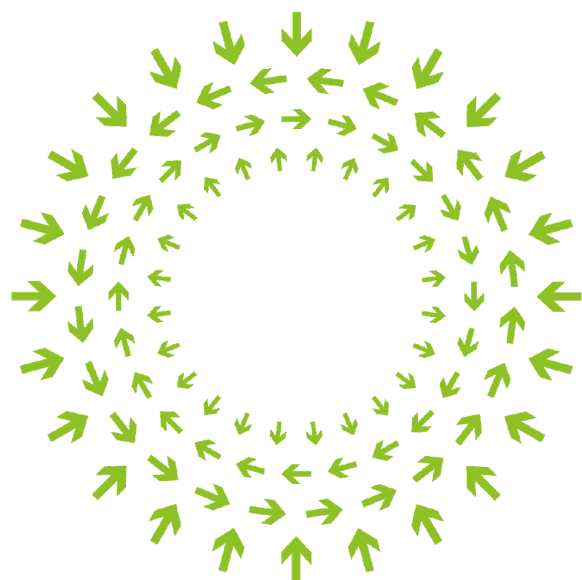




OPEN

Compute Project

Grand Teton Platform Specification

Version 1.0

Effective September 5, 2023

Authors:

Matt Bowman, Hardware Engineer, Meta
Jeremy Baumgartner, Hardware Engineer, Meta
Hao Shen, Hardware Engineer, Meta
Anthony Chan, Power Engineer, Meta
Devika Vishwanath, Power Engineer, Meta
Enrique Beuses, Mechanical Engineer, Meta
Jaret Wyatt, Mechanical Engineer, Meta
Cheng Chen, Thermal Engineer, Meta
Kalpak Dhake, Thermal Engineer, Meta

Tanmoy Roy, Signal Integrity Engineer, Meta
Ben Kim, Compliance Engineer, Meta
Jeremy Yang, Hardware System Engineer, Meta
Gada Badeer, Hardware System Engineer, Meta
Nitish Vanne, Hardware System Engineer, Meta
Anil Agrawal, Hardware System Engineer, Meta
Amithash Prasad, Software Engineer, Meta
Yuvraj Madhaya Weerasinghe, Product Quality Engineer, Meta
Jia Xin (Maggie) Han, Technical Program Manager, Meta

Table of Contents

1. License	5
2. Compliance with OCP Tenets	6
2.1 Openness	6
2.2 Efficiency	6
2.3 Impact	6
2.4 Scale	6
2.5 Sustainability	6
3. Change Log	7
4. Scope	7
5. Overview	8
5.1 Base Specification	9
6. Environmental and Regulatory Compliance	10
6.1 Environmental	10
6.2 Regulatory Compliance	10
7. Physical Specifications	11
7.1 Platform Block Diagram	11
7.2 CPU Tray	11
7.3 Switch Tray	12
7.4 Accelerator Tray	14
7.5 SCM	15
7.6 Horizontal Power Distribution Board (HPDB)	18
7.7 Vertical Power Distribution Board (VPDB)	21
7.8 Fan Boards	23
7.9 Fans	24
7.10 High Speed Cables	25
7.11 Power Cables	27
7.12 Mechanical	28
7.13 Rack Compatibility	32
8. Electrical Requirements	32
8.1 Input Voltage	33
8.2 Power	33
9. Thermal Design Requirements	33
9.1 Thermal Architecture	33
9.2 System Thermal Specification	33

9.3 Fan Configuration	34
9.4 Fan Control Target	34
9.5 Thermal Sensors	34
9.6 Fan Control Algorithm	34
9.7 Special Conditions	34
10. Interfaces	35
10.1 Signal List	35
10.2 Cable Backplane	35
11. Onboard Power System	36
12. System Firmware	36
13. Hardware Management	37
13.1 Out-Of-Band Management	37
13.2 BMC	37
13.3 Remote Upgradeability	37
13.4 Sensors	37
14. Security	41
Appendix A - Checklist for IC approval of this Specification (to be completed by contributor(s) of this Spec)	42

1. License

Contributions to this Specification are made under the terms and conditions set forth in Modified OWF-CLA-1.0.2 (As of June 1, 2023) (“Contribution License”) by:

Meta Platforms, Inc.

Usage of this Specification is governed by the terms and conditions set forth in **Modified OWFa1.0.2 Final Specification Agreement (FSA) (As of June 1, 2023) (“Specification License”)**.

You can review the applicable Specification License(s) referenced above by the contributors to this Specification on the OCP website at <http://www.opencompute.org/participate/legal-documents/>. For actual executed copies of either agreement, please contact OCP directly.

Notes:

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

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

2. Compliance with OCP Tenets

2.1 Openness

The Grand Teton platform design exemplifies openness by defining a platform that can support multiple use cases and workloads. This includes modular swappable trays for all system components and open-source BMC/BIC firmware, which allows users to optimize the system for their own needs.

2.2 Efficiency

Compared to past generations, the Grand Teton platform offers efficiency gains across the power and thermal systems by upgrading to the 51V ORv3 rack architecture. There are also efficiency gains through the system architecture, which utilizes trays that can be swapped to improve performance or optimize the hardware for evolving workloads.

2.3 Impact

The Grand Teton platform represents a number of firsts in large-scale AI infrastructure, including next-gen CPUs, PCIe Gen5, DDR5, and NVIDIA HGX. These technologies, along with the flexibility of the system architecture, support high-performance workloads that can scale out and utilize other OCP solutions, such as ORv3, DC-SCM, OCP NIC, OCP E1.S, Wedge400, and Minipack2.

2.4 Scale

The Grand Teton platform utilizes the latest features to support large-scale deployment, including tool-less FRUs, extensive remote management, error reporting, future upgradeability, and on-premises serviceability enhancements. Offering flexible scale capabilities, Grand Teton can be used as a single system up through hyper-scale deployment. The system also uses OpenBMC and OpenBIC open-source software.

2.5 Sustainability

The Grand Teton platform follows environmental compliance standards and best practices. The modular design allows for more targeted repairs, reducing waste and allowing the system to be used across multiple generations. The platform is fully compliant with major environmental standards, such as RoHS 2 and REACH. The system fans are optimized to sufficiently cool components while minimizing power consumption.

3. Change Log

Date	Version #	Author	Description
07AUG2023	0.1	Matt Bowman Jeremy Baumgartner	Initial draft
23AUG2023	0.9	Matt Bowman Jeremy Baumgartner	Added content to several sections
05SEPT2023	1.0	Matt Bowman Jeremy Baumgartner	Added content

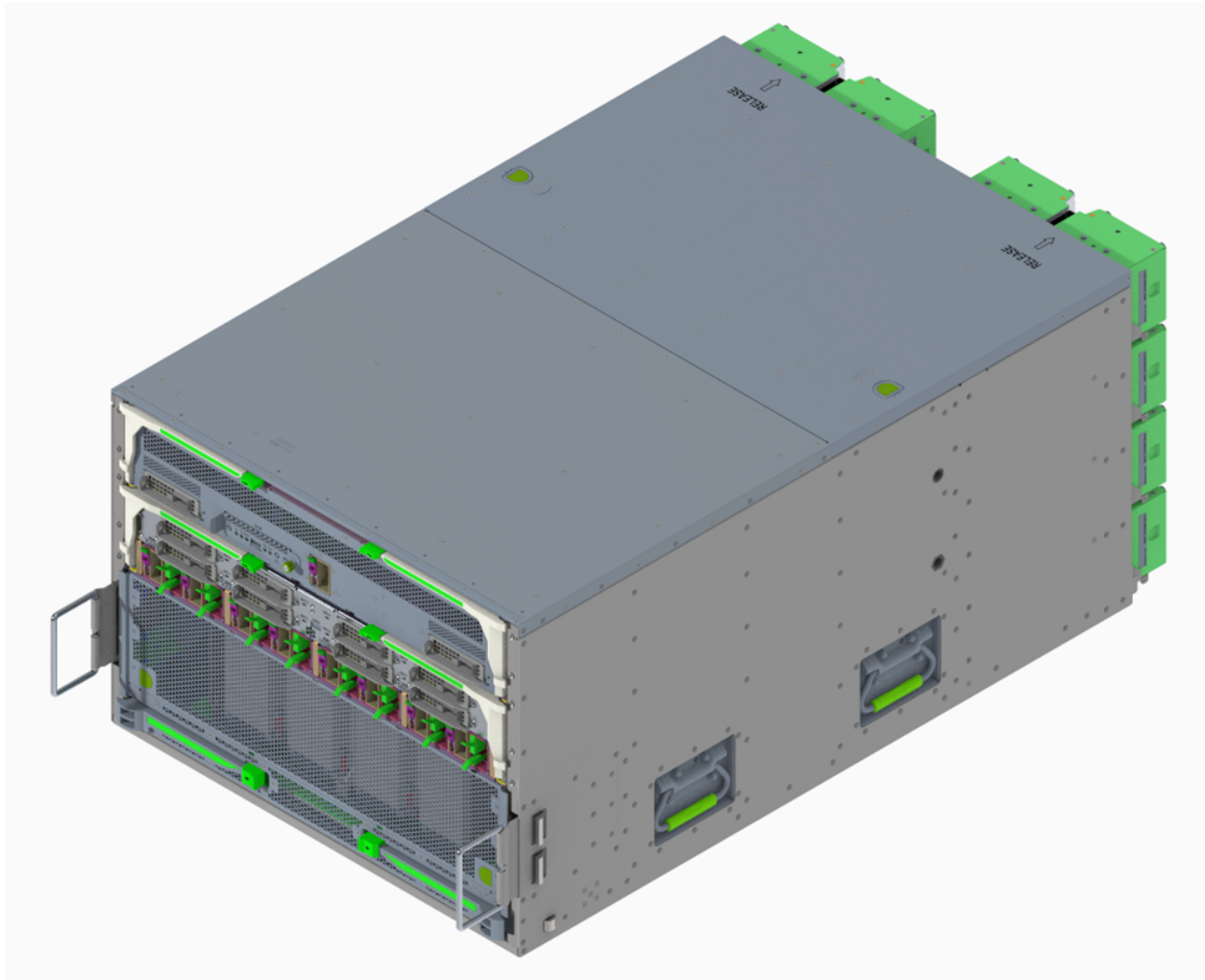
4. Scope

This specification describes the design of the Grand Teton platform which supports AI/ML training and inference use cases.

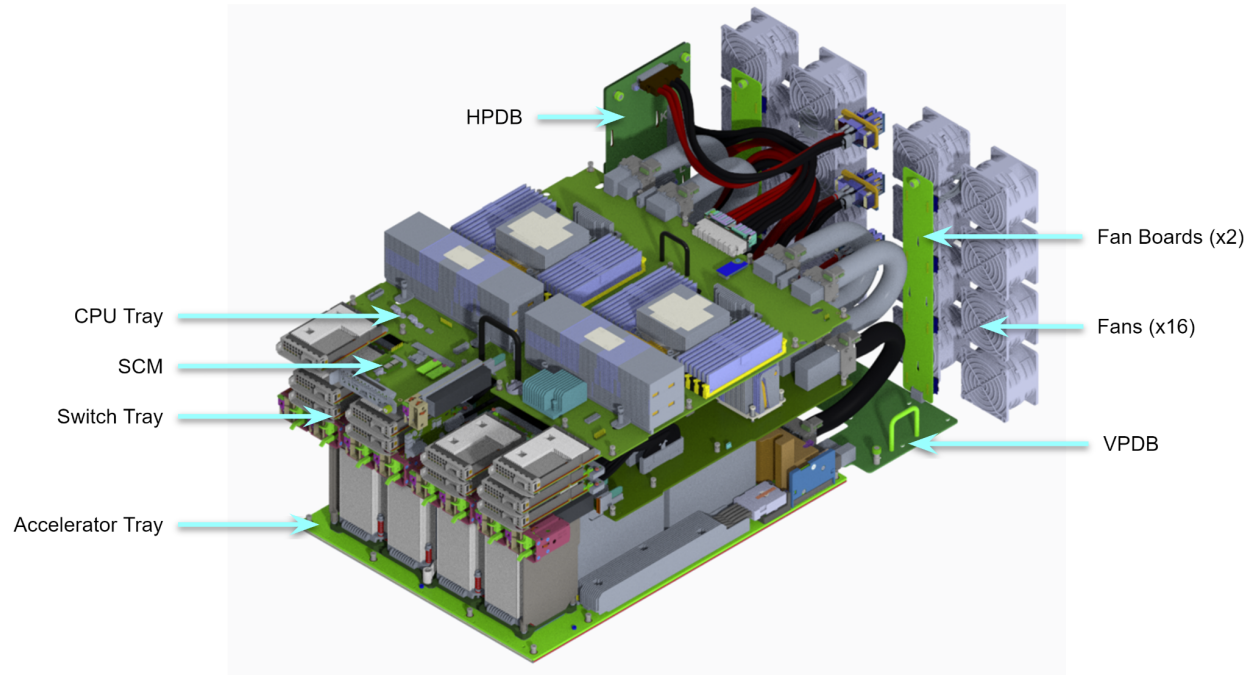
5. Overview

Grand Teton is an 8OU chassis consisting of three interconnected trays: CPU Tray, Switch Tray, and Accelerator Tray. Depending on the necessary usage (ex. training or inference) different tray configurations can be installed to best meet the workload requirements.

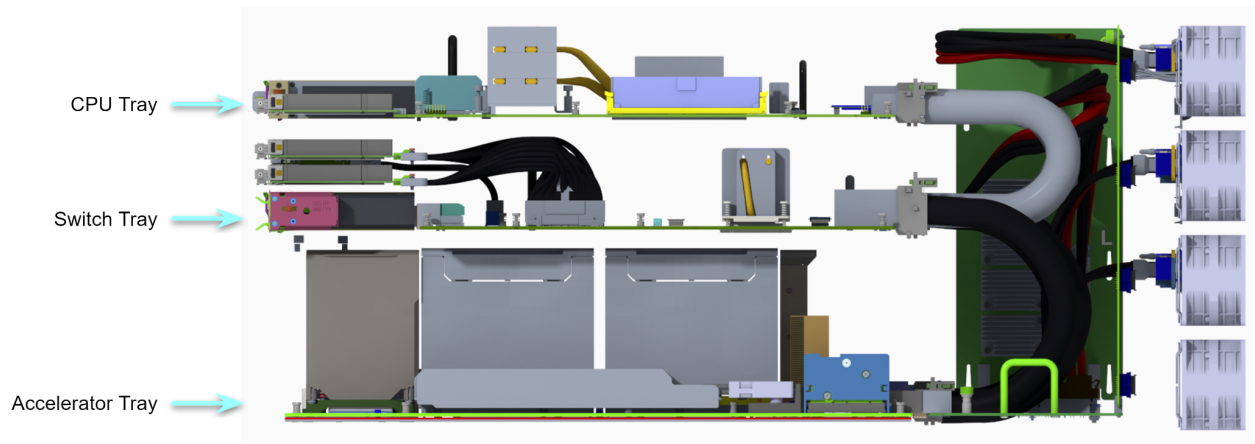
This specification covers the overall Grand Teton platform, including power, mechanical, thermal, and interconnect details. The CPU Tray and Switch Tray have their own OCP specifications, but some specific design details may be mentioned in this specification to serve as examples or improve clarity.



Overall view of the Grand Teton system



Overall view of a Grand Teton system without sheet metal



Side view of a Grand Teton system without sheet metal

5.1 Base Specification

This specification describes a new platform called Grand Teton. More detail and specifics can be found in these specifications:

Grand Teton Intel-based CPU Tray Specification

Grand Teton AMD-based CPU Tray Specification

Grand Teton Switch Tray Specification

6. Environmental and Regulatory Compliance

6.1 Environmental

- Gaseous Contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Ambient operating temperature range: +5°C to +35°C
- Operating and Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C
- Transportation temperature range: -55°C to +85°C (short-term storage)
- Operating altitude with no de-ratings: 6000 feet

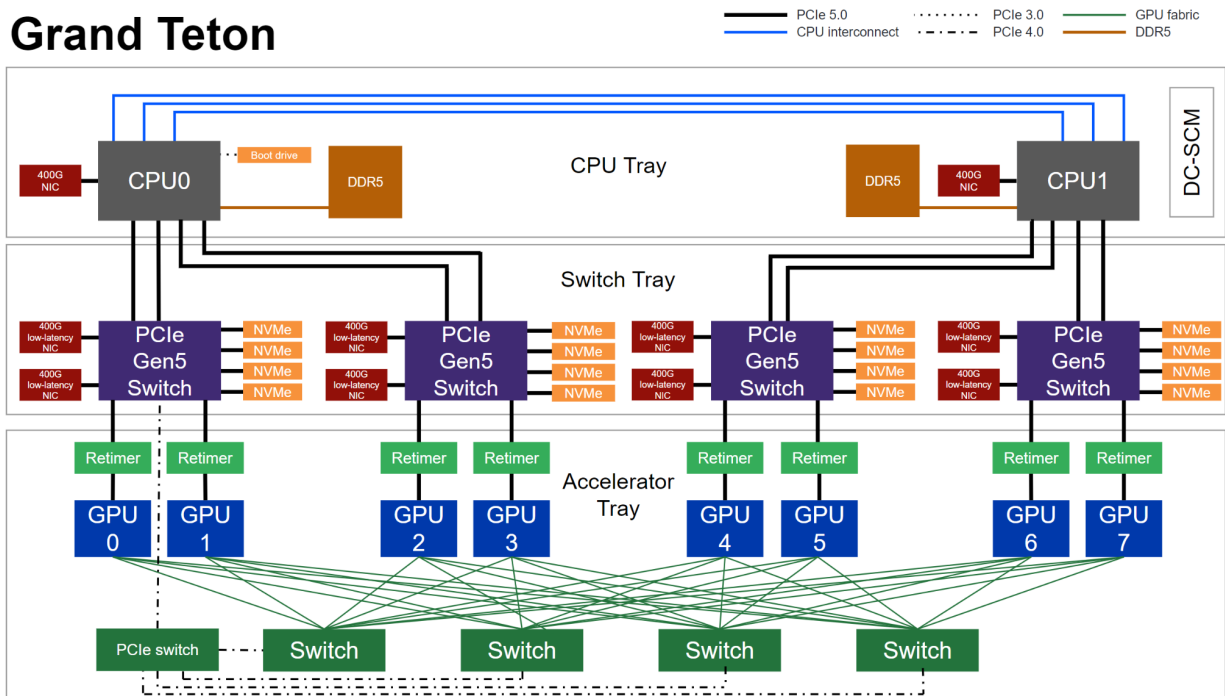
6.2 Regulatory Compliance

- CE and NRTL (UL listed) marks in accordance with the listing and surveillance requirements of the NRTL
- EU Low Voltage Directive (2014/35/EC)
- EU EMC Directive (2014/30/EC)
- EN55032/CISPR32, Class A criteria
- EN55035/CISPR35
- CISPR16
- ANSI C63.4
- EN61000-4-2
- EN61000-4-3
- EN61000-4-4
- EN61000-4-5
- EN61000-4-6
- EN61000-4-11
- FCC CFR47 Part 15, Subpart B, Class A criteria
- RoHS Directive (2015/863/EU)
- WEEE Directive (2012/19/EU)
- REACH regulation (EC 1097/2006)
- US EPA 40 CFR751
- IEC/EN 62368-1, hazard-based performance standard for audio video, and ITE

7. Physical Specifications

7.1 Platform Block Diagram

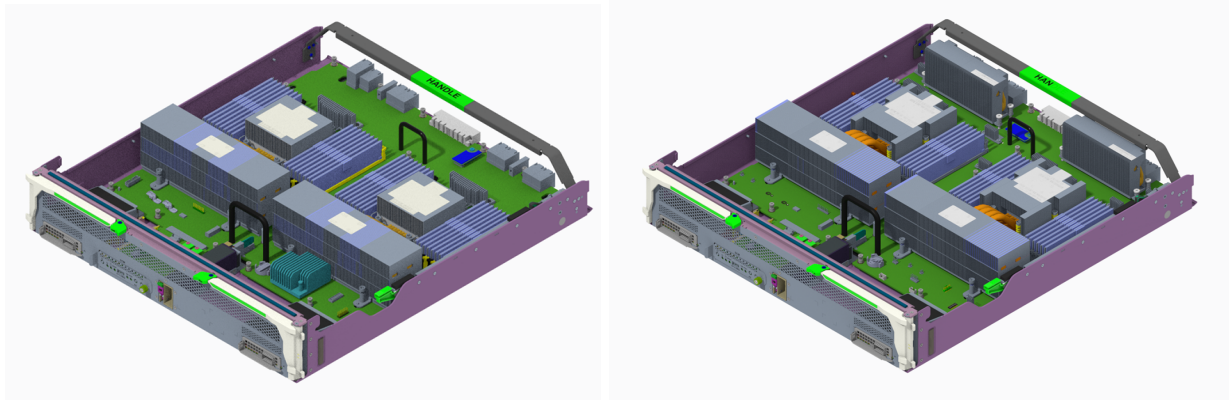
Grand Teton



High-level block diagram of the Grand Teton system

7.2 CPU Tray

The CPU Tray contains a 2-socket CPU motherboard, two OCP NICs, two boot drives, and one SCM with BMC.



Overall view of the Intel-based (left) and AMD-based (right) CPU Trays

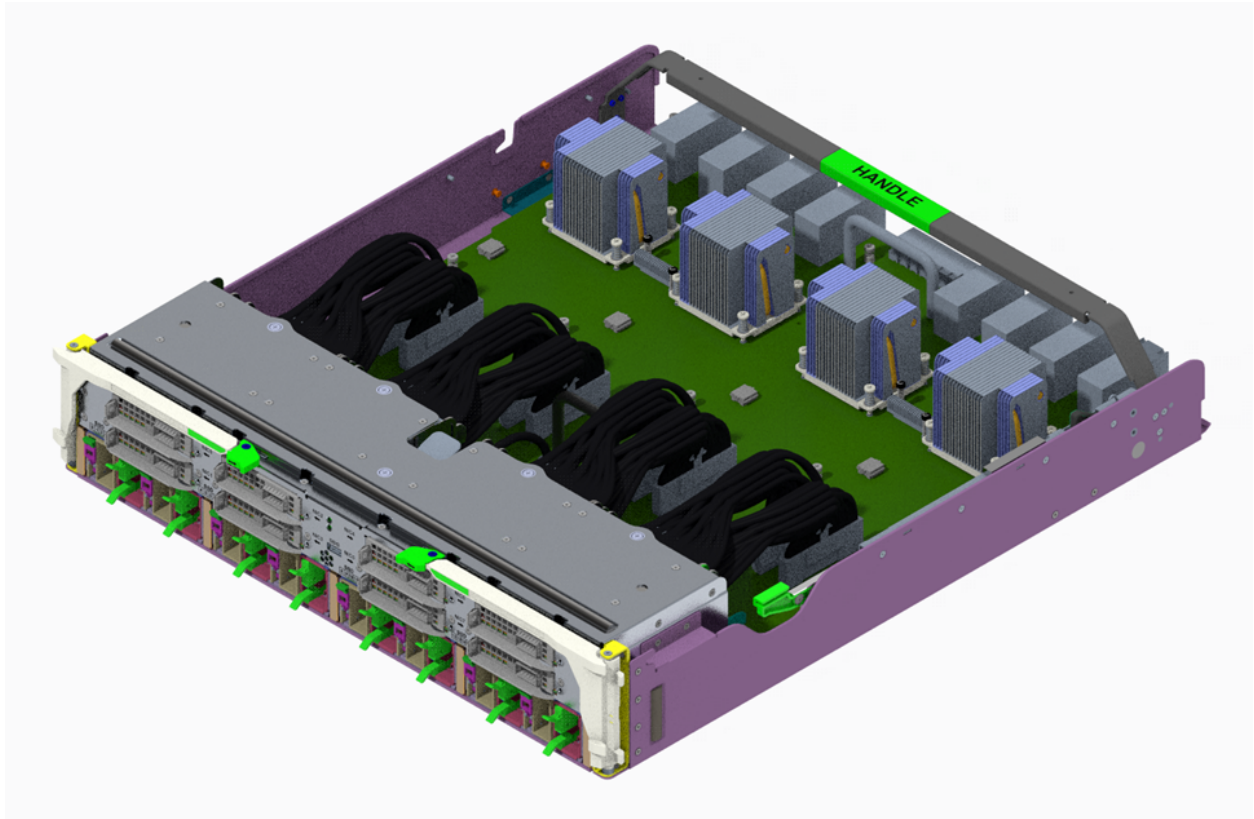
For more details, see the following specifications:

Grand Teton Intel-based CPU Tray Specification

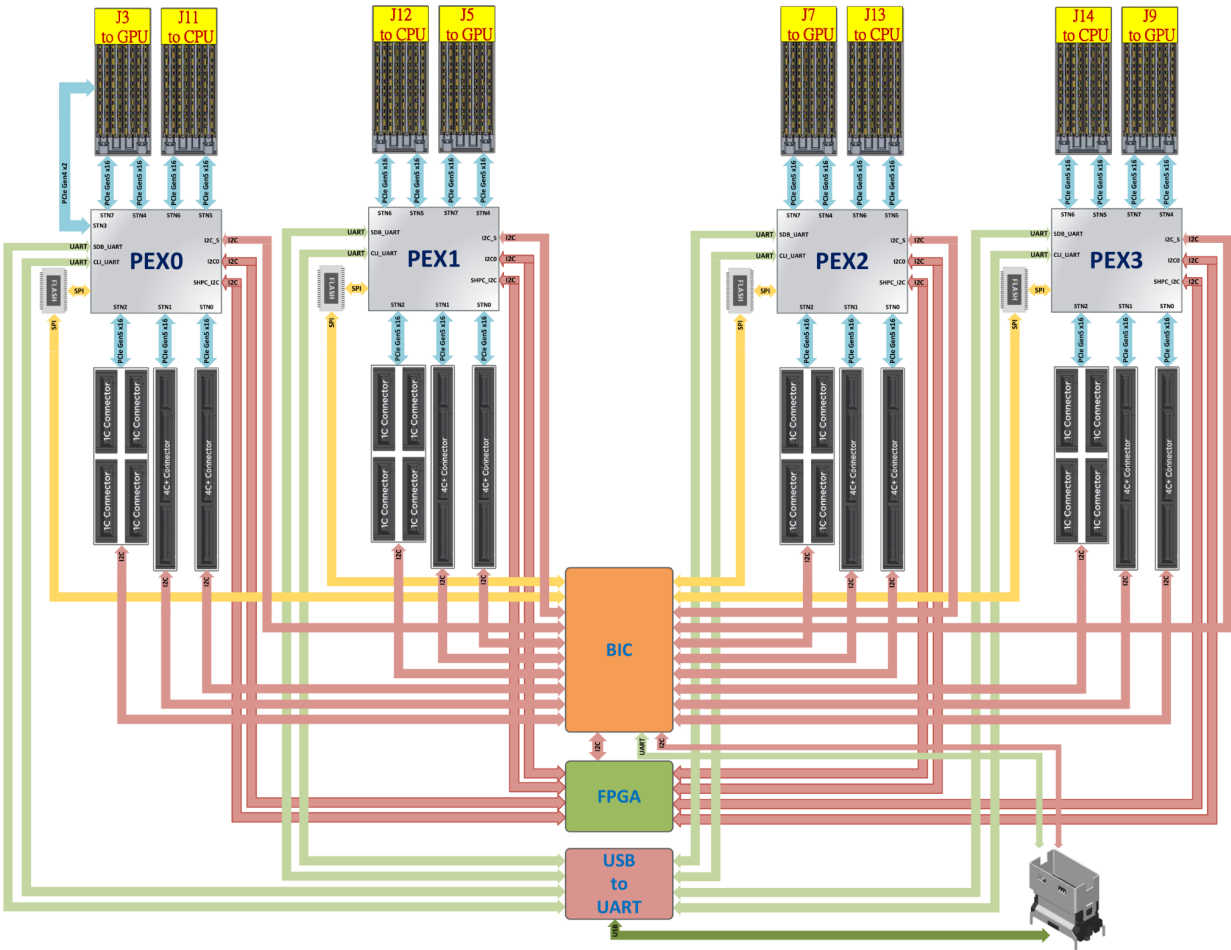
Grand Teton AMD-based CPU Tray Specification

7.3 Switch Tray

The Switch Tray contains four 144-lane Broadcom PCIe 5.0 switches that interconnect the CPUs, NICs, SSDs, and accelerators.



Overall view of the Switch Tray



Block diagram of high-speed and other signals on the Switch Tray

The Switch Tray supports:

- Eight (8) 400G low-latency OCP NIC 3.0
- Sixteen (16) E1.S SSDs

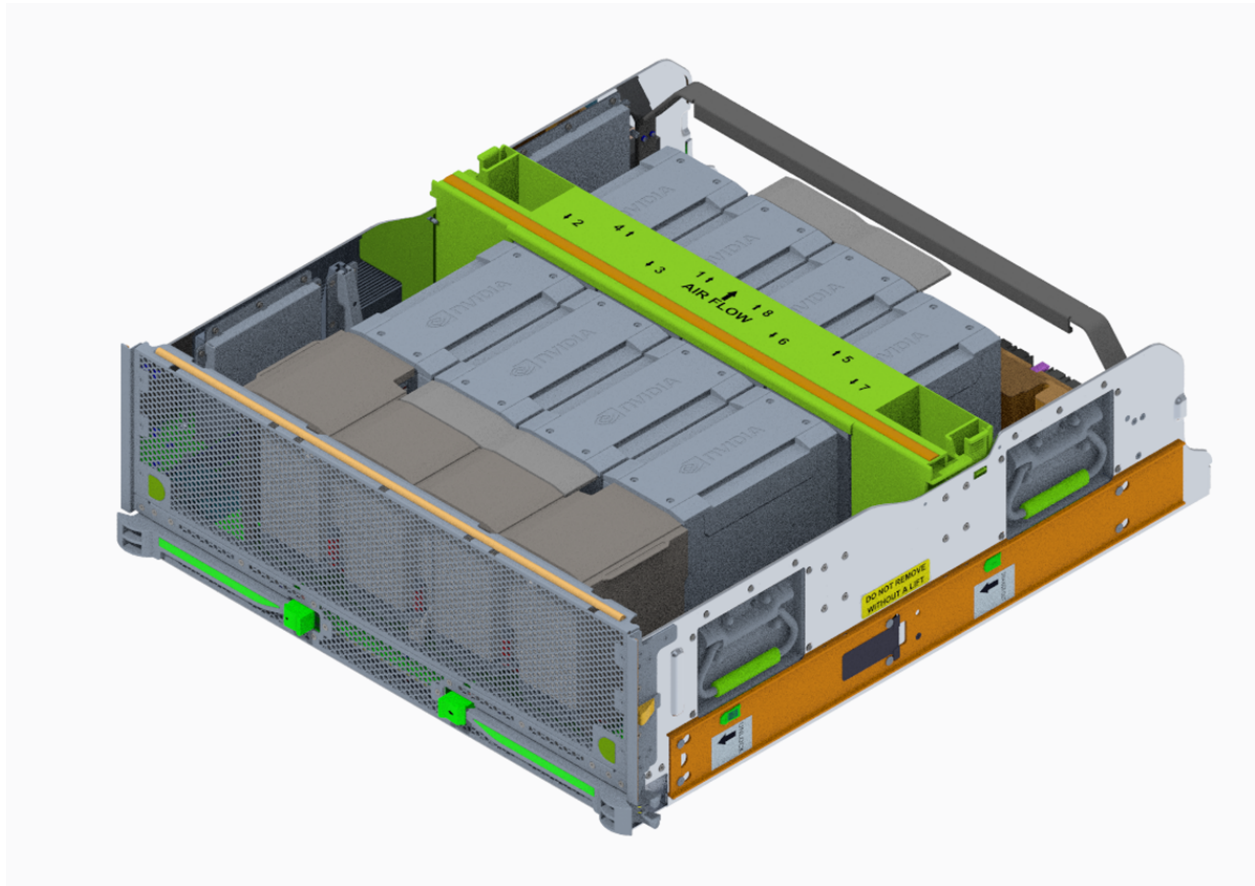
The Switch Tray also contains a cabled extension board called the Front I/O Board, which has a USB 3.0 connector for use with OCP Debug Card.

For more details, see *Grand Teton Switch Tray Specification*.

7.4 Accelerator Tray

The Grand Teton platform is designed to accommodate an Accelerator Tray that follows the *HGX Form Factor Specification*:

<https://www.opencompute.org/documents/open-compute-specification-hgx-baseboard-contribution-r1-v0-1-pdf>



Overall view of an Accelerator Tray

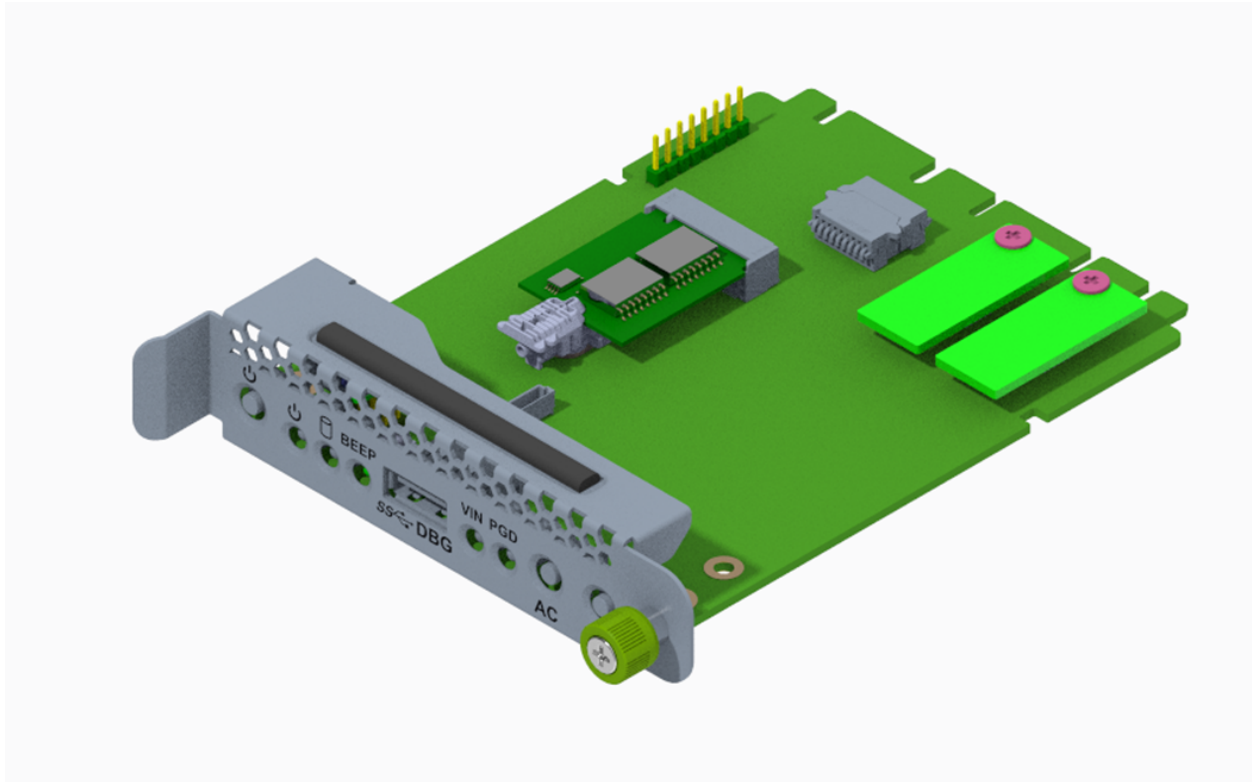
7.5 SCM

The SCM follows the DC-SCM v1.0 specification, with pinout exceptions permitted to include features not covered in the base specification.

The system uses an ASPEED AST2600 BMC for baseboard management with the following connectivity:

- Single x16 8Gb DDR4 DRAM memory interface
- NCSI/RMII network interface on NIC0 to allow out-of-band management
- I2C management of all system sensors for power, thermal and environment
- I2C/I3C for management of Switch Tray BIC
- PCIe 2.0 x1 interface for management of Accelerator Tray management device

- USB interface for management of Accelerator Tray management devices
- JTAG to CPU subsystem (with physical access to system)



Overall view of the SCM

The BMC operates independently of the host system and any other programmable devices within the system with the exception of NIC0 required for network access. The BMC is capable of recovering any firmware in the system regardless of the state of the host system.

The SCM contains removable TPM2.0 for both BIOS and BMC through SPI interface, a SPI flash for BIOS, and a removable board storage module (BSM) containing all BMC non-volatile storage.

The SCM follows the DC-SCM v1.0 specification with the following deviations to pinout to allow features not covered in the base v1.0 specification:

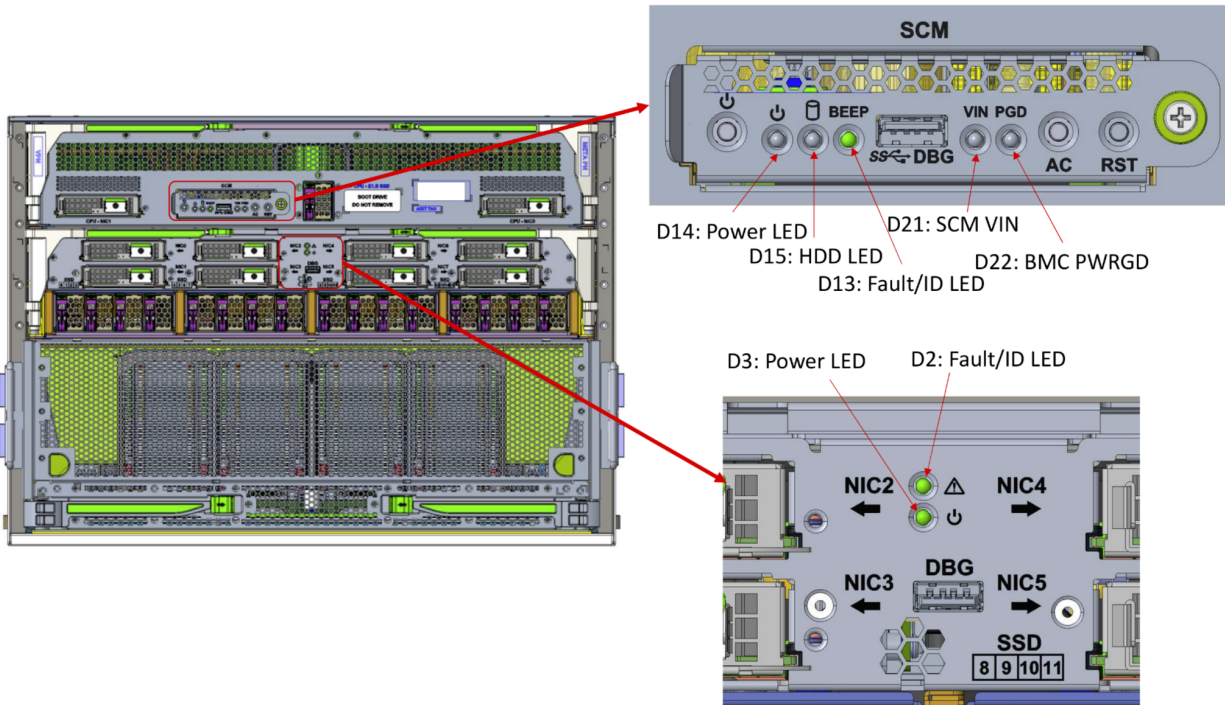
Pin	V1.0 spec	Grand Teton	Description
A52	SPARE1	FM_BMC_UARTSW_MSB_N	Debug UART mux channel
A55	SPARE0	FM_BMC_UARTSW_LSB_N	Debug UART mux channel

Open Compute Project • Grand Teton Platform Specification

A59	PCIE_HPM_RXP[0]	USB3_FPNL_RXP	PCH USB 3.0 on front panel
A60	PCIE_HPM_RXN[0]	USB3_FPNL_RXN	PCH USB 3.0 on front panel
A65	PCIE_HPM_RXP[2]	P2E_GPU_RX_DP	PCIe (BMC Root Complex) to allow management of PCIe Endpoint on Accelerator Tray
A66	PCIE_HPM_RXN[2]	P2E_GPU_RX_DN	PCIe (BMC Root Complex) to allow management of PCIe Endpoint on Accelerator Tray
A68	PCIE_HPM_RXP[3]	CLK_100M_GPU_DP	Reference clock for PCIe RC
A69	PCIE_HPM_RXN[3]	CLK_100M_GPU_DP	Reference clock for PCIe RC
OB11	SPI0_CLK	RST_BMC	Allow BMC reset from system as recovery method
OB13	SPI0_MOSI	CPU_RMCA_CATERR	Error signal from CPU, routed directly for fast response
OB14	SPI0_MISO	SOL_UART_CH_SEL	OCP debug card UART channel select signal
B9	RSVD3	LPC_ESPI_SEL	LPC/ESPI select strap which must be set before power up
B35	RSVD2	SGPIO_CLK_1	Independent clocking for SGPIO0 and SGPIO1
B42	QSPI0_CS1_N	AC_CYCLE	Signal back to system HSC to allow hard AC cycle
B56	RSVD0	USB_FPNL_DP	PCH USB 2.0 on front panel
B57	RSVD1	USB_FPNL_DN	PCH USB 2.0 on front panel
B59	PCIE_HPM_TXP[0]	USB3_FPNL_TXP	PCH USB 3.0 on front panel
B60	PCIE_HPM_TXN[0]	USB3_FPNL_TXN	PCH USB 3.0 on front panel
B65	PCIE_HPM_TXP[2]	P2E_GPU_TX_DP	PCIe (BMC Root Complex) to allow management of PCIe Endpoint on Accelerator Tray
B66	PCIE_HPM_TXN[2]	P2E_GPU_TX_DN	PCIe (BMC Root Complex) to allow management of PCIe Endpoint on Accelerator Tray

The front panel of the SCM has the following:

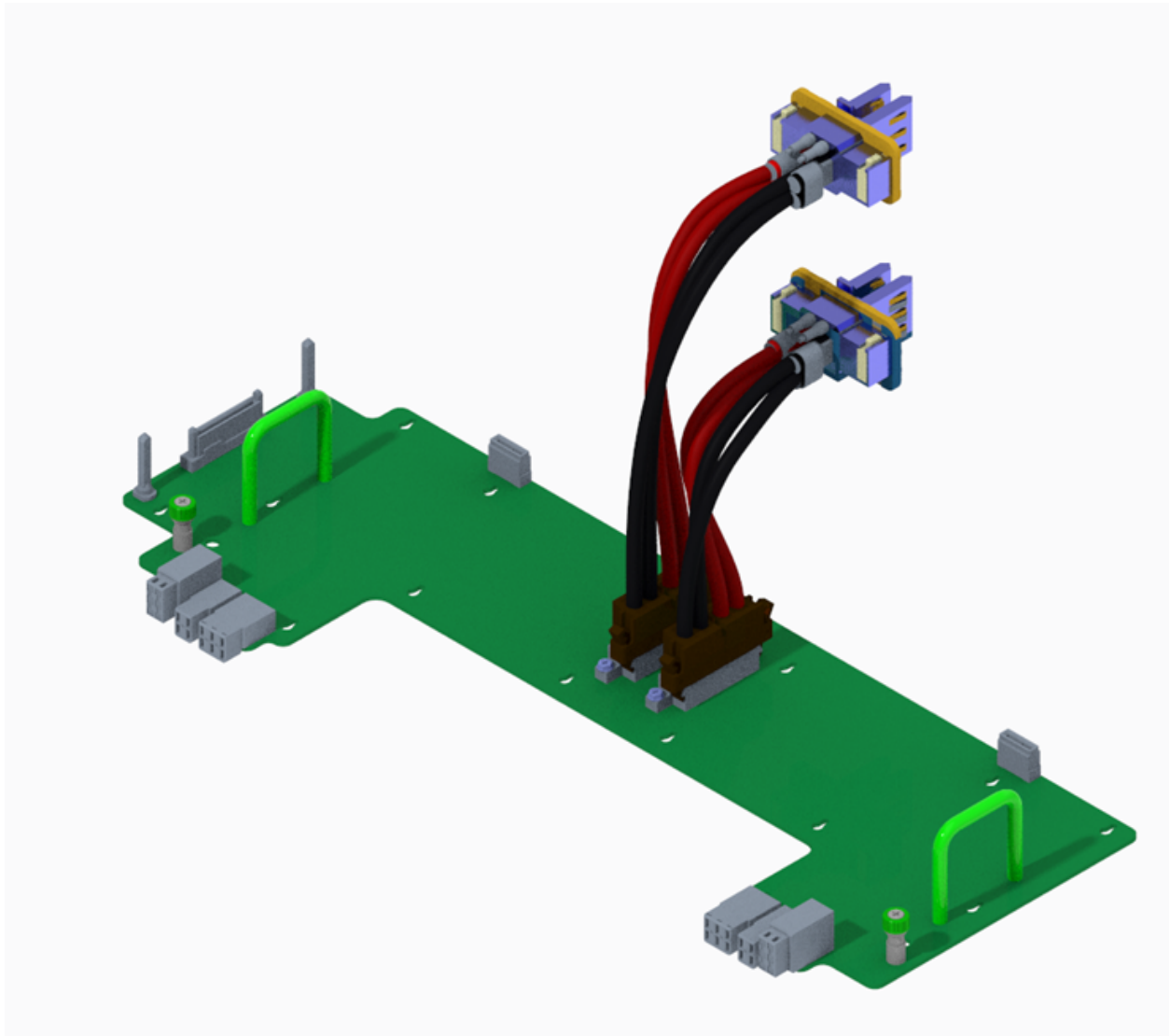
- Power-on pushbutton
- Power LED
- HDD LED
- Fault/ID LED
- USB 3.0 debug port
- SCM voltage status LED
- SCM power good status LED
- AC power hard-cycle pushbutton
- Reset pushbutton



Front view of the Grand Teton system buttons and indicators

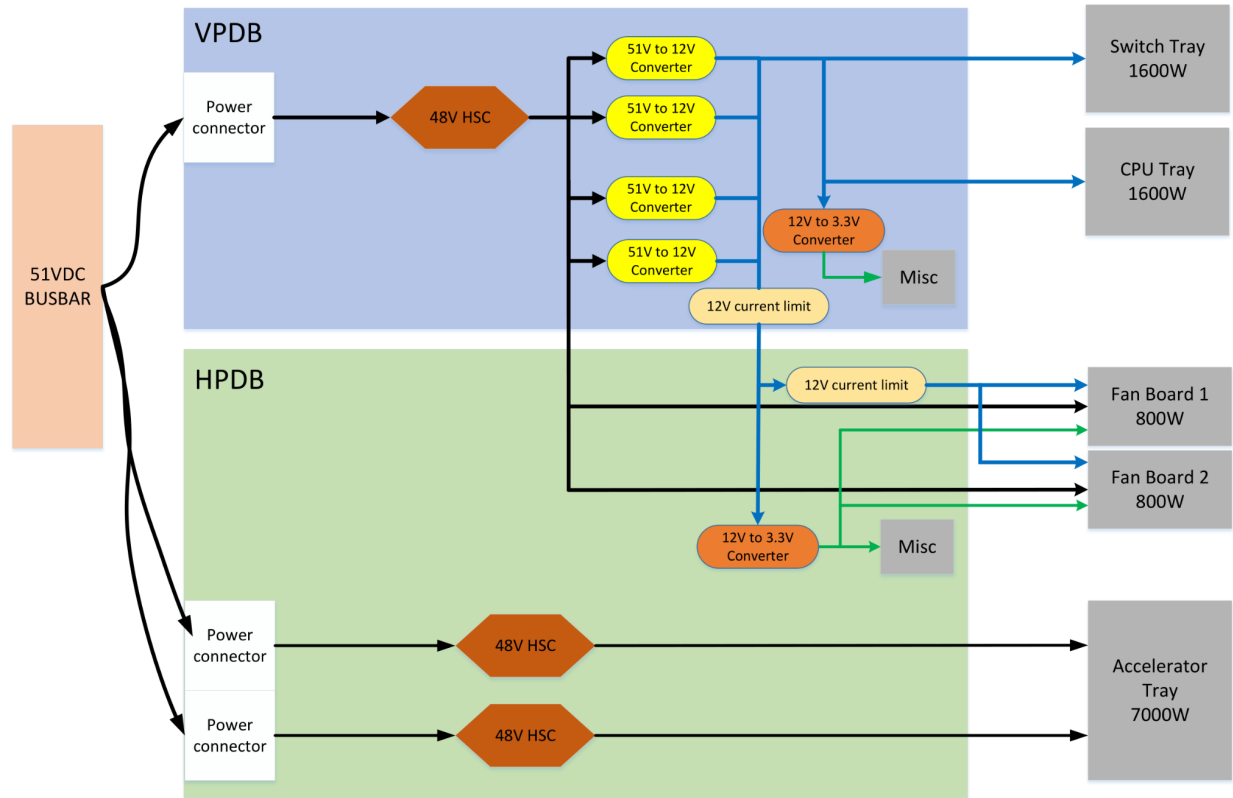
7.6 Horizontal Power Distribution Board (HPDB)

The primary purpose of the VPDB is to convert the ORv3 bus bar voltage (51VDC) to the native voltage of the Accelerator Tray (12VDC).



Overall view of the HPDB with attached bus bar power cables

The HPDB supports hot-swap into a live 51VDC bus bar and full control and monitoring of the output voltage rail by the system BMC. The HPDB contains a high efficiency 51V to 12V buck converter capable of operating over the full ORv3 bus bar input voltage range with separate telemetry provided to the system BMC.



High-level block diagram of the Grand Teton power system and boards

The HPDB has multiple connections:

- Power In from bus bar (2 connectors)
- Power and data from VPDB
 - Fan power passed through to Fan Boards
 - I2C, control, and status signals from BMC
- Power and data to Fan Board (2 connectors)

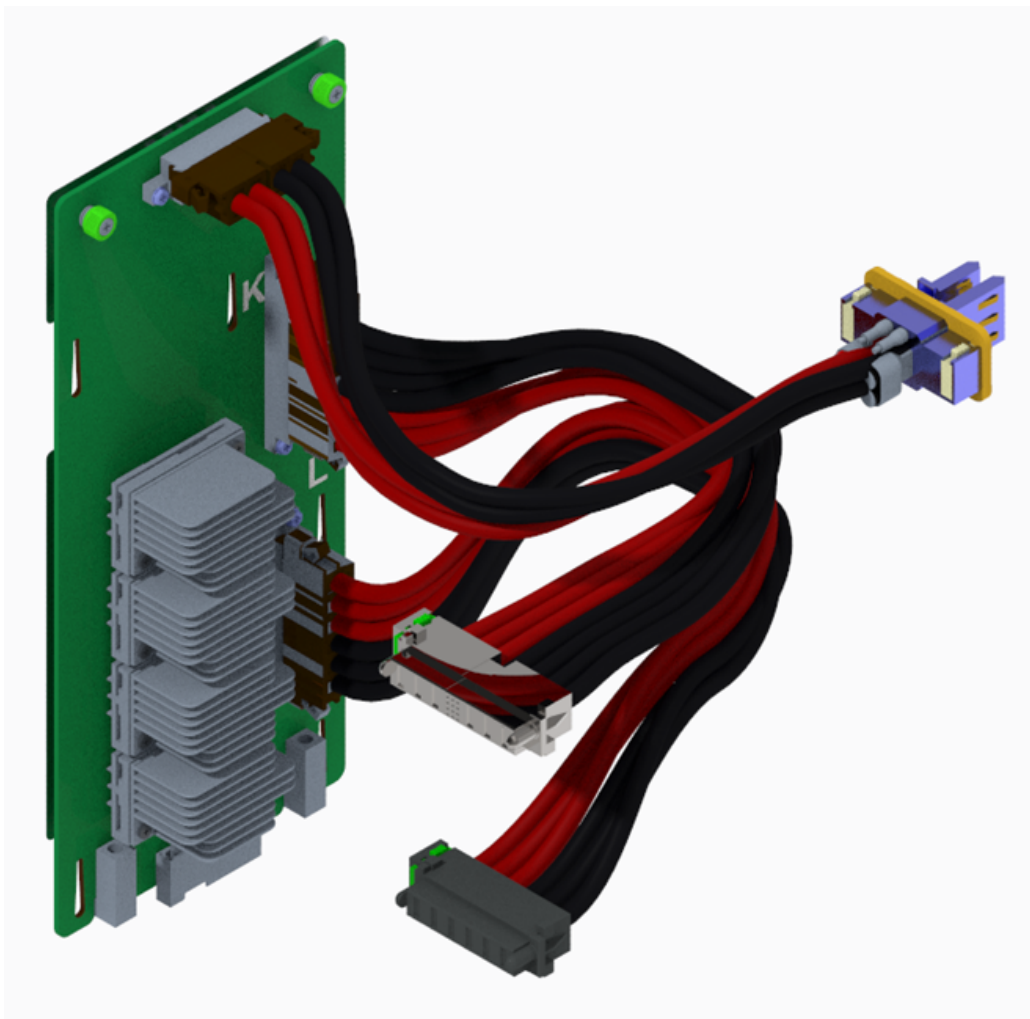
The following status and control signals are present between the VPDB and HPDB.

Signal name	Direction	PU/PD	Description
CPU_P3V3_PGOOD	IN	PD	Output from CPU CPLD to indicate CPU Tray 3V3 is good and HPDB IO buffers can be enabled. Before this signal is asserted, all signals on HPDB are defined by the Hardware defaults (e.g. HPDB_HSC_EN).

FAN_EN	IN	none	Global enable for Fan 48V eFuse passed through from VPDB to Fan boards.
HPDB_AUX_EN	IN	PD	Enable HPDB 3V3 VR and 12V eFuse for Fan board. Asserted by VPDB when 12V is ready.
HPDB_HSC_EN	IN	PD	Enable signal for HPDB HSC controlled by CPU CPLD logical AND GPU_PRSENT signal (from ExaMax) to ensure HSC can only be enabled if GPU tray is present. Before CPU_P3V3_PGOOD, buffer is disabled, PD keeps HSC disabled. After CPU_P3V3_PGOOD, signal under BMC control.
HPDB_HSC_PWRGD	OUT	PD	Power good signal from GPU HSC indicating both 48V HSC are enabled and within spec.
HPDB_HSC_SCL/SDA	IN/OUT	PU	I2C interfaces for HPDB HSC
PDB_SCL/SDA	IN/OUT	PU	I2C interface for HPDB EEPROM, IO Expander and Fan board
PDB_SMB_RST_N	IN	none	Reset supported SMBus/I2C devices on HPDB/FAN, e.g. I2C mux
P12V	PWR_IN	none	12V for HPDB 3V3 VR and Fan LED supply
P52V	PWR_IN	none	52V passthrough to Fan board

7.7 Vertical Power Distribution Board (VPDB)

The primary purpose of the VPDB is to convert the ORv3 bus bar voltage (51VDC) to the native voltage of the CPU Tray and Switch Tray (12VDC).



Overall view of the VPDB with attached power cables

The VPDB supports hot-swap into a live 51VDC bus bar and full control and monitoring of the output voltage rail by the system BMC. The VPDB contains a high efficiency 51V to 12V buck converter capable of operating over the full ORv3 bus bar input voltage range with separate telemetry provided to the system BMC.

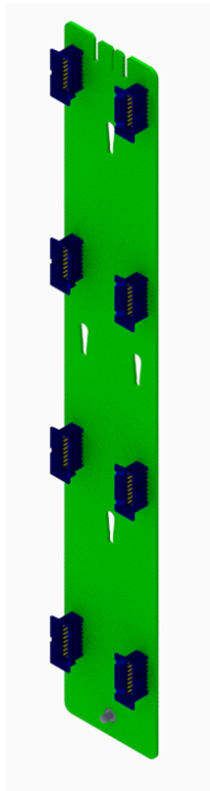
Voltage Input	Voltage Output	Design Current	Over Current Protect
46-52V	12.2V +/-5%	186A	360A

The following signals are buffered from CPU Tray to HPDB with defaults listed below.

Signal name	Direction	PU/PD	Description
CPU_P3V3_PGOOD	CPU to HPDB	PD	If the CPU Tray is not present or P3V3 is not ready, the buffer will be tri-state.
FAN_EN	CPU to HPDB	PD	If the CPU Tray is not present or P3V3 is not ready, the buffer will be tri-state.
PDB_SMB_RST_N	CPU to HPDB	PU	If the CPU Tray is not present or P3V3 is not ready, the buffer will be tri-state.

7.8 Fan Boards

There are two Fan Boards at the rear of the system, each are capable of powering, controlling, and monitoring 8 dual rotor fans for a total of 16 system fans. Fan Boards are installed on the inside of the chassis and expose eight blind mate fan connectors to the rear of the chassis for hot aisle service of fan assemblies.

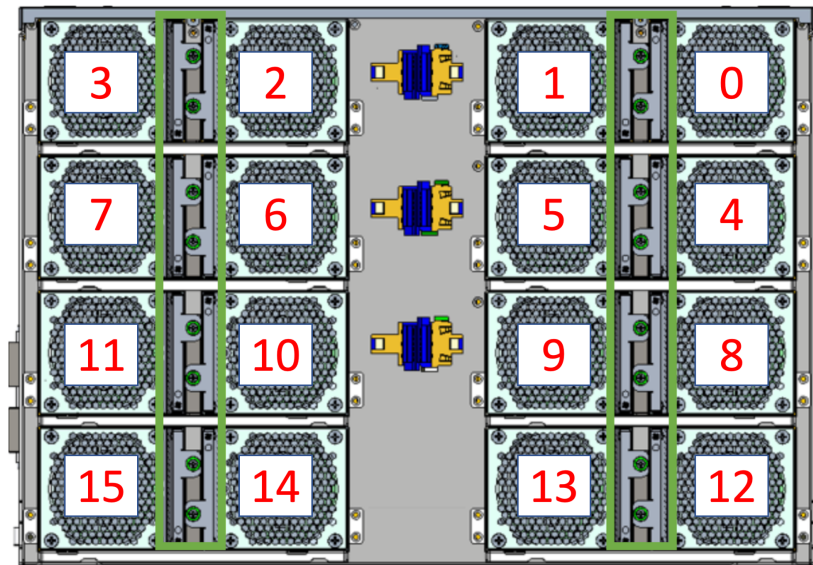


Overall view of a single fan board

Each fan location has an individual current limit circuit to limit inrush to the fan at startup and provide protection in case of fan failure.

Signal name	Direction	PU/PD	Description
FAN_SCL/SDA	IN/OUT	none	I2C interfaces for Fan Board
FAN_EN	IN	none	Global enable for Fan 48V eFuse
PDB_SMB_RST_N	IN	PU	Reset supported SMBus/I2C devices on FAN, e.g. I2C mux
P3V3	PWR_IN	none	3V3 supply for IC (e.g. fan controller, IO expander)
P12V	PWR_IN	none	12V supply for Fan LED
P48V	PWR_IN	none	Main fan power supply

The following diagram shows the arrangement of the fans in the chassis:

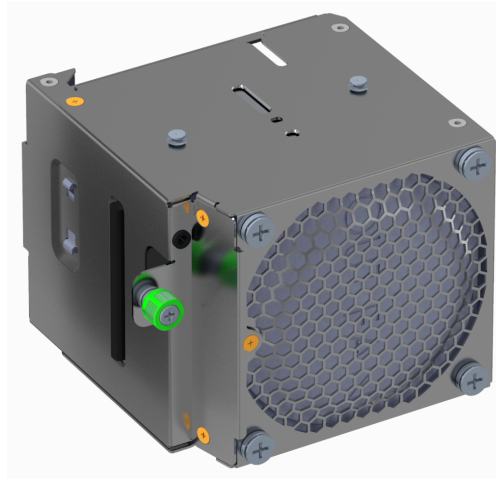


Rear view of the Grand Teton system, with fan module numbering

7.9 Fans

Each fan has the following specifications:

- 46-52V operating voltage
- Dual rotor design
- Common PWM signal
- Independent TACH for monitoring fan speed
- Color LED for fan status
- I2C interface for Fan FRU EEPROM

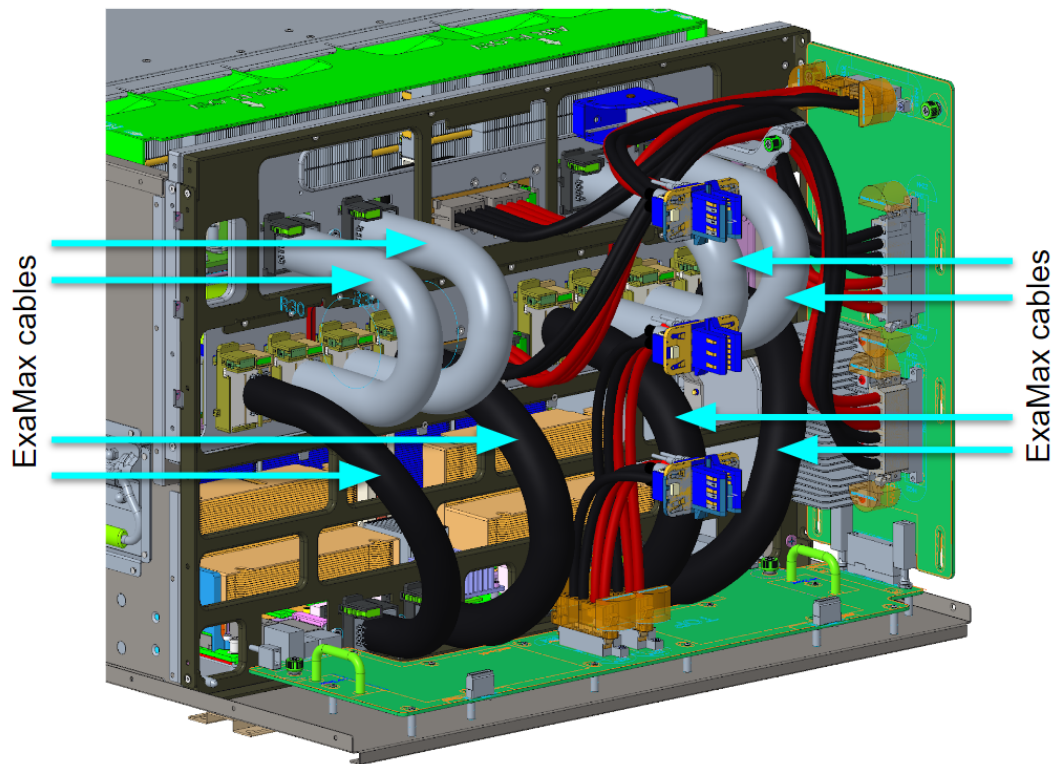


Overall view of a fan module

For more details, see Thermal section.

7.10 High Speed Cables

All inter-tray connectivity is achieved through an ExaMax high speed cable backplane. Four cables connect the CPU Tray to the Switch Tray. Four cables connect the Switch Tray to the Accelerator Tray.



Overall diagram of the ExaMax cable backplane

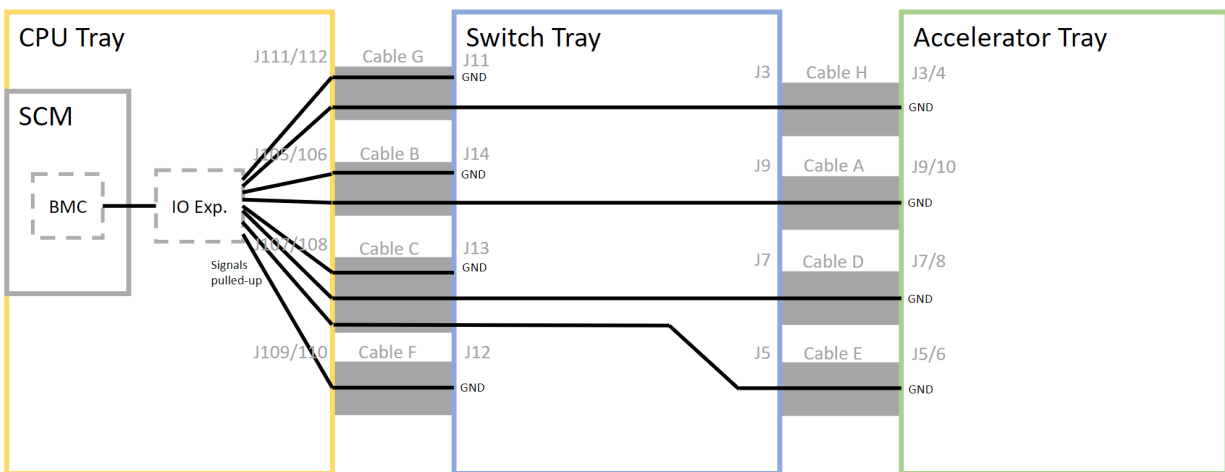
CPU Tray		J111+J112 4x8+6x8	J109+J110 4x8+4x8			J107+J108 4x8+4x8	J105+J106 4x8+4x8	
Switch Tray	J3 8x10	J11 8x10	J12 8x10	J5 8x10	J7 8x10	J14 8x10	J13 8x10	J9 8x10
Accelerator Tray	J4+J3 4x8+6x8			J6+J5 4x8+4x8	J8+J7 4x8+4x8			J10+J9 4x8+4x8

Diagram of the ExaMax cable connectivity. Each color block is one cable (8 total).

The cable types are as follows:

- 6 ExaMax cables, 4x8+4x8 connector to 8x10 connector
 - High-speed signals for PCIe 5.0 x16
 - 16 GPIO
 - Length: 350mm
- 2 ExaMax cables, 4x8+6x8 connector to 8x10 connector
 - High-speed signals for PCIe 5.0 x16, clock, and USB
 - Low speed control and status signals
 - Length: 350mm

Each ExaMax cable includes a present signal to allow detection at the BMC. The following diagram shows the present signal paths across each ExaMax cable in the system. These are normally-high signals that get pulled-down (active-low) when the cable and board are present and connected.



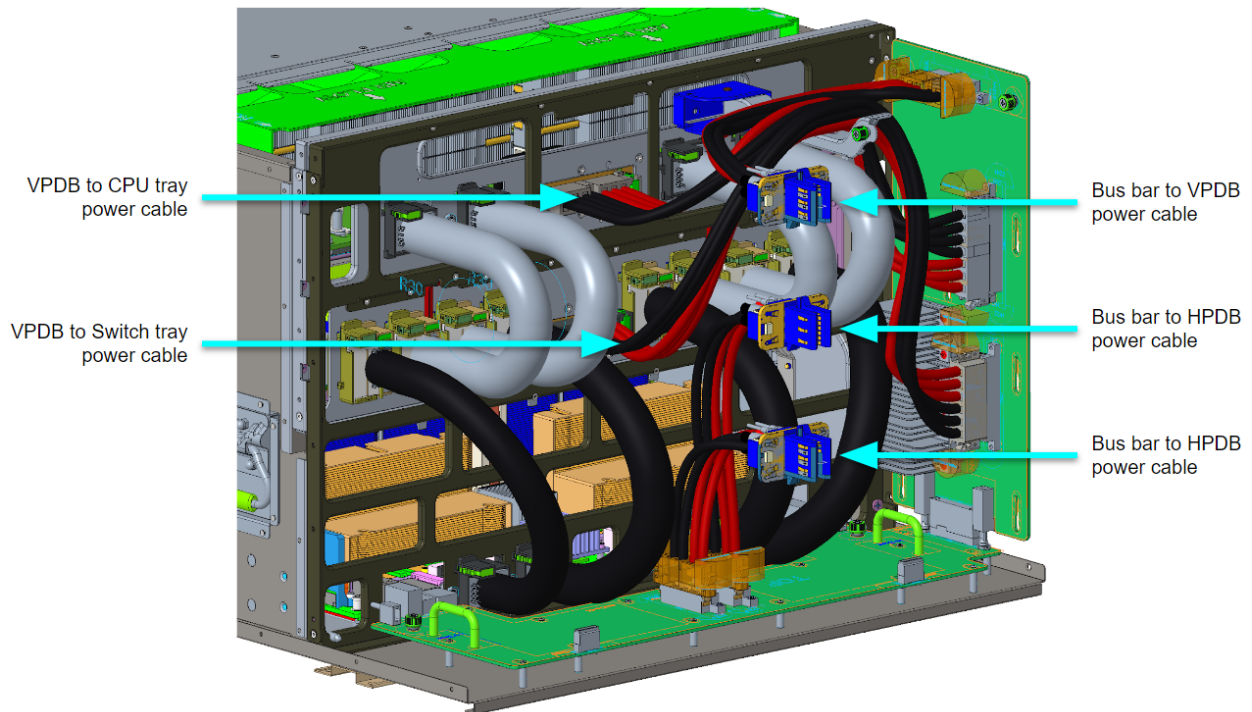
Overall diagram of the present signals through the ExaMax cables

For presence signals that start at the GPU board and pass through all boards, BMC firmware logic can identify the problematic cable by reading the other cable present status.

7.11 Power Cables

Three power cables are used to connect the 51VDC ORv3 bus bar to power distribution boards. The three bus bar clips are stacked vertically at the rear-center of the chassis to mate with the ORv3 bus bar. The VPDB uses one cable, and the HPDB uses two cables.

Additionally, two 12V cables connect the VPDB to the interior backplane to provide power to the CPU Tray and Switch Tray.



Overall diagram of the power cables backplane

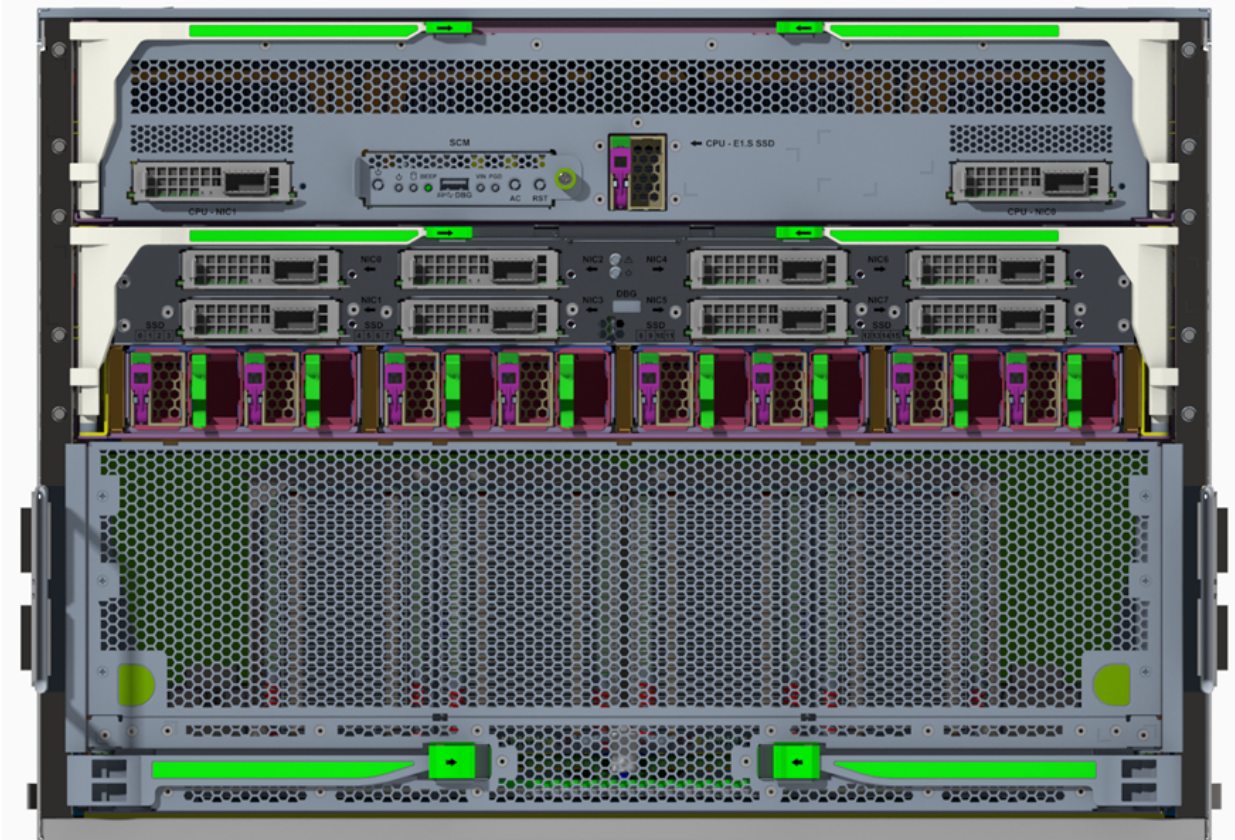
The CPU Tray cable also carries present, status, and control signals for the BMC and CPU to control circuits on VPDB/HPDB and Fan boards.

The Switch Tray cable only carries 12V power. Status, control, and present signals are located in the ExaMax cables.

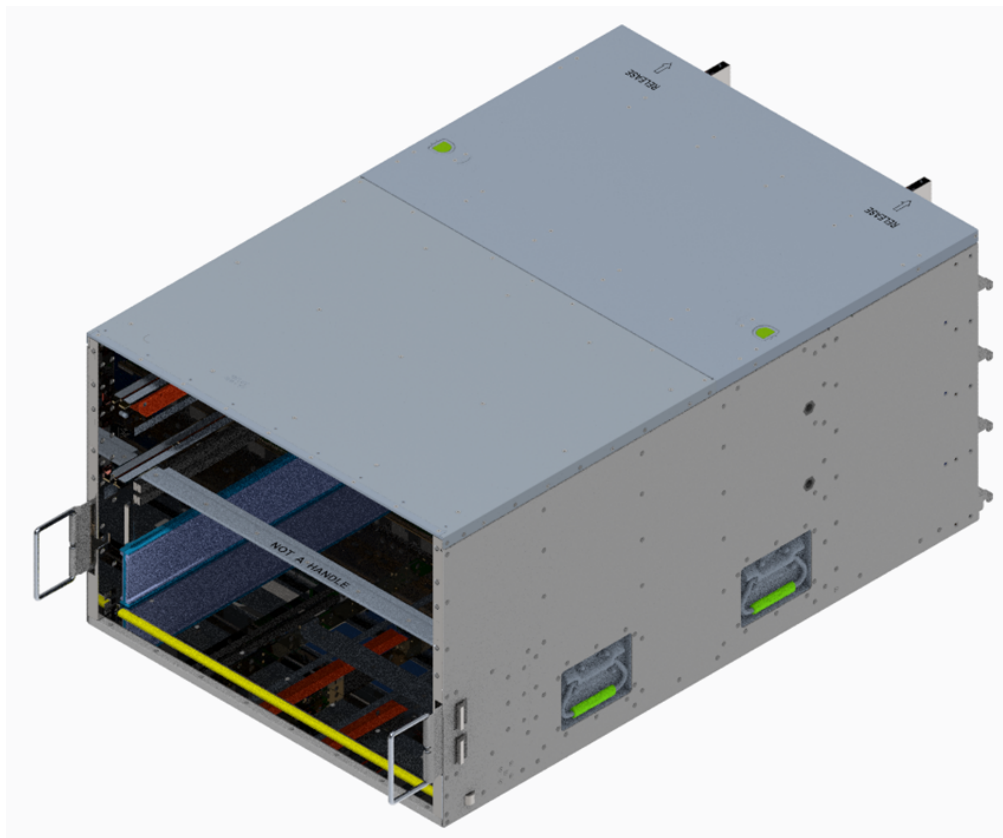
Power to the Accelerator Tray is provided through coplanar connectors from the HPDB. Power to the Fan Boards is provided through vertical power connectors on the HPDB.

7.12 Mechanical

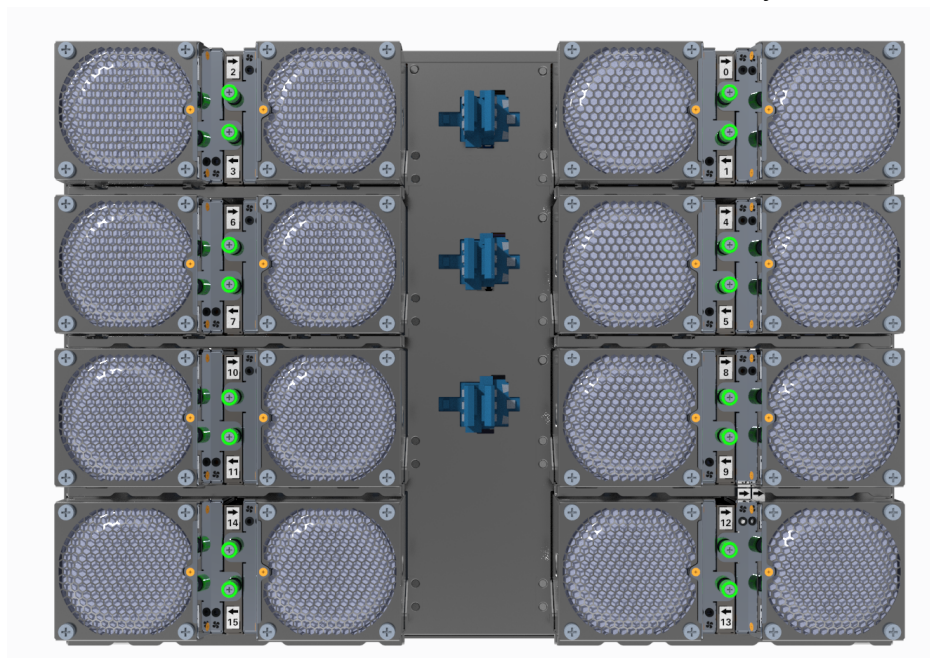
The Grand Teton design consists of three front accessible trays in an 8OU chassis. The three trays from top to bottom are the CPU Tray, the Switch Tray, and the Accelerator Tray. These trays are approximately 20U, 20U, and 40U respectively.



Front view of the Grand Teton system

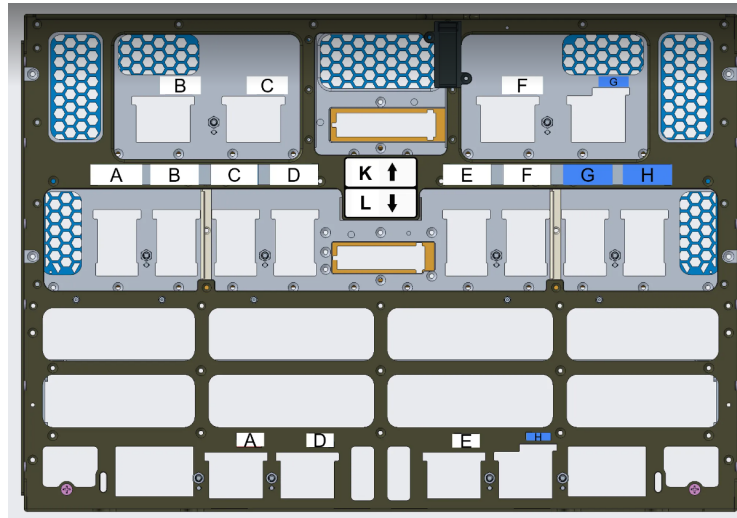


Overall view of the Grand Teton chassis with all three trays removed



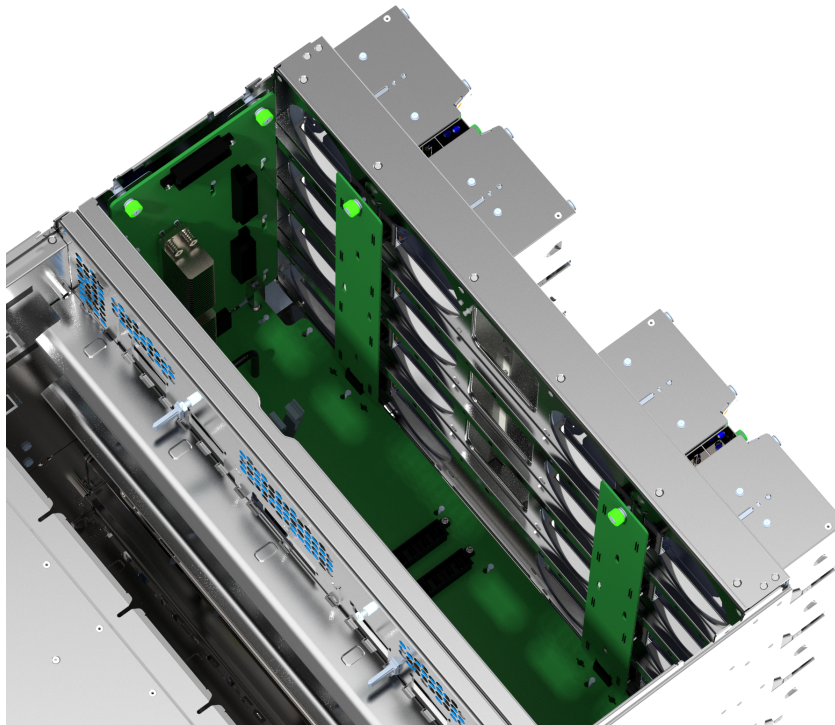
Rear view of the Grand Teton system

The CPU Tray and the Accelerator Tray connect to the Switch Tray using a blind mate cable backplane of ExaMax cables.



Isolated view of the Grand Teton interior cable backplane

There are 16x rear-accessible, hot-swappable 80x80x56mm fans in the back of the system. The fans connect to two Fan Boards with a blind mate floating cable connector.



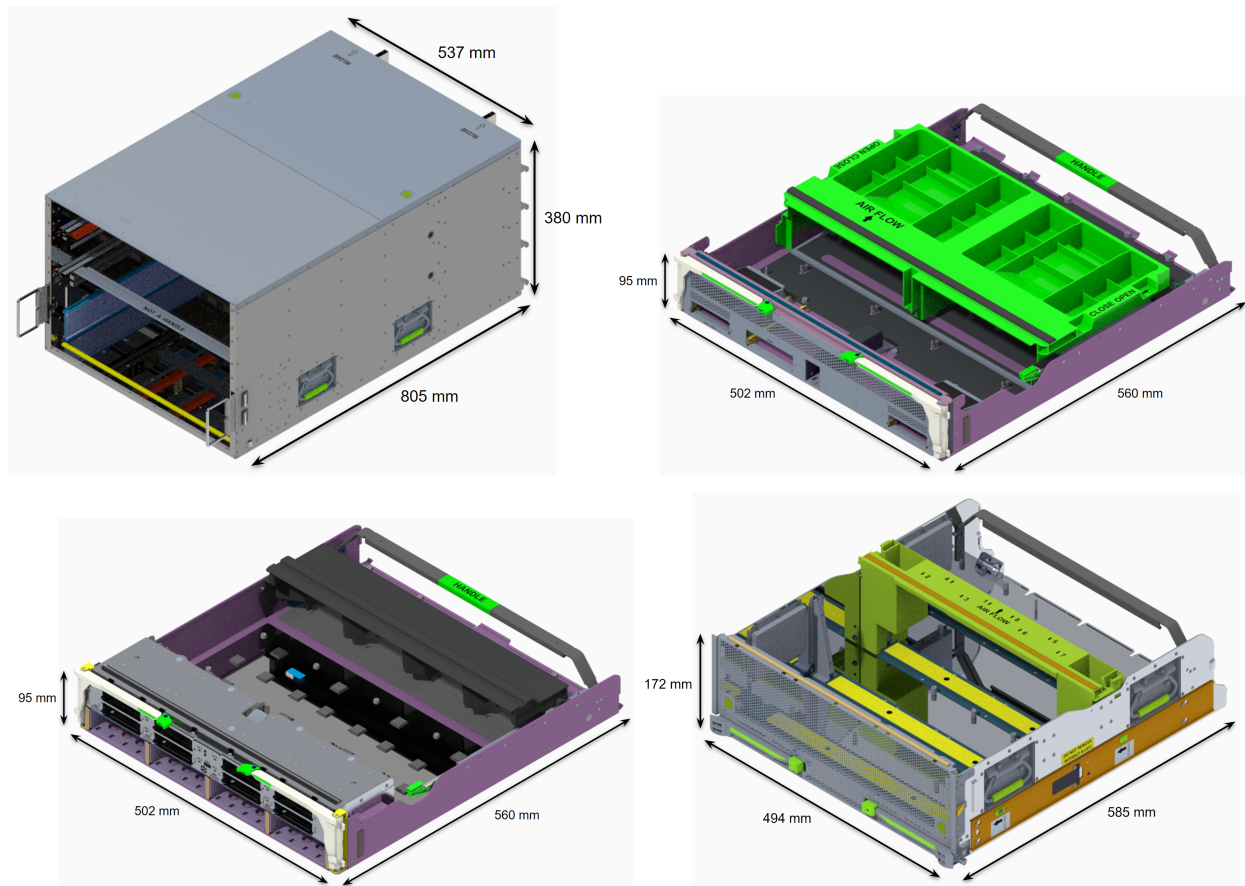
Angled interior view of the Grand Teton fan boards and fan modules (without cables)

All server components remain inside ORV3 standard 8OU space during operation.

The design is vanity-free and simple, with no expense or design effort paid to parts or processes that do not contribute to function. The part count is minimized to reduce component and manufacturing costs.

The system has the following dimensions:

- Chassis
 - Dimensions: 380 mm (H) x 537 mm (W) x 805 mm (D)
 - Empty chassis weight: 42.8 kg / 94.4 lb
 - Fully loaded chassis weight: 118.7 kg / 261.7 lb
- CPU Tray
 - Dimensions: 95 mm (H) x 502 mm (W) x 560 mm (D)
 - Weight: 12.1 kg / 26.6 lb
- Switch Tray
 - Dimensions: 95 mm (H) x 502 mm (W) x 560 mm (D)
 - Weight: 17.6 kg / 38.9 lb
- Accelerator Tray
 - Dimensions: 172 mm (H) x 494 mm (W) x 585 mm (D)
 - Weight: 32.9 kg / 72.7 lb



Mechanical dimensions of the chassis and 3 trays

All server and rack components are easily accessible from the front side of the rack (cold aisle), with the exception of low-failure rate redundant components such as fans, which are serviced from the hot aisle.

Trays and other system components are field-replaceable and tool-less with green touch points and free of sharp edges.

7.13 Rack Compatibility

Grand Teton is designed for use in Open Rack V3 (ORv3) racks.

8. Electrical Requirements

The Grand Teton platform is compliant with ORv3 specifications.

8.1 Input Voltage

Grand Teton is compliant with the ORv3 specification for 51VDC input voltage. The system will operate normally with an input voltage tolerance range between 46V and 52V.

8.2 Power

The Grand Teton system power consumption depends on the implementation and workload.

The power distribution boards are capable of supplying:

- 6kW from the VPDB
- 8kW from the HPDB

9. Thermal Design Requirements

9.1 Thermal Architecture

Grand Teton is an 8-OU system with 3 trays in parallel: CPU Tray, Switch Tray and Accelerator Tray. Airflow to all 3 trays is delivered by a fan array at the rear side of the chassis, with a shared plenum. All components are air-cooled, using passive heatsinks where applicable.

9.2 System Thermal Specification

	Normal Operation	Corner Case	Remain Power on
Inlet Temperature	Up to 30°C	Up to 35°C	Up to 45°C
Altitude (no inlet temp deration)	Up to 6000 ft	Up to 6000 ft	Up to 6000 ft
Cold/Hot Aisle Δp	Within +/- 0.005 InH2O	Within +/- 0.1 InH2O	Within +/- 0.1 InH2O
Relative Humidity	10% ~ 90%	10% ~ 90%	10% ~ 90%
Fan Status	All Healthy	Single fan/rotor failure	Single fan/rotor failure
Air side ΔT (CFM/W)	More than 12.5°C (22°F) Less than 0.145 CFM/W at sea level [Max loading condition only]	NA	NA
Thermal Throttling	Not allowed	Not allowed	Allowed

Minimum Thermal Margin	7% of component temperature spec	NA	NA
Fan Power Ratio	Below 5% of system total power [Max loading condition only]	NA	NA

9.3 Fan Configuration

Grand Teton uses sixteen (16) 80x56mm counter rotating fans. A shared fan wall is used for all trays, sharing the same PWM control algorithm. Single rotor failure is supported in corner cases.

9.4 Fan Control Target

Under normal operation conditions, minimum margin (critical component) shall be at least 7% of component temperature spec, while fan power consumption shall be less than 5% of system total power (account for max loading power only).

9.5 Thermal Sensors

The BMC monitors all applicable temperature, power, and fan speed sensors, and adopts selective temperature readings for fan speed control algorithms when applicable. The maximum allowable tolerance of the inlet temperature sensor is $\pm 2^{\circ}\text{C}$. The maximum allowable tolerance of other thermal sensors is $\pm 5^{\circ}\text{C}$.

9.6 Fan Control Algorithm

Look up tables are applied to the inlet temperature sensor, and temperature sensors of all components with thermal risk under all possible operation conditions, if the fans operate at the minimum duty specified (20%). PID control tables should be applied to temperature sensors of all components that may be thermally critical under all possible operation conditions.

9.7 Special Conditions

For devices/components/parts that do not have built-in temperature sensors or cannot be included in fan speed control, their thermal solution design should enable the inlet-temperature based FSC to ensure their cooling at idling/low loading conditions. The PID control FSCs should ensure their cooling at heavy load conditions.

Special fan modes shall be designed for the following scenarios:

- AC Cycle
- DC Power Off/Cycle
- Single rotor Failure

- Fan Board Failure
- Sensor failure or exceed LCR/UCR limits

10. Interfaces

10.1 Signal List

The system includes the following general signals:

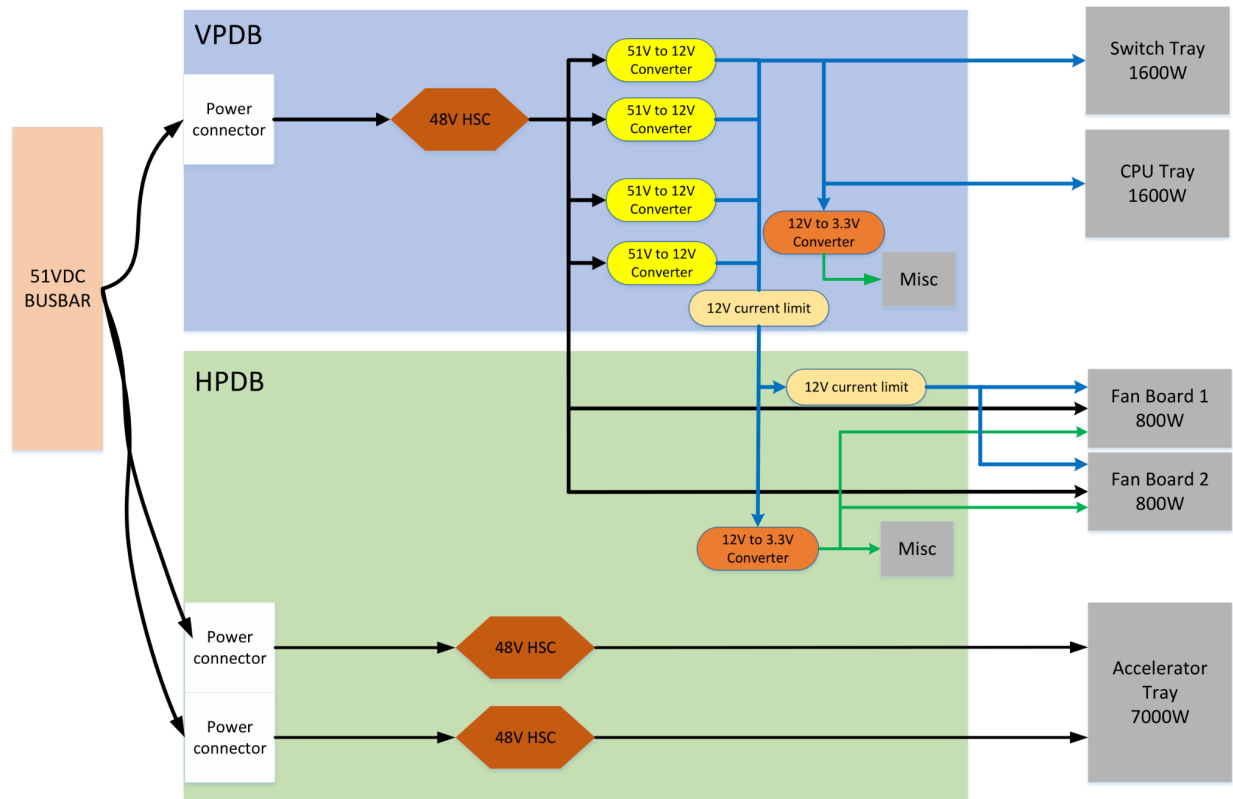
- High-speed signals
 - PCIe 5.0 between the CPUs, switches, NICs, SSDs, accelerators, and backplane connectors
 - Clocks
- Low-speed signals
 - I2C/I3C/SPI for management, sensors, I/O expanders, and other components
 - USB 3.0 for debug connections
 - UART
 - GPIO
 - JTAG

For more detail, see the Grand Teton CPU Tray and Switch Tray specifications.

10.2 Cable Backplane

The CPU, Switch, and Accelerator Trays slide into the chassis to blind mate into the ExaMax and power cable backplane at the rear interior of the chassis. These connections include 12VDC power from the power distribution boards and high-speed signals between all three trays. All ExaMax cables include a presence detect signal.

11. Onboard Power System



High-level block diagram of the Grand Teton power system

The Grand Teton system takes 51VDC power from the ORv3 bus bar and cables it to the power distribution boards. 51VDC bus bar voltage supplies the Accelerator Tray and fans. 51VDC is converted down to 12VDC to supply the CPU Tray and Switch Tray. All incoming rails are protected by hot swap controllers. Lower voltage regulators handle further down conversion on each board.

For more details about the power system, refer to the HPDB and VPDB sections.

12. System Firmware

Grand Teton uses the following open source firmware:

- OpenBMC, an open-source firmware stack for the BMC on the SCM:
<https://github.com/facebook/openbmc/tree/helium/meta-facebook/meta-grandteton>

- OpenBIC, an open-source firmware for the BIC in the Switch Tray
<https://github.com/facebook/OpenBIC/tree/main/meta-facebook/gt-cc>

13. Hardware Management

13.1 Out-Of-Band Management

Grand Teton has a dedicated 400G NIC for out-of-band (OOB) management (NIC0). OOB access is always online as long as standby power is provided. The NIC attaches to the CPU Tray and is OCP NIC 3.0 compliant.

The BMC can update NIC firmware through OOB to ensure NIC update is available in case the OS is not accessible and in-band update fails.

13.2 BMC

This system uses an ASPEED AST2600 BMC with one 8Gb x16 DDR4 SDRAM and 65MB flash for various platform management services. The BMC interfaces with hardware, BIOS, and host firmware.

The BMC resides on the SCM board, which follows the DC-SCM v1.0 specification (with board pinout exceptions).

The BMC is a standalone system in parallel to the host. The health status of the host system does not affect the normal operation and network connectivity of BMC.

The BMC is remote upgradeable either by in-band across the network or out-of-band through the NIC.

13.3 Remote Upgradeability

The system is capable of remotely upgrading the VRs, CPLDs, BIC, switches, and FRU EEPROMs. The BMC can access each device, update code from BMC to the device, and verify the code.

The BMC is capable of recovering all CPLD and BIC in the system from image corruption or an accidental erase. The BMC does not rely on proper operation of the BIC or CPLD for booting or to be accessible remotely.

13.4 Sensors

BMC has access to all analog sensors directly or through host management. Below is a table of all sensors:

Open Compute Project • Grand Teton Platform Specification

Sensor Name	Sensor Name
GT_MB_E1S_P3V3_VOLT_V	GT_MB_CPU1_DIMM_B3_PWR_W
GT_MB_E1S_P12V_VOLT_V	GT_MB_CPU1_DIMM_D3_PWR_W
GT_MB_E1S_P12V_CURR_A	GT_MB_CPU1_DIMM_B4_PWR_W
GT_MB_E1S_P12V_PWR_W	GT_MB_CPU1_DIMM_D4_PWR_W
GT_MB_E1S_TEMP_C	GT_MB_CPU1_DIMM_B5_PWR_W
GT_MB_POWER_FAIL	GT_MB_CPU1_DIMM_D5_PWR_W
GT_MB_PROCESSOR_FAIL	GT_MB_CPU1_DIMM_B6_PWR_W
GT_MB_HSC_VOLT_V	GT_MB_CPU1_DIMM_D6_PWR_W
GT_MB_HSC_CURR_A	GT_MB_CPU1_DIMM_B7_PWR_W
GT_MB_HSC_PWR_W	GT_NIC0_P3V3_VOLT_V
GT_MB_HSC_TEMP_C	GT_NIC0_P12V_VOLT_V
GT_MB_HSC_PEAK_PIN_W	GT_NIC0_P12V_CURR_A
GT_MB_INLET_R_TEMP_C	GT_NIC0_P12V_PWR_W
GT_MB_INLET_L_TEMP_C	GT_NIC0_TEMP_C
GT_MB_OUTLET_R_TEMP_C	GT_NIC1_P3V3_VOLT_V
GT_MB_OUTLET_L_TEMP_C	GT_NIC1_P12_VOLT_V
GT_MB_CPU0_TEMP_C	GT_NIC1_P12_CURR_A
GT_MB_CPU1_TEMP_C	GT_NIC1_P12_PWR_W
GT_MB_CPU0_THERM_MARGIN	GT_NIC1_TEMP_C
GT_MB_CPU1_THERM_MARGIN	GT_PDBV_HSC0_VIN
GT_MB_CPU0_TJMAX	GT_PDBV_HSC0_CURR_A
GT_MB_CPU1_TJMAX	GT_PDBV_HSC0_PWR_W
GT_MB_CPU0_PKG_PWR_W	GT_PDBV_HSC0_TEMP_C
GT_MB_CPU1_PKG_PWR_W	GT_PDBV_HSC0_PEAK_PIN_W
GT_MB_CPU0_DIMM_GRP_A_TEMP_C	GT_PDBV_BRICK0_VOLT_V
GT_MB_CPU0_DIMM_GRP_B_TEMP_C	GT_PDBV_BRICK0_CURR_A
GT_MB_CPU0_DIMM_GRP_C_TEMP_C	GT_PDBV_BRICK0_TEMP_C
GT_MB_CPU0_DIMM_GRP_D_TEMP_C	GT_PDBV_BRICK0_PWR_W
GT_MB_CPU0_DIMM_GRP_E_TEMP_C	GT_PDBV_BRICK1_VOLT_V
GT_MB_CPU0_DIMM_GRP_F_TEMP_C	GT_PDBV_BRICK1_CURR_A
GT_MB_CPU0_DIMM_GRP_G_TEMP_C	GT_PDBV_BRICK1_TEMP_C
GT_MB_CPU0_DIMM_GRP_H_TEMP_C	GT_PDBV_BRICK1_PWR_W
GT_MB_CPU1_DIMM_GRP_A_TEMP_C	GT_PDBV_BRICK2_VOLT_V
GT_MB_CPU1_DIMM_GRP_B_TEMP_C	GT_PDBV_BRICK2_CURR_A
GT_MB_CPU1_DIMM_GRP_C_TEMP_C	GT_PDBV_BRICK2_TEMP_C

GT_MB_CPU1_DIMM_GRPD_TEMP_C	GT_PDBV_BRICK2_PWR_W
GT_MB_CPU1_DIMM_GRPE_TEMP_C	GT_PDBV_P3V3_AUX_IN6_VOLT_V
GT_MB_CPU1_DIMM_GRPF_TEMP_C	GT_PDBH_HSC1_VOLT_V
GT_MB_CPU1_DIMM_GRPV_TEMP_C	GT_PDBH_HSC1_CURR_A
GT_MB_CPU1_DIMM_GRPW_TEMP_C	GT_PDBH_HSC1_PWR_W
GT_MB_VR_CPU0_VCCIN_VOLT_V	GT_PDBH_HSC1_TEMP_C
GT_MB_VR_CPU0_VCCIN_TEMP_C	GT_PDBH_HSC1_PEAK_PIN_W
GT_MB_VR_CPU0_VCCIN_CURR_A	GT_PDBH_HSC2_VOLT_V
GT_MB_VR_CPU0_VCCIN_PWR_W	GT_PDBH_HSC2_CURR_A
GT_MB_VR_CPU0_VCCFA_FIVRA_VOLT_V	GT_PDBH_HSC2_PWR_W
GT_MB_VR_CPU0_VCCFA_FIVRA_TEMP_C	GT_PDBH_HSC2_TEMP_C
GT_MB_VR_CPU0_VCCFA_FIVRA_CURR_A	GT_PDBH_HSC2_PEAK_PIN_W
GT_MB_VR_CPU0_VCCFA_FIVRA_PWR_W	GT_BP0_FAN0_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCIN_FAON_VOLT_V	GT_BP0_FAN0_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCIN_FAON_TEMP_C	GT_BP0_FAN1_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCIN_FAON_CURR_A	GT_BP0_FAN1_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCIN_FAON_PWR_W	GT_BP0_FAN4_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCFA_VOLT_V	GT_BP0_FAN4_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCFA_TEMP_C	GT_BP0_FAN5_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCFA_CURR_A	GT_BP0_FAN5_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCFA_PWR_W	GT_BP0_FAN8_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCD_HV_VOLT_V	GT_BP0_FAN8_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCD_HV_TEMP_C	GT_BP0_FAN9_INLET_SPEED_RPM
GT_MB_VR_CPU0_VCCD_HV_CURR_A	GT_BP0_FAN9_OUTLET_SPEED_RPM
GT_MB_VR_CPU0_VCCD_HV_PWR_W	GT_BP0_FAN12_INLET_SPEED_RPM
GT_MB_P12V_AUX_IN0_VOLT_V	GT_BP0_FAN12_OUTLET_SPEED_RPM
GT_MB_P5V_IN3_VOLT_V	GT_BP0_FAN13_INLET_SPEED_RPM
GT_MB_P3V3_IN4_VOLT_V	GT_BP0_FAN13_OUTLET_SPEED_RPM
GT_MB_P3V3_AUX_IN5_VOLT_V	GT_BP1_FAN2_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_VOLT_V	GT_BP1_FAN2_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_TEMP_C	GT_BP1_FAN3_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_CURR_A	GT_BP1_FAN3_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_PWR_W	GT_BP1_FAN6_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_FIVRA_VOLT_V	GT_BP1_FAN6_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_FIVRA_TEMP_C	GT_BP1_FAN7_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_FIVRA_CURR_A	GT_BP1_FAN7_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_FIVRA_PWR_W	GT_BP1_FAN10_INLET_SPEED_RPM

Open Compute Project • Grand Teton Platform Specification

GT_MB_VR_CPU1_VCCIN_FAON_VOLT_V	GT_BP1_FAN10_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_FAON_TEMP_C	GT_BP1_FAN11_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_FAON_CURR_A	GT_BP1_FAN11_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCIN_FAON_PWR_W	GT_BP1_FAN14_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_VOLT_V	GT_BP1_FAN14_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_TEMP_C	GT_BP1_FAN15_INLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_CURR_A	GT_BP1_FAN15_OUTLET_SPEED_RPM
GT_MB_VR_CPU1_VCCFA_PWR_W	GT_SCM_P12V_V
GT_MB_VR_CPU1_VCCD_HV_VOLT_V	GT_SCM_P5V_V
GT_MB_VR_CPU1_VCCD_HV_TEMP_C	GT_SCM_P3V3_V
GT_MB_VR_CPU1_VCCD_HV_CURR_A	GT_SCM_P2V5_V
GT_MB_VR_CPU1_VCCD_HV_PWR_W	GT_SCM_P1V8_V
GT_MB_P3V_BAT_V	GT_SCM_PGPPA_V
GT_MB_PCH_TEMP_C	GT_SCM_P1V2_V
GT_MB_CPU0_DIMM_A0_PWR_W	GT_SCM_P1V0_V
GT_MB_CPU0_DIMM_C0_PWR_W	GT_SCM_P12V_VOLT_V
GT_MB_CPU0_DIMM_A1_PWR_W	GT_SCM_P12V_CURR_A
GT_MB_CPU0_DIMM_C1_PWR_W	GT_SCM_P12V_PWR_W
GT_MB_CPU0_DIMM_A2_PWR_W	GT_SCM_TEMP_C
GT_MB_CPU0_DIMM_C2_PWR_W	
GT_MB_CPU0_DIMM_A3_PWR_W	
GT_MB_CPU0_DIMM_C3_PWR_W	
GT_MB_CPU0_DIMM_A4_PWR_W	
GT_MB_CPU0_DIMM_C4_PWR_W	
GT_MB_CPU0_DIMM_A5_PWR_W	
GT_MB_CPU0_DIMM_C5_PWR_W	
GT_MB_CPU0_DIMM_A6_PWR_W	
GT_MB_CPU0_DIMM_C6_PWR_W	
GT_MB_CPU0_DIMM_A7_PWR_W	
GT_MB_CPU0_DIMM_C7_PWR_W	
GT_MB_CPU1_DIMM_B0_PWR_W	
GT_MB_CPU1_DIMM_D0_PWR_W	
GT_MB_CPU1_DIMM_B1_PWR_W	
GT_MB_CPU1_DIMM_D1_PWR_W	
GT_MB_CPU1_DIMM_B2_PWR_W	
GT_MB_CPU1_DIMM_D2_PWR_W	

14. Security

The Grand Teton platform includes the following security features:

- SCM with OpenBMC
 - Verified boot with signed firmware
- Intel-based CPU Tray
 - Converged Boot Guard and Intel Trusted Execution Technology
- AMD-based CPU Tray
 - Platform Secure Boot
- Switch Tray
 - PEX switch secure boot and signed firmware
- Accelerator Tray
 - Vendor-implemented, typically based on a root-of-trust

Appendix A - Checklist for IC approval of this Specification (to be completed by contributor(s) of this Spec)

Complete all the checklist items in the table with links to the section where it is described in this spec or an external document .

Item	Status	Link to detailed explanation
Has this contribution been presented to an OCP Project group during a project call or engineering workshop?	Yes or No	If “No”, please state the reason.
Approval by Project Leads	Yes or No	If “No”, please state the reason.
Is this contribution entered into the OCP Contribution Portal?	Yes or No	If “No”, please state the reason.
Was it approved in the OCP Contribution Portal?	Yes or No	If “No”, please state the reason.