FM) block Diagram

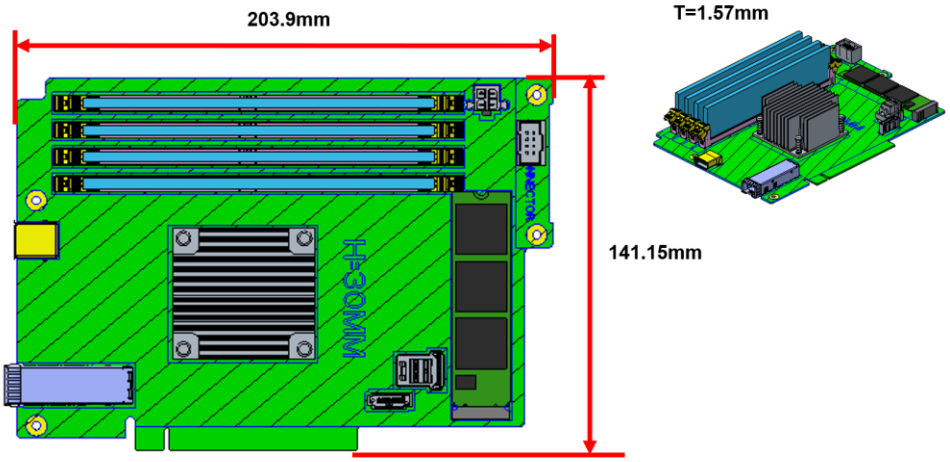


Figure 41 Fabric Manager Board

The FM is an SoC on a PCIe Card (Dimension: 111.15mm x 201.65mm).

As Figure 43 shows, the FM SoC couples to a dedicated, 8-ported PCIe Switch which resides on the Chassis Manager Board (CM). (x4 to FM, 4 x1 to the internal PCIe Switches, and 3 x4 ports to three External HGX-1s)

Figure 39 shows that this Switch forms a dedicated PCIe network so that one FM may manage the PCIe Fabric of up to four HGX-1s.

Figure 42 shows FM (vertical board) and CM (horizontal board) in the rear-view of the HGX-1. The FM is a double-width PCIe add-in board. The CM board may support several PCIe Slots in the rear of the Expansion Box.

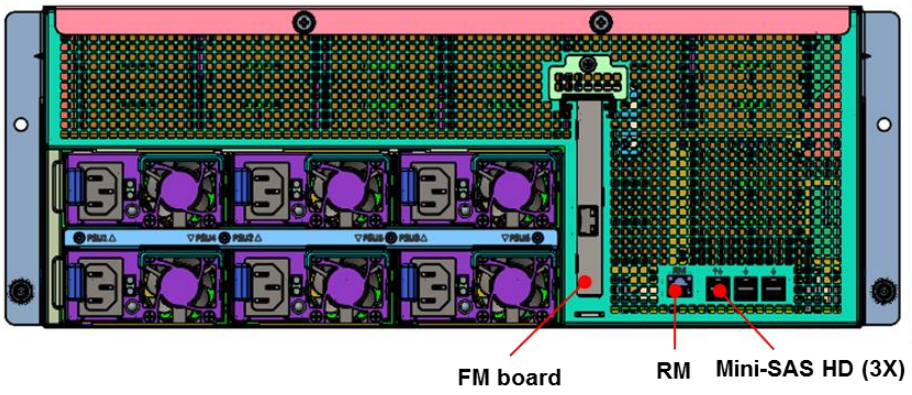


Figure 42 FM and CM at the rear of the HGX-1

12.2 In-Chassis Head Node

This HGX-1 typically couples to an external Head Node such as Project Olympus Server. However, since the SoC PCIe Card shown in Figure 40 may also serve as a general-purpose Server, it may plug into a PCIe Slot and perform as the Head Node for the HGX-1. This PCIe SoC Card of Figure 40 has the following components:

- Double-width PCIe Card
- 4 standard-height DIMMs
- Cable-connected to an external BMC (I^2C, UART, GPIO)
- x16 PCIe Card Edge (bifurcates to two x8s or four x4 Ports)
- SATA M.2 SSD as boot device
- 10GE Port
- SATA connectors
- USB connectors

Each double-width PCIe Slot of the HGX-1 may support one such SoC Card.

12.3 BMC (Baseboard Management Controller)

The Baseboard Management Controller circuitry (BMC) for the HGX-1 uses the ASPEED AST2500 or equivalent. This section describes the requirements for management of the HGX-1.

Primary features include.

- BMC - ASPEED AST2500 (or equivalent)
- BMC dedicated GbE LAN for communication with Rack Manager
- Connection to the Fabric Manager SoC to support in-band management
- Out of band environmental controls for power and thermal management
- FRUID EEPROM for storage of manufacturing data and events (I^2C)
- Thermal sensors for inlet and exhaust temperature monitoring (I^2C)
- Power monitoring through the 12V Hot Swap Controller circuitry (I^2C)
- Service LEDs

Figure 43 shows the Chassis Management (CM) block diagram.

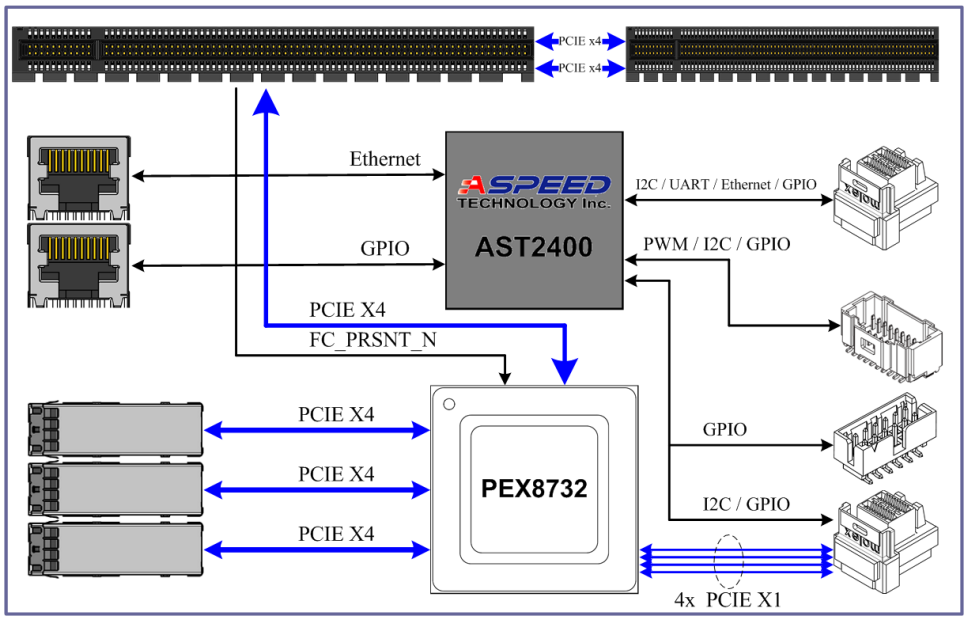


Figure 43. Chassis Manager Block Diagram

The design for the BMC is based on the ASPEED 2500 family. Primary features include:

- Embedded ARM Processor
- Minimum 256MiB of DDR3/4 memory

- Redundant NOR/NAND/SPI flash memory
- I²C /SMBus
- UART
- LPC Bus Interface
- GPIO pins

Figure 44 shows BMC control signals and the I²C topology diagram.

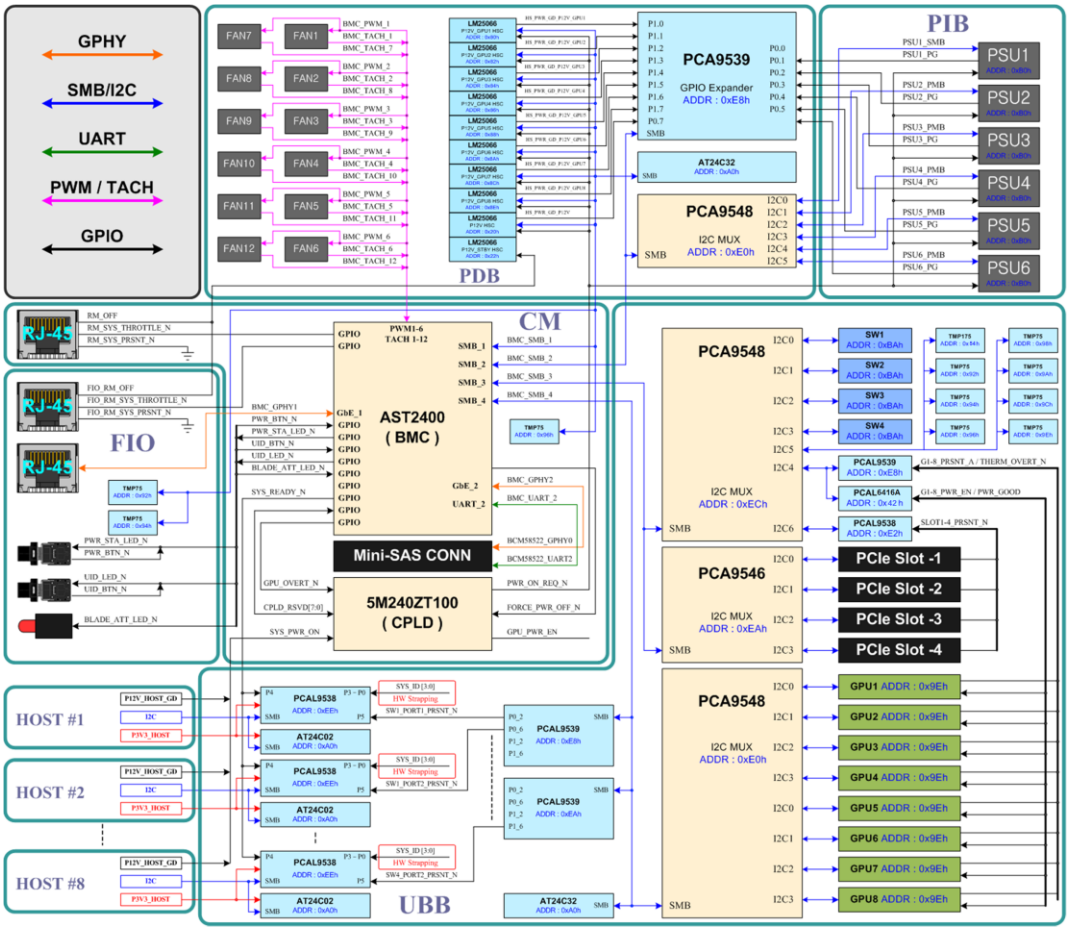


Figure 44 BMC Control Signals and I²C Topology Diagram

12.3.1 DRAM

The BMC supports a minimum of 256MiB of DDR3/4 memory.

12.3.2 BMC Boot Flash

The BMC boots from a flash memory device located on the SPI bus. The device size shall be 256Mb (32MiB) minimum and shall also be used to store FPGA and CPLD recovery images. Recommend using Windbond W25Q256 or equivalent. A secondary device shall be supported to provide BMC recovery. The design shall support a socket for the primary device during preproduction. It is not required to support a socket for the secondary device.

12.3.3 BIOS Flash

The BIOS utilizes one 256Mb (32MiB) Flash BIOS device. The device shall be Windbond MPN W25Q256 or equivalent. The BIOS shall be recoverable from the BMC in the event the chipset is inaccessible.

12.3.4 1GbE PHY

The HGX-1 is managed through a 1GbE PHY connected to the BMC. An RJ45 connector located at the front of the server provides 1GbE connectivity to an external management switch. The PHY shall be Broadcom BCM54612E or equivalent.

12.4 Thermal Monitoring

The BMC shall support a mechanism to query thermal information from various HGX-1 components to provide optimized fan speed control.

12.5 I^2C

To optimize telemetry gathering for power and thermal management as well as general management of the HGX-1 components. The following functions shall support I^2C access from the BMC.

- All PCIe slots and devices
- Local hot swap controllers
- Power Supplies
- Key voltage regulators
- Temperature sensors
- FRUID PROM

Electrically isolate components that are powered from separate power domains even if they are located on the same I^2C bus.

12.6 UART

The HGX-1 provides a UART header connected to the BMC for debug.

12.7 JTAG Master

The HGX-1 supports JTAG programming of programmable logic devices using the BMC's JTAG master controller. The JTAG programming shall be supported to any required CPLDs and every PCIe Slot/Device. The HGX-1 shall contain mux circuitry controlled by the BMC to switch between the two programming

paths. While, the BMC provides 6 bits as Mux Select to control up to 64 devices, Figure 45 shows a simple block diagram as an example of controlling on CPLD and one PCIe Slot.

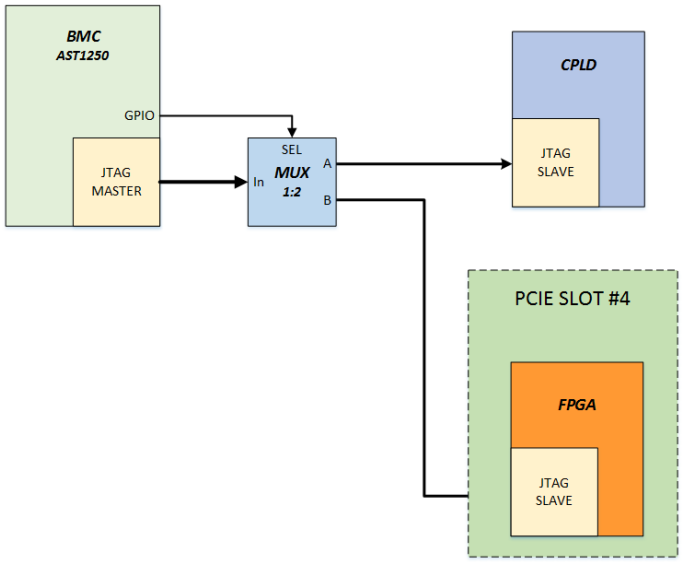


Figure 45. JTAG Programming

When the Debug Header is not populated, the BMC is the primary JTAG Root. In addition to a JTAG Root Port (TMS#, TCLK#, TRST#, TDI, TDO), the BMC provides 6 bits of Mux Control to individually select up to 64 leaves of the JTAG Tree.

A CPLD implements isolated JTAG signals to be routed to each PCIe Slot and Device within HGX-1.

Once populated, the Debug Header selects a Mux (within the CPLD) and takes over the control of the JTAG Tree. This Header provides debug JTAG Root, Root Select Override based on the Header-populated indication, and 6 bits for Device Select.

12.8 Voltage Regulators

Voltage regulators that support I^2C or PMBus should be available to the BMC.

Clock circuitry that supports I^2C should be available to the BMC.

12.9 FRUID PROM

The HGX-1 shall include a 64Kb serial EEPROM MPN AT24C64 or equivalent for storing manufacturing data. This I^2C EEPROM is accessible via the local BMC.

UBB provides eight x16 external PCIe Links. Each external x16 PCIe Link may connect to a different Host. Along with each one of the eight x16 PCIe Links, there are a few sideband signals. Each Host may see one FRUID via the I^2C bus carried over the PCIe Cable.

12.10 Temperature Sensors

The HGX-1 shall include I^2C support for a minimum of two temperature sensors, MPN TMP411 or equivalent, for monitoring the inlet and outlet air temperatures of the Upper Baseboard. The UBB shall include provisions to support temperature monitoring of all PCIe slots and devices. For accurate temperature reading, care shall be taken to not place these temperature sensors close to excessive heat sources on the board.

12.11 Hot Swap Controllers

To enable the e-fuses (via Hot-swap Controller) to detect over-current, the UBB has several power islands. Figure 21 shows UBB Power Islands. Separate e-fuses (HSC) monitor and control each power island.

The HGX-1 includes I^2C support for hot swap controllers (HSC) for power monitoring of each Power Zone (power island).

12.12 LEDs

The following sections describe the light-emitting diodes (LEDs) used as indicators on the motherboard. Table 4 lists the minimum LEDs required and provides a brief description. Greater detail for some LEDs is included in subsequent sections below. The visible diameter, color (λ) and brightness requirements of the LEDs are TBD. All LEDs shall be visible at the front of the motherboard (cold aisle).

Table 4. LEDs

LED Name	Color	Description
UID LED	Blue	Unit Identification LED
Attention LED	Red	Indicates that Server requires servicing
Power Status LED	Amber/Green	Indicates Power Status of the Server
Catastrophic Error	Red	Indicates that a CPU catastrophic error has occurred
BMC Heartbeat	Green	Blinks to indicate BMC is alive
GbE Port 0 Activity	Green	Indicates activity on 10GbE Port 0 (not supported for production)
GbE Port 0 Speed	Green/Orange	Green=high speed, Orange=Low speed (not supported for production)

12.12.1 UID LED

The HGX-1 shall support a blue UID (unit ID) LED used to help visually locate a specific server within a datacenter.

12.12.2 Power Status LED

As the HGX-1 is powered up, the Power Status LEDs shall turn amber if 12V is present at the output of any Hot Swap Controller. This indicates that the 12V power is connected and present at the HGX-1 and that the Hot-swap Controllers (HSC) are enabled, but it does not indicate that all PSUs are functioning!

When the server management software turns on the system power (BladeEn# aka PWR_En#), the Power Status LED turns green. Note that the power status LED may be driven by an analog resistor network tied directly to a power rail. It is not an indication of the health of the server. Table 5 describes the operation of the Power Status LED.

Table 5. Power Status LED Description

LED status	Condition
Off	12V power is absent or PWR_EN# is de-asserted
Solid Amber ON	PWR_EN# is asserted and 12V power output of the Hot Swap Controller is present.
Solid Green ON	Indicates that the management (BMC or IE) is booted and HGX-1 power is enabled.

12.12.3 Attention LED

The Attention LED directs the service technicians to the HGX-1 that requires service. When possible, HGX-1 diagnostics are used to direct repairs. Alternatively, the Microsoft scale-out management software can be used. In both cases, event logs of the repair work are available.

The LED is driven by a single BMC GPIO. Table 6 describes the operation of the Attention LED.

Table 6. Attention LED Description

LED status	Condition
Off	No attention indicated
Solid RED	System needs attention

12.13 Fan Control

The HGX-1 shall support control of twelve 60mm fans via 12V power, a single PWM, and twelve TACH signals for controlling the 12 fans in a single zone.

12.14 Power Management

As Figure 12 and Figure 14 showed, the HGX-1 Expansion Chassis receives six AC feeds (C14) via six Power Supply Units (PSUs). HGX-1 shall provide a power distribution board (PDB) and a power supply interface

board (PIB) to provide 12V power to the Upper Baseboard. The HGX-1 shall also provide a rear management connector (RJ45) for enabling external control of server power. Refer to Project Olympus Standalone Rack Manager specification for the pinout definition of this RJ45.

Figure 46 shows an example block diagram of the interface. In this example, the Blade Motherboard connects to Project Olympus PSU as described in Project Olympus Server Specification. Similarly, the HGX-1 receives Blade_En# and Blade_Throttle signals from the external Project Olympus Standalone Rack Manager (via a sideband RJ45 connector) and provides HGX-1 Presence# to the Standalone Rack Manager.

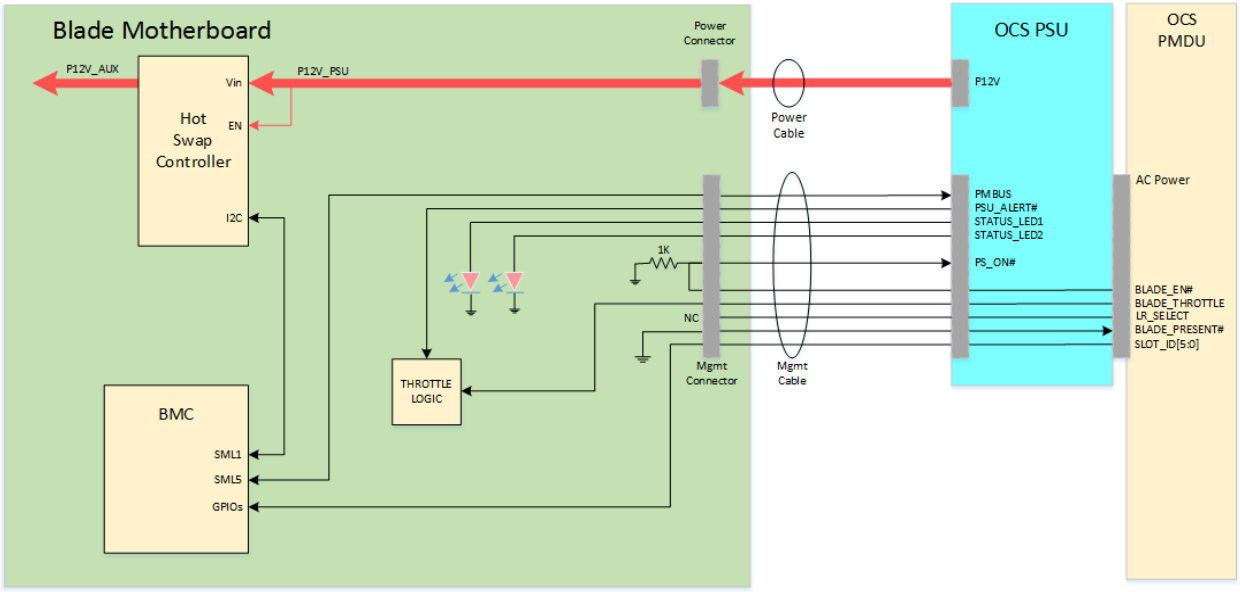


Figure 46. Power Management Block Diagram

Figure 47 and Figure 48 show Power Interface Board (PIB) and Power Distribution Board (PDB) respectively. PIB receives six power supplies (PSU), and PDB provide connectors for cabling power to UBB and to PCIe Add-in Cards.

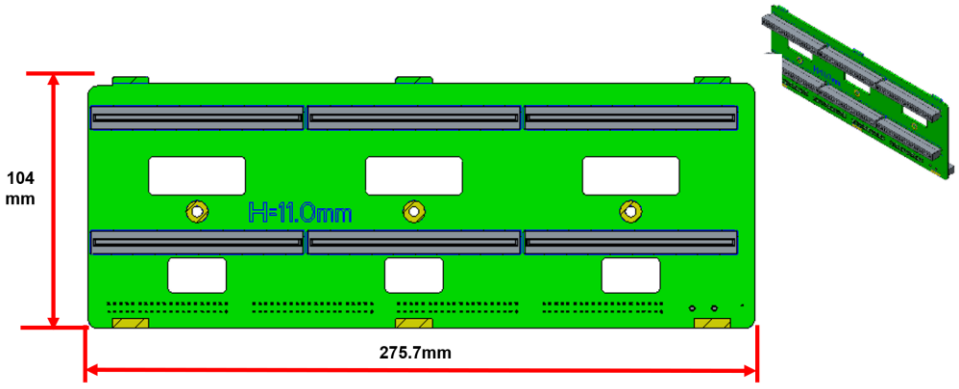


Figure 47 Power Interface Board (PIB)

Figure 48 Power Distribution Board (PDB)

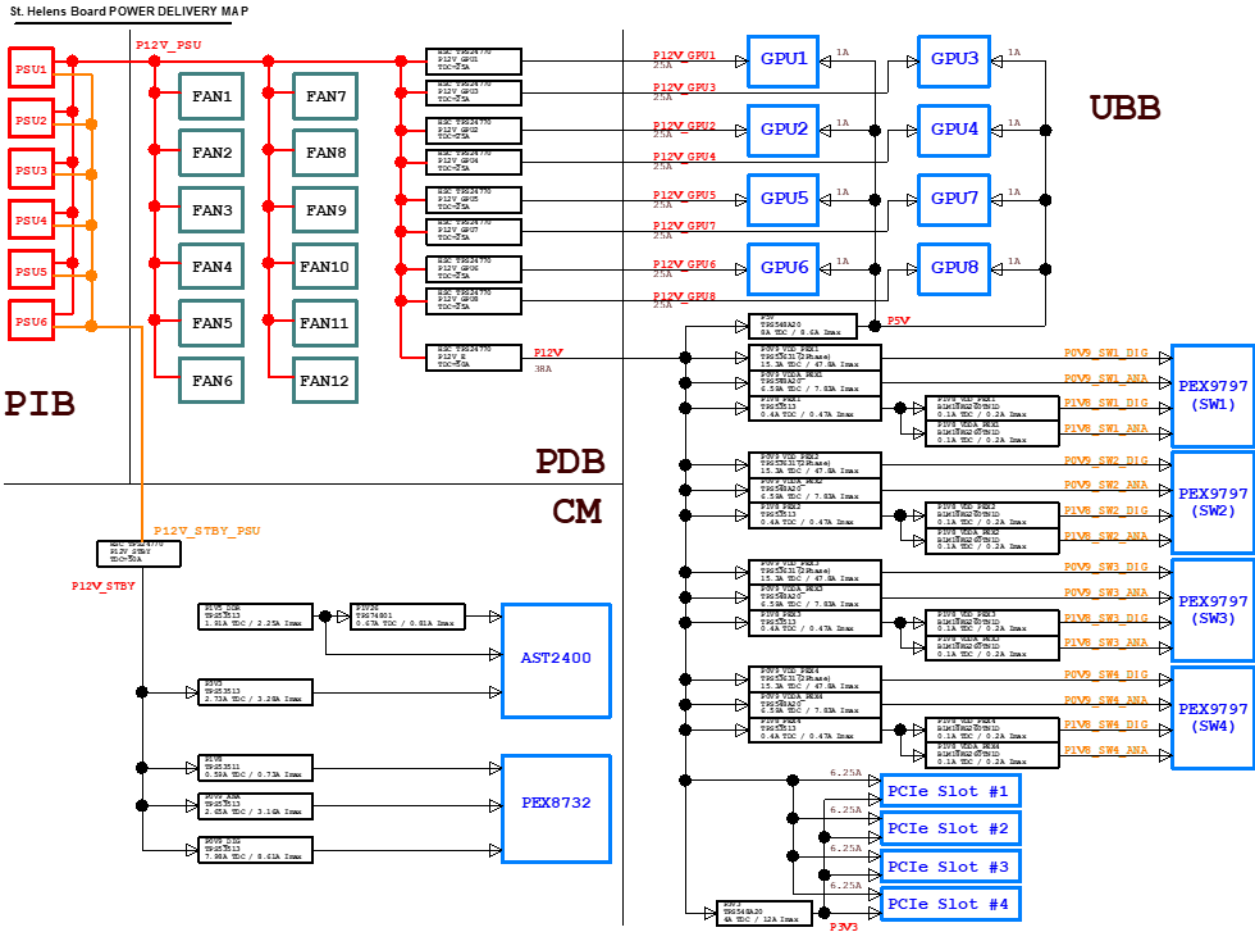


Figure 49 Power Distribution Diagram

12.14.1 Rack Management

The Standalone Rack Manager communicates with the HGX-1 CM using one RJ45 for 1GE and one RJ45 for sideband signals. The following describes the management interfaces.

- PWR_EN# (aka BladeEn#) - Active low signal used to enable/disable power to the PSUs. A 1K ohm pulldown resistor is used on the Chassis Manager board to ensure a default low state (active) if the Rack Manager is not present. This signal connects to the PS_ON# signal of the PSUs. When in high state (inactive), this signal disables output power from the PSUs.
- SERVER_PRESENT# - Active low signal used to communicate physical presence of the HGX-1 to the Rack Manager. This signal should be tied to GND on the Chassis Manager Board (CM).
- SERVER_THROTTLE - Active high signal used to put the HGX-1 into a low power (power cap) state. If implemented, this signal should default low (inactive) if the Rack Manager is not present. This signal is fanned out from the Rack Manager to multiple servers and therefore the circuit design must support electrical isolation of this signal from the motherboard power planes.

12.14.2 PSU Management

The HGX-1 shall support management of the PSUs. Below is a description of associated signals.

- PS_ON# - Active low signal used to enable/disable power to the PSU. This signal is driven by the PWR_EN# signal from the Rack Manager. A 1K ohm pulldown resistor is used on the motherboard to ensure a default low state if the RM is not present.
- PSU_ALERT# - Active low signal used to alert the HGX-1 that a fault has occurred in the associated PSU. This signal is connected to the BMC for monitoring of PSU status.
- PMBUS – I^2C interface to the PSUs. The BMC uses this signal to read the status of the PSU.
- STATUS_LED – Controls LED to provide visual indication of a PSU fault.

12.14.3 Hot Swap Controller

The motherboard shall support five hot swap controllers (HSC) for in-rush current protection. The HSC shall include support for the PMBUS interface. Recommend using the ADM1278 or equivalent.

12.14.4 Power Capping (optional)

If implemented, the HGX-1 may enable power capping from different trigger sources. Each of these triggers can be disabled by the BMC. BMC should be able to trigger these events for debugging.

- SERVER_THROTTLE - Throttle signal driven by the Rack Manager.
- PSU_ALERT# - Alert signal driven by the PSU. The PSU will be programmed to assert ALERT# in the event the PSU transitions its power source from AC to battery backup.
- FM_THROTTLE# - Test signal that allows BMC to assert power cap

12.14.5 Overcurrent Protection

Hot swap controllers (HSC) are responsible for detecting current levels that indicate catastrophic failure. In this event, the HSC should disable 12V typically by disabling the HSC's input FETS.

12.14.6 Service Requirements

12.14.6.1 LED Visibility

HGX-1 LEDs determined to be important for communicating status to service personnel shall be made visible at the front (cold aisle) of the server. This should shall include the following LEDs at a minimum.

- UID LED
- Power Status LED
- Attention LED

13 Upper Baseboard (UBB) Interfaces

This section describes the connector interfaces to the UBB.

13.1 Dual-device Card (e.g., FPGA) Support

The HGX-1 supports a dual-device, Full-Height, Half-Length (FHHL) x16 PCIe Card form factor. The card installs in a x16 Riser board which installs in a standard PCIe x16 slot. The Riser board includes a x8 OCuLink connector for cabling a x8 PCIe Link from the Slot connector to the FPGA card.

Figure 50 shows that Project Olympus HGX-1 supports a dual-device PCIe Card via a x16 Riser and a x8 OCuLink Cable.

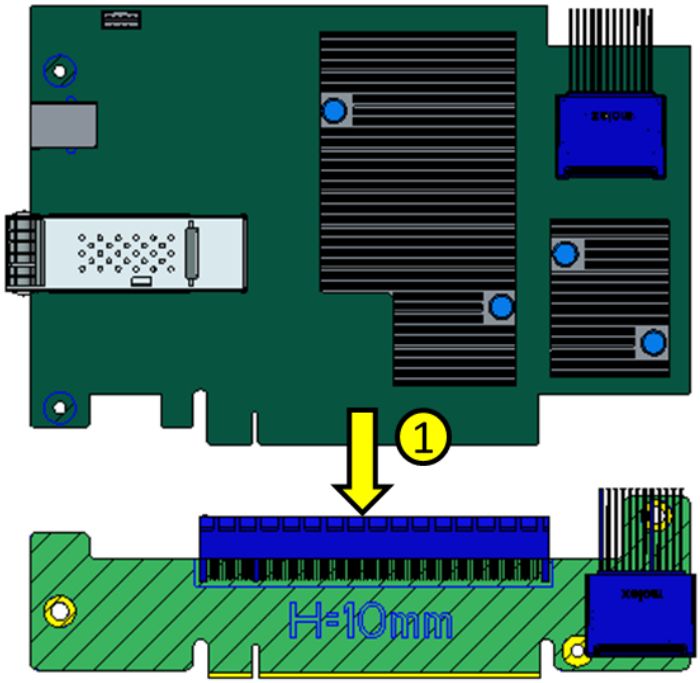


Figure 50 Supporting a dual-device PCIe Card via a x16 Riser and a x8 OCuLink Cable

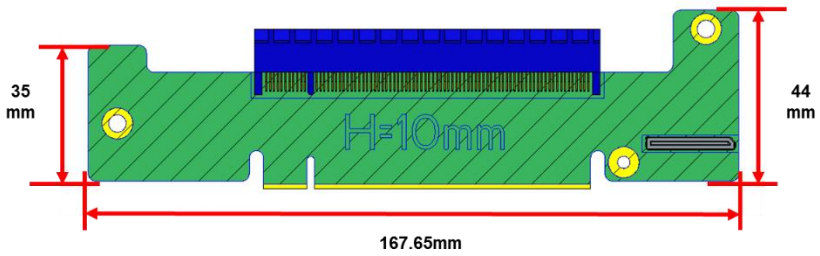


Figure 51 Dual-device PCIe Riser

13.1.1 OCuLink x8 Connector

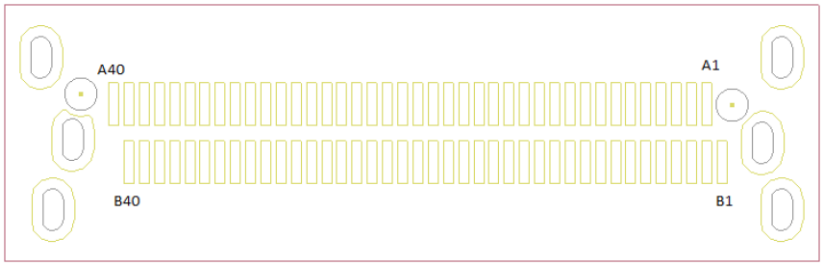
The Riser board shall include a connector for cabling PCIe x8 from to the FPGA Card. The connector for supporting the cable shall be an 80-pin vertical Molex Nanopitch™ series connector or equivalent (Molex part number 173162-0334). Table 7 describes the connector pinout. Figure 52 shows a top view of the physical pin numbering.

Table 7. OCuLink x8 Connector Pinout

Row A			Row B		
Pin	Signal Name	Description	Pin	Signal Name	Description
A1	GND	Ground	B1	GND	Ground
A2	PERp0	PCIe Receive Data to CPU	B2	PETp0	PCIe Transmit Data from CPU

A3	PERn0	PCIe Receive Data to CPU	B3	PETn0	PCIe Transmit Data from CPU
A4	GND	Ground	B4	GND	Ground
A5	PERp1	PCIe Receive Data to CPU	B5	PETp1	PCIe Transmit Data from CPU
A6	PERn1	PCIe Receive Data to CPU	B6	PETn1	PCIe Transmit Data from CPU
A7	GND	Ground	B7	GND	Ground
A8	NC	No Connect	B8	NC	No Connect
A9	NC	No Connect	B9	NC	No Connect
A10	GND	Ground	B10	GND	Ground
A11	NC	No Connect	B11	NC	No Connect
A12	NC	No Connect	B12	CPRSNT#	Cable Present
A13	GND	Ground	B13	GND	Ground
A14	PERp2	PCIe Receive Data to CPU	B14	PETp2	PCIe Transmit Data from CPU
A15	PERn2	PCIe Receive Data to CPU	B15	PETn2	PCIe Transmit Data from CPU
A16	GND	Ground	B16	GND	Ground
A17	PERp3	PCIe Receive Data to CPU	B17	PETp3	PCIe Transmit Data from CPU
A18	PERn3	PCIe Receive Data to CPU	B18	PETn3	PCIe Transmit Data from CPU
A19	GND	Ground	B19	GND	Ground
A20	RSVD	Reserved	B20	RSVD	Reserved
A21	RSVD	Reserved	B21	RSVD	Reserved
A22	GND	Ground	B22	GND	Ground
A23	PERp4	PCIe Receive Data to CPU	B23	PETp4	PCIe Transmit Data from CPU
A24	PERn4	PCIe Receive Data to CPU	B24	PETn4	PCIe Transmit Data from CPU
A25	GND	Ground	B25	GND	Ground
A26	PERp5	PCIe Receive Data to CPU	B26	PETp5	PCIe Transmit Data from CPU
A27	PERn5	PCIe Receive Data to CPU	B27	PETn5	PCIe Transmit Data from CPU
A28	GND	Ground	B28	GND	Ground
A29	NC	No Connect	B29	NC	No Connect
A30	NC	No Connect	B30	NC	No Connect
A31	GND	Ground	B31	GND	Ground
A32	NC	No Connect	B32	NC	No Connect
A33	NC	No Connect	B33	NC	No Connect
A34	GND	Ground	B34	GND	Ground
A35	PERp6	PCIe Receive Data to CPU	B35	PETp6	PCIe Transmit Data from CPU
A36	PERn6	PCIe Receive Data to CPU	B36	PETn6	PCIe Transmit Data from CPU
A37	GND	Ground	B37	GND	Ground
A38	PERp7	PCIe Receive Data to CPU	B38	PETp7	PCIe Transmit Data from CPU
A39	PERn7	PCIe Receive Data to CPU	B39	PETn7	PCIe Transmit Data from CPU
A40	GND	Ground	B40	GND	Ground

Figure 52. OCuLink x8 Pin Numbering



13.2 PCIe x16 Connectors

The PCIe x16 connector interface is designed to support a standard PCIe x16 full-height card. The pinout for supporting PCIe x16 described in Table 8. This interface also supports the PWRBRK# power reduction feature. Note that this signal is declared on pins B12 and B30. This enables support for the feature on existing platforms (B12) and meets the latest PCI SIG definition (B30). For further information, refer to the PCI Express® Card Electromechanical Specification.

Table 8. PCIe x16 connector pinout

Pin	Side B Connector		Side A Connector	
#	Name	Description	Name	Description
1	+12v	+12-volt power	PRSNT#1	Hot plug presence detect
2	+12v	+12-volt power	+12v	+12-volt power
3	+12v	+12-volt power	+12v	+12-volt power
4	GND	Ground	GND	Ground
5	SMCLK	SMBus clock	JTAG2	TCK
6	SMDAT	SMBus data	JTAG3	TDI
7	GND	Ground	JTAG4	TDO
8	+3.3v	+3.3-volt power	JTAG5	TMS
9	JTAG1	+TRST#	+3.3v	+3.3-volt power
10	3.3Vaux	3.3v volt power	+3.3v	+3.3-volt power
11	WAKE#	Link Reactivation	PWRGD	Power Good
Mechanical Key				
12	PWRBRK#	Power Reduction	GND	Ground
13	GND	Ground	REFCLK+	Reference Clock
14	PETP(0)	Transmitter Lane 0,	REFCLK-	Differential pair
15	PETN(0)	Differential pair	GND	Ground
16	GND	Ground	PERP(0)	Receiver Lane 0,
17	PRSNT#2	Presence detect	PERN(0)	Differential pair
18	GND	Ground	GND	Ground
19	PETP(1)	Transmitter Lane 1,	RSVD	Reserved

20	PETN(1)	Differential pair	GND	Ground
21	GND	Ground	PERP(1)	Receiver Lane 1,
22	GND	Ground	PERN(1)	Differential pair
23	PETP(2)	Transmitter Lane 2,	GND	Ground
24	PETN(2)	Differential pair	GND	Ground
25	GND	Ground	PERP(2)	Receiver Lane 2,
26	GND	Ground	PERN(2)	Differential pair
27	PETP(3)	Transmitter Lane 3,	GND	Ground
28	PETN(3)	Differential pair	GND	Ground
29	GND	Ground	PERP(3)	Receiver Lane 3,
30	PWRBRK#	Power Reduction	PERN(3)	Differential pair
31	PRSNT#2	Hot-plug detect	GND	Ground
32	GND	Ground	RSVD	Reserved
33	PETP(4)	Transmitter Lane 4,	RSVD	Reserved
34	PETN(4)	Differential pair	GND	Ground
35	GND	Ground	PERP(4)	Receiver Lane 4,
36	GND	Ground	PERN(4)	Differential pair
37	PETP(5)	Transmitter Lane 5,	GND	Ground
38	PETN(5)	Differential pair	GND	Ground
39	GND	Ground	PERP(5)	Receiver Lane 5,
40	GND	Ground	PERN(5)	Differential pair
41	PETP(6)	Transmitter Lane 6,	GND	Ground
42	PETN(6)	Differential pair	GND	Ground
43	GND	Ground	PERP(6)	Receiver Lane 6,
44	GND	Ground	PERN(6)	Differential pair
45	PETP(7)	Transmitter Lane 7,	GND	Ground
46	PETN(7)	Differential pair	GND	Ground
47	GND	Ground	PERP(7)	Receiver Lane 7,
48	PRSNT#2	Hot-plug detect	PERN(7)	Differential pair
49	GND	Ground	GND	Ground
50	PETP(8)	Transmitter Lane 8,	RSVD	Reserved
51	PETN(8)	Differential pair	GND	Ground
52	GND	Ground	PERP(8)	Receiver Lane 8,
53	GND	Ground	PERN(8)	Differential pair
54	PETP(9)	Transmitter Lane 9,	GND	Ground
55	PETN(9)	Differential pair	GND	Ground

56	GND	Ground	PERP(9)	Receiver Lane 9,
57	GND	Ground	PERN(9)	Differential pair
58	PETP(10)	Transmitter Lane 10,	GND	Ground
59	PETN(10)	Differential pair	GND	Ground
60	GND	Ground	PERP(10)	Receiver Lane 10,
61	GND	Ground	PERN(10)	Differential pair
62	PETP(11)	Transmitter Lane 11,	GND	Ground
63	PETN(11)	Differential pair	GND	Ground
64	GND	Ground	PERP(11)	Receiver Lane 11,
65	GND	Ground	PERN(11)	Differential pair
66	PETP(12)	Transmitter Lane 12,	GND	Ground
67	PETN(12)	Differential pair	GND	Ground
68	GND	Ground	PERP(12)	Receiver Lane 12,
69	GND	Ground	PERN(12)	Differential pair
70	PETP(13)	Transmitter Lane 13,	GND	Ground
71	PETN(13)	Differential pair	GND	Ground
72	GND	Ground	PERP(13)	Receiver Lane 13,
73	GND	Ground	PERN(13)	Differential pair
74	PETP(14)	Transmitter Lane 14,	GND	Ground
75	PETN(14)	Differential pair	GND	Ground
76	GND	Ground	PERP(14)	Receiver Lane 14,
77	GND	Ground	PERN(14)	Differential pair
78	PETP(15)	Transmitter Lane 15,	GND	Ground
79	PETN(15)	Differential pair	GND	Ground
80	GND	Ground	PERP(15)	Receiver Lane 15,
81	PRSNT#2	Hot plug present detect	PERN(15)	Differential pair
82	RSVD#2	Hot Plug Detect	GND	Ground

13.3 Fan Control Connector

The HGX-1 shall support header connectors for enabling control of 12 fans.

Table 9. Fan Control Signals

Fan Management Signals				
Pin	**Signal**	**I/O**	**Voltage**	**Description**
	FAN[12:1]_PWM	O	5V	Fan PWM

	FAN[12:1]_TACH	I	5V	Fan Tachometer
several	P12V	O	12V	12V Fan Power
several	GND	O	0V	Ground

13.4 Connector Quality

The Project Olympus system is designed for use in datacenters with a wide range of humidity. The connectors for these deployments shall be capable of withstanding high humidity during shipping and installation. The baseline for plating DIMMs and PCIe connectors will be 30µ"-thick gold. DIMM connectors shall also include lubricant/sealant applied by the connector manufacturer that can remain intact after soldering and other manufacturing processes. The sealant is required to displace any voids in the connector gold plating.

14 Regulations and Certifications

Project Olympus and its associated HGX-1 and Modules are to be used in cloud-scale datacenters which meet various environmental, safety, reliability, and emissions regulations. The following outline the relevant specifications.

14.1 Environmental

The HGX-1 is to be deployed in an environmentally controlled location. The inlet to the server will be exposed to the environment described in Table 10. The server must have the capability to provide full functional operation under the conditions given.

Table 10. Environmental Requirement

Specification		Requirement
Inlet temperature	Operating	• 50°F to 95°F (10°C to 35°C) • Maximum rate of change: 18°F (10°C)/hour • Allowable derating guideline of 1.6°F/1000ft (0.9°C/304m) above 3000 ft.
	Non-operating	• -40°F to 140°F (-40°C to 60°C) • Rate of change less than 36°F (20°C)/hour

14.2 Safety

14.3 Reliability

14.4 FCC Certification