Open CloudServer blade specification V1.0

Authors:

Mark Shaw, Director of Hardware Engineering, Microsoft

Martin Goldstein, Principal Systems Architect, Microsoft

1 Revision History

Date	Name	Description
1/28/2014		Version 1.0

2 Scope

This document provides the technical specifications for the design of the Open CloudServer blade.

3 Contents

1 Revision History 2
2 Scope 4
3 Contents 4
4 Overview........ 6
5 Software and Firmware Specification 6
5.1 Operating System Version Requirements 7
5.2 Certification Requirements 7
5.3 Device Driver Requirements 7
5.4 Firmware Requirements........ 8
5.5 BIOS and UEFI Requirements 9
5.6 BIOS Debug........ 11
5.7 BIOS Recovery 11
5.8 Event Log Requirements 12
5.9 Network Systems Requirements........ 13
5.10 Storage Controller 13
6 Blade Physical Specifications........ 14
6.1 Volumetric Specifications........ 14
6.2 Weight and Structure 16
6.3 Guiding and Latching Features........ 16
6.4 Electromagnetic Interference Mitigation........ 20
7 Blade Interconnects........ 20
7.1 Power 22
7.2 Signal Interface........ 26
8 Mezzanine I/O 33
9 PCI-Express Expansion Slot........ 33
10 Blade Management 37

10.1 Overview of Blade Management 37
10.2 Blade LEDs 39
11 Appendix: Commonly Used Acronyms 42

4 Overview

The Open CloudServer blades are highly configurable, and are usually compute blades or storage "just a bunch of disks" (JBOD) blades. Figure 1 shows an example of a compute blade.

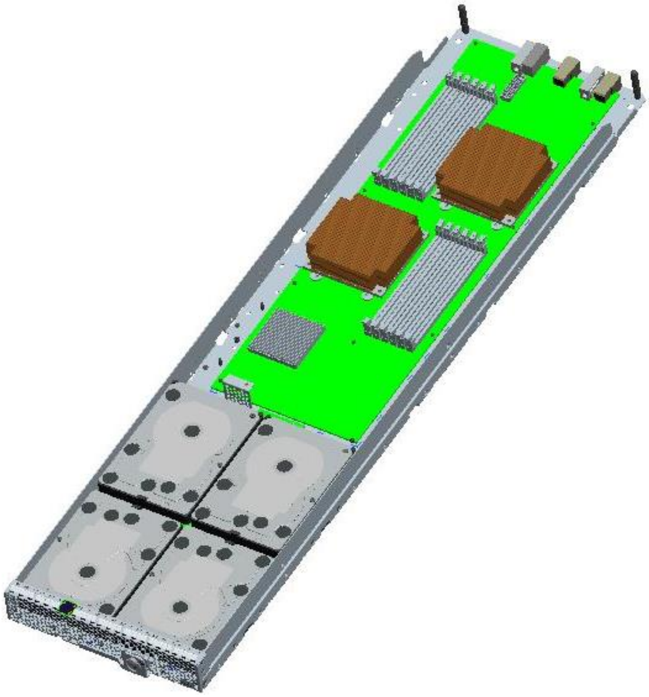


Figure 1. Compute blade

Power, management, and networking signals are delivered to the blade through blind-mate connectors to the tray backplane (TB). Networking and storage signals are routed to attachments on the TB at the rear of the chassis, enabling service without touching cables.

5 Software and Firmware Specification

The following sections define the software/firmware requirements a server must meet to be deployable in cloud-scale environments. These requirements are targeted for supplier partners.

5.1 Operating System Version Requirements

Any reference to an operating system in this specification refers to the Windows Server operating system. Unless otherwise specified, "Windows" refers to the Windows Server 2012 operating system (or later versions). The term "driver" refers to the Windows Server device driver. Note that the only supported version of Windows Server is x64.

Windows Preinstallation Environment (Windows PE) is not used for production; however, it is used for operating system deployment (imaging), for recovery, and for some specific maintenance/manageability scenarios (for example, for issuing secure erase on hard-disk drives [HDDs] and solid-state drives [SSDs]). For this reason, Windows PE 4.0 (64-bit versions) should boot on the server, and all devices should be accessible after Windows PE boots.

5.2 Certification Requirements

The server and any associated devices must pass the Windows Hardware Certification program tests. The most recent Windows Hardware Certification Kit (HCK) should be used unless otherwise specified.

If the server fails to pass certain tests because of required hardware design tradeoffs, failures must be documented and approved by Open Compute Certification and Integration team.

5.3 Device Driver Requirements

Table 1 lists the requirements that all device and device-driver updates should meet.

Table 1. Device Driver Requirements

Device driver and driver update requirements
• Support x64 versions of Windows Server 2012 (and later) and Windows PE 4.0 (or later)
• Pass Windows Hardware Certification (must be signed)
• Support silent option to allow for unattended (automated) installation
• Have consistent return codes across all operating system versions
• Include symbols (.pdb files) for debugging purposes
• Driver update package should be accompanied by a detailed list of all bugs fixed

Device driver and driver update requirements
• Driver updates should meet and pass all performance and functional tests
• Driver upgrade tools should run on x64 versions of Windows Server 2012 (or later), and Windows PE 4.0 (or later)
• Driver upgrade tools should be scriptable and support silent install mode (to support automated install/upgrade mode)
• Driver upgrade tools should include version checking to prevent older-version drivers from being installed in place of newer drivers
• Driver upgrade tools should support automation from Perl or Visual Basic Scripting Edition (VBScript) for automated upgrade and install operations
• Install/upgrade tools should provide detailed logging information to help identify install/upgrade failures
• Include dedicated path/secure location for driver packages (can be an file transfer protocol [FTP] or website); location should be refreshed regularly during every release period
• Driver updates should be shipped a regular and predictable release plan
• Driver update tool should use a fixed path for the error log
• Driver updates should maintain the plug-and play (PNP) ID numbers

5.4 Firmware Requirements

Table 2 lists the requirements that all firmware and firmware updates, including basic input/output system (BIOS) and unified extensible firmware interface (UEFI), should meet.

Table 2. Firmware Requirements

Firmware and firmware update requirements
• Support x64 versions of Windows Server 2012 (and later) and Windows PE 4.0 (and later)
• Support silent option to allow for unattended (automated) install
• Have consistent return codes across all operating system versions
• Firmware update package should be accompanied by a detailed list of all bugs fixed
• Meet and pass all related performance and functional tests
• Firmware upgrade tools should run on x64 versions of Windows Server 2012 (and later) and Windows PE 4.0 (and later)
• Firmware upgrade tools should have version checking to prevent older-version drivers from being installed in place of newer drivers

Firmware and firmware update requirements
• Firmware upgrade tools should be scriptable and support silent install mode to support automated install/upgrade mode
• Firmware upgrade tools should support automation from Perl or VBScript for automated upgrade and install operations.
• Firmware upgrade tools should provide detailed logging information to help identify install/upgrade failures
• Include dedicated path/secure location for firmware packages (can be FTP or website); location should be refreshed regularly during every release period
• Firmware updates should be shipped a regular and predictable release plan
• Firmware update tool should use a fixed path for the error log
• Firmware updates should maintain the plug and play (PNP) ID numbers

5.5 BIOS and UEFI Requirements

The BIOS and UEFI are provided by the supplier. Table 3 lists the requirements that, in addition to those listed in Table 2, the BIOS and UEFI should meet.

Table 3. BIOS and UEFI Requirements

BIOS and UEFI requirements
• BIOS/UEFI chip uses the PCH/ Wellsburg Platform Controller Hub (PCH) (Intel/AMD) serial peripheral interface (SPI) • 16MB or higher is required • Supplier is responsible for selecting the BIOS chip that meets the required functionality
• BIOS/UEFI source code is based on AMI Aptio 5.x reference BIOS for Intel Grantley CPU and Wellsburg PCH for the Mayan City reference motherboard I • Supplier is responsible for porting and maintaining the BIOS to make sure it has the latest version from AMI, AMD, or Intel
• BIOS should have a setting to boot in native UEFI or UEFI/BIOS compatibility mode
• UEFI/BIOS compatibility mode should be the default option
• The BIOS should be tuned to minimize power consumption: ○ Disable unused devices such peripheral component interconnect express (PCIe) lanes, USB ports, and serial AT attachment (SATA)/ serial-attached SCSI (SAS) ports ○ Tune CPU/chipset settings to read the best power/performance mix ○ A modified SPECpower that takes disk I/O into account should be used as the benchmark for testing/validating the optimal power/performance settings

BIOS and UEFI requirements
• BIOS/UEFI should expose the following settings: o Memory speed, Intel QuickPath Interconnect (QPI)/hyperthreading (HT) link speed, Speed-Step/Turbo mode, and CPU Cx power states o HT on/off o Enabling/disabling a socket or a core within a socket o Non-uniform memory access (NUMA) on/off switch o Advanced host controller interface (AHCI)/integrated development environment (IDE) mode switch o Implementation of a watchdog timer that is compatible with Windows Server (see http://msdn.microsoft.com/en-us/windows/hardware/gg463368.aspx) ▪ Initial state should be "enabled/stopped" and the initial timeout value should be 15 minutes ▪ When the motherboard has a baseboard management controller (BMC), then the intelligent platform management interface (IPMI) watchdog timer should be used o Default power setting should be "power on" after power failure o BIOS/UEFI should implement a random delay for power on after power failure ▪ The delay should be a random value ranging from 1 to 15 seconds ▪ BIOS will provide enable/disable switch for this feature ▪ If a BMC is available, then the delay should be implemented by the BMC; BIOS enable/disable switch will still control this BMC feature o Console redirection should be set to PCH/SP5100 virtual COM with baud rate 115200, no flow control, and terminal type VT100 o Setting for error-correcting code (ECC) error threshold; available setting should be 1, 4, 10, and 1000
• If a complementary metal–oxide–semiconductor (CMOS) checksum errors occurs (during BIOS/UEFI update, for example), the BIOS/UEFI should load the system defaults automatically.
• BIOS/UEFI settings and configuration tool should permit saving and updating the BIOS/UEFI settings to an XML or text configuration file
• BIOS/UEFI should implement a BIOIS/UEFI recovery mechanism that can be used when a BIOS/UEFI upgrades fails • A recovery mechanism that used a USB storage device to recover the BIOS is highly recommended
• BIOS should support Windows Hardware Error Architecture (WHEA) and complete WHEA memory structures.

5.6 BIOS Debug

The blade contains a serial pin header on the main board for local serial console redirect. This functionality is in addition to the baseboard management controller (BMC) serial console redirect functionality. Serial console redirect should be capable of providing VT100 MRC debug, BIOS console debug, and Microsoft Windows EMS for host debug.

5.7 BIOS Recovery

The blade contains a USB connector for BIOS image recovery. When the BIOS is forced into recovery mode, it will automatically look for a recovery image on FAT32-formatted USB media connected to the USB pins.

The recovery process can be initiated by setting the recovery jumper. The BIOS will detect the recovery jumper set and start to execute the recovery code.

The BIOS consists of three parts: the main BIOS section, the Non-volatile random-access memory (NVRAM) section, and the boot block recovery section. The main BIOS section is updated during recovery process, but the NVRAM and the boot block sections are preserved.

BIOS recovery is performed using the FAT32-formatted USB removable drive containing the BIOS image file (REC.ROM).

Following is the recovery flow:

1. Plug in the USB recovery media containing the file REC.ROM.
2. Short the recovery jumper.
3. Power on the system.
4. The BIOS will automatically boot into the recovery page in BIOS the set up menu.
5. Wait until recovery progress is complete.
6. Set the recovery jumper back.
7. Reboot the system with the new BIOS.

5.8 Event Log Requirements

A hardware event log that keeps track of various system events and errors is required. This log is used for managing and diagnosing servers. Table 4 lists event log requirements.

Table 4. Event Log Requirements

Event log requirements
• If a BMC is not available, the BIOS/UEFI is responsible for logging all events in a log compliant with systems management BIOS (SMBIOS) 2.6 • If there is no hardware management device (HMD), BIOS/UEFI has to implement a log that is compliant with SMBIOS specification Rev 2.6
• The log should hold at least 500 events and follow the SMBIOS event log organization format if an HMD is not present.
• Each event record includes enhanced information identifying the error source device's vendor ID and device ID
• In addition to being accessible through BIOS/UEFI, the supplier must provide a system access interface through the Windows Management Interface (WMI) • Driver should be compatible with 64-bit Windows Server 2012 (or a later version)
• BIOS/UEFI, as well as the WMI interface, should allow for reading as well clearing the log
• Logged errors should include: o CPU/memory errors: ▪ Both correctable and uncorrectable ECC errors should be logged in the event log ▪ Error categories include DRAM, Link, and L3 cache o QPI/HT link errors: ▪ Any errors that have a status register should be logged into the event log o PCIe errors: ▪ Any errors that have a status register should be logged into the event log, including root complex, endpoint device, and any switch upstream/downstream ports if available ▪ Link disable on errors should also be logged ▪ Fatal, non-fatal, or correctable classification follows the chipset vendor's recommendation o Power-on self-test (POST) errors: ▪ All POST errors detected by the BIOS during POST are logged into the event log o Power errors; two power errors are logged: ▪ 12V DC input power failure that causes all power rails on motherboard to lose power including standby power

Event log requirements
▪ Unexpected system shutdown during system S0/S1, while 12V DC input is still valid
• An error threshold setting must be enabled for both correctable errors • Once the programmed threshold is reached, an event should be triggered and logged ○ Memory correctable ECC: ▪ The default threshold value is 1000 ▪ When the threshold is reached, the BIOS should log the event including dual inline memory module (DIMM) location ○ QPI/HT Link errors: ▪ Follow the chipset vendor's suggestion ○ PCIe errors: ▪ Follow the chipset vendor's suggestion

5.9 Network Systems Requirements

Table 5 lists the systems requirements the network must meet, in addition to the requirements for device drivers listed in Table 1. Application specific requirements will be provided as needed during networking card development.

Table 5. Network Systems Requirements

Network systems requirements
• Any firmware or driver upgrade should keep all static IP configurations to prevent the server from going offline after the upgrade
• If multiple network interface cards (NICs) are present, they should PXE boot capable; by default, NIC 1 should be the network (PXE) NIC
• Any firmware or driver upgrade should maintain NIC team configuration if present

5.10 Storage Controller

If used, a storage controller must meet the requirements listed in Table 6 and those for firmware listed in Table 2.

Table 6. Storage Controller Requirements

Storage controller update requirements
• Storage controller settings and configuration tool should allow for saving the firmware settings to a TXT, INI, or XML configuration file • Redundant array of independent disks (RAID) set information should be included in the

output (such as RAID configuration, drives, and sizes)
• Storage controller settings and Configuration Tool should allow for updating the firmware settings from a XML or text configuration file
• Storage controller should provide a command line-interface to allow for an automated configuration
• Storage controller should support secure erase on individual disks

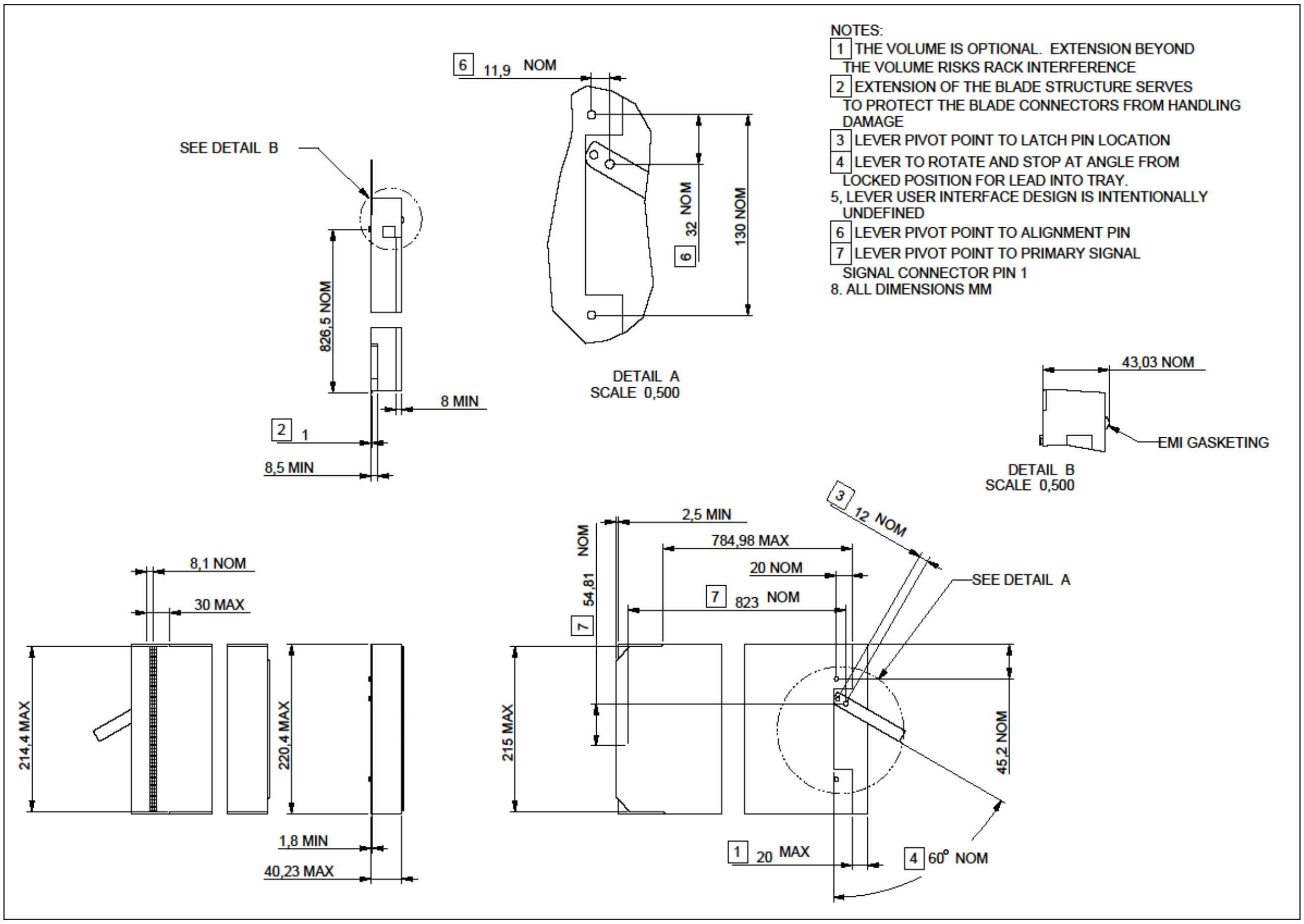
6 Blade Physical Specifications

The physical specifications for the blades include the dimensions of the space into which the blade is installed, the blade weight limits, a description of the guiding and latching features, and information about electromagnetic interference (EMI) shielding.

6.1 Volumetric Specifications

The outer surfaces of the blade must fit into the volume defined in Figure 2, with all tolerances and manufacturing methods accounted for. No features of the blade (including the latch and the handle) can extend beyond the specified distance from the front column of the rack so that the front door of the rack can be closed without interference.

This specification assumes a coplanar power/network distribution printed circuit board (PCB); if another structure is used, the connectors need to be form, fit, and functionally compatible with this specification. Note that if necessary, an EMI enclosure can be built to fit within the volumetric restrictions.



NOTES:
1 THE VOLUME IS OPTIONAL. EXTENSION BEYOND THE VOLUME RISKS RACK INTERFERENCE
2 EXTENSION OF THE BLADE STRUCTURE SERVES TO PROTECT THE BLADE CONNECTORS FROM HANDLING DAMAGE
3 LEVER PIVOT POINT TO LATCH PIN LOCATION
4 LEVER TO ROTATE AND STOP AT ANGLE FROM LOCKED POSITION FOR LEAD INTO TRAY.
5, LEVER USER INTERFACE DESIGN IS INTENTIONALLY UNDEFINED
6 LEVER PIVOT POINT TO ALIGNMENT PIN
7 LEVER PIVOT POINT TO PRIMARY SIGNAL SIGNAL CONNECTOR PIN 1
8. ALL DIMENSIONS MM

Figure 2. Dimensions of the volume that holds the blade

Figure 3 shows the dimensions of the blade-mounting envelope, which holds the blade on the tray.

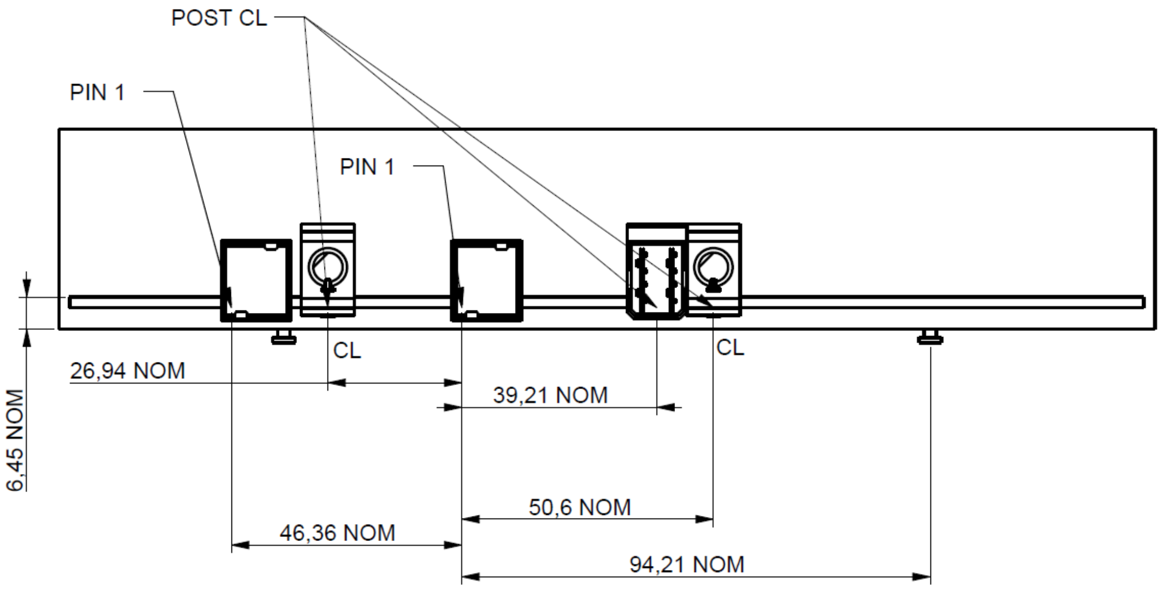


Figure 3. Blade-mounting envelope, front view

Note that the only serviceable components in the blade are the hard drives, which can be replaced without tools. All other repairs to the blade are made in a bench-top environment. Components such as heat sinks, hard drives, and motherboards can be attached so that removal requires a tool, if this option is less expensive.

6.2 Weight and Structure

Because the tray can support a total weight of 50 pounds, a standard blade should not exceed a static weight of 25 pounds when loaded with all options.

The blade must be mechanically stiff enough to support its own weight during shipping and handling. The side walls of the blade should be made as tall as possible within the given envelope to maximize the stiffness of the blade structure and allow for damage-free stacking during service.

6.3 Guiding and Latching Features

The following sections describe the guiding and latching features for the blade.

6.3.1 Latching Feature Identification

Latches, as well as thumb screws and other components used to lock, unlock, or remove a subassembly from the chassis, are colored blue (Pantone code 285 C blue) to make them easy to identify.

If the identity and function of a latching feature for a particular field replaceable unit (FRU) is clear, coloring might not be required. For example, a lock/unlock indicator is sufficient for a blade lever.

6.3.2 Guiding and Latching

Because a single tray design without a center rail is used for both half-width and full-width blades, guides and alignment features are located in both the front and back of the enclosure. These features are especially important when inserting a half-width blade into an empty tray.

The blade enclosure interfaces with the side walls of the tray and the guide pin in the tray to make sure the blade is aligned within the ±3.50mm (±0.138”) tolerance necessary to engage the connector. The guide pin in the tray aligns the half-width blades with the correct side of the tray. The blade connectors are protected from damage by the guide pin by the blade sheet metal.

Figure 4 shows the blade guiding and latching features associated with the tray. Blade guide pins slide into slots on the front of the tray for alignment. A latch attached to the blade fits into a notch in the tray to secure the blade in place.

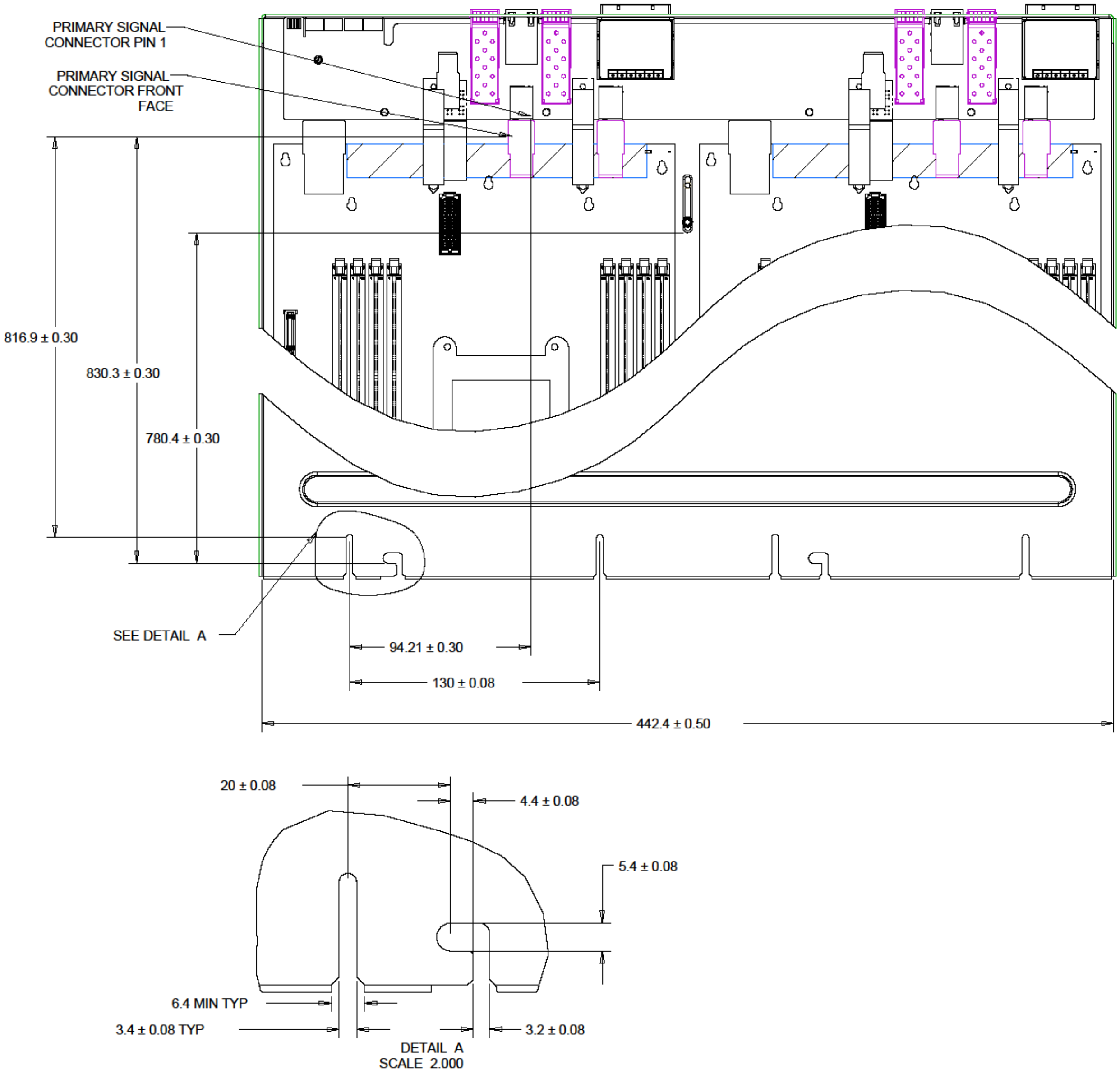


Figure 4. Guide and latch details, top view

Figure 5 and Figure 6 show examples of a front-blade guide and latch and of a rear-blade guide.

Note that a lever is included at the front of the blade to provide additional guidance when the blade is locking into the tray, and to help with the force required to install and remove the blade from the tray.

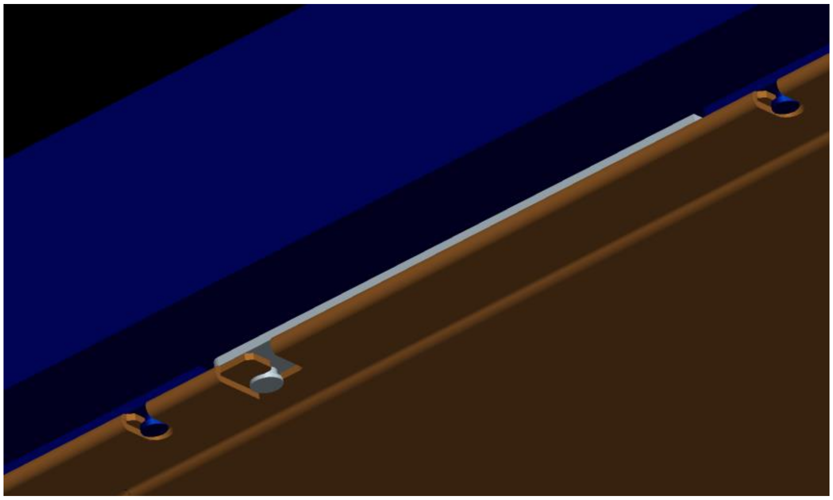


Figure 5. Example of a front-blade guide and latch

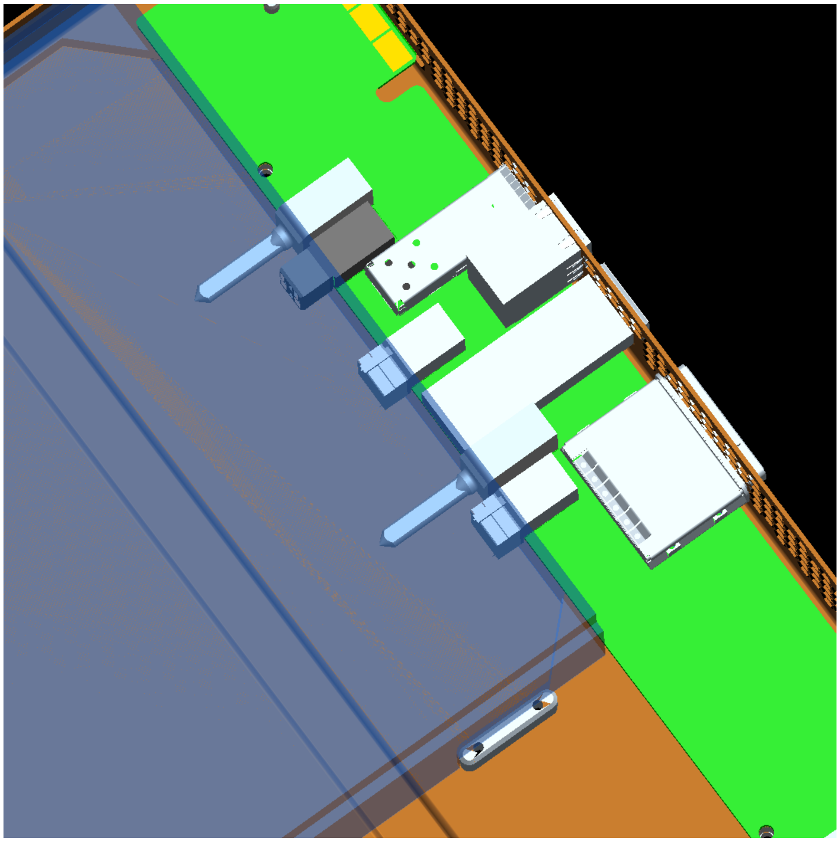


Figure 6. Example of a rear-blade guide pin

6.4 Electromagnetic Interference Mitigation

For EMI containment, an EMI shield is added to the top rear edge of the tray, as shown in Figure 7.

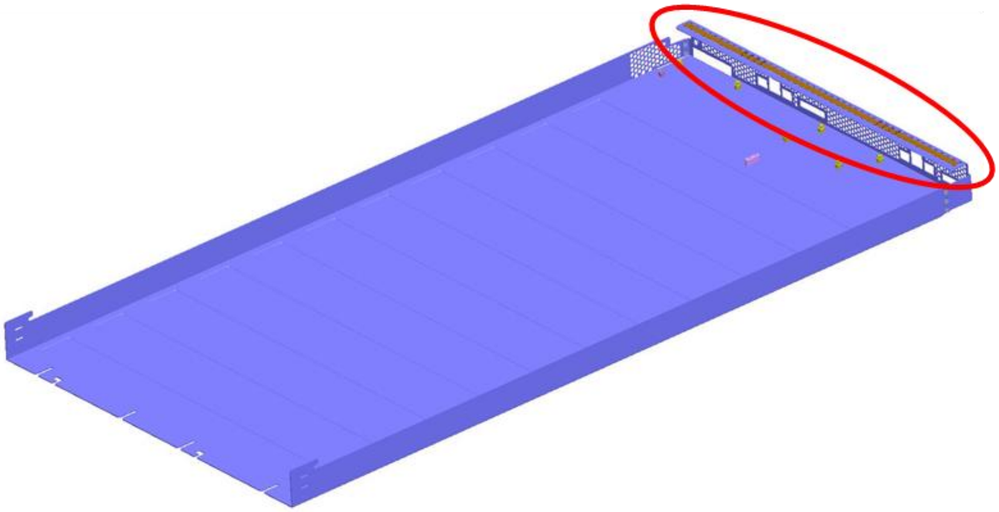


Figure 7. Tray with EMI enclosure, blade volume shown

When all trays are in place, they are electrically stitched together to prevent leakage of electromagnetic fields. The blade also has a gasket on the front of the top surface that provides electrical sealing, as shown in Figure 8. Dimensions for this gasket are shown in Figure 2.

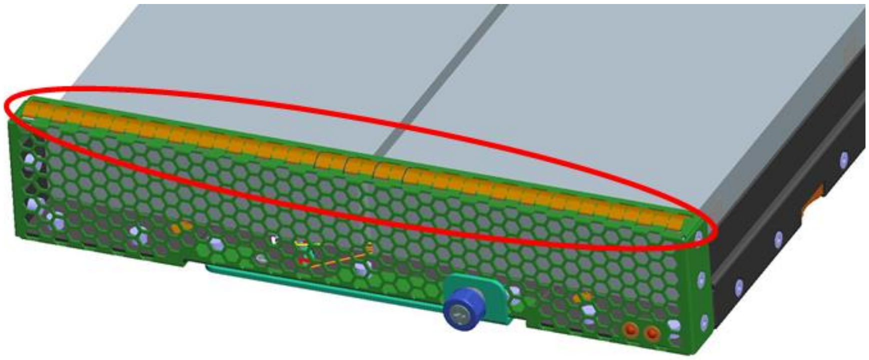


Figure 8. Blade EMI seal

7 Blade Interconnects

The tray (or other supporting infrastructure) provides the electrical interface to the blade using the connectors listed in Table 7 (or their functional equivalents). Note that the choice of these connectors is based on a coplanar PCB for power and network distribution.

Table 7 lists the connectors used for mating the blade to the tray backplane.

Table 7. Blade-to-Tray Backplane Connectors List

Assembly mounting	Assembly attachment	Connector description	Manufacturing part number
Blade	Tray backplane	Power header AirMax VS, 2x2	(FCI) 10028918-001LF
Blade	Tray backplane	Signal connector—three pair, 54 contact, 2mm spacing, 17mm pitch, six column	(FCI) 10039851-101LF
Blade	Tray backplane	Guide pin receptacle—10.8mm right angle, 0° key	(FCI) 10037912-101LF

Figure 9 shows the connector interface on the blade.

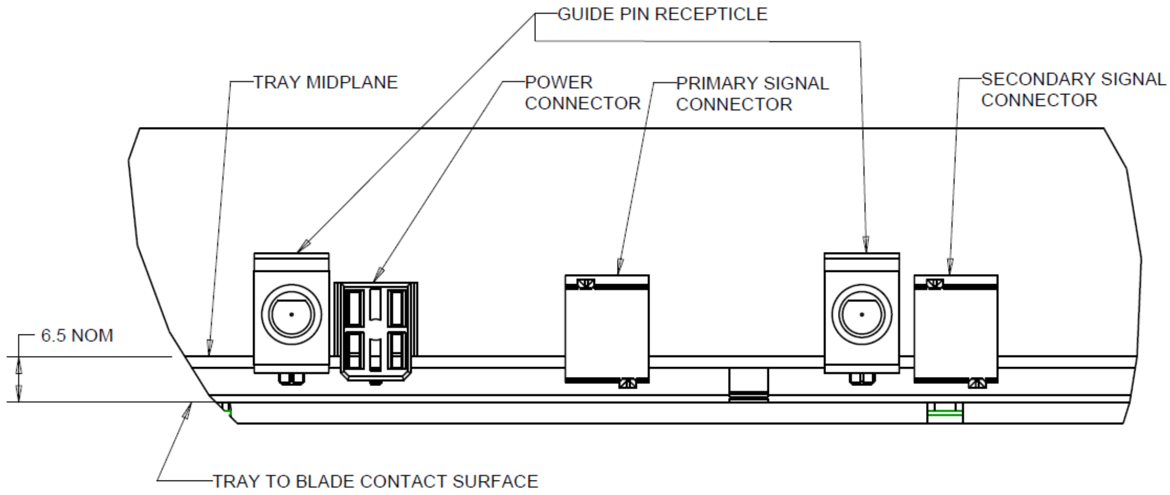


Figure 9. Blade electrical connector interface

Figure 10 shows the placement for the power and network distribution connectors.

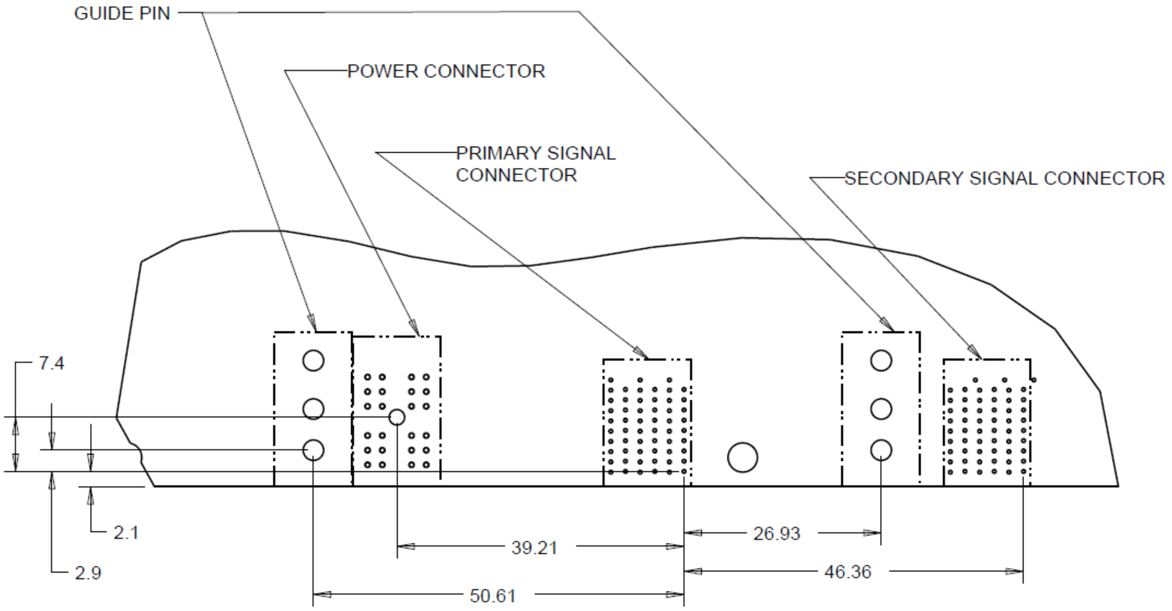


Figure 10. Tray Backplane connector placement

The total amount of force required to mate the blade to the tray backplane should not exceed 18.6 pounds throughout the expected service life of the connector set. With the leverage provided by the latch at the face of the blade, the force required should not exceed 3.15 pounds. The retention force of the connectors is a minimum of 5.66 pounds, which equates to 0.94 pounds minimum force at the latch.

7.1 Power

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

7.1.1 Power Connector

The power connector used is in the FCI AirMax VS connector family.

Figure 11 shows the power connector layout, and Table 8 shows the power receptacle pinout.

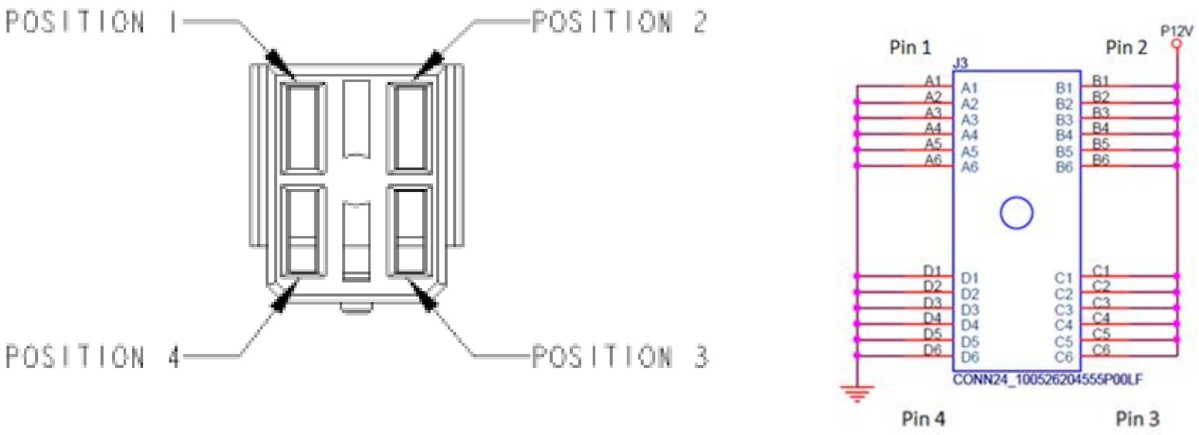


Figure 11. AirMax power receptacle pin out arrangement

Table 8. Power Receptacle Pinout

Pin	Signal name	Capacity (in A)
Position 1	12V return – first mate	18A
Position 2	12V supply	18A
Position 4	12V supply	18A
Position 3	12V return	18A

The maximum power capacity for the blade with a single connector is 345W (12V x 2 x 18A x 80 percent de-rating). Note that depending on the implementation, the maximum might not be achievable.

7.1.2 Input Voltage, Power, and Current

Table 9 lists the nominal, maximum, and minimum values for the blade input voltage. The maximum and minimum voltages include the effects of connector temperature, age, noise/ripple, and dynamic loading.

Table 9. 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 blade is defined during system power allocation. The number of blades in a chassis might be limited by the capacity of the AC power cord or by the cooling capacity of the deployment. Table 10 lists the input power allocation for a low-power blade.

Table 10. Input Power Allocation and Maximum Current

	Nominal voltage	**System power allocation (in W)**	**Maximum current (in A)**
Low-power blade (1x power connector)	12.25VDC	300W	27.3A

The blade provides in-rush current control through the 12V bus rail; return-side in-rush control is not used. The in-rush current rises linearly from 0A to the load current over a 5 millisecond (ms) period (this time period must be no longer than 200ms).

7.1.3 In-Rush Power Control Signals

The sections that follow describe how the Blade_Enable (or Blade_EN) and Blade_Mated signals work.

7.1.3.1 Blade Enable Signal

The active high blade interrupts the 12Vbus when the Blade_EN[1..2] signal is toggled to ground. The Blade_EN signal driver from the chassis management card (CMC) is an open collector; in normal operation, it will be pulled high through a resistive termination when the blade is provided with 3.3V.

To disable the power, the CMC pulls the Blade_EN signal to reference ground. The Blade_EN signal should not be pulled low on the blade and driven back to the Chassis Manager.

7.1.3.2 Blade Mated Signal

The active low Blade[1:2]_Mated_N signal is used to qualify Blade_EN to indicate that the connectors have properly mated during blade insertion before the 12V in-rush power is turned on. The mated signal is pulled to ground on the tray backplane through a 100 ohm (Ω) resistor. It is pulled high through a resistive termination based on the incoming 12V on the blade.

7.1.3.3 Interpretation of Signals

Table 11 describes the Blade_EN and Blade_Mated signals.

Table 11. Interpretation of Blade_EN and Blade_Mated Signals

Blade_EN on blade	Blade_Mated on blade	Blade/hot swap controller status
1	1	During blade removal: • No 12V power, no signal connection to backplane • Blade was either not inserted or recently removed • Blade 12V should be DISABLED
1	1	During blade insertion: • Blade installation, 12V connector hits before signal • Signal connector is not mated • Hot swap controller is DISABLED
1	0	• Blade is installed, 12V is present • Signal connector is mated • Chassis Manager is either not present or is driving high • Hot swap controller is ENABLED
0	1	• Invalid or faulty state • The only way to drive Blade_Enable low is from the CM through the signal connector • If Blade_Mated is high, the connection was not made • Hot swap controller is DISABLED
0	0	• Blade is installed, 12V is present • Signal connector is mated • Chassis Manager drives Blade_Enable low • Hot swap controller is DISABLED

7.1.4 Current Interrupt Protection and Power, Voltage, and Current Monitoring

The blade provides a cost-effective way to measure and report blade voltage and current consumption, and to make instance reporting available to the system. The V1.0 blades include power consumption measurements at the in-rush controller.

(Accuracy of the measurement will be decided upon during the implementation of the design when costs are known.)

The blade also provides a way to interrupt current flow within 1 microsecond (µs) of exceeding three times the maximum current load.

7.1.5 Filtering and Capacitance

The blade provides sufficient filtering to operate with a 0.25A/µSec DC source.

Maximum capacitance allowed per blade is 12,000µFarads (µF). Minimum capacitance allowed per blade is 300µF.

7.1.6 Grounding and Return

The blade chassis grounding/return is provided by the assembly guide pin. The logic return, or 12V bus return, is provided through the 12V bus power connectors and signal connectors. The logic return and chassis return are not attached, bonded, or connected within the blade assembly.

7.2 Signal Interface

The sections that follow define the signaling interface to the blade.

7.2.1 Signal Connectors

The signal connector used is in the FCI AirMax VSe family. Figure 12 shows an example of coplanar connector pair.

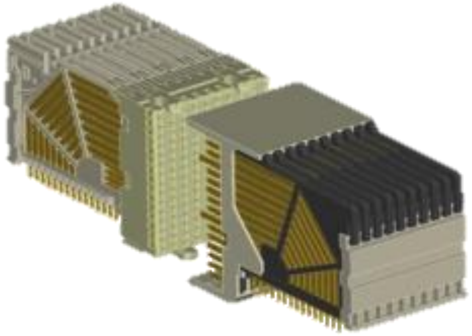


Figure 12. Example of a coplanar blade signal connector

This primary connector supports two 10G Ethernets, two 1G Ethernets, management signals, two serial buses, and management signals.

The second optional connector supports eight SAS channels for the storage expansion connector and signals reserved for an OEM management network.

The connector is organized as a grid, with rows A through F (three groups: A-C, D-F, and G-I) and columns 1 through 6. Note that the columns flip from header to receptacle so that the mated pairs match.

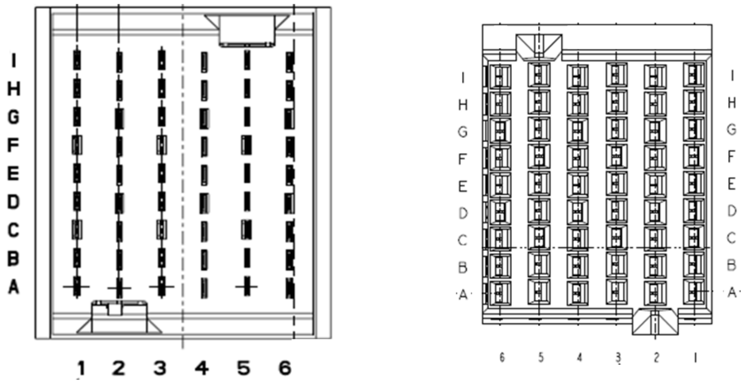


Figure 13. AirMax VS pinout arrangement (blade header on left, and backplane receptacle on right)

For the 10G Ethernet drivers to operate properly, the controller must be able to directly read the serial electrically erasable programmable read-only memory (EEPROM) on the 10G SFP+ cables. Serial data (SDA) signals and serial clock (SCK) signals are driven from the blade for each small form-factor pluggable + (SFP+) connector because the EEPROMs located on the SFP+ cables are hard-wired to address 0x0.

7.2.2 Signal Definitions

Table 12 below shows the pin definitions for the primary connector (Ethernet) J22.

Table 12. Primary Connector (Ethernet) J22

Signal	Ball pin #	Pin type	Description	Use
ENET1G_1_MDI0P	A1	I/O	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_1_MDI0N	B1	I/O		

Signal	Ball pin #	Pin type	Description	Use
GND	C1			
ENET10G_1_RDP	D1	O	10GbE Receive lane	10Gb from SFP+ to NIC
ENET10G_1_RDN	E1	O		
GND	F1			
ENET10G_1_TDP	G1	I	10GbE Transmit lane	10Gb from NIC to SFP+
ENET10G_1_TDN	H1	I		
GND	I1			
GND	A2			
ENET1G_1_MDI2P	B2	I/O	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_1_MDI2N	C2	I/O		
GND	D2	O		
ENET1G_1_MDI1P	E2	I/O	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_1_MDI1N	F2	I/O		
GND	G2			
BLADE_EN1	H2	O	Node1 Enable for blade	Two nodes per sled
SFP1_PRESENT_N	I2	O	Indicates SFP+ cable is present	Pullup on blade
ENET1G_1_MDI3P	A3	I/O	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_1_MDI3N	B3	I/O		
GND	C3			
SFP1_I2C_DATA1	D3	I/O	I2C for SFP+ connector	Pullups on blade
SFP1_I2C_CLK1	E3	I		
TRAY_BP_SKU_ID0	F3	O	SKU ID Bit for tray backplane	Pullup/pulldown on tray BP
NODE0_RXD	G3	I	UART for CM communication	

Signal	Ball pin #	Pin type	Description	Use
NODE0_TXD	H3	O		
SFP_P3V3	I3	I	3.3V power for SFP+ cables	3.3V Power for SFP+ cables
GND	A4			
ENET1G_2_MDI0P	B4	I	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_2_MDI0N	C4	I		
GND	D4			
SFP2_I2C_DATA2	E4	I	I2C for SFP+ connector	Pullups on blade
SFP2_I2C_CLK2	F4			
NODE1_RXD	G4		1GbE MDI lane	1GbE MDI Lane
NODE1_TXD	H4			
TRAY_BP_SKU_ID1	I4		SKU bit for tray backplane	Pullup/Pulldown on tray BP
ENET1G_2_MDI1P	A5	G	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_2_MDI1N	B5	G		
GND	C5	G		
ENET1G_2_MDI2P	D5	G	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_2_MDI2N	E5	G		
GND	F5	G		
BLADE_EN2	G5	G	Node2 Enable for blade	Two nodes per sled
SFP2_PRESENT_N	H5	G	Indicates SFP+ cable is present	Pullup on blade
BLADE1_MATED_N	I5	G	Indicates connectors are mated	Pullup on blade
GND	A6	G		
ENET1G_2_MDI3P	B6	G	1GbE MDI lane	1Gb from NIC to 1G Base-T (unused)
ENET1G_2_MDI3N	C6	G		

Signal	Ball pin #	Pin type	Description	Use
GND	D6	G		
ENET10G_2_RDP	E6	G	10GbE Receive lane	10Gb RX from SFP+ to NIC
ENET10G_2_RDN	F6	G		
GND	G6	G		
ENET10G_2_TDP	H6	G	10GbE Transmit lane	10GB TX from NIC to SFP+
ENET10G_2_TDN	I6	G		

Table 13 shows the pin definitions for the secondary connector (SAS) J8.

Table 13. Pin Definition for Secondary Connector (SAS) J8

Signal	Ball pin #	Pin type	Sch name	Description	Use
SAS Interface					
SAS_BLAD1_RX1_P	D1	O	SAS_MEZZ_RX1_C_P	8 lanes of SAS differential receive	Connect to SAS Mezz conn with 0.01uF CAP
SAS_BLAD1_RX2_P	H2	O	SAS_MEZZ_RX2_C_P		
SAS_BLAD1_RX3_P	B2	O	SAS_MEZZ_RX3_C_P		
SAS_BLAD1_RX4_P	D3	O	SAS_MEZZ_RX4_C_P		
SAS_BLAD1_RX5_P	H4	O	SAS_MEZZ_RX5_C_P		
SAS_BLAD1_RX6_P	D5	O	SAS_MEZZ_RX6_C_P		
SAS_BLAD1_RX8_P	B6	O	SAS_MEZZ_RX8_C_P		
SAS_BLAD1_RX7_P	H6	O	SAS_MEZZ_RX7_C_P		
SAS_BLAD1_RX1_N	E1	O	SAS_MEZZ_RX1_C_N		
SAS_BLAD1_RX2_N	I2	O	SAS_MEZZ_RX2_C_N		
SAS_BLAD1_RX3_N	C2	O	SAS_MEZZ_RX3_C_N		
SAS_BLAD1_RX4_N	E3	O	SAS_MEZZ_RX4_C_N		

Signal	Ball pin #	Pin type	Sch name	Description	Use
SAS_BLAD1_RX5_N	I4	O	SAS_MEZZ_RX5_C_N		
SAS_BLAD1_RX6_N	E5	O	SAS_MEZZ_RX6_C_N		
SAS_BLAD1_RX8_N	C6	O	SAS_MEZZ_RX8_C_N		
SAS_BLAD1_RX7_N	I6	O	SAS_MEZZ_RX7_C_N		
SAS_BLAD1_TX2_P	G1	I	SAS_MEZZ_TX2_P	8 lanes of SAS differential transmit	Connect to SAS Mezz conn
SAS_BLAD1_TX1_P	A1	I	SAS_MEZZ_TX1_P		
SAS_BLAD1_TX3_P	E2	I	SAS_MEZZ_TX3_P		
SAS_BLAD1_TX4_P	G3	I	SAS_MEZZ_TX4_P		
SAS_BLAD1_TX5_P	E4	I	SAS_MEZZ_TX5_P		
SAS_BLAD1_TX6_P	A5	I	SAS_MEZZ_TX6_P		
SAS_BLAD1_TX7_P	G5	I	SAS_MEZZ_TX7_P		
SAS_BLAD1_TX8_P	E6	I	SAS_MEZZ_TX8_P		
SAS_BLAD1_TX2_N	H1	I	SAS_MEZZ_TX2_N		
SAS_BLAD1_TX1_N	B1	I	SAS_MEZZ_TX1_N		
SAS_BLAD1_TX3_N	F2	I	SAS_MEZZ_TX3_N		
SAS_BLAD1_TX4_N	H3	I	SAS_MEZZ_TX4_N		
SAS_BLAD1_TX5_N	F4	I	SAS_MEZZ_TX5_N		
SAS_BLAD1_TX6_N	B5	I	SAS_MEZZ_TX6_N		
SAS_BLAD1_TX7_N	H5	I	SAS_MEZZ_TX7_N		
SAS_BLAD1_TX8_N	F6	I	SAS_MEZZ_TX8_N		
AIRMAX_RSVD3	A3		TP	Reserved	TP
AIRMAX_RSVD4	B3		TP	Reserved	

Signal	Ball pin #	Pin type	Sch name	Description	Use
AIRMAX_RSVD1	B4		TP	Reserved	
AIRMAX_RSVD2	C4		TP	Reserved	
Power and ground					
GND	C1	G	Gnd	Gnd	Gnd
GND	F1	G	Gnd		
GND	I1	G	Gnd		
GND	A2	G	Gnd		
GND	D2	G	Gnd		
GND	G2	G	Gnd		
GND	C3	G	Gnd		
GND	F3	G	Gnd		
GND	I3	G	Gnd		
GND	A4	G	Gnd		
GND	D4	G	Gnd		
GND	G4	G	Gnd		
GND	C5	G	Gnd		
GND	F5	G	Gnd		
GND	I5	G	Gnd		
GND	A6	G	Gnd		
GND	D6	G	Gnd		
GND	G6	G	Gnd		

8 Mezzanine I/O

The motherboard supports a network mezzanine and a Serial Attach SCSI (SAS) mezzanine card. Details, including connector and connector pinouts, are provided in the mezzanine NIC and mezzanine SAS specifications. Figure 14 shows a block diagram of the mezzanine connections. The mezzanines attach to connectors on the tray backplane through the AirMax connectors, as described in Section 7.2

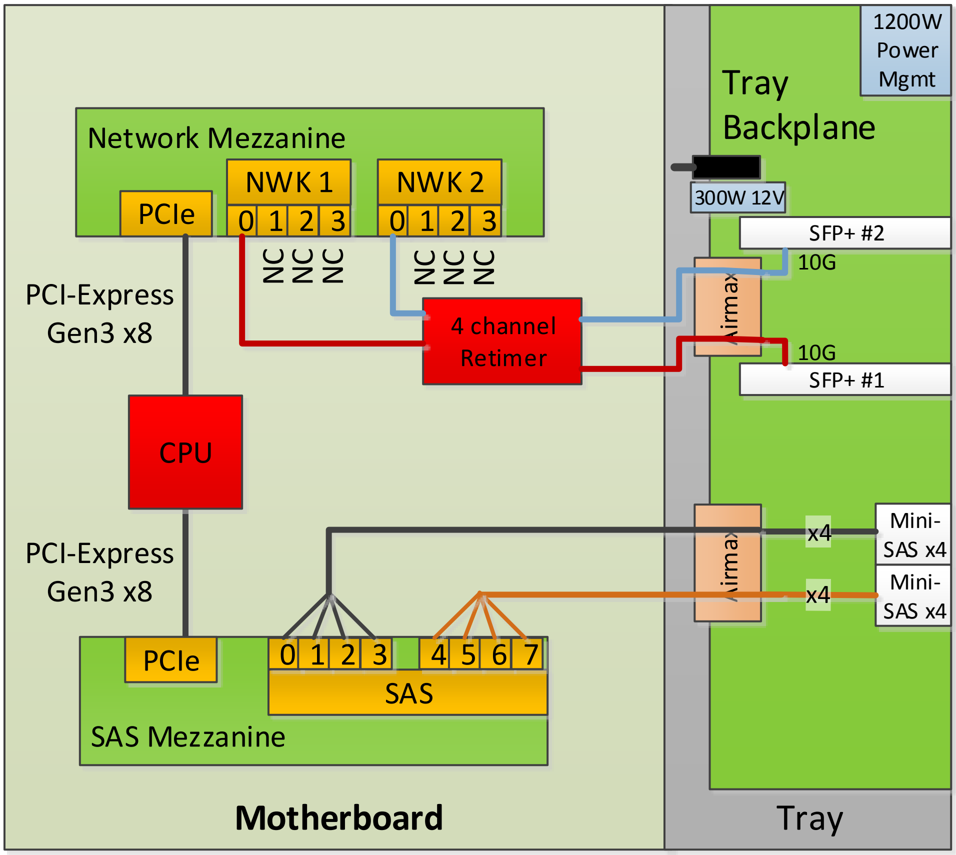


Figure 14. Mezzanine I/O

9 PCI-Express Expansion Slot

A PCI-Express Gen3 x16 Riser interface is supported on a 164-pin golden finger connector.

Table 14 shows the PCI-Express x16 Riser pin definition.

Table 14. PCI-Express Gen3 x16 Riser Pin Definition

	Side B	Side A
1	P12V_BP_PCIE	GND
2	P12V_BP_PCIE	P12V_BP_PCIE
3	P12V_BP_PCIE	P12V_BP_PCIE
4	GND	GND
5	SMB_RSER_R_CLK	CLK_100M_SLOTB_DP
6	SMB_RSER_R_DAT	CLK_100M_SLOTB_DN
7	GND	GND
8	P3V3	
9	PD_SLOT6_JTAG1	P3V3
10	P3V3_AUX	P3V3
11	IRQ_LVC3_WAKE_N	RST_PERST0_N
12	RISER_TYPE0	GND
13	GND	CLK_100M_SLOTA_DP
14	P3E_CPU0_SLOT1_TX_C_DP<0>	CLK_100M_SLOTA_DN
15	P3E_CPU0_SLOT1_TX_C_DN<0>	GND
16	GND	P3E_CPU0_SLOT1_RX_DP<0>
17	PCIE_SLOT1_CPRSNT1_N	P3E_CPU0_SLOT1_RX_DN<0>
18	GND	GND
19	P3E_CPU0_SLOT1_TX_C_DP<1>	RISER_TYPE1
20	P3E_CPU0_SLOT1_TX_C_DN<1>	GND
21	GND	P3E_CPU0_SLOT1_RX_DP<1>
22	GND	P3E_CPU0_SLOT1_RX_DN<1>
23	P3E_CPU0_SLOT1_TX_C_DP<2>	GND

	Side B	Side A
24	P3E_CPU0_SLOT1_TX_C_DN<2>	GND
25	GND	P3E_CPU0_SLOT1_RX_DP<2>
26	GND	P3E_CPU0_SLOT1_RX_DN<2>
27	P3E_CPU0_SLOT1_TX_C_DP<3>	GND
28	P3E_CPU0_SLOT1_TX_C_DN<3>	GND
29	GND	P3E_CPU0_SLOT1_RX_DP<3>
30		P3E_CPU0_SLOT1_RX_DN<3>
31		GND
32	GND	
33	P3E_CPU0_SLOT1_TX_C_DP<4>	
34	P3E_CPU0_SLOT1_TX_C_DN<4>	GND
35	GND	P3E_CPU0_SLOT1_RX_DP<4>
36	GND	P3E_CPU0_SLOT1_RX_DN<4>
37	P3E_CPU0_SLOT1_TX_C_DP<5>	GND
38	P3E_CPU0_SLOT1_TX_C_DN<5>	GND
39	GND	P3E_CPU0_SLOT1_RX_DP<5>
40	GND	P3E_CPU0_SLOT1_RX_DN<5>
41	P3E_CPU0_SLOT1_TX_C_DP<6>	GND
42	P3E_CPU0_SLOT1_TX_C_DN<6>	GND
43	GND	P3E_CPU0_SLOT1_RX_DP<6>
44	GND	P3E_CPU0_SLOT1_RX_DN<6>
45	P3E_CPU0_SLOT1_TX_C_DP<7>	GND
46	P3E_CPU0_SLOT1_TX_C_DN<7>	GND
47	GND	P3E_CPU0_SLOT1_RX_DP<7>
48		P3E_CPU0_SLOT1_RX_DN<7>

	Side B	Side A
49	GND	GND
50	P3E_CPU0_SLOT1_TX_C_DP<8>	
51	P3E_CPU0_SLOT1_TX_C_DN<8>	GND
52	GND	P3E_CPU0_SLOT1_RX_DP<8>
53	GND	P3E_CPU0_SLOT1_RX_DN<8>
54	P3E_CPU0_SLOT1_TX_C_DP<9>	GND
55	P3E_CPU0_SLOT1_TX_C_DN<9>	GND
56	GND	P3E_CPU0_SLOT1_RX_DP<9>
57	GND	P3E_CPU0_SLOT1_RX_DN<9>
58	P3E_CPU0_SLOT1_TX_C_DP<10>	GND
59	P3E_CPU0_SLOT1_TX_C_DN<10>	GND
60	GND	P3E_CPU0_SLOT1_RX_DP<10>
61	GND	P3E_CPU0_SLOT1_RX_DN<10>
62	P3E_CPU0_SLOT1_TX_C_DP<11>	GND
63	P3E_CPU0_SLOT1_TX_C_DN<11>	GND
64	GND	P3E_CPU0_SLOT1_RX_DP<11>
65	GND	P3E_CPU0_SLOT1_RX_DN<11>
66	P3E_CPU0_SLOT1_TX_C_DP<12>	GND
67	P3E_CPU0_SLOT1_TX_C_DN<12>	GND
68	GND	P3E_CPU0_SLOT1_RX_DP<12>
69	GND	P3E_CPU0_SLOT1_RX_DN<12>
70	P3E_CPU0_SLOT1_TX_C_DP<13>	GND
71	P3E_CPU0_SLOT1_TX_C_DN<13>	GND
72	GND	P3E_CPU0_SLOT1_RX_DP<13>

	Side B	Side A
73	GND	P3E_CPU0_SLOT1_RX_DN<13>
74	P3E_CPU0_SLOT1_TX_C_DP<14>	GND
75	P3E_CPU0_SLOT1_TX_C_DN<14>	GND
76	GND	P3E_CPU0_SLOT1_RX_DP<14>
77	GND	P3E_CPU0_SLOT1_RX_DN<14>
78	P3E_CPU0_SLOT1_TX_C_DP<15>	GND
79	P3E_CPU0_SLOT1_TX_C_DN<15>	GND
80	GND	P3E_CPU0_SLOT1_RX_DP<15>
81		P3E_CPU0_SLOT1_RX_DN<15>
82		GND

Note that the connector can support two PCIe daughter cards.

10 Blade Management

The blade is managed with a customized, low-cost hardware monitoring device (HMD), which communicates to the Chassis Manager through a serial port. JBOD blades use serial ports already on the chipset for communicating to the Chassis Manager.

10.1 Overview of Blade Management

Two Blade_Enable signals are provided to the blade, in support of up to two servers or devices per blade. Note that the Blade_Enable signals require a complete shutdown of power to the server and any associated management hardware.

There are two dedicated serial links routed to the blade, for use by up to two servers or devices per blade. These links use 3.3V. If only one server or device is supported, serial port #1 should be used. Both compute and JBOD blades can use the serial link.

Figure 15 shows a block diagram of the blade, Chassis Manager, and other components.

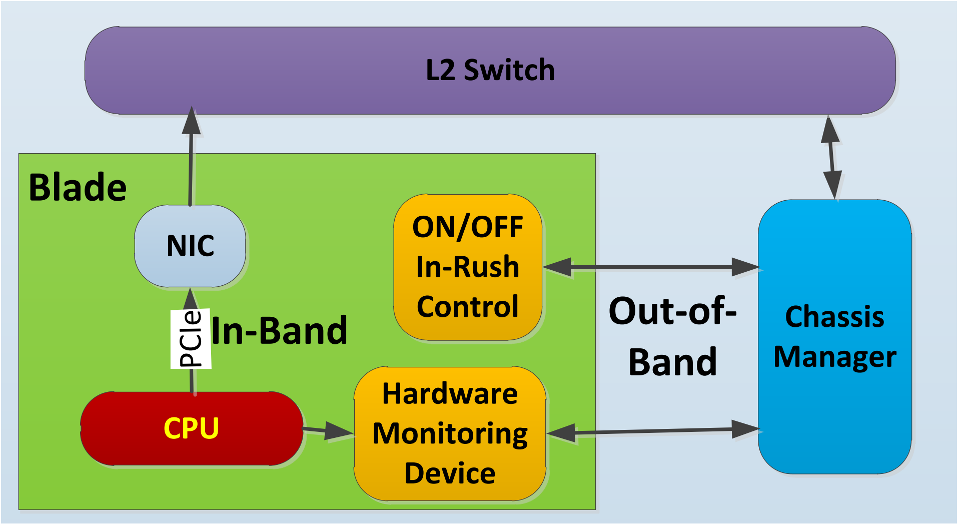


Figure 15. Server management

Figure 16 shows details of the blade management hardware. Primary features include:

- Low pin count (LPC) connection to the chipset to support in-band management
- Serial numbers, asset information, and event logs stored in electrically erasable programmable read-only memory (EEPROM)
- Thermal sensors for inlet and exhaust
- Power monitoring through the in-rush controller circuitry
- Service LEDs
- Debugging paths for the second Blade_EN signal and serial port

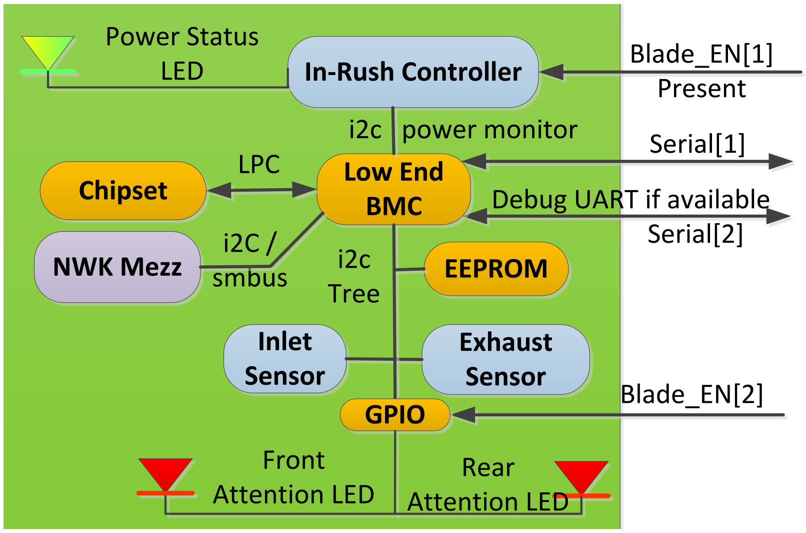


Figure 16. Blade management hardware

JBOD blades are managed using the Blade_EN signal to the in-rush controller, managed using a serial port attached to the SAS fanout expander, and managed in-band through the operating system. Only the exhaust temperature sensor is required because the inlet temperature information can be gathered from the attached server. Table 15 lists the blade numbers.

Table 15. Blade Numbers

U number	Blades when viewed from front		Blades when viewed from rear	
12	12	24	24	12
11	11	23	23	11
10	10	22	22	10
9	9	21	21	9
8	8	20	20	8
7	7	19	19	7
6	6	18	18	6
5	5	17	17	5
4	4	16	16	4
3	3	15	15	3
2	2	14	14	2
1	1	13	13	1

10.2 Blade LEDs

The following sections describe the LEDs used as indicators on the blades and chassis. (Also see Section 7.1.3 for information about the in-rush power control signals.)

Each blade has two LEDs: a power status LED that is green/amber and an attention LED that is red. Figure 17 shows the location of these LEDs. Because of variations in the blade design and functional requirements, the location, visible diameter, and brightness of the LEDs can vary. Additional LEDs can be used to provide more information as long as they will not be confused with the required power status and attention LEDs.

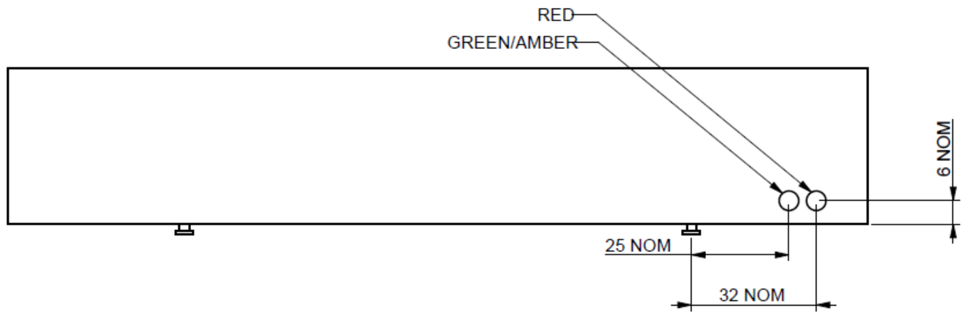


Figure 17. Locations of blade LEDs

10.2.1 Blade Power Status LED

When a blade is first inserted, the LED will turn amber if the backplane has 12V power, if the alignment pin is mated, and if the Blade_EN signal is asserted. These conditions ensure that the high-speed signal has made connection, and are more reliable indicators of proper insertion than simply the 12V power connector.

When the blade's management software turns on the primary 12V power, the power status LED turns green. Note that the power status LED is driven by an analog resistor network tied directly to the power rails, and is not an indication of the health of the blade. Table 16 describes the operation of the blade power status LED.

Table 16. Blade Power Status LED Description

LED status	Condition
Off	• Blade is not fully inserted, 12V power is absent, or Blade_EN is de-asserted • Standby and CPU power are off
Solid Amber ON	• Blade is inserted, 12V power is available, and Blade_EN is asserted • Standby power is on, but CPU power is off
Solid Green ON	• Standby and CPU power are turned on

10.2.2 Blade Attention LED

The blade attention LED directs the service technicians to the blade that requires repair. The technician can remove the blade from the rack and replace it with an operational blade. When possible, blade diagnostics are used to direct repairs; alternately, the scale-out cloud management software can be used. In both cases, logs of the repair work are available.

The attention LED is driven by a single General Purpose I/O (GPIO) bit off the blade's management i2c tree. The attention LED is stateless; when the blade is removed or when the Blade_EN signal is de-asserted and turns off the inrush controller power, chassis management sends an "off" command to the management subsystem and the attention LED is cleared. Table 17 describes the operation of the blade attention LED.

Table 17. Blade Attention LED Description

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

If a future blade design supports repair without blade removal (such as hot-swap HDDs), the red LED will be used to signal repair.

Note that a second red LED is located at the rear of the blade so that it is visible through the tray when the fan door is opened. This LED is set at the same time as the front red attention LED, and indicates which blade needs attention from the rear of the chassis.

11 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

AHJ – authority having jurisdiction

ANSI – American National Standards Institute

API – application programming interface

ASHRAE – American Society of Heating, Refrigerating and Air Conditioning Engineers

ASIC – application-specific integrated circuit

BCD – binary-coded decimal

BIOS – basic input/output system

BMC – baseboard management controller

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

CM – Chassis Manager

CMOS – complementary metal–oxide–semiconductor

COLO – co-location

CTS – clear to send

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

DSR – data set ready

DTR – data terminal ready

ECC – error-correcting code

EEPROM - electrically erasable programmable read-only memory

EIA – Electronic Industries Alliance

EMC – electromagnetic compatibility

EMI – electromagnetic interference

FRU – field replaceable unit

FTP – file transfer protocol

GPIO – general purpose input output

GUID – globally unique identifier

HBI – high business intelligence

HCK – Windows Hardware Certification Kit

HMD – hardware monitoring device

HT – hyperthreading

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

LPC – low pin count

LS – least significant

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

NUMA – non-uniform memory access

OOB – out of band

OSHA - Occupational Safety & Health Administration

OTS – off the shelf

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

PNP – plug and play

POST – power-on self-test

PSU – power supply unit

PWM – pulse-width modulation

PXE – preboot execution environment

QDR – quad data rate

QFN – quad flat package no-lead

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

RTS – request to send

RU – rack unit

RxD – received data

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

SATA – serial AT attachment

SCK – serial clock

SCSI – small computer system interface

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

TB – tray backplane

TDP – thermal design power

TOR – top of rack

TPM – trusted platform module

TxD – transmit data

U – rack unit

UART – universal asynchronous receiver/transmitter

UEFI – unified extensible firmware interface

UL – Underwriters Laboratories

UPS – uninterrupted power supply

Vpp – voltage peak to peak

WMI – Windows Management Interface