CG-OpenRack-19
Sled and Rack Specification
Version 0.95

Authors:

Matt St Peter, Hardware Architect, Radisys Corporation

Nilan Naidoo, Software Architect, Radisys Corporation

1 Revision History

Date	Name	Description
4/25/2016	Radisys Corp	Version 0.50, Initial internal review copy
5/17/2016	Radisys Corp	Version 0.81, Internal copy with review edits
6/30/2016	Radisys Corp	Version 0.82, draft for OCP IC review. Updated name.
7/21/2016	Radisys Corp	Version 0.84, submission copy with external review edits
8/4/2016	Radisys Corp	Version 0.95, final review copy for external review

2 Contents

3 License

As of 8-10-2016, the following persons or entities have made this Specification available under the OWFa1.0, which is available at [To be replaced by permanent link after license has been decided].

Radisys Corporation

You can review the signed copies of the Contributor License for this Specification on the OCP website http://opencompute.org/, 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.

4 Scope

CG-OpenRack-19 is scalable carrier-grade rack level system that integrates high performance compute, storage and networking in a standard rack. As such, there are many possible system configurations. However, certain rules and requirements ensure that different modules will operate properly in compliant systems.

This document provides the detailed interoperability requirements of a CG-OpenRack-19 sled as well as the rack system. This document describes the physical makeup of the rack and sled with a focus on the interface between the sled and the supporting rack-level infrastructure. This includes mechanical requirements as well as requirements related to electrical power feeds, optical and electrical interfaces, airflow and cooling, and environmental robustness.

5 Overview

This section describes the physical components of a rack-level system and their internal and external connections. In general, the system is made up of compute- and storage-centric sleds, Ethernet switches, a power source, and supporting rack infrastructure as shown in Figure 1.

There are five major system component types: the system rack, power conversion/distribution, Top-of-Rack (ToR) switches, and two types of pluggable modules (sleds) which are differentiated by form factor (width).

- System Rack: The rack provides physical structure for the system, airflow control, security, and network interconnects.
- Power Source: Rack-mount power conversion and distribution infrastructure that provides a nominal 12VDC to two bus bar pairs mounted at the rear of the system. This specification only defines the 12VDC power delivery from the converters to the sleds.
- Half-Width Sled: A 2U enclosure that takes up half of the shelf width. A common use of this sled size is to provide high density compute resources.
- Full-Width Sled: A 2U enclosure that takes up the full shelf width. A common use of this sled size is to provide high capacity/throughput storage resources.
- Top-of-Rack (ToR) Switch: Each rack can be equipped with a variable number of management plane and data plane switches, also referred to as Top of Rack switches. Each of these aggregates management and/or data traffic to internal network switching planes; for example:
    - Device Management Switch: 1GbE downlink management traffic interfaces (e.g., to BMCs) and 10GbE uplink interfaces.
    - Application Management Switch: 1GbE downlink management traffic interfaces (e.g., to payload processors) and 10GbE uplink interfaces.
    - Primary Dataplane Switch: 10bE downlink data plane traffic interfaces (e.g., to payload processors) and 40bGE uplink interfaces.
    - Secondary Dataplane Switch: 10GbE downlink data plane traffic interfaces (e.g., to payload processors) and 40GbE uplink interfaces.

Although the switch configuration is not a fixed system requirement, it is recommended that a rack system include at least a Device Management Switch and Primary Dataplane Switch.

Figure 1. Rack with Half-Width and Full-Width Sleds

6 Sled Physical Specifications

The sleds are based on a uniform form factor that allows interoperability with the rack, and specifically the CG-OpenRack-19 shelves. They must have specific external dimensions, with external connections and features at specified locations in order to interoperate with the rack.

Sleds are positioned at a 2RU vertical pitch. The number of shelves within a rack is flexible, and can be defined by the required density or application. The vertical bus bars must be set to cover the sled area, or can extend beyond the sled area.

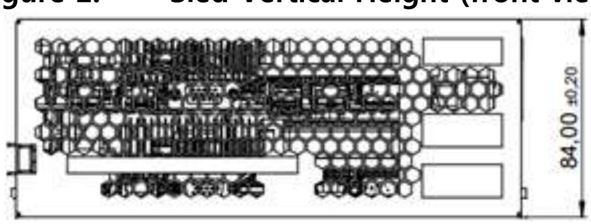
All dimensional tolerances are per DIN ISO 2768-mK, except sled height, width, and depth, which is +/-0.20mm.

6.1 Sled Common Mechanical Dimensions

All shelves **shall** occupy a vertical spacing of 2 Rack Units (RU), equivalent to 3.50 inches.

All sleds **shall** have a vertical height of 84.0mm ±0.20mm as measured from the bottom face to the top face.

Figure 2. Sled Vertical Height (front view)

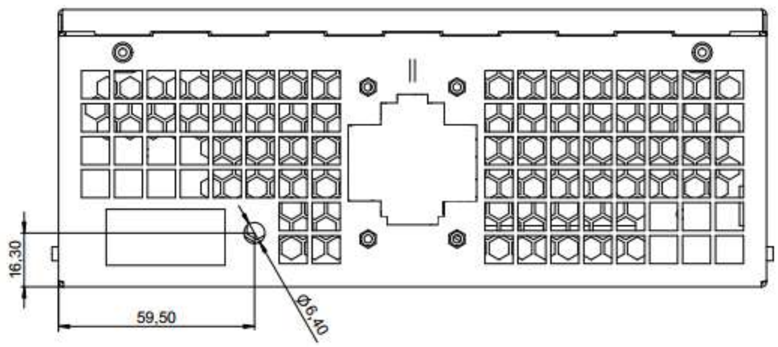


All sleds **shall** have a horizontal depth of 780.0mm ±0.20mm as measured from the front face to the rear face.

Figure 3. Sled Horizontal Depth (side view)

All sleds **shall** have a 6.4mm hole on the rear sled face to receive alignment pin. The hole **shall** be 16.3mm above the bottom face of the sled enclosure and 59.5mm from the right face of the sled enclosure.

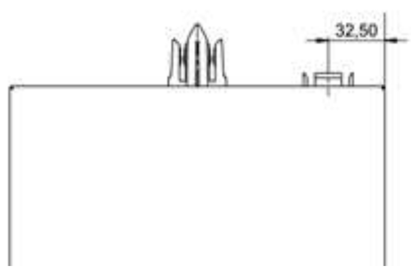
Figure 4. Sled Alignment Pin Hole (rear view)



6.2 Sled Optical Connector Placement

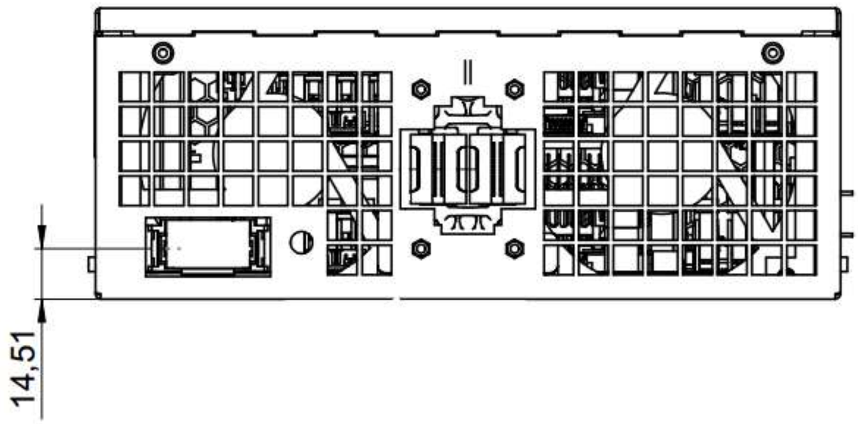
All sled optical connectors **shall** have a nominal centerline position 32.50mm from the right face of the sled enclosure.

Figure 5. Sled optical connector horizontal position (top-down view)



All sled optical connectors **shall** have a nominal centerline position 14.51mm from the bottom face of the sled enclosure.

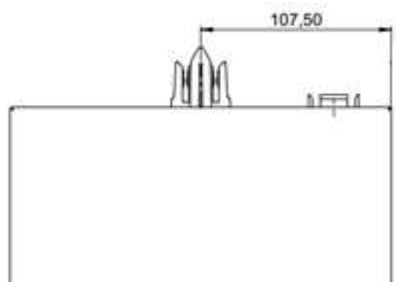
Figure 6. Sled optical connector vertical position (rear view)



6.3 Sled Power Connector Placement

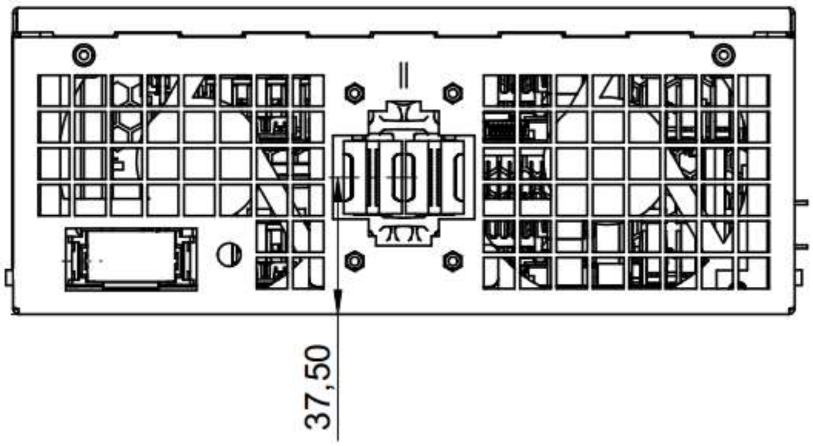
All sled DC power connectors **shall** have a nominal centerline position 107.50mm from the right face of the sled enclosure.

Figure 7. Sled Power connector horizontal position (top-down view)



All sled DC power connectors **shall** have a nominal centerline position 37.50mm from the bottom face of the sled enclosure.

Figure 8. Sled power connector vertical position (rear view)

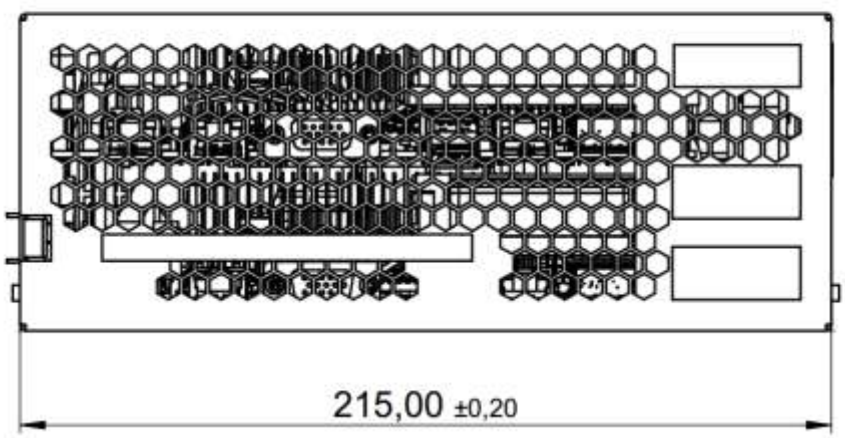


6.4 Half-Width Sled Mechanical Dimensions

The half-width sleds share a common height and depth with full-width sleds. They occupy half of a shelf (left or right side), and fit two-across.

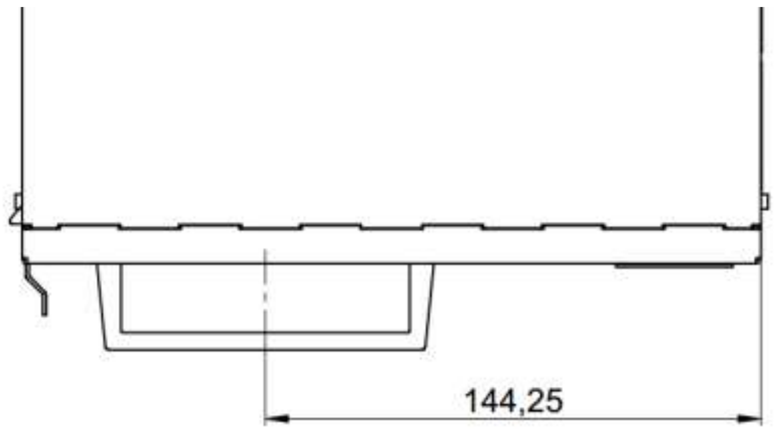
Half-width sleds **shall** have a horizontal width of 215.0mm ±0.20mm as measured from the left face to the right face.

Figure 9. Half-width sled horizontal width (front view)



Half-width sled front handles **shall** have a nominal centerline position 144.25mm from the right face of the sled enclosure.

Figure 10. Half-width sled handle horizontal position (top-down view)

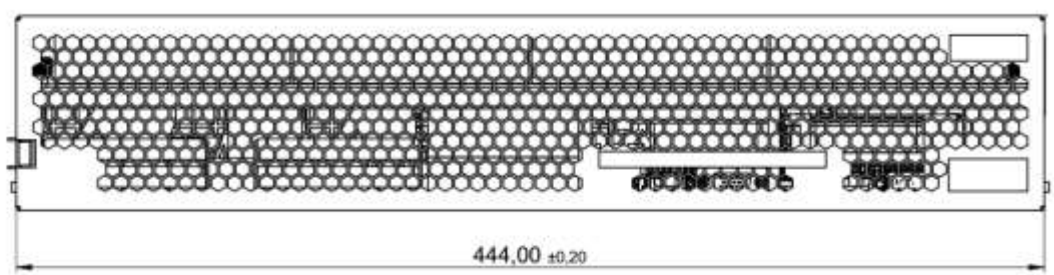


6.5 Full-Width Sled Mechanical Dimensions

The full-width sleds share a common height and depth with half-width sleds. They occupy the full width of a shelf.

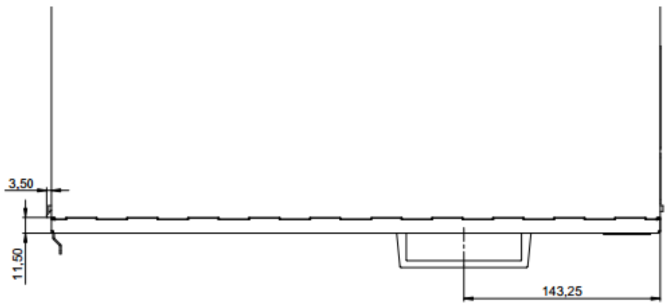
Full-width sleds **shall** have a horizontal width of 444.00mm ±0.20mm.

Figure 11. Full-width sled horizontal width (front view)



Full-width sled front handles **shall** have a nominal centerline position 143.25mm from the right face of the sled enclosure.

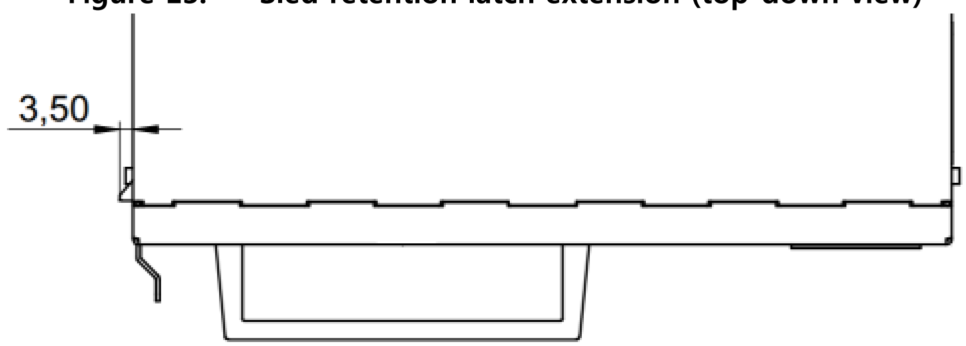
Figure 12. Full-width sled handle horizontal position (top-down view)



6.6 Mechanical Retention and Handle Features

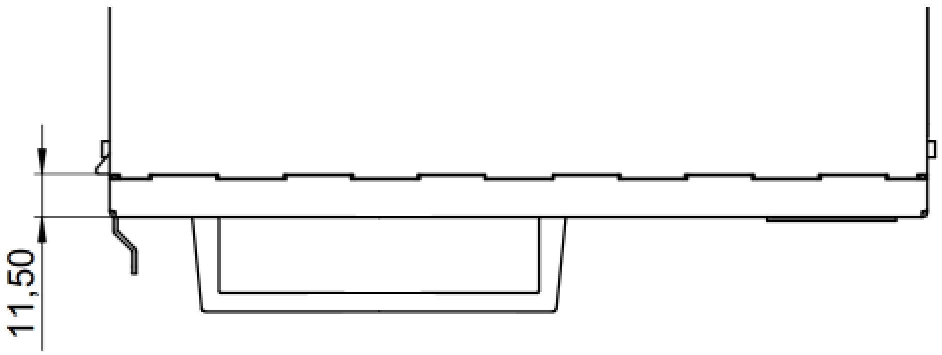
All sled retention latches **shall** extend a minimum of 3.50mm from the left face of the sled enclosure.

Figure 13. Sled retention latch extension (top-down view)



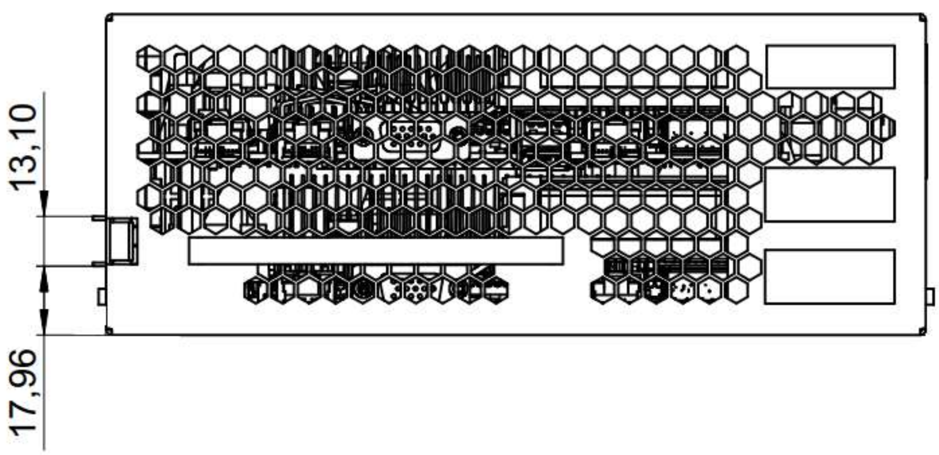
All sled retention latch forward faces **shall** be set back 11.50mm from the front face of the sled enclosure.

Figure 14. Sled retention latch face position (top-down view)



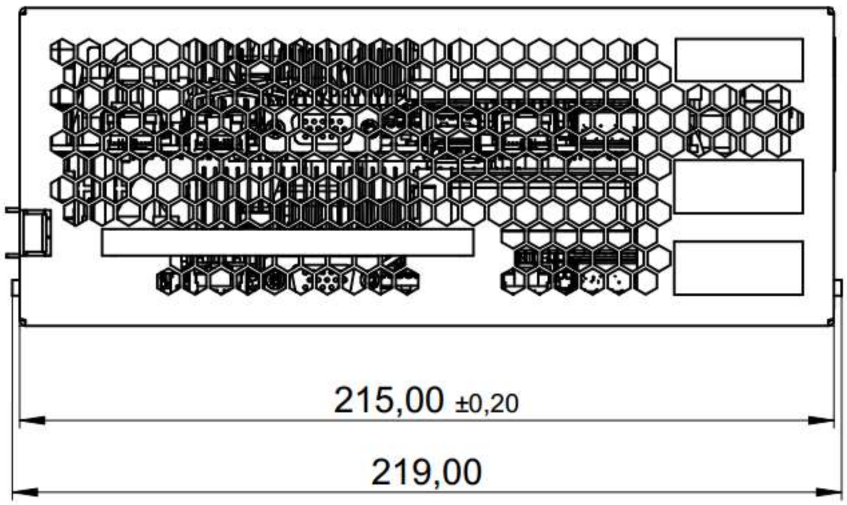
All sled retention latches **shall** have a height of 13.10mm and be 17.96mm above the bottom face of the sled enclosure.

Figure 15. Sled retention latch vertical position (front view)



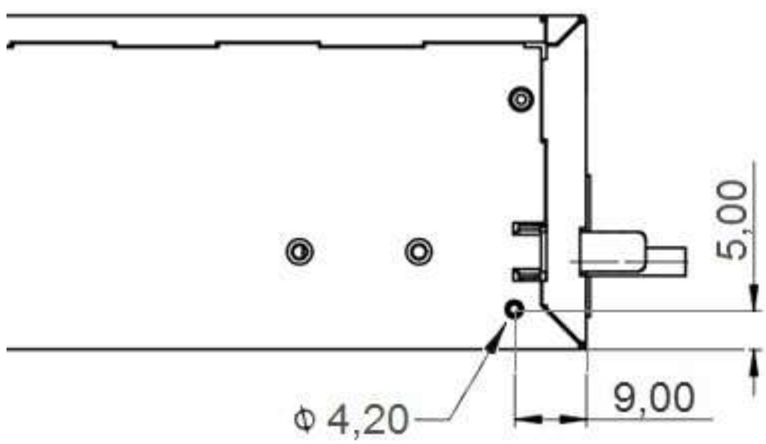
All sleds **shall** include retention pins extending 2.0mm +/-01.mm from the left and right enclosure faces.

Figure 16. Sled retention pin horizontal extension (front view)



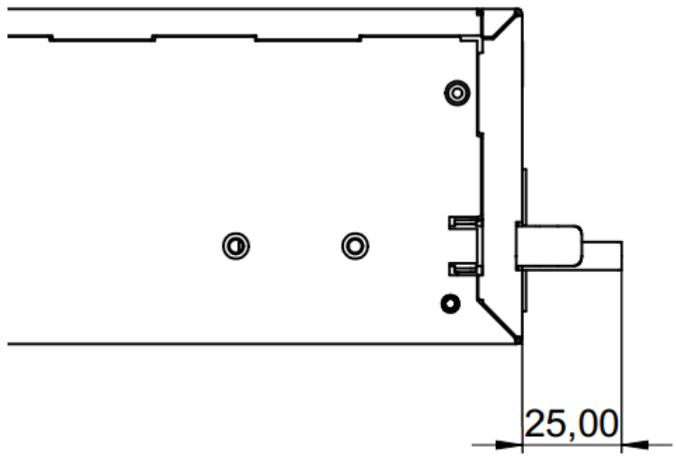
Sled retention pins **shall** be 4.20mm +/-0.1mm in diameter, and shall be 9mm from the enclosure front face and 5.00 mm from the enclosure bottom face.

Figure 17. Sled retention pin horizontal extension (front view)



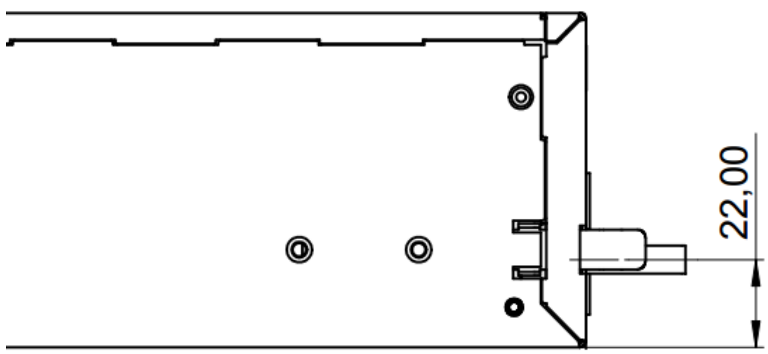
All sled front handles **should** extend 25.00mm from the front face of the enclosure.

Figure 18. Sled handle horizontal extension (side view)



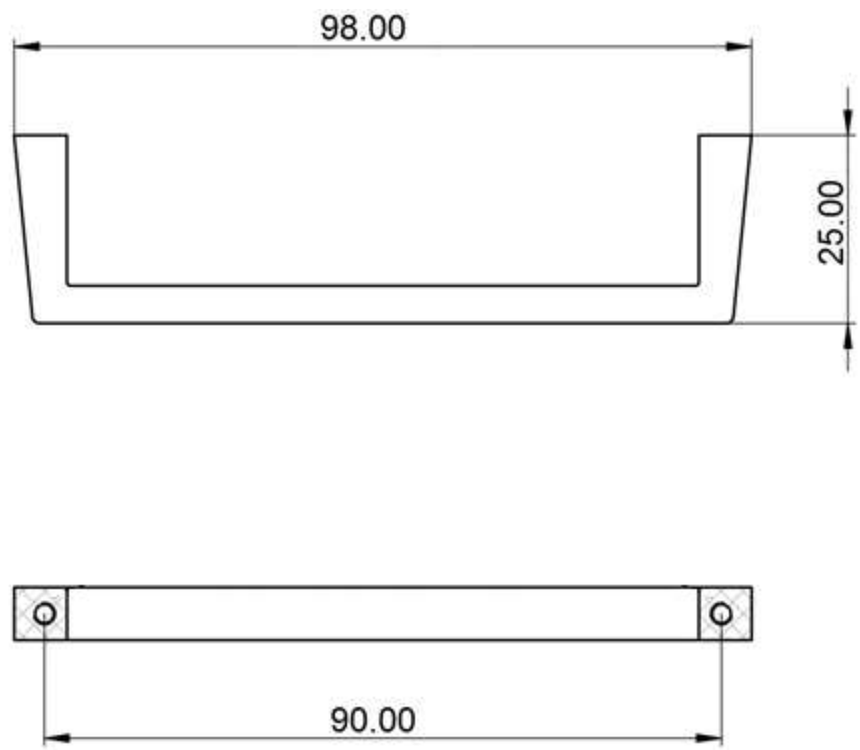
All sled front handles **should** be centered 22.00mm from the bottom face of the enclosure.

Figure 19. Sled handle vertical position (side view)



All sled front handles **should** comply with the following dimensions.

Figure 20. Sled handle dimensions

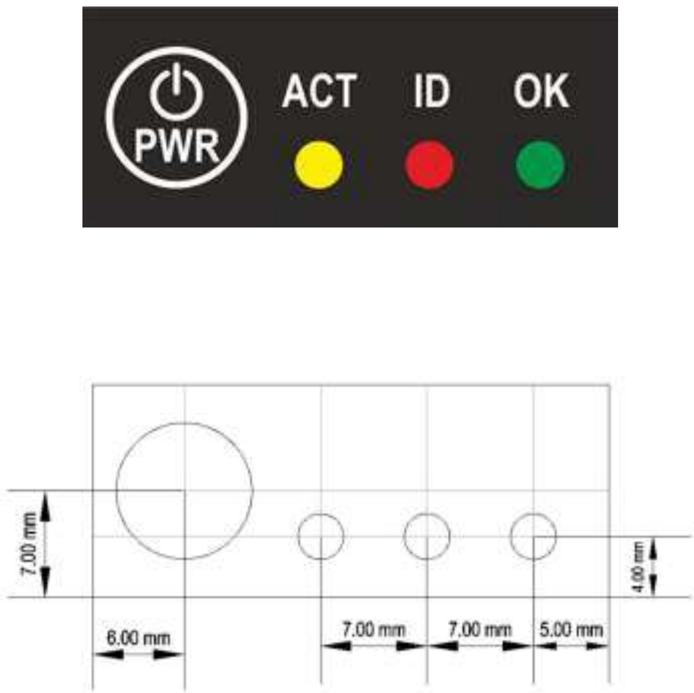


6.7 Sled Front Interface

The front of each sled provides status information and allows operator input by way of indicator(s) and button(s). In order to give the operator a consistent interface, the following dimensions and functions should be followed.

All sleds **should** include one operator panel for each separately managed server node. The dimensions each panel are 34mm width and 14mm height.

Figure 21. Sled Operator Panel Art and Dimensions



6.8 Weight and Structure

CG-OpenRack-19 sled weights may vary over a wide range depending on the composition of the enclosure and the type of internal components.

Table 1. Sled expected nominal and maximum weights

Sled Type	Nominal Weight	Expected Max. Weight
Half-Width	25lbs, 11.4kg	30lbs, 13.6kg
Full-Width	65lbs, 29.5kg	75lbs, 34.0kg

7 Rack Physical Specifications

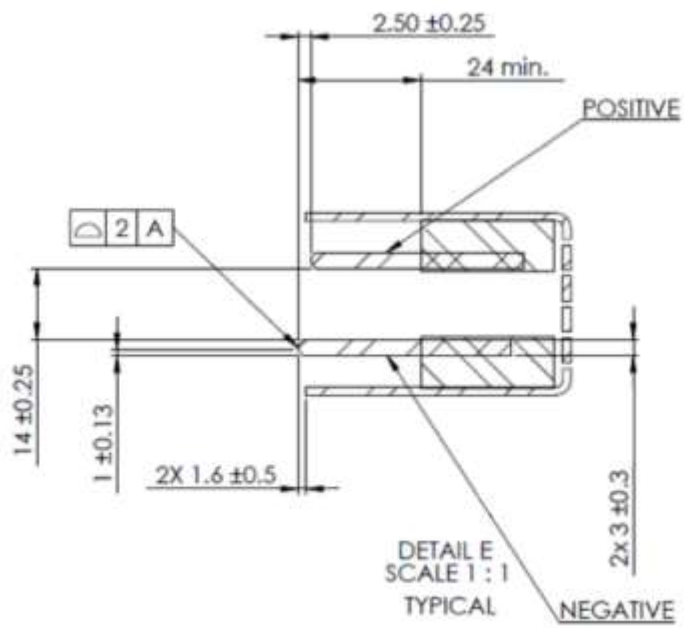
In order to achieve interoperability between sleds and the rack, certain dimensions of the rack, specifically the shelves, must be followed.

7.1 Bus Bar Pair Construction

The rear bus bar pairs provide 12VDC to the installed sleds.

The vertical bus bar pairs **shall** be constructed in compliance with the following dimensions.

Figure 22. Bus bar pair dimensions (top view)



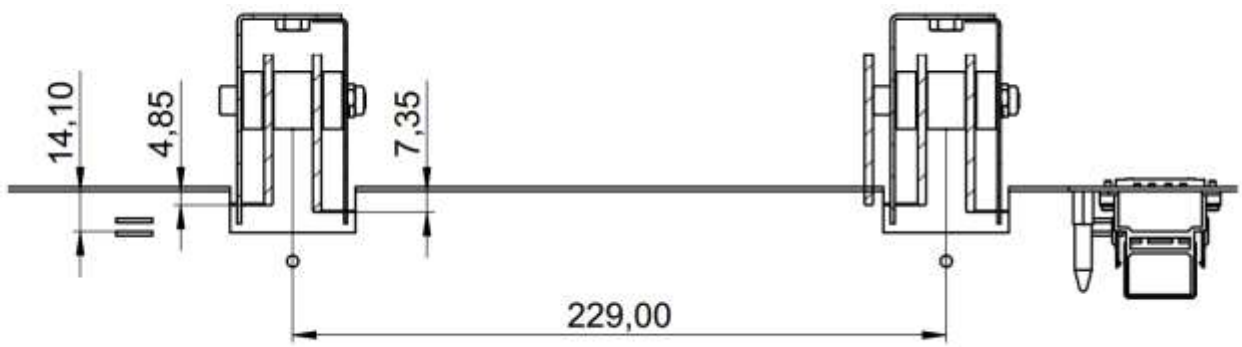
7.2 Bus Bar Location

The rack **shall** have two vertical bus bar pairs to the rear of all shelves.

Each right-hand (facing front) vertical bus bar pair **shall** be centered 108mm ±0.20mm from the right side of the sled bay wall. Each left-hand vertical bus bar pair

shall be centered 229mm ±0.20mm from the right-hand pair. The bus bar pairs **shall** be positioned as follows relative to the shelf back wall and sled stop.

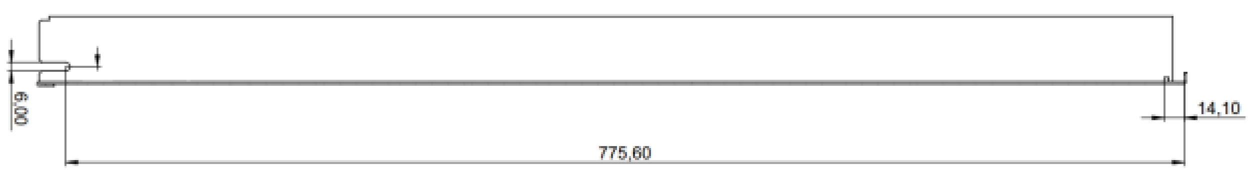
Figure 23. Bus bar pair horizontal location (top view)



7.3 Shelf Retention Feature

Each shelf **shall** have two retention slots at the front of each bay with the following dimensions.

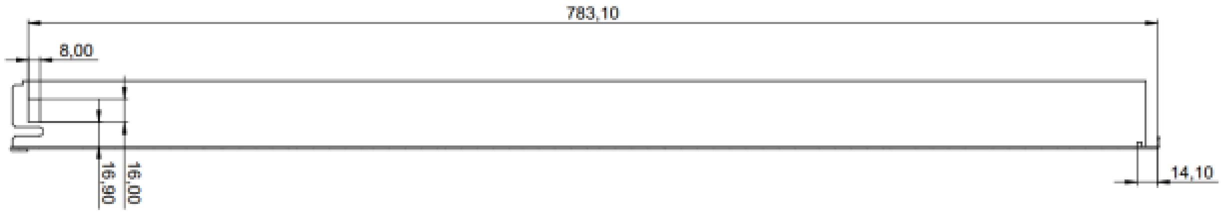
Figure 24. Shelf retention slot locations



7.4 Shelf Latch Feature

Each shelf **shall** have two latch openings at the front of each bay with the following dimensions.

Figure 25. Shelf latch opening locations



8 Sled Optical Data Port Requirements

All sleds have a rear-mounted blind-mate optical housing that supports data-plane and management plane connections to the spine switches. In order to interoperate with the connections made in the rack, the fiber positions in the rear connector must match the defined positions of the shelf connector.

All data and management interconnects between the sleds and the rack infrastructure pass through a single blind-mate connector housing set. Sleds have a rear-mounted blind-mate “daughtercard” optical housing that mate with a compatible “backplane” housing on the rear of the CG-OpenRack-19 shelves. In order to interoperate, both the shelf and the sleds must use the same ferrule and fiber position assignments.

The optical blind-mate connector housing supports four ferrules. Standard ferrules are available with 24 fiber positions (or more). The typical number of fiber positions in CG-OpenRack-19 is 12, and 8 fibers are loaded on each ferrule. The ferrule positions are color-coded, with each color corresponding to one of four network types.

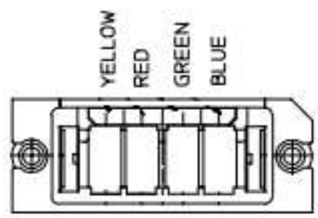
8.1 Sled Common Data Port Requirements

Each compute or storage shelf has two or one “backplane” housings (respectively) that will blind-mate with the “daughtercard” version of the housing that is part of the installed sled.

All interfaces between server nodes and the rack infrastructure pass through the blind-mate optical connector at the rear of the compute and storage sleds.

The housing has four sections, each of which capable of holding a single fiber optic ferrule. The sections are color coded

Figure 26. Sled blind-mate section assignments (rear view)



Sleds **should** use the green section of the rear optical connector for Device Management links.

Sleds **should** use the blue section of the rear optical connector for Application Management links.

Sleds **should** use the yellow section of the rear optical connector for Primary Data-plane links.

Sleds **should** use the red section of the rear optical connector for Secondary Data-plane links.

8.2 Sled to Shelf Interconnects

All interfaces between server nodes and the rack infrastructure pass through the blind-mate optical connector at the rear of the compute and storage sleds.

Table 2. Sled Interconnect Components

Assembly mounting	Connector description	Manufacturing part number
Sled	HBMT daughtercard housing	(Molex) 106105-2100
Sled	Power rail connector	(Methode) 5313-07458-00107

Table 3. Shelf Interconnect Components

Assembly mounting	Connector description	Manufacturing part number
Shelf	HBMT backplane housing	(Molex) 106105-2000

8.3 Interconnect Topology Example

To illustrate how the interconnect scheme defined in this specification may be used, this section provides an example of an interconnect topology.

In this example, each rack can be equipped with a variable number of management plane and data plane switches (also referred to as Top of Rack switches). Each of these aggregate management and data traffic to internal network switch functions such as:

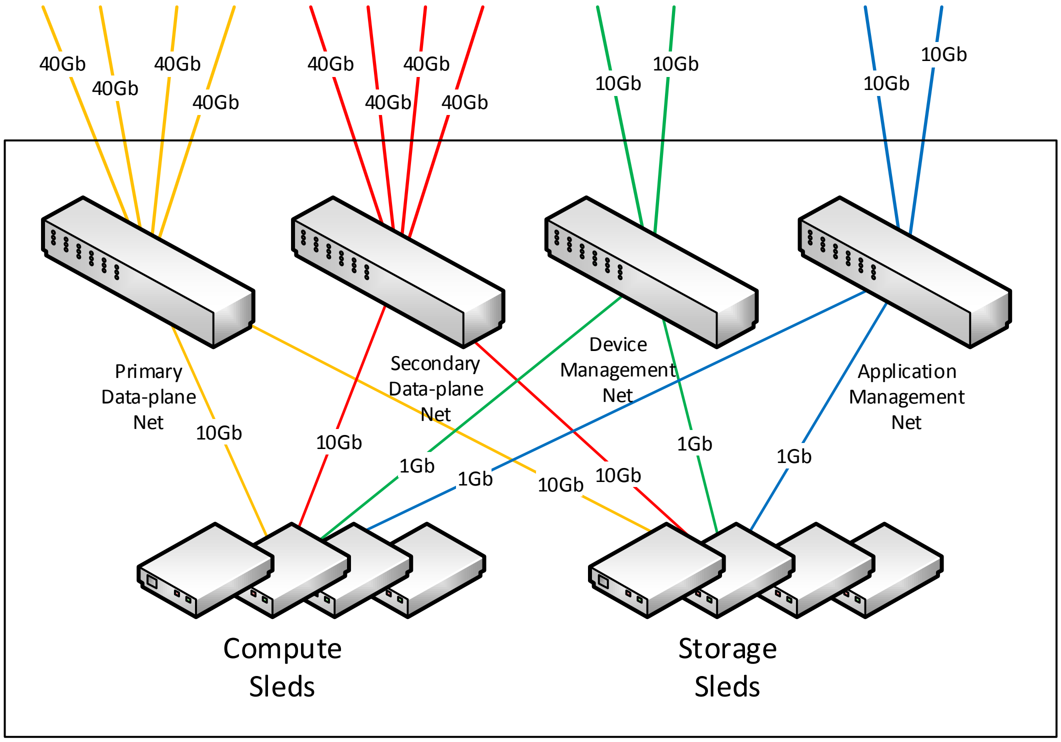
- Device management Switch: The 1GbE IPMI management ports (i.e. BMC port) of each of the rack component (i.e. servers, switches, power control, etc.) are connected to the downlink ports on the platform management switch. The uplink ports (10GbE) can be connected to a cluster or end-of-row (EoR) aggregation switch in the datacenter.

- Application management switch: All servers in the rack connect to this switch using a lower speed 1GbE port. This switch provides connectivity between the rack servers and external cluster or EoR switches to an application management network. The uplink ports (10GbE) connect to the application management spine switches.

- Primary Dataplane Switch: All servers in the rack connect to the downlinks of the primary dataplane switch using their first 10GbE port. The switch uplink ports (40GbE) provide external connectivity to the cluster aggregation switches.

- Secondary Dataplane Switch: All servers in the rack connect to the downlinks of the secondary dataplane switch using their second 10GbE port. This switch uplink ports (40GbE) provide external connectivity to the cluster aggregation switches.

Figure 27. Example Data Plane and Management Plane Architecture



9 Electrical Specifications

9.1 Power

The sections that follow provide information about the power specifications for the sled.

9.1.1 Power Connector

All sled power **shall** be supplied via a single DC power connector.

9.1.2 Input Voltage, Power, and Current

Table 4 lists the nominal, maximum, and minimum values for the sled input voltage. The maximum and minimum voltages include the effects of connector temperature, age, noise/ripple, and dynamic loading.

Table 4. Input Voltage Range

Nominal voltage	Maximum voltage	Minimum voltage
12.25V DC	12.95V DC	11.65V DC

The maximum amount of power allowed per sled is defined during system power allocation. The number of sleds in a chassis might be limited by the capacity of the AC power cord or by the cooling capacity of the deployment. Table 5 lists the input power allocation for a low-power sled.

Table 5. Input Power Allocation and Maximum Current

Connector	Nominal voltage	System power allocation (in W)	Maximum current (in A)
Methode Mini-MCC Bus Connector	12.25VDC	1000W	105A

All sleds **shall** provide in-rush current control through the 12V bus rail; return-side in-rush control is not required. The in-rush current **shall** rise linearly from 0A to the full load current at a rate not to exceed 20A/ms. The ramp-up to full load **shall not** take more than 200ms).

9.1.3 Current Interrupt Protection

All sleds **shall** provide overcurrent protection at the point of power entry to the primary power distribution circuit.

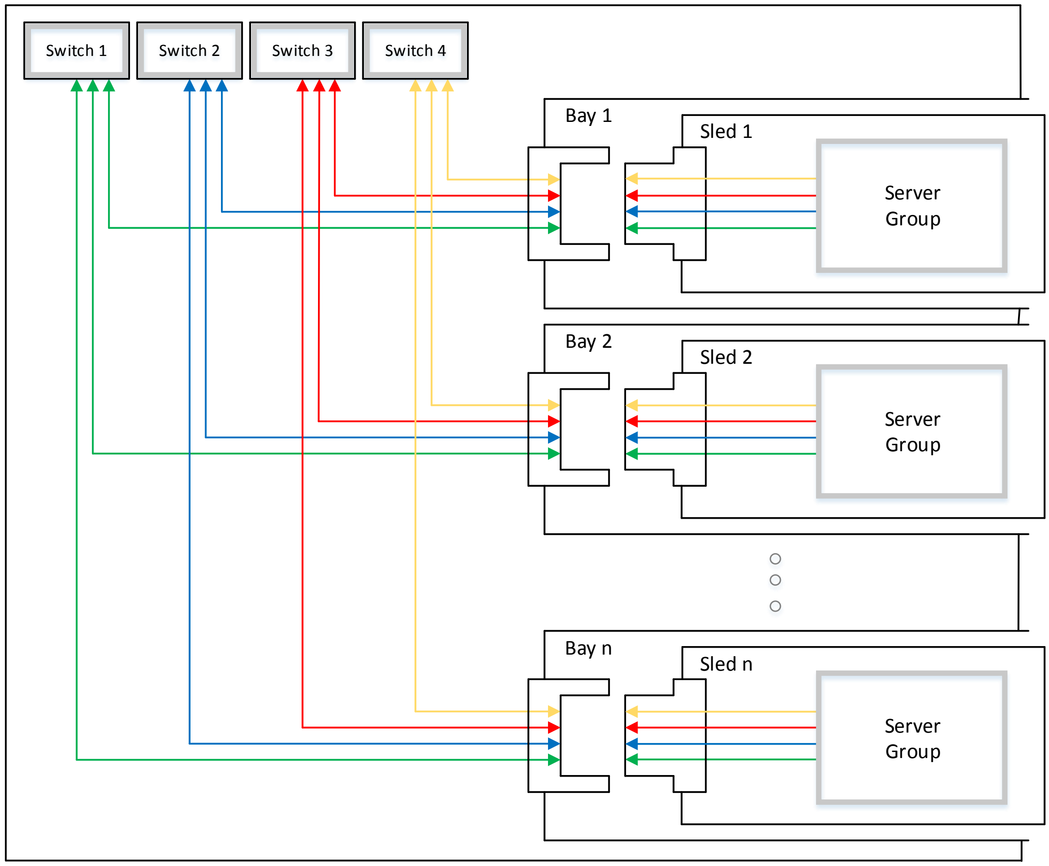
9.1.4 Grounding and Return

All sleds **shall** directly connect the DC supply return to the enclosure case as part of a chassis grounding method.

9.2 Signal Interface

All interfaces between server nodes and the rack infrastructure pass through the blind-mate optical connector at the rear of the compute and storage sleds.

Figure 28. Network connection to sleds

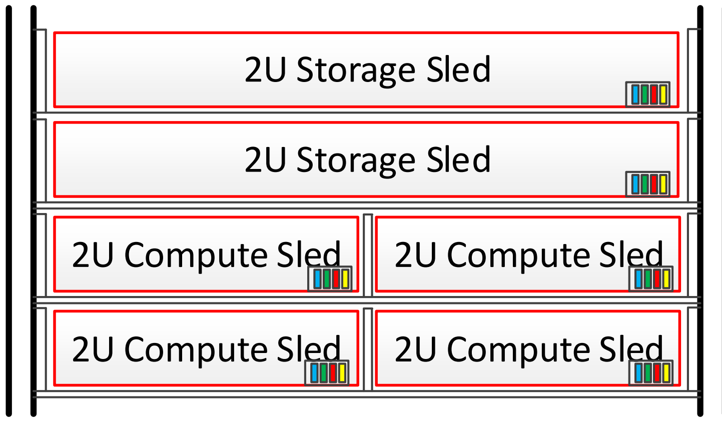


9.2.1 Sled Data Connections

Each sled connects to the rack-level networks via a set of up to four optical ferules in a blind-mate housing system. This connector system is positioned at the rear face of each sled and at the back of each shelf. When the sled is installed on the shelf, all of the ferules automatically align and connect, which ultimately connects the on-sled network interfaces to the rack-level network switches.

The optical blind-mate system and the in-rack optical cables provide the end-to-end network connections to all installed sleds.

Figure 29. Sled Blind-Mate Optical Connections, Front View



9.2.2 Signal Definitions

Table 6 below shows the fiber naming and assignments for the blind-mate optical interconnect.

Table 6. Optical Connector Asignments

Signal	Housing Section	Fiber Location	Description	Example Use
ENET_PMGMT1_TX	Green	1	1GbE (1000BASE-SX) lane	1Gb Primary Management Plane connection (BMC IPMI link)
ENET_PMGMT1_RX	Green	12		
ENET_PMGMT2_TX	Green	2	1GbE (1000BASE-SX) lane	1Gb Primary Management Plane connection (BMC IPMI link)
ENET_PMGMT2_RX	Green	11		
ENET_SMGMT1_TX	Blue	1	1GbE (1000BASE-SX) lane	1Gb Secondary Management Plane connection
ENET_SMGMT1_RX	Blue	12		
ENET_SMGMT2_TX	Blue	2	1GbE (1000BASE-SX) lane	1Gb Secondary Management Plane connection
ENET_SMGMT2_RX	Blue	11		
ENET_PDATA1_TX	Yellow	1	10GbE (10GBASE-SR) lane	10Gb Primary Data Plane connection
ENET_PDATA1_RX	Yellow	12		
ENET_PDATA2_TX	Yellow	2	10GbE (10GBASE-SR) lane	

Signal	Housing Section	Fiber Location	Description	Example Use
ENET_PDATA2_RX	Yellow	11		10Gb Primary Data Plane connection
ENET_SDATA1_TX	Red	1	10GbE (10GBASE-SR) lane	10Gb Secondary Data Plane connection
ENET_SDATA1_RX	Red	12		
ENET_SDATA2_TX	Red	2	10GbE (10GBASE-SR) lane	10Gb Secondary Data Plane connection
ENET_SDATA2_RX	Red	11		

9.3 Electromagnetic Interference Mitigation

The emission level of CG-OpenRack-19 sleds will be tested as part of a complete system. Because of this, sleds must be tested as part of a fully loaded rack-level configuration, including power distribution, power conversion, switching networks, and a maximum population of sleds.

Sleds **shall** meet FCC Class A when tested as a maximum population in a fully loaded rack-level system operating with a representative computational and data transmission workload.

10 Management

Each CG-OpenRack-19 server **shall** have a dedicated BMC for various out-of-band platform management services.

The health status of host processor **should not** affect the normal operation and network connectivity of the BMC.

The BMC **shall not** share memory with host processor.

The BMC management connectivity **should** work independently from host, and have no NIC driver dependency for Out-of-Band communication if using a shared-NIC.

The BMC **shall** support a remote console, also known as Serial-Over-LAN (SOL) through the management network.

POST codes **should** be displayed in SOL console during system POST.

Before the system has the first screen, POST codes **should** be stored and displayed in the SOL console in sequence. For example, display as "[00] [01] [02] [E0]..." etc. After the system has the first screen in the SOL console, the last POST code received on port 80 is displayed in the lower right corner of the SOL console screen.

The BMC **should** support high-performance remote KVM and remote storage over the management interface.

The BMC **shall** be fully IPMI 2.0 and DCMI 1.5 Compliant.

BMC firmware **shall** support remote system power on/off/cycle and warm reboot through In-Band or Out-of-Band IPMI commands

BMC firmware **shall** support a power-on policy selectable to be last-state, always-on, and always-off. The default setting is last-state.

A change of power policy **should** be supported by IPMI command and take effect without BMC a firmware cold reset or a system reboot.

Vendors **should** provide tool(s) to implement remote BMC firmware update, which will not require any physical access to the system. Remote update means either through Out-of-Band by management network or through In-Band by logging into local OS with data network.

11 Appendix

11.1 Appendix: Commonly Used Acronyms

This section provides definitions of acronyms used in the system specifications.

ACPI – advanced configuration and power interface

AHCI – advanced host controller interface

ANSI – American National Standards Institute

API – application programming interface

BIOS – basic input/output system

BMC – baseboard management controller

CFM – cubic feet per minute (measure of volume flow rate)

CG – Carrier Grade

CM – Chassis Manager

COLO – co-location

DCMI – Data Center Manageability Interface

DDR3 – double data rate type 3

DHCP – dynamic host configuration protocol

DIMM – dual inline memory module

DPC - DIMMs per memory channel

DRAM – dynamic random access memory

ECC – error-correcting code

EEPROM - electrically erasable programmable read-only memory

EIA – Electronic Industries Alliance

EMC – electromagnetic compatibility

EMI – electromagnetic interference

EOR – end of row

FRU – field replaceable unit

GUID – globally unique identifier

HCK – Windows Hardware Certification Kit

HMD – hardware monitoring device

I^2C – inter-integrated circuit

IBC – international building code

IDE – integrated development environment

IEC – International Electrotechnical Commission

IOC – I/O controller

IPMI – intelligent platform management interface

IPsec – IP security

ITPAC – IT pre-assembled components

JBOD – "just a bunch of disks"

KCS – keyboard controller style

L2 – layer 2

LAN – local area network

LFF – large form factor

LUN – logical unit number

MAC – media access control

MDC – modular data center containers

MLC – multi-level call

MTBF – mean time between failures

MUX – multiplexer

NIC – network interface card

OTS – off the shelf

OU – Open Compute Rack Unit (48mm)

PCB – printed circuit board

PCIe – peripheral component interconnect express

PCH – platform control hub

PDB – power distribution backplane

PDU – power distribution unit

Ph-ph – phase to phase

Ph-N – phase to neutral

POST – power-on self-test

PSU – power supply unit

PWM – pulse-width modulation

PXE – preboot execution environment

QDR – quad data rate

QPI – Intel QuickPath Interconnect

QSFP – Quad small form-factor pluggable

RAID – redundant array of independent disks

REST - representational state transfer

RM – Rack Manager

RMA – remote management agent

ROC – RAID-on-chip controller

RSS – receive-side scaling

RU – rack unit (1.75”)

SAS – serial-attached small computer system interface (SCSI)

SATA – serial AT attachment

SDA – serial data signal

SDR – sensor data record

SFF – small form factor

SFP - small form-factor pluggable

SMBUS – systems management bus

SMBIOS – systems management BIOS

SOL – serial over LAN

SPI – serial peripheral interface

SSD – solid-state drive

TDP – thermal design power

TOR – top of rack

TPM – trusted platform module

U – rack unit

UART – universal asynchronous receiver/transmitter

UEFI – unified extensible firmware interface

UL – Underwriters Laboratories

UPS – uninterrupted power supply