Intel® Server Board S7200AP
Intel® Compute Module HNS7200AP

Technical Product Specification

A document providing an overview of product features, functions, architecture, and support specifications

Revision 1.0
Jun, 2016
Intel® Server Boards and Systems

<This page is intentionally left blank.>

Revision History

Date	Revision Number	Modifications
June 2015	0.5	Initial Release- Alpha
January 2016	0.75	Illustration update; usage update; node feature update
March 2016	0.80	Update riser slot information; update CPU TDP
June 2016	1.0	Update environmental specifcation table; update air flow table; update 2130w PSU section; update product weight; update BMC sensor table

Document Disclaimer Statements

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest TPS.

The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request.

Intel, and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others

Table of Contents

List of Figures

List of Tables

1 Introduction

This *Technical Product Specification (TPS)* provides specific information detailing the features, functionality, and high-level architecture of the Intel® Server Board S7200AP product family and the Intel® Compute Module HNS7200AP product family.

Design-level information related to specific server board components and subsystems can be obtained by ordering *External Product Specifications* (EPS) or *External Design Specifications* (EDS) related to this server generation. EPS and EDS documents are made available under NDA with Intel and must be ordered through your local Intel representative. See the Reference Documents section for a list of available documents.

1.1 Chapter Outline

This document is divided into the following chapters:

- Chapter 1 – Introduction
- Chapter 2 – Product Features Overview
- Chapter 3 – Processor Support
- Chapter 4 – Memory Support
- Chapter 5 – Server Board I/O
- Chapter 6 – Connector and Header
- Chapter 7 – Configuration Jumpers
- Chapter 8 – Intel® Light-Guided Diagnostics
- Chapter 9 – Platform Management
- Chapter 10 – Thermal Management
- Chapter 11 – System Security
- Chapter 12 – Environmental Limits Specification
- Chapter 13 – Power Supply Specification Guidelines
- Appendix A – Integration and Usage Tips
- Appendix B – Integrated BMC Sensor Tables
- Appendix C – BIOS Sensors and SEL Data
- Appendix D – POST Code Diagnostic LED Decoder
- Appendix E – POST Code Errors
- Appendix F – Statement of Volatility
- Appendix G - Glossary
- Appendix H – Reference Documents

1.2 Server Board Use Disclaimer

Intel Corporation server boards contain a number of high-density VLSI (Very Large Scale Integration) and power delivery components that need adequate airflow to cool. Intel ensures through its own chassis development and testing that when Intel server building blocks are used together, the fully integrated

system will meet the intended thermal requirements of these components. It is the responsibility of the system integrator who chooses not to use Intel developed server building blocks to consult vendor datasheets and operating parameters to determine the amount of air flow required for their specific application and environmental conditions. Intel Corporation cannot be held responsible if components fail or the server board does not operate correctly when used outside any of their published operating or non-operating limits.

2 Product Features Overview

The Intel® Server Board S7200AP product family is a monolithic printed circuit board (PCB) assembly with features designed to support the high performance and high density computing markets. This server board is designed to support the Intel® Xeon™ Phi™ processor family. Previous generation of Intel® Xeon™ Phi™ processor are not supported.

The Intel® Server Board S7200AP product family contains two server board options. Many of the features and functions of the server board family are common. A board will be identified by its name which has described features or functions unique to it.

- S7200AP – With bootable Intel® Xeon™ Phi™ processor (KNL-D)
- S7200APF With bootable Intel® Xeon™ Phi™ processor (KNL-F) with embedded Intel® Omni-Path Host Fabric Interface (HFI).

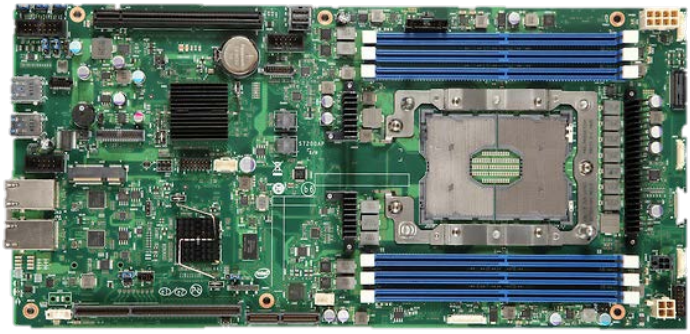


Figure 1. Intel® Server Board S7200AP (demo picture)

Figure 2. Intel® Compute Module HNS7200AP (demo picture)

The following table provides a high-level product feature list.

Table 1. Intel® Server Board S7200AP Product Family Feature Set

Feature	Description
Processor Support	Intel® Xeon™ Phi™ processor (KNL-D-72 Core) Bootable ▪ Single processor socket P (3647 pins) ▪ Thermal Design Power (TDP) up to 215W ▪ 36 lanes of Integrated PCI Express® 3.0 low-latency I/O Intel® Xeon™ Phi™ processor (KNL-F-72 Core) Host Fabric Interface Bootable ▪ Single processor socket P (3647 pins) ▪ Integrated 4x25 Gb/s Host Fabric Interface ▪ Thermal Design Power (TDP) up to 230W ▪ 4 lanes of Integrated PCI Express® 3.0 low-latency I/O
Memory Support	▪ Six DIMM slots in total across six memory channels ▪ Registered DDR4 (RDIMM), Load Reduced DDR4 (LRDIMM) ▪ Memory DDR4 data transfer rates of 1866/2133/2400 MT/s ▪ 1 DIMM per channel ▪ Max memory 384GB
Chipset	Intel C610 “Wellsburg” Platform Controller Hub (PCH)
External I/O Connections	▪ Two USB 3.0 connectors ▪ Two RJ-45 10/100/1000 Mbit Network Interface Controller (NIC) ports
Internal I/O connectors/headers	▪ One USB 2.0 Header ▪ One TPM Header ▪ One Intel® Omni-Path I Fabric Signal Connector ▪ One mSATA Connector ▪ One Bridge Board Connector ▪ One 2x7 pin header for system fan module ▪ One Aux Front Panel Connector ▪ Three 8-pin fan headers for third-party chassis support ▪ One 4 pin CPU Fan or Water Pump header ▪ One PSU Control Header ▪ One RMII header for Intel® RMM4 Lite ▪ One internal RGB Video Header ▪ One Serial Port A Header
PCIe Support	PCIe* 3.0 (2.5, 5, 8 GT/s)
Power Connections	• Two sets of 2x3 pin connector • One 8 pin Power control connector • One 4 pin Power connector for Disk Drive power
System Fan Support	▪ Three 40x56mm double rotor fans ▪ One 4 pin CPU Fan or Water Pump header
Video	▪ Integrated 2D video graphics controller ▪ 128MB DDR3 memory
Riser Support	▪ One PCIe Gen3 x16 standard riser connector o Supports a low-profile adapter in Riser slot 1 ▪ One PCIe Gen3 x20 HSEC-8 fine-pitch riser connector o Supports a x16 low-profile adapter in Riser slot 2 ▪ Supports a x4 low-profile adapter in Riser 2 when fabric is used
On-board storage controllers and options	▪ Integrated 9-port SATA o 4 ports to bridge board, o 1 port to mSATA o 4 ports to MiniSAS HD connector
Fabric	▪ Dual port Intel® Omni-Path Fabric via KNL-F Processor or ▪ Single Port Intel® Omni-Path Fabric via x16 Gen 3 PCIe Adapter

Feature	Description
Network (LAN)	▪ Dual i210 Springvilles ▪ Dual 10/100/1000Gbe RJ45 connectors ▪ NC_SI sideband to BMC. Option to host share or dedicate a Network port to management traffic.
RAID Support	▪ Intel® Embedded Server RAID Technology 2 (ESRT2)
Server Management	▪ Onboard Emulex* Pilot III* Controller ▪ Support for Intel® Remote Management Module 4 Lite solutions ▪ Support for Intel® System Management Software ▪ Support for Intel® Intelligent Power Node Manager (Need PMBus*-compliant power supply)

Table 2. Intel® Compute Module HNS7200AP Product Family Feature Set

Feature[1]	Description
Server Board	Intel® Server Board S7200AP product family Intel® Xeon™ Phi™ processor (KNL-D-72 Core) Bootable ▪ Intel® Xeon™ Phi™ processor (KNL-D-72 Core) Bootable ▪ Intel® Xeon™ Phi™ processor (KNL-F-72 Core) Host Fabric Interface Bootable
Processor Support	Maximum supported Thermal Design Power (TDP) of up to 230W
Heatsink	▪ One 80x107mm 1U Heatsink
Fan	Three 40x56mm dual rotor system fans
Riser Support	▪ One PCIe Gen3 x16 standard riser connector o Supports a low-profile adapter in Riser slot 1 ▪ One PCIe Gen3 x20 HSEC-8 fine-pitch riser connector o Supports a x16 low-profile adapter in Riser slot 2 Supports a x4 low-profile adapter in Riser 2 when fabric is used
Compute Module Board	▪ Bridge boards: o 6G SATA Bridge Board (Default) ▪ One compute module power docking board
Air Duct	One transparent air duct
Form Factor	Length 14.17" (360m), width 6.81" (173mm)

Notes:

1. *The table only lists features that are unique to the compute module or different with the server board.*

2.1 Components and Features Identification

This section provides a general overview of the server board and compute module, identifying key features and component locations. The majority of the items identified are common in the product family.

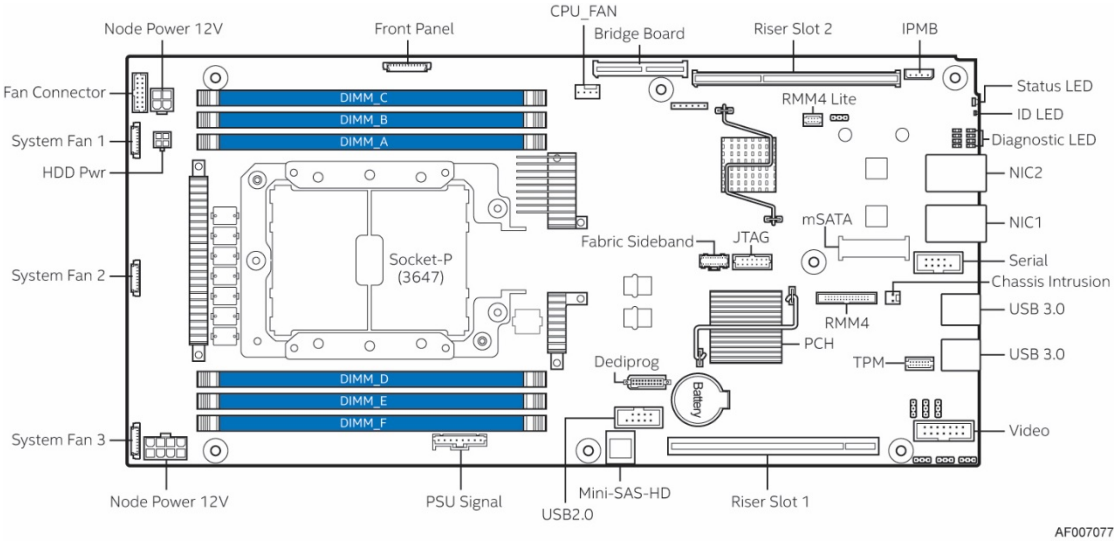


Figure 3. Server Board S7200AP Components

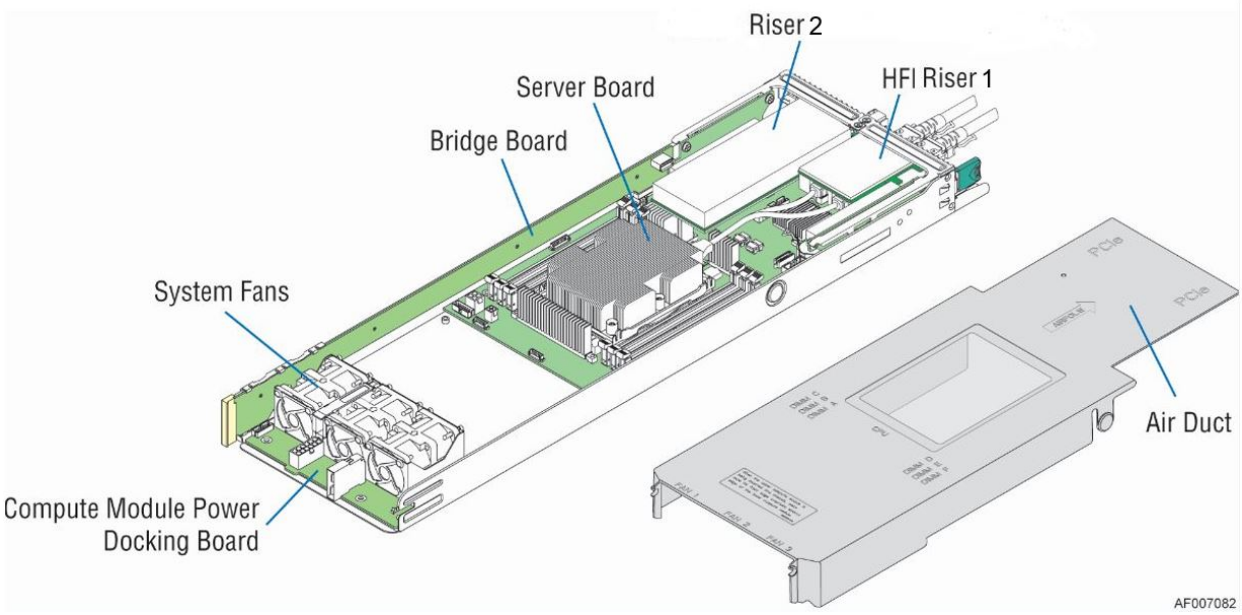


Figure 4. Compute Module Components

2.2 Rear Connectors and Back Panel Feature Identification

The Intel® Server Board S7200AP product family has the following board rear connector placement.

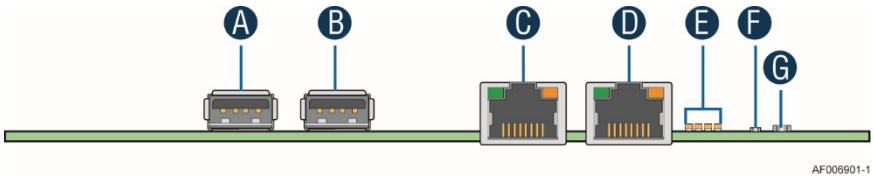


Figure 5. Server Board Rear Connectors

Table 3. Rear Connector Descriptions

	Description		Description
A	USB 3.0 Port 1	B	USB 3.0 Port 2
C	NIC port 1 (RJ45)	D	NIC port 2 (RJ45)
E	POST Code LEDs (8 LEDs)	F	ID LED
G	Status LED		

The Intel® Compute Module HNS7200AP product family has the following back panel features.

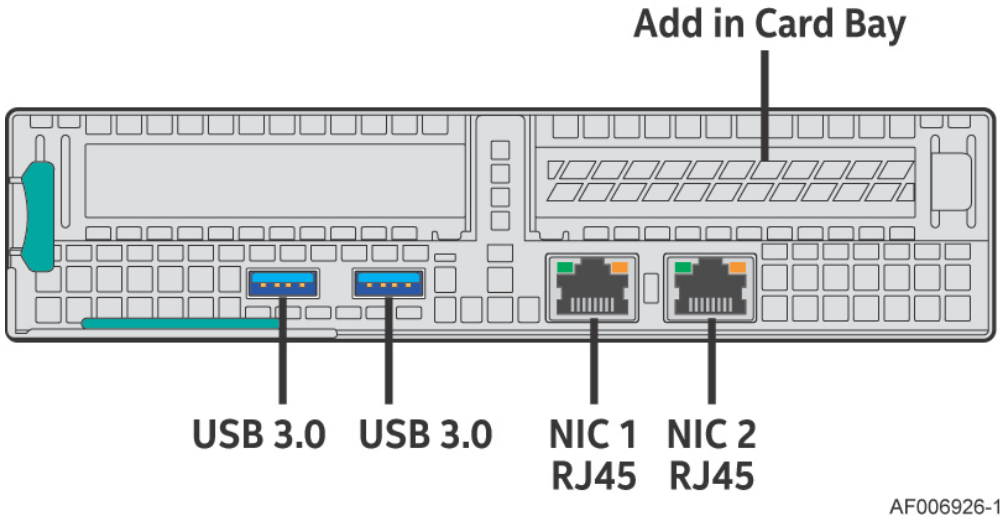


Figure 6. Compute Module Back Panel

2.3 Intel® Light Guided Diagnostic LED

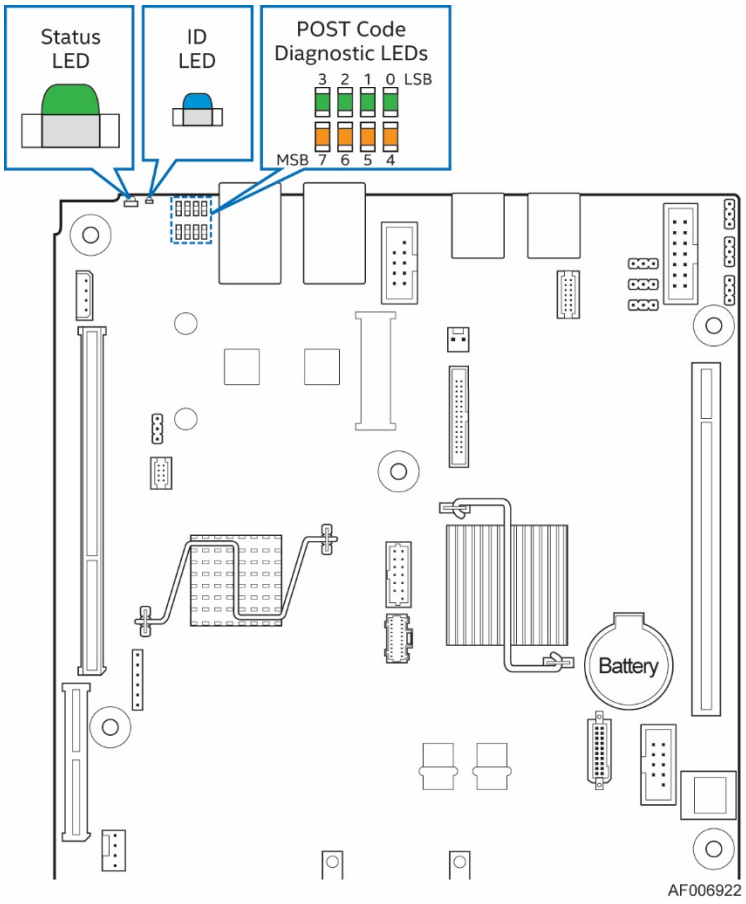


Figure 7. Intel® Light Guided Diagnostic LED

2.4 Jumper Identification

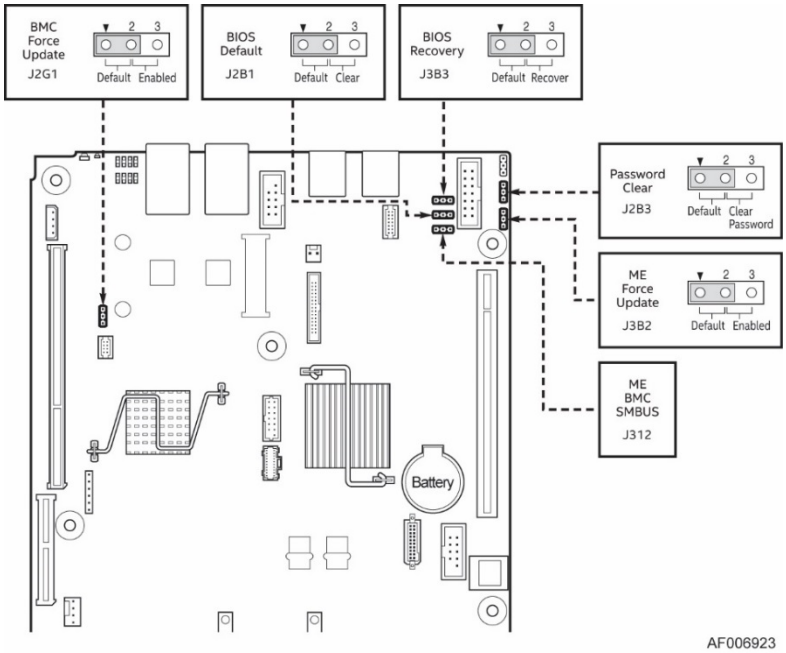


Figure 8. Jumper Identification

2.5 Mechanical Dimensions and Weight

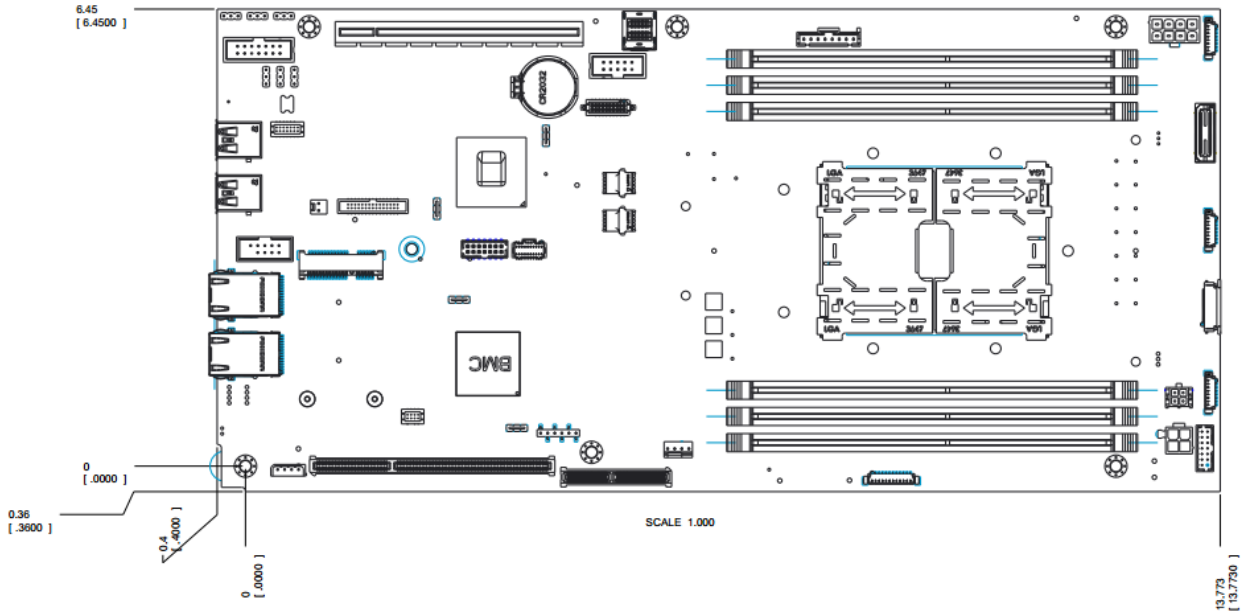


Figure 9. Server Board Dimension

The dimensions of the HNS7200AP compute module:

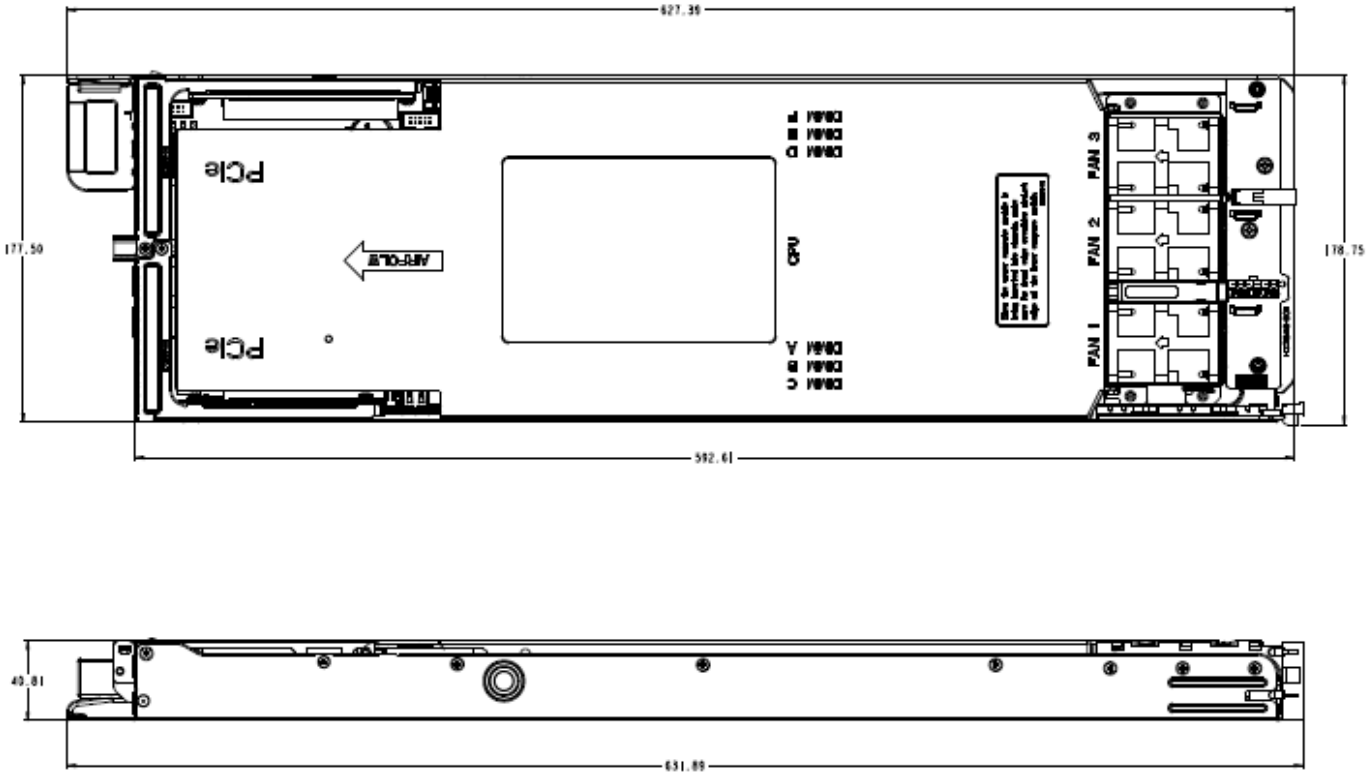


Figure 10. Compute Module Dimension

Approximate product weight is listed in the following table for reference. Variations are expected with real shipping products.

Table 4. Product Weight and Packaging

Product Code	Quantity per Box	Box Dimension (mm)	Net Weight	Package Weight
HNS7200AP	1	716X269X158	2.62kg	3.86kg

2.6 Product Architecture Overview

The Intel® Server Board S7200AP product family is a purpose built, rack-optimized, liquid cooling friendly server board used in a high-density rack system. It is designed around the integrated features and functions of the Intel® Xeon™ Phi™ processor product family, the Intel® C610 chipset, and other supporting components. This platform model provides significant leap in scalar and vector performance improvement and innovative memory architecture for high bandwidth and high capacity over previous server generations.

The half-width board size allows four boards reside in a standard multi-compute module 2U Intel® Server Chassis H2000G product family, for high-performance and high-density computing platforms.

The following diagram provides an overview of the server board architecture, showing the features and interconnects of each of the major subsystem components.

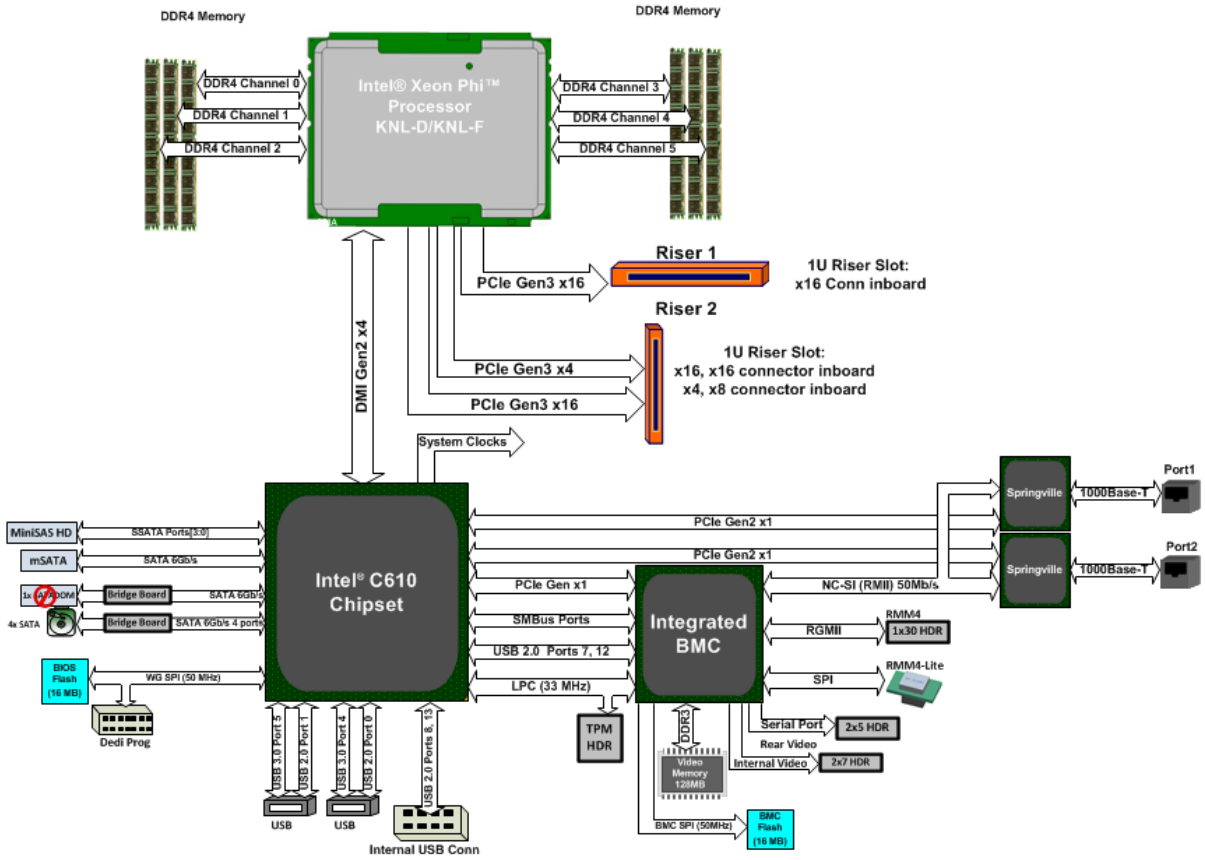


Figure 11. Intel® Server Board S720000AP Block Diagram

The Intel® Compute Module HNS7200AP product family provides a series of features including the power docking board, bridge boards, riser cards, fans, and the air duct.

2.7 Power Docking Board

The power docking board provides hot swap docking of 12V main power between the compute module and the server. It supports three dual rotor fan connections, 12V main power hot swap controller, and current sensing. The power docking board is intended to support the usage of the compute module with the Intel® Server Board S7200AP product family.

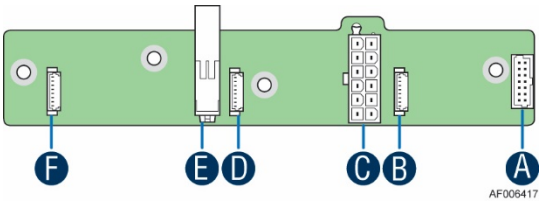


Table 5. Power Docking Pin-out

Label	Description
A	2x7-pin fan control connector
B	8-pin connector for fan 1
C	2x6-pin main power output connector
D	8-pin connector for fan 2
E	12-pin connector for main power input
F	8-pin connector for fan 3

2.8 Bridge Board

The platform supports a 6G SATA bridge board as the default option. There are no other bridge boards supported at this time.

The 6G SATA bridge board provides hot swap interconnect of all electrical signals to the backplane of the server chassis (except for main 12V power). It supports up to 4x lanes of SATA, and a type-A USB connector for USB flash device. One bridge board is used per one compute module. The bridge board is secured with screws to the compute module.

Table 6. Bridge Board Pinout

Label	Description
A	2x40-pin card edge connector (to the backplane)
B	USB 2.0 Type-A connector
C	2x40-pin card edge connector (to the bridge board connector on the server board)

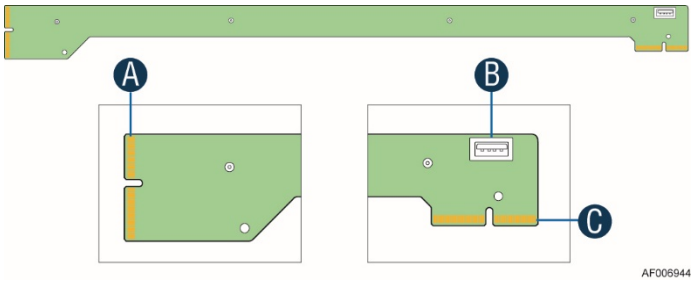


Figure 12. 6G SATA Bridge Board Overview

2.9 Riser Card Support

There are three types of riser cards:

- Riser Slot 1 riser card (for Riser slot 1 only)
- Riser Slot 2 riser card for Riser slot 2 only. Intended for use with Intel® Xeon™ Phi™ processor (KNL-**D**-72 Core) Bootable
- Riser Slot 2 riser card for Riser slot 2 only. Intended for use with Intel® Xeon™ Phi™ processor (KNL-**F**-72 Core) **Host Fabric Interface** Bootable

2.9.1 Riser Slot 1 x 16 Riser Card

The riser card for riser slot 1 has one PCIe* Gen 3.0 x16 slot.

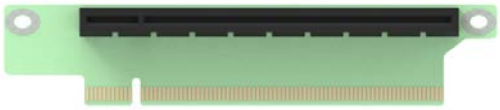


Figure 13. Riser Card for Riser Slot #1

2.9.2 Riser Slot 2 x16 Riser Card

This riser card for riser slot 2 has one PCIe* 3.0 x16 slot.

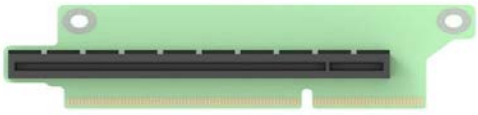


Figure 14. Riser Card for Riser Slot #2

2.9.3 Riser Slot 2 x8 Riser Card

This riser card for riser slot 2 has one PCIe* 3.0 x8 slot (x4 electrical- x8 mechanical)

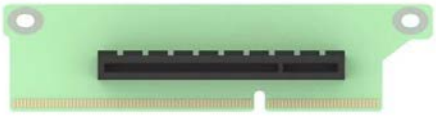


Figure 15. Riser Card for Riser Slot #2

Note: Riser Cards for Slot 2 have connectors which have been rotated 180°. The intent of this design is to prevent accidental insertion of previous generation riser boards into the platform.

2.10 Compute Module Fans

The cooling subsystem for the compute module consists of three 40 x 40 x 56 dual rotor fans. These components provide the necessary cooling and airflow.

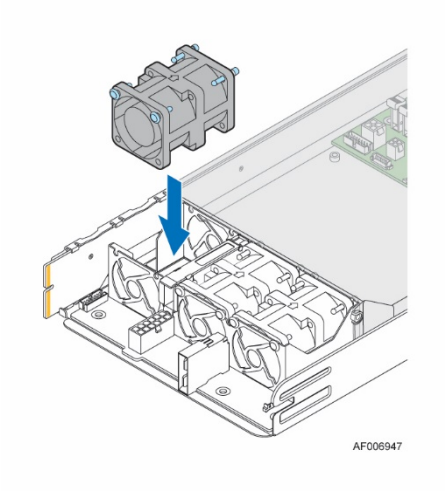


Figure 16. Compute Module Fans

Note: The Intel® Compute Moduel HNS7200AP product family does not support redundant cooling. If one of the compute module fan fails, it is recommended to replace the failed fan as soon as possible.

Each fan within the compute module can support multiple speeds. Fan speed changes automatically when internal ambient temperature of the system or processor temperature changes. The fan speed control algorithm is programmed into the server board's BIOS.

Each fan connector within the module supplies a tachometer signal that allows the BMC to monitor the status of each fan. If one of the fans fails, the status LED on the server board will light up.

The fan control signal is from the BMC on the mother board to the power docking board and then is distributed to three sets of dual rotor fans.

2.11 Air Duct

Each compute module requires the use of a transparent plastic air duct to direct airflow over critical areas within the compute module. To maintain the necessary airflow, the air duct must be properly installed. Before sliding the compute module into the chassis, make sure the air duct is installed properly.

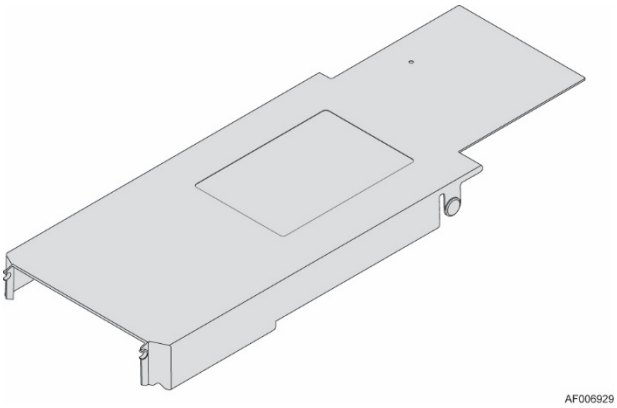


Figure 17. Air Duct

2.12 Intel® Remote Management Module 4 (Intel® RMM4) Lite

The optional Intel® RMM4 Lite is a small board that unlocks the advanced management features when installed on the server board.

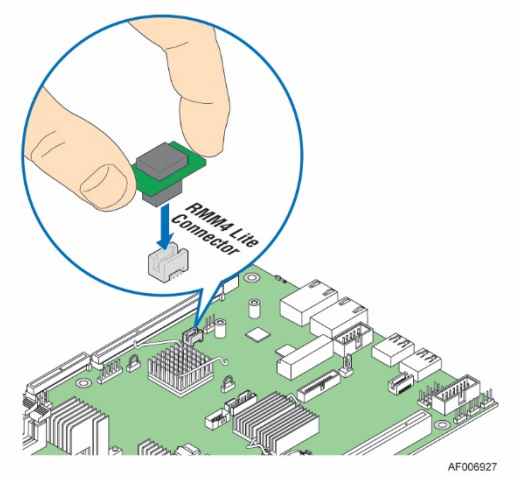


Figure 18. Intel® RMM4 Lite

2.13 System Software Overview

The server board includes an embedded software stack to enable, configure, and support various system functions. This software stack includes the System BIOS, Baseboard Management Controller (BMC) Firmware, Management Engine (ME) Firmware, and management support data including Field Replaceable Unit (FRU) data and Sensor Data Record (SDR) data.

The system software is pre-programmed on the server board during factory assembly, making the server board functional at first power-on after system integration. Typically, as part of the initial system integration

process, FRU and SDR data will have to be installed onto the server board by the system integrator to ensure the embedded platform management subsystem is able to provide best performance and cooling for the final system configuration. It is also not uncommon for the system software stack to be updated to later revisions to ensure the most reliable system operation. Intel makes periodic system software updates available for download at the following Intel website: http://downloadcenter.intel.com.

System updates can be performed in a number of operating environments, including the uEFI Shell using the uEFI-only System Update Package (SUP), or under different operating systems using the Intel® One Boot Flash Update Utility (OFU).

Reference the following Intel documents for more in-depth information about the system software stack and their functions:

- *Intel® Server System BIOS External Product Specification for Intel® Servers Systems supporting the* Intel® Xeon™ Phi™ processor *product family*
- *Intel® Server System BMC Firmware External Product Specification for Intel® Servers Systems supporting the Intel® Xeon® processor E5 V3 product family*

2.13.1 System BIOS

The system BIOS is implemented as firmware that resides in flash memory on the server board. The BIOS provides hardware-specific initialization algorithms and standard compatible basic input/output services, and standard Intel® Server Board features. The flash memory also contains firmware for certain embedded devices.

This BIOS implementation is based on the Extensible Firmware Interface (EFI), according to the Intel® Platform Innovation Framework for EFI architecture, as embodied in the industry standards for Unified Extensible Firmware Interface (UEFI).

The implementation is compliant with all Intel® Platform Innovation Framework for EFI architecture specifications, as further specified in the *Unified Extensible Firmware Interface Reference Specification*, Version 2.3.1.

In the UEFI BIOS design, there are three primary components: the BIOS itself, the Human Interface Infrastructure (HII) that supports communication between the BIOS and external programs, and the Shell which provides a limited OS-type command-line interface. This BIOS system implementation complies with HII Version 2.3.1, and includes a Shell.

2.13.1.1 BIOS Revision Identification

The BIOS Identification string is used to uniquely identify the revision of the BIOS being used on the server. The BIOS ID string is displayed on the Power On Self-Test (POST) Diagnostic Screen and in the <F2> BIOS Setup Main Screen, as well as in System Management BIOS (SMBIOS) structures.

The BIOS ID string for S7200AP series server boards is formatted as follows:

BoardFamilyID.OEMID.MajorVer.MinorVer.RelNum.BuildDateTime

Where:

- **BoardFamilyID** = String name to identify board family.

- "***S72C610***" is used to identify BIOS builds for Intel® S7200AP series Server Boards, based on the Second Generation Intel® Xeon Phi™ Processor Product Family and the Intel® C610 chipset family.

- **OEMID** = Three-character OEM BIOS Identifier, to identify the board BIOS "owner".
  - "***86B***" is used for Intel Commercial BIOS Releases.
- **MajorVer** = Major Version, two decimal digits 01-99 which are changed only to identify major hardware or functionality changes that affect BIOS compatibility between boards.
  - "***01***" is the starting BIOS Major Version for all platforms.
- **MinorVer** = Minor Version, two decimal digits 00-99 which are changed to identify less significant hardware or functionality changes which do not necessarily cause incompatibilities but do display differences in behavior or in support of specific functions for the board. For example, the Minor Version might be incremented when the Memory Reference Code changes, or when there is a minor change in the Firmware Volume layout.
  - *"**01**"* is the starting Minor Version for change in the BIOS ID Major Version. It resets to 01 whenever the Major Version is changed. The Minor Version may be changed at the discretion of the BIOS Development Team Leads.

    The sequence will be as in the following examples for Minor Version and Major Version:

    Major Version **1**, initially _____ = Major/Minor Version/RelNum "**.**0**1.**01**.**0001"
    Major Version **1**, Minor Version 1st change _ = Major/Minor Version/RelNum "**.**0**1.**02**.**0002"
    Major Version **1**, Minor Version 2nd change _ = Major/Minor Version/RelNum "**.**0**1.**03**.**0002"
    Major Version **2**, initially _____ = Major/Minor Version/RelNum "**.**0**2.**01**.**0001"
    Major Version **2**, Minor Version 1st change _ = Major/Minor Version/RelNum "**.**0**2.**02**.**0002"
- **RelNum** = Release Number, four decimal digits which are changed to identify distinct BIOS Releases. BIOS Releases are collections of fixes and/or changes in functionality, built together into a BIOS Update to be applied to a Server Board. However, there are "Full Releases" which may introduce many new fixes/functions, and there are "Point Releases" which may be built to address very specific fixes to a Full Release.

  The Release Numbers for Full Releases increase by 1 for each release. For Point Releases, the first digit of the Full Release number on which the Point Release is based is increased by 1. That digit is always 0 (zero) for a Full Release.

  Note *that Point Releases are* ***not*** *Standard Operating Procedure, and are built at the discretion of BIOS Development only if and when needed to address a significant fix that is required for one or more specific customers.*
  - *"**0001**"* is the starting Release Number for all platform BIOS Full Releases, for each distinct BoardFamilyID and OEMID. This number increases by 1 for each BIOS Full Release.

    The Release Number does not reset when the Major or Minor Version changes – it remains as a strictly monotonically increasing sequence from the initial platform BIOS Release until the platform goes to "End of Life" and BIOS releases cease.

    The Release Number sequence will be as in the following examples for Full Releases Point Releases:

    Full Release **1** _____ = RelNum "**.**000**1.**"
    Full Release **1** Point Release **1** _____ = RelNum "**.1**00**1.**"
    Full Release **1** Point Release **2** _____ = RelNum "**.2**00**1.**"
    Full Release **1** Point Release **3** _____ = RelNum "**.3**00**1.**"
    Full Release **2** _____ = RelNum "**.**000**2.**"

Full Release **2** Point Release **1** _____ = RelNum "**.1**00**2.**"

Full Release **3** _____ = RelNum "**.**000**3.**"

- **BuildDateTime** = Build timestamp – date and time in MMDDYYYYHHMM format:
  - **MM** = Two-digit month.
  - **DD** = Two-digit day of month.
  - **YYYY** = Four-digit year.
  - **HH** = Two-digit hour using 24-hour clock.
  - **MM** = Two-digit minute.

For example, the following BIOS ID string is displayed on the POST diagnostic screen for BIOS Major Version 01, Minor Version 01, and Full Release 0003 that is generated on August 13, 2011 at 8:56 AM:

S72C610.86B.01.01.0003.081320110856

The BIOS version in the Setup Utility Main Screen is displayed without the time/date timestamp, which is displayed separately as "Build Date":

S72C610.86B.01.01.0003

For the SMBIOS Type 0 BIOS Version field, the full BIOS ID string is used, including the complete timestamp.

2.13.1.2 Hot Keys Supported During POST

Certain "Hot Keys" are recognized during POST. A Hot Key is a key or key combination that is recognized as an unprompted command input, that is, the operator is not prompted to press the Hot Key and typically the Hot Key will be recognized even while other processing is in progress.

The BIOS recognizes a number of Hot Keys during POST. After the OS is booted, Hot Keys are the responsibility of the OS and the OS defines its own set of recognized Hot Keys.

The following table provides a list of available POST Hot Keys along with a description for each.

Table 7. POST Hot-Keys

HotKey Combination	Function
<F2>	Enter the BIOS Setup Utility
<F6>	Pop-up BIOS Boot Menu
<F12>	Network boot
<Esc>	Switch from Logo Screen to Diagnostic Screen
<Pause>	Stop POST temporarily

2.13.1.3 POST Logo/Diagnostic Screen

The Logo/Diagnostic Screen appears in one of two forms:

- If Quiet Boot is enabled in the <F2> BIOS setup, a "splash screen" is displayed with a logo image, which may be the standard Intel Logo Screen or a customized OEM Logo Screen. By default, Quiet Boot is enabled in BIOS setup, so the Logo Screen is the default POST display. However, if the logo is displayed during POST, the user can press <Esc> to hide the logo and display the Diagnostic Screen instead.
- If a customized OEM Logo Screen is present in the designated Flash Memory location, the OEM Logo Screen will be displayed, overriding the default Intel Logo Screen.

- If a logo is not present in the BIOS Flash Memory space, or if Quiet Boot is disabled in the system configuration, the POST Diagnostic Screen is displayed with a summary of system configuration information. The POST Diagnostic Screen is purely a Text Mode screen, as opposed to the Graphics Mode logo screen.
- If Console Redirection is enabled in Setup, the Quiet Boot setting is disregarded and the Text Mode Diagnostic Screen is displayed unconditionally. This is due to the limitations of Console Redirection, which transfers data in a mode that is not graphics-compatible.

2.13.1.4 BIOS Boot Pop-Up Menu

The BIOS Boot Specification (BBS) provides a Boot Pop-up menu that can be invoked by pressing the **<F6>** key during POST. The BBS Pop-up menu displays all available boot devices. The boot order in the pop-up menu is not the same as the boot order in the BIOS setup. The pop-up menu simply lists all of the available devices from which the system can be booted, and allows a manual selection of the desired boot device.

When an Administrator password is installed in Setup, the Administrator password will be required in order to access the Boot Pop-up menu using the **<F6>** key. If a User password is entered, the Boot Pop-up menu will not even appear – the user will be taken directly to the Boot Manager in the Setup, where a User password allows only booting in the order previously defined by the Administrator.

2.13.1.5 Entering BIOS Setup

To enter the BIOS Setup Utility using a keyboard (or emulated keyboard), press the <F2> function key during boot time when the OEM or Intel Logo Screen or the POST Diagnostic Screen is displayed.

The following instructional message is displayed on the Diagnostic Screen or under the Quiet Boot Logo Screen:

Press <F2> to enter setup, <F6> Boot Menu, <F12> Network Boot

---

Note: With a USB keyboard, it is important to wait until the BIOS “discovers” the keyboard and beeps – until the USB Controller has been initialized and the USB keyboard activated, key presses will not be read by the system.

When the Setup Utility is entered, the Main screen is displayed initially. However, in the event a serious error occurs during POST, the system will enter the BIOS Setup Utility and display the Error Manager screen instead of the Main screen.

2.13.1.6 BIOS Update Capability

In order to bring BIOS fixes or new features into the system, it will be necessary to replace the current installed BIOS image with an updated one. The BIOS image can be updated using a standalone IFLASH32 utility in the uEFI shell, or can be done using the OFU utility program under a given operating system. Full BIOS update instructions are provided when update packages are downloaded from the Intel web site.

2.13.1.7 BIOS Recovery

If a system is completely unable to boot successfully to an OS, hangs during POST, or even hangs and fails to start executing POST, it may be necessary to perform a BIOS Recovery procedure, which can replace a defective copy of the Primary BIOS.

The BIOS introduces three mechanisms to start the BIOS recovery process, which is called Recovery Mode:

- Recovery Mode Jumper – This jumper causes the BIOS to boot in Recovery Mode.
- The Boot Block detects partial BIOS update and automatically boots in Recovery Mode.
- The BMC asserts Recovery Mode GPIO in case of partial BIOS update and FRB2 time-out.

The BIOS Recovery takes place without any external media or Mass Storage device as it utilizes a Backup BIOS image inside the BIOS flash in Recovery Mode.

The Recovery procedure is included here for general reference. However, if in conflict, the instructions in the BIOS Release Notes are the definitive version.

When the *BIOS Recovery Jumper* is set, the BIOS begins by logging a "Recovery Start" event to the System Event Log (SEL). It then loads and boots with a Backup BIOS image residing in the BIOS flash device. This process takes place before any video or console is available. The system boots to the embedded uEFI shell, and a "Recovery Complete" event is logged to the SEL. From the uEFI Shell, the BIOS can then be updated using a standard BIOS update procedure, defined in Update Instructions provided with the system update package downloaded from the Intel web site. Once the update has completed, the recovery jumper is switched back to its default position and the system is power cycled.

If the BIOS detects a partial BIOS update or the BMC asserts Recovery Mode GPIO, the BIOS will boot up with Recovery Mode. The difference is that the BIOS boots up to the Error Manager Page in the BIOS Setup utility. In the BIOS Setup utility, boot device, Shell or Linux for example, could be selected to perform the BIOS update procedure under Shell or OS environment.

2.13.2 Field Replaceable Unit (FRU) and Sensor Data Record (SDR) Data

As part of the initial system integration process, the server board/system must have the proper FRU and SDR data loaded. This ensures that the embedded platform management system is able to monitor the appropriate sensor data and operate the system with best cooling and performance. The BMC supports automatic configuration of the manageability subsystem after changes have been made to the system's hardware configuration. Once the system integrator has performed an initial SDR/CFG package update, subsequent auto-configuration occurs without the need to perform additional SDR updates or provide other user input to the system when any of the following components are added or removed.

- Processors
- I/O Modules (dedicated slot modules)
- Storage modules such as a SAS module (dedicated slot modules)
- Power supplies
- Fans
- Fan options (e.g. upgrade from non-redundant cooling to redundant cooling)

- Intel® Xeon Phi™ co-processor cards
- Hot Swap Backplane
- Front Panel

Note: The system may not operate with best performance or best/appropriate cooling if the proper FRU and SDR data is not installed.

2.13.2.1 Loading FRU and SDR Data

The FRU and SDR data can be updated using a standalone FRUSDR utility in the uEFI shell, or can be done using the OFU utility program under a given operating system. Full FRU and SDR update instructions are provided with the appropriate system update package (SUP) or OFU utility which can be downloaded from the Intel web site.

2.13.3 Baseboard Management Controller (BMC) Firmware

See Platform Management.

3 Processor Support

The server board includes a single Socket-P1 (LGA 3647-1) processor socket that supports Intel® Xeon™ Phi™ processor (Knights Landing) product family, with a Thermal Design Power (TDP) of up to 230W.

Note: Previous generation Intel® Xeon™ Phi™ processors are not supported on the Intel® Server Boards described in this document.

Visit http://www.intel.com/support for a complete list of supported processors.

3.1 Intel® Xeon™ Phi™ processor (Knights Landing) Overview

Intel® Xeon™ Phi™ processor (Knights Landing) is a HPC optimized processor that is binary compatible with Intel® Xeon® processors. It has 72 Cores with each core supporting 4 threads giving a total of 288 threads in a socket. It doesn't have any coherent interfaces such as QPI/FSB and only supports single socket systems. The Cores are organized as Tiles. Each Tile contains 2 Cores, 4 Vector Processing Units (VPU), L2 Cache and CHA, providing a total of 36 Tiles in an Intel® Xeon™ Phi™ processor. Intel® Xeon™ Phi™ processor Uncore is an LLC-less design that uses the mesh architecture (instead of ring) to interconnect all the Tiles and Cores.

Intel® Xeon™ Phi™ processor is the first CPU to use the mesh architecture. Intel® Xeon™ Phi™ processor also integrates the Home Agent (HA) functionality into the Caching Agent (CA). The combined unit is called a CHA. Like the previous generation's processors' CBo (Cbox), the Intel® Xeon™ Phi™ processor CHA is a distributed caching architecture. Each CHA provides the Caching Agent and Home Agent functionality for a subset of the physical address. Since there is no coherent interface links like QPI, there is no QPI Phy/Link layer functionality in the Uncore. Intel® Xeon™ Phi™ processor does have the Ubox that serves as the centralized interrupt handler among other functionality similar to previous generation processors. Intel® Xeon™ Phi™ processor supports DDR4 memory technology. Like the Intel® Xeon® processors, the Intel® Xeon™ Phi™ processor (Knights Landing) also has Integrated IO (IIO). The Intel® Xeon™ Phi™ processor IIO supports x4 Gen2 DMI connection to PCH and up to x36 Gen3 PCIe.

3.1.1 New Technology MCDRAM

Intel® Xeon™ Phi™ processor introduces new memory technology called Multi Channel DRAM (MCDRAM). Intel® Xeon™ Phi™ processor is the first product to have this two level memory hierarchy. MCDRAM is a high band width memory designed to meet the bandwidth demand due to high number of threads present in this CPU. MCDRAM is an on-package memory that is connected to the Intel® Xeon™ Phi™ processor Embedded DRAM Controller (EDC) using high speed interface called On-Package IO (OPIO). The capacity of each of the memory type supported and even their presence varies depending on form factor and platform. These two memory types can be configured to operate in different modes namely Cache Mode, Flat Mode, and Hybrid Mode.

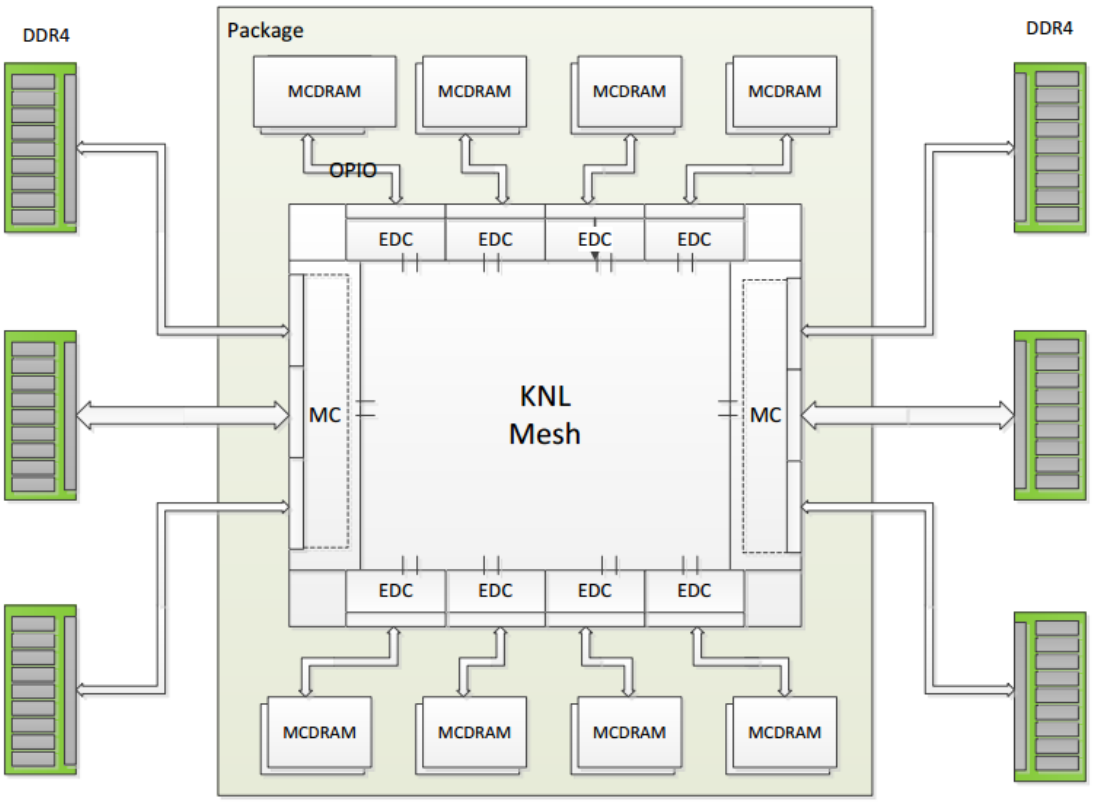


Figure 19. MCDRAM Block Diagram

MCDRAM is a separate die but it shares the package with the Intel® Xeon™ Phi™ processor die. The MCDRAM is connected to the Intel® Xeon™ Phi™ processor die using OPIO interface. The MCDRAM die is manufactured by other vendors and packaged into the Intel® Xeon™ Phi™ processor.

OPIO is a high speed interface designed to meet the high memory bandwidth needs of the large number of threads supported by the Intel® Xeon™ Phi™ processor. It is used to connect processor die to the MCDRAM die. The OPIO interface supports 8GT/s, 7.2GT/s and 6.4GT/s speeds. Data transfers may occur in both directions simultaneously.

Intel® Xeon™ Phi™ processor supports 3 different Memory Models: All2All, Sub NUMA Cluster 2/4, and Hemishpere /Quadrant. It also supports 3 different Memory Modes for 2LM namely Cache, Flat and Hybrid. All the permutations and combinations of these Memory Models and Modes are supported.

All members of the Second Generation Intel® Xeon Phi™ processor Product Family support the following Intel Technologies at a minimum.

- Intel® 64 Architecture
- Enhanced Intel SpeedStep® Technology
- Intel® Turbo Boost Technology
- Intel® Hyper-Threading Technology
- Intel® Virtualization Technology (Intel® VT-x)
- Intel® Virtualization Technology for Directed I/O (Intel® VT-d)
- Execute Disable Bit

- Intel® Advanced Vector Extensions (Intel® AVX)
- Advanced Encryption Standard New Instructions (AES-NI)
- Intel® Advanced Vector Extensions 2 (Intel® AVX2)
- Intel® Node Manager
- Intel® Quick Data Technology

Table 8. Intel® Xeon™ Phi™ processor Features

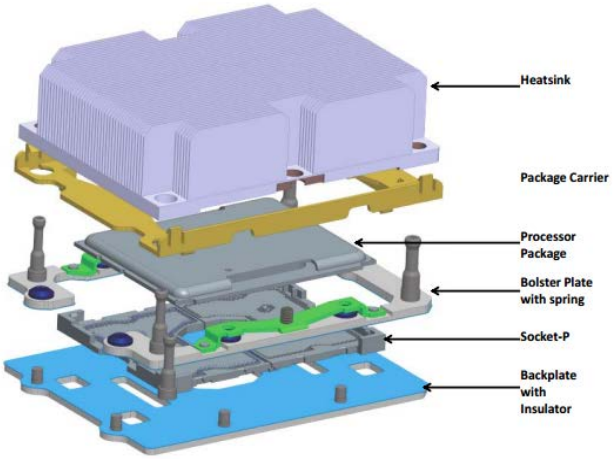
Feature	Description
Processor details	• 36 Tiles, 72 Cores, 288 Threads • 46-bit physical address, 48-bit virtual address • 32KB L1 Instruction Cache Per Core • 32KB L1 Data Cache per Core • 1MB L2 Unified Cache per Tile • No LLC
Technologies	• Intel® Virtualization Technology • Intel® Virtualization Technology (Intel® VT) for Directed I/O (Intel® VT-d) • Intel® Hyper-Threading with 4 threads per Core • Execute Disable Bit • Intel® Turbo Boost Technology (Intel® TBT)
Integrated Memory Controller	• 2 Integrated Memory Controllers (iMC) • 3 DDR4 channels per iMC • 1 DIMM per channel • RDIMM/LRDIMM • 2133 and 2400 MT/s • ECC • Standard DDR4 • 4GB, 8GB, 16GB, 32GB, 64GB • Up to 4 ranks per channel. 1, 2, or 4 ranks per DIMM • Integrated SMBus controller (in PCU, not in iMC) • Demand and Patrol Scrubbing • SDDC • Memory thermal throttling
EDC/MCDRAM	• Up to 8 EDC/OPIO channels • OPIO speed 6.4, 7.2 and 8GT/s • 1GB, 2GB, 4GB MCDRAM per OPIO • Thermal Throttling • MemBIST • Demand and Patrol Scrubbing • Error handling • ECC

Feature	Description
Integrated IO	• Fully compliant to PCIe 3.0 • Gen 1, Gen 2, and Gen 3 • Up to 36 lanes at Gen 3 speed • ATS 1.0
DMI2	• interface to Wellsburg PCH • x4 link width • Gen2 or Gen1 speed • APIC and MSI • Virtual Channel VC0
Power Management	• S State S0, S5 • C State: C0, C1, C1E, C6 • Memory CKE • PCIe ASPM L0s and L1
Thermal Management	• on die DTS • THERMTRIP and PROCHOT • OLTT and CLTT • Fan Speed Control • Memory thermal throttling with MEM_HOT

3.2 Processor Socket and Mechanical Retention Assembly

Each Intel® Server Board S7200AP supports the Intel® Xeon™ Phi™ processor socket, back plate assembly, bolster plate assembly, Processor Heatsink Module (PHM) consisting of the CPU package, Carrier assembly, TIM, and heatsink. The Intel® Xeon™ Phi™ processor does not use an Integrated Loading Mechanism (ILM). The Processor Heatsink Module (PHM), Bolster Plate, and Back Plate allow for secure placement of the processor and heatsink to the server board.

The following illustration identifies each sub-assembly component.



e

Figure 20. Processor Socket Assembly

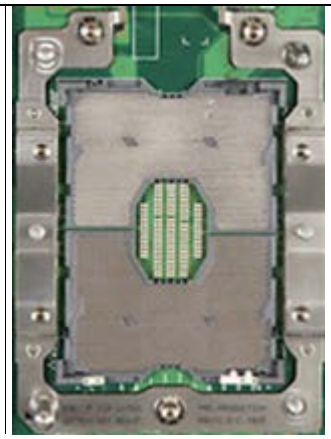


Figure 21. Processor Socket

Note: The pins inside the CPU socket are extremely sensitive. Other than the CPU, no object should make contact with the pins inside the CPU socket. **A damaged CPU Socket pin may render the socket inoperable, and will produce erroneous CPU or other system errors if used.**

3.3 Processor Thermal Design Power (TDP) Support

To allow optimal operation and long-term reliability of Intel processor-based systems, the processor must remain within the defined minimum and maximum case temperature (TCASE) specifications. Thermal solutions not designed to provide sufficient thermal capability may affect the long-term reliability of the processor and system. The server board described in this document is designed to support Intel® Xeon™ Phi™ processor (KNL-D Core) Bootable product family TDP guidelines up to and including 215W. The compute module described in this document is designed to support the Intel® Xeon™ Phi™ processor (KNL-F Core) Bootable with Host Fabric Interface (HFI) product family TDP guidelines up to and including 230W.

Disclaimer Note: Intel Corporation server boards contain a number of high-density VLSI and power delivery components that need adequate airflow to cool. Intel ensures through its own chassis development and testing that when Intel server building blocks are used together, the fully integrated system will meet the intended thermal requirements of these components. It is the responsibility of the system integrator who chooses not to use Intel developed server building blocks to consult vendor datasheets and operating parameters to determine the amount of airflow required for their specific application and environmental conditions. Intel Corporation cannot be held responsible if components fail or the server board does not operate correctly when used outside any of their published operating or non-operating limits.

3.4 Processor Initialization Error Summary

The following information describes conditions and recommended actions for all Intel® server boards and Intel server systems designed around the Intel® Xeon™ Phi™ processor family and Intel® C610 chipset product family architecture. The errors fall into one of the following categories:

- **Fatal:** If the system can boot, POST will halt and display the following message:

  "Unrecoverable fatal error found. System will not boot until the error is resolved
  Press <F2> to enter setup"

When the <F2> key on the keyboard is pressed, the error message is displayed on the Error Manager screen, and an error is logged to the System Event Log (SEL) with the POST Error Code.

This operation will occur regardless of whether the BIOS Setup option "Post Error Pause" is set to Enable or Disable.

If the system is not able to boot, the system will generate a beep code consisting of 3 long beeps and 1 short beep. The system cannot boot unless the error is resolved. The faulty component must be replaced.

The System Status LED will be set to a steady Amber color for all Fatal Errors that are detected during processor initialization. A steady Amber System Status LED indicates that an unrecoverable system failure condition has occurred.

- **Major:** If the BIOS Setup option for "Post Error Pause" is Enabled, and a Major error is detected, the system will go directly to the Error Manager screen in BIOS Setup to display the error, and logs the POST Error Code to SEL. Operator intervention is required to continue booting the system.

  If the BIOS Setup option for "POST Error Pause" is Disabled, and a Major error is detected, the Post Error Code may be displayed to the screen, will be logged to the BIOS Setup Error Manager, an error event will be logged to the System Event Log (SEL), and the system will continue to boot.
- **Minor:** An error message may be displayed to the screen, the error will be logged to the BIOS Setup Error Manager, and the POST Error Code is logged to the SEL. The system continues booting in a degraded state. The user may want to replace the erroneous unit. The POST Error Pause option setting in the BIOS setup does not have any effect on this error.

3.5 Processor Heatsink

One heatsink is included in the compute module package.

- 1U 80mm x 107mm Heatsink

Intel® Xeon™ Phi™ processor heatsink is made of is made of a copper base with aluminum fins. There are 50 fins in total, each 0.3 mm thick. The heatsink is integrated into the PHM which is attached to the bolster plate springs via four captive screws (T-30 Torx bit) on either side of the heatsink. The bolster plate is held in place around the socket by the back plate.

This heatsink is designed for optimal cooling and performance. To achieve better cooling performance, you must properly attach the heatsink bottom base with TIM (thermal interface material). The mechanical performance of the heatsink must satisfy mechanical requirement of the processor. To keep chipsets and VR temperature at or below maximum temperature limit, the heatsink is required if necessary.

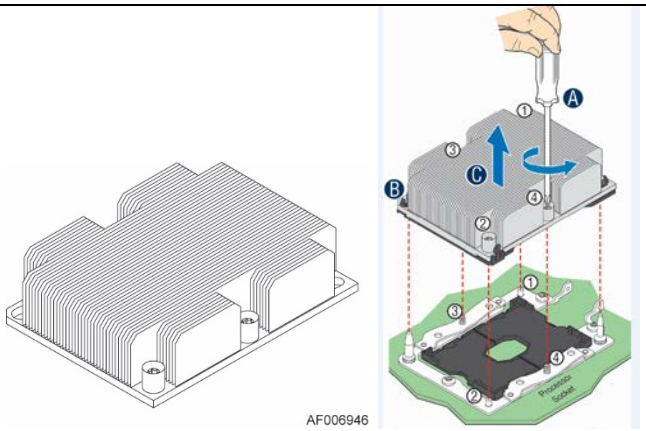


Figure 22. Processor Heatsink Overview

4 Memory Support

This chapter describes the architecture that drives the memory subsystem, supported memory types, memory population rules, and supported memory RAS features.

4.1 Memory Subsystem Architecture

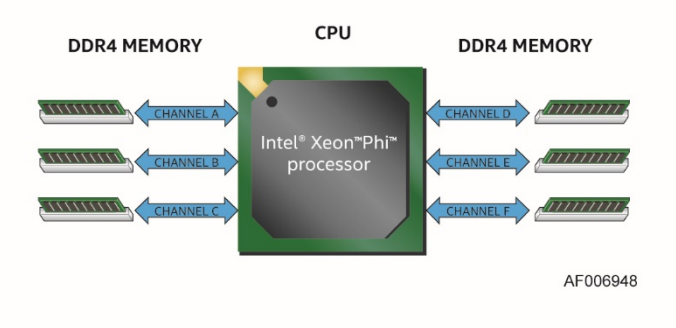


Figure 23. Integrated Memory Controller Functional Block Diagram

Note: This generation server board has support for DDR4 DIMMs only. DDR3 DIMMs are not supported on this generation server board.

Each processor includes two integrated memory controllers (IMC) capable of supporting three memory channels each. Each memory channel is capable of supporting one DIMM per channel. The processor IMC supports the following:

Registered DIMMs (RDIMMs), and Load Reduced DIMMs (LRDIMMs) are supported

DIMMs of different types may not be mixed – this is a Fatal Error in memory initialization

DIMMs composed of 4 Gb or 8 Gb Dynamic Random Access Memory (DRAM) technology

DIMMs using x4 or x8 DRAM technology

DIMMs organized as Single Rank (SR), Dual Rank (DR), or Quad Rank (QR)

Maximum of 8 ranks per channel

- There are 10 physical chip select signals per channel, 4 each for the first 2 DIMM slots and 2 more for the last DIMM slot. Since the IMC supports fewer logical ranks than the total number of chip selects and physical ranks, the address decode must map chip selects to logical ranks.

- DIMM sizes of 2 GB, 4 GB, 8 GB, or 16 GB depending on ranks and technology
- DIMM speeds of 2133 or 2400 MT/s (Mega Transfers/second)
- Only Error Correction Code (ECC) enabled RDIMMs or LRDIMMs are supported
- Only RDIMMs and LRDIMMs with integrated Thermal Sensor On Die (TSOD) are supported
- Memory RASM Support:
  - DRAM Single Device Data Correction (SDDCx4)
  - Error Correction Code (ECC)
  - Memory Disable and Map out for FRB
  - Data scrambling with command and address

- DDR4 Command/Address parity check and retry
- Intra-socket memory mirroring
- Memory demand and patrol scrubbing
- HA and IMC corrupt data containment

4.1.1 IMC Modes of Operation

The Integrated Memory Controller (IMC) in the Intel® Xeon Phi™ Processor Product Family includes 3 DDR4 memory channels, one DIMM slot per channel. Each Intel® Xeon Phi™ Processor has two IMCs providing a total of 6 total channels.

The two IMCs support DDR4 channels **0-5** and are silkscreened as channels A, B, C, D and E, and F.

iMC0 = Channel A/B/C = DIMM_A – DIMM_C

iMC1 = Channel D/E/F = DIMM_D – DIMM_F

4.1.2 Memory RASM Features

- **DRAM Single Device Data Correction (SDDC):** SDDC provides error checking and correction that protects against a single x4 DRAM device failure (hard-errors) as well as multi-bit faults in any portion of a single DRAM device on a DIMM (require lockstep mode for x8 DRAM device based DIMM).
- **Memory Disable and Map out for FRB:** Allows memory initialization and booting to OS even when a memory fault occurs.
- **Memory Demand and Patrol Scrubbing:** Demand scrubbing is the ability to write corrected data back to the memory once a correctable error is detected on a read transaction. Patrol scrubbing proactively searches the system memory, repairing correctable errors. It prevents accumulation of single-bit errors.
- **Error Correction Code (ECC) Memory:** ECC uses “extra bits” – 64-bit data in a 72-bit DRAM array – to add an 8-bit calculated “Hamming Code” to each 64 bits of data. These bits of data are distributed across the DRAM array so that even if an entire DRAM device fails, there will be enough “code bits” left to compute what the data was. This is the essence of the SDDCx4 capability. This additional SDDC/ECC encoding enables the memory controller to detect and report single or multiple bit errors when data is read, and to correct single-bit errors. There is a specific step in memory initialization for ECC memory in which all of memory is cleared to zeroes before the ECC function is enabled, in order to bring the ECC codes into agreement with memory contents.

  During operation, in the process of every fetch from memory, the data and ECC bits are examined for each 64-bit data + 8-bit ECC group. If the ECC computation indicates that a single bit Correctable Error (CE) has occurred, it is corrected and the corrected data is passed on to the processor. If a multi-bit error is detected, it cannot be corrected and it is handled as an Uncorrectable Error (UCE).

  Memory errors are processed in System Management Mode (SMM) by the SMI Handler. The SMI Handler attempts to log the error and pass control on to the Operating System error handlers before the termination of operations.

  Correctable ECC Errors are counted, and when a certain threshold value is reached, a Correctable Error event occurs. This is handled by the SMI Handler, much like an Uncorrectable Error except that it is not fatal and execution continues unless the Operating System error handlers terminate execution. In cases when a Correctable or Uncorrectable ECC Error event is generated, it is logged via the BMC’s SEL log.

4.2 Supported Memory

Table 9. DDR4 DIMM Support Guidelines

Memory Technology	DRAM Organization		DIMM Capacity		DIMM Capacity		DIMM Capacity		DIMM Capacity	
			SR	SR	DR	DR	QR/LR	QR/LR	8R/LR	8R/LR
			X4	X8	X4	X8	X4	X8	X4	X8
2Gb	512Mb x4	256Mb x8	4GB	2GB	8GB	4GB	16GB	8GB	32GB	16GB
4Gb	1024MB x4	512Mb x8	8GB	4GB	16GB	8GB	32GB	16GB	64GB	32GB
8Gb	2Gb x4	1Gb x8	16GB	8GB	32GB	16GB	64GB	32GB	*	64GB
16Gb	4Gb x4	2Gb x8	32GB	16GB	64GB	32GB	*	64GB	*	*

Note: Intel® Xeon™ Phi™ processor does not support DIMM capacities greater that 64GB

4.3 Memory Slot Identification and Population Rules

Note: Although mixed DIMM configurations are supported, Intel only performs platform validation on systems that are configured with identical DIMMs installed.

- Each installed processor provides six channels of memory. On the Intel® Server Board S7200AP product family, each memory channel supports one memory slot, for a total possible six DIMMs installed.
- All populated channels must have the same memory size.
- If both IMCs have populated channels, the channel population must be the same on both IMCs.
- All DIMMs must be DDR4 DIMMs.
- Mixing of LRDIMM with any other DIMM type is not allowed per platform.
- Mixing of DDR4 operating frequencies is not validated within a socket or across sockets by Intel. If DIMMs with different frequencies are mixed, all DIMMs run at the common lowest frequency.
- DIMM slots on any memory channel must be filled following the “farthest fill first” rule.
  - The DIMM slot farthest away from the processor socket must be filled first on any channel. This will always be designated on the board as Slot 1 for the channel.
- A maximum of 8 logical ranks can be used on any one channel, as well as a maximum of 10 physical ranks loaded on a channel.
- DIMM types (RDIMM, LRDIMM) must not be mixed within or across processor sockets.
  - This is a Fatal Error Halt in Memory Initialization.
- Mixing DIMMs of different frequencies and latencies is not supported within or across processor sockets.
  - If a mixed configuration is encountered, the BIOS will attempt to operate at the highest common frequency and the lowest latency possible.

- LRDIMM Rank Multiplication Mode and Direct Map Mode must not be mixed within or across processor sockets.
  - This is a Fatal Error Halt in Memory Initialization.
- In order to install 3 QR LRDIMMs on the same channel, they must be operated with Rank Multiplication as RM = 2.
  - This will make each LRDIMM appear as a DR DIMM with ranks twice as large.

4.3.1 Recommendations for Installing, Adding, or Upgrading Memory

- Try to balance the number and size of DIMMs on the channels of the board.
- Try to balance the number and size of DIMMs between two iMCs on the board.

4.3.2 S7200AP Series DIMM Population Sample Matrix

This section describes examples of various DIMM configurations on current Intel® S7200AP family Server Boards. This is a sampling, not an exhaustive enumeration. It uses examples of S7200AP family Server Boards.

The following is a description of the columns in below table:

A/B/C/D/E/F are DIMM slot labels.

Table 10. Sample DIMM Populations for S7200AP Family Server Boards

Configuration number	Total number of DIMMs	iMC0			iMC1		
		A	B	C	D	E	F
1	1	X					
2	2	X	X				
3	2				X	X	
4	2	X			X		
5	3	X	X	X			
6	3				X	X	X
7	4	X	X		X	X	
8	6	X	X	X	X	X	X
9	0						

Note:

1. DIMM size can be 4/8/16/32/64GB. No support for DIMMs larger than 64GB.
2. DIMMs_**A/B/C** belong to iMC0; DIMMs_**D/E/F** belong to iMC1.
3. For configuration 9, there could be no DDR memory installed as long as there is working MCDRAM and in flat mode.
4. Each memory slot should be populated with identical DDR4 DIMMs. Population of 5 DIMMs is not supported.

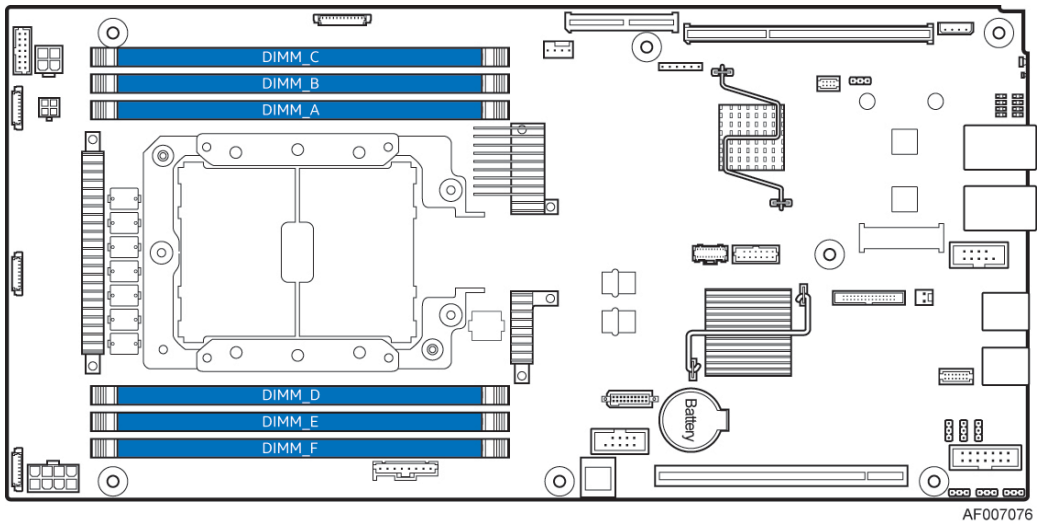


Figure 24. DIMM Slot Identification

4.4 System Memory Sizing and Publishing

The address space configured in a system depends on the amount of actual physical memory installed, on the RAS configuration, and on the PCI/PCIe configuration. RAS configurations reduce the memory space available in return for the RAS features. PCI/PCIe devices which require address space for Memory Mapped IO (MMIO) with 32-bit or 64- bit addressing, increase the address space in use, and introduce discontinuities in the correspondence between physical memory and memory addresses.

The discontinuities in addressing physical memory revolve around the 4GB 32-bit addressing limit. Since the system reserves memory address space just below the 4GB limit, and 32-bit MMIO is allocated just below that, the addresses assigned to physical memory go up to the bottom of the PCI allocations, then "jump" to above the 4GB limit into 64-bit space.

4.4.1 Effects of Memory Configuration on Memory Sizing

This BIOS supports 1 memory configuration – Independent Channel Mode. In some modes, memory reserved for RAS functions reduce the amount of memory available.

- **Independent Channel mode:** In Independent Channel Mode, the amount of installed physical memory is the amount of effective memory available. There is no reduction.

4.4.2 Publishing System Memory

There are a number of different situations in which the memory size and/or configuration are displayed. Most of these displays differ in one way or another, so the same memory configuration may appear to display differently, depending on when and where the display occurs. The address space configured in a system depends on the amount of actual physical memory installed, on the RAS configuration, and on the PCI/PCIe configuration. RAS configurations reduce the memory space available in return for the RAS features. PCI/PCIe devices which require allocations address space for Memory Mapped IO (MMIO) with 32-bit or 64- bit addressing increase the address space in use, and introduce discontinuities in the correspondence between physical memory and memory addresses.

- The BIOS displays the “Total Memory” of the system during POST if Quiet Boot is disabled in BIOS setup. This is the total size of memory discovered by the BIOS during POST, and is the sum of the individual sizes of installed DDR4 DIMMs in the system.
- The BIOS displays the “Effective Memory” of the system in the BIOS Setup. The term Effective Memory refers to the total size of all DDR4 DIMMs that are active (not disabled) and not used as redundant units (see Note below).
- The BIOS provides the total memory of the system in the main page of BIOS setup. This total is the same as the amount described by the first bullet above.
- If Quiet Boot is disabled, the BIOS displays the total system memory on the diagnostic screen at the end of POST. This total is the same as the amount described by the first bullet above.
- The BIOS provides the total amount of memory in the system by supporting the EFI Boot Service function.
- The BIOS provides the total amount of memory in the system by supporting the INT 15h, E820h function. For details, see the Advanced Configuration and Power Interface Specification.

Note: Some server operating systems do not display the total physical memory installed. What is displayed is the amount of physical memory minus the approximate memory space used by system BIOS components. These BIOS components include but are not limited to:

- ACPI (may vary depending on the number of PCI devices detected in the system)
- ACPI NVS table
- Processor microcode
- Memory Mapped I/O (MMIO)
- Manageability Engine (ME)
- BIOS flash

4.5 Memory Initialization

Memory Initialization at the beginning of POST includes multiple functions, including “discovery”, channel training, a certain amount of validation that the DIMM population is acceptable and functional, initialization of IMC and other hardware settings, and initialization of RAS configurations as applicable.

These are generally encapsulated in specialized Memory Reference Code (MRC) which is provided by the “silicon enabling” teams who work directly with the hardware designers and engineers responsible for the processor and Integrated Memory Controller (IMC) silicon.

There are several errors which can be detected in different phases of initializations. At this point in early POST, before system memory is available, serious errors that would prevent a system boot with data integrity will cause a System Halt with a beep code and a memory error code displayed in the Diagnostic LEDs.

Less fatal errors will cause a POST Error Code to be generated as a Major Error. This POST Error Code will be displayed in the Error Manager screen, and will also be logged to the System Event Log (SEL) to record the problem for later viewing or for accessing through Remote Management.

4.5.1 MCDRAM Initialization

BIOS is required to initialize the MCDRAM with the help of pCode and to train the OPIO channels without pCode involvement.

MCDRAM Init steps include powering up the MCDRAM and bringing it out of reset among other steps. OPIO channel training requires the read/write/request/response channels to be trained. BIOS will use the CPGC engine to do the OPIO training. There is no BIOS intervention needed in training the LMI channels. Also during OPIO training the MCDRAM memory is not accessible for data pattern storage/retrieval. So BIOS has to rely on MCDRAM loop back capability for OPIO training. MCDRAM initialization will be carried out by BIOS in two phases namely Cold Reset phase and Warm Reset phase. MCDRAM/OPIO training will be done only once for a given cold reset cycle and subsequent warm resets of a given cold reset cycle will use the previously trained values. However MCDRAM/OPIO will be retrained across cold reset cycles. The separation of initialization steps into Cold vs. Warm phase is determined based on if the step requires reset to take effect or not.

4.5.2 DDR4 Memory Initialization

The Memory Initialization process is essentially the same on all boards in the S7200AP family of Intel® Server Boards.

1. Memory initialization begins by determining for each channel which DIMM slots have DIMMs installed in them. By reading the Serial Presence Detect (SPD) information from an SEEPROM on the DIMM, the type, size, latency, and other descriptive parameters for the DIMM can be acquired.

   Potential Error Cases:

   - DIMM SPD does not respond – The DIMM will not be detected, which could result in a "No usable memory installed" *Fatal Error Halt **0xE8*** if there are no other detectable DIMMs in the system. The undetected DIMM could result later in an invalid configuration if the "no SPD" DIMM is in Slot 1 or 2 ahead of other DIMMs on the same channel.
   - DIMM SPD read error – This DIMM will be disabled. *POST Error Codes **856x** "SPD Error"* and ***854x** "DIMM Disabled"* will be generated. If all DIMMs are failed, this will result in a *Fatal Error Halt **0xE8***.

     All DIMMs on the channel in higher-numbered sockets behind the disabled DIMM will also be disabled with a *POST Error **Code 854x** "DIMM Disabled"* for each. This could also result in a "No usable memory installed" *Fatal Error Halt **0xE8***.
   - No usable memory installed – If no usable (not failed or disabled) DIMMs or MCDRAMs can be detected as installed in the system, this will result in a *Fatal Error Halt **0xE8***. Other error conditions which cause DIMMs to fail or be disabled so they are mapped out as unusable may result in causing this error when no usable DIMM remains in the memory configuration.

2. For each memory channel, once the DIMM SPD parameters have been read, they are checked to verify that the DIMMs on the channel are installed in a valid configuration. This involves checking DIMM type, DRAM type and organization, DRAM rank organization, DIMM speed and size, ECC capability, and in which memory slots the DIMMs are installed. An invalid configuration may cause the system to halt.

   Potential Error Cases:

   - Invalid DIMM (type, organization, speed, size) – If a DIMM is found that is not a type supported in this system, this will result in a *POST Error Code **8501** "DIMM Population Error"*, and a "Population Error" *Fatal Error Halt **0xED***.

- Invalid DIMM Installation – The DIMMs are installed incorrectly on a channel, not following the "Fill Farthest First" rule (Slot 1 must be filled before Slot 2, Slot 2 before Slot 3). This will result in a *POST Error Code **8501** "DIMM Population Error"* with the channel being disabled, and all DIMMs on the channel will be disabled with a *POST Error **Code 854x** "DIMM Disabled"* for each. This could also result in a "No usable memory installed" *Fatal Error Halt **0xE8***.
- Invalid DIMM Population – A QR RDIMM, or a QR LRDIMM in Direct Map mode which is installed in Slot3 on a 3 DIMM per channel server board is not allowed. This will result in a *POST Error Code **8501** "DIMM Population Error"* and a "Population Error" *Fatal Error Halt **0xED***.
  - ⇨ However, 3 QR LRDIMMs on a channel *is* an acceptable configuration if operating in Rank Multiplication mode with RM = 2 or 4. In this case each QR LRDIMM appears to be a DR or SR DIMM.
- Mixed DIMM Types – A mixture of RDIMMs and/or LRDIMMs is not allowed. A mixture of LRDIMMs operating in Direct Map mode and Rank Multiplication mode is also not allowed. This will result in a *POST Error Code **8501** "DIMM Population Error"* and "Population Error" *Fatal Error Halt **0xED***.
  - ⇨ Mixed DIMM Parameters – Within an RDIMM or LRDIMM configuration, mixtures of valid DIMM technologies, sizes, speeds, latencies, etc., although not supported, will be initialized and operated on a best efforts basis, if possible.
- No usable memory installed – If no enabled and available memory remains in the system, this will result in a *Fatal Error Halt **0xE8***.

3. The Integrated Memory Controller registers are programmed at the controller level and the memory channel level. Using the DIMM operational parameters read from the SPD of the DIMMs on the channel, each channel is trained for optimal data transfer between the IMC and the DIMMs installed on the channel.

   Potential Error Cases:

   - Channel Training Error – If the Data/Data Strobe timing on the channel cannot be set correctly so that the DIMMs can become operational, this results in a momentary *Error Display **0xEA***, and the channel is disabled. All DIMMs on the channel are marked as disabled, with *POST Error **Code 854x** "DIMM Disabled"* for each. If there are no populated channels which can be trained correctly, this becomes a *Fatal Error Halt **0xEA***.

4. Thermal (CLTT) and power throttling are initialized.

   Potential Error Cases:

   - CLTT Structure Error – The CLTT initialization fails due to an error in the data structure passed in by the BIOS. This results in a *Fatal Error Halt **0xEF***.

5. Once the memory is functional, a memory test is executed. This is a hardware-based Built In Self-Test (BIST) which confirms minimum acceptable functionality. Any DIMMs which fail are disabled and removed from the configuration.

   Potential Error Cases:

   - Memory Test Error – The DIMM has failed BIST and is disabled. *POST Error Codes **852x** "Failed test/initialization"* and ***854x** "DIMM Disabled"* will be generated for each DIMM that fails. Any DIMMs installed on the channel behind the failed DIMM will be marked as disabled, with *POST Error Code **854x** "DIMM Disabled"*. This results in a momentary *Error Display **0xEB***, and if all DIMMs have failed, this will result in a *Fatal Error Halt **0xE8***.
   - No usable memory installed – If no enabled and available memory remains, this will result in a *Fatal Error Halt **0xE8***.

6. Next, the IMC registers and other hardware settings are finalized. The memory map and interleaving are set up.
7. The ECC functionality is enabled after all of memory has been cleared to zeroes to make sure that the data bits and the ECC bits are in agreement.
8. The RAS mode, if any, is initialized. The DIMM configuration is validated for the RAS mode which has been selected. If the enabled DIMM configuration is acceptable for the RAS mode selected, then the necessary register settings are done and the RAS mode is started into operation.

   Potential Error Cases:

   - RAS Configuration Failure – If the DIMM configuration is not valid for the RAS mode which was selected, then the operating mode falls back to Independent Channel Mode, with *POST Error Code **8500** "Selected RAS Mode could not be configured"*. In that case there will also be a *"RAS Configuration Disabled"* SEL entry for "RAS Configuration Status" (BIOS Sensor 02/Type 0Ch/Generator ID 01).

At this point, the discovery, test/verification, and initialization of memory and memory RAS mode is complete.

4.5.3 Memory Initialization Error Summary

This section has a summary of the Memory Initialization Fatal Error codes for each of the two different Memory Reference Code initializations used by the BIOS. Their Fatal Error lists differ considerably.

4.5.3.1 Memory Initialization Error Summary – S7200AP

There are three kinds of error indications from the memory and memory RAS initialization process. There are POST Error Codes, SEL log entries, and Fatal Error Halts.

The SEL log entries are either logging POST Error Codes from the table below or logging RAS Configuration Status.

The POST Error Codes either show system conditions, like "DIMM Population Error", or indication conditions specific to a DIMM, like "DIMM Disabled". A short list of the relevant POST Error Codes follows. Complete information about POST Error Codes in the appendix.

Table 11. POST Error Codes in Memory Initialization

Error Code	Error Message	Response
8500	Memory component could not be configured in the selected RAS mode	Major
8501	DIMM Population Error	Major
8520-853F 85C0-85CF	DIMM_xy failed test/initialization	Major
8540-855F 85D0-85DF	DIMM_xy disabled	Major
8560-857F 85E0-85EF	DIMM_xy encountered a Serial Presence Detection (SPD) failure	Major

The Fatal Errors for memory initialization for S7200AP series server boards, which use processors from the Second Generation Intel® Xeon Phi™ Processor Product Family, are shown in the table below.

These Fatal Errors terminate POST in a Halt with a memory error beep code sounded and the Error Code displayed in the Diagnostic LEDs.

Fatal Error Halts do not change the System Status LED, and they do not get logged as SEL Events. The system is halted because memory was not properly configured or initialized, so there is no opportunity to communicate status or logging information to the BMC.

Table 12. MRC Fatal Error Halts

Error Code	Fatal Error Code Explanation (with MRC Internal Minor Code)
0xE8	No Usable Memory Error: 01h = No memory was detected via SPD read, or invalid config that causes no operable memory. 02h = Memory DIMMs on all channels of all sockets are disabled due to hardware memtest error. 03h = No memory installed. All channels are disabled.
0xE9	Memory is locked by Intel® Trusted Execution Technology and is inaccessible.
0xEA	DDR4 Channel Training Error: 01h = Error on read DQ/DQS (Data/Data Strobe) init 02h = Error on Receive Enable 03h = Error on Write Leveling 04h = Error on write DQ/DQS (Data/Data Strobe)
0xEB	Memory Test Failure: 01h = Software memtest failure. 02h = Hardware memtest failed. 03h = Hardware Memtest failure in Lockstep Channel mode requiring a channel to be disabled. *This is a fatal error which requires a reset and calling MRC with a different RAS mode to retry.*
0xED	DIMM Configuration/Population Error: 01h = Different DIMM types (RDIMM, LRDIMM) are detected installed in the system. 02h = Violation of DIMM population rules. 03h = The 3rd DIMM slot cannot be populated when QR DIMMs are installed. 04h = UDIMMs are not supported. 05h = Unsupported DIMM Voltage.
0xEF	Indicates a CLTT table structure error.

4.5.4 Memory Thermal, Acoustic, and Power Management

The memory subsystem is a major consumer of power and producer of heat in any system. As such, it may be closely managed with respect to both power consumption and heat generation. Either or both of these may be enabled, depending of the parameters of the workload. For example, in a case where computing performance were more important than controlling power consumption, memory power management would probably be disabled.

4.5.4.1 Thermal and Acoustic Management

For management of the system's thermal state, there is a balance to configure between fan speed in the chassis and memory power throttling for thermal control. The system's acoustic behavior is also factored in, because raising the fan speed produces more sound. Throttling the memory power instead reduces heat dissipation with no sound produced but at the cost of reducing memory performance.

The parameters for Fan Speed Control and memory thermal throttling are configured in a separate Setup screen.

4.5.4.2 Power Management

An important feature in this generation of servers is the Intel® Intelligent Power Node Manager. Working in concert with the Memory Reference Code (MRC) and the Manageability Engine (ME) and the BMC, the Node manager can monitor and control power consumption across a group of systems, adjusting memory power throttling across the entire group, using algorithms like Running Average Power Limiting.

An important feature in this generation of servers is the Intel® Intelligent Power Node Manager. Working in concert with the Memory Reference Code (MRC) and the Manageability Engine (ME) and the BMC, the Node manager can monitor and control power consumption across a group of systems, adjusting memory power throttling across the entire group, using algorithms like Running Average Power Limiting.

The selection of whether or not to allow the system to participate in power management is controlled by a BIOS setting, “Memory Power Optimization”.

5 Server Board I/O

The server board input/output features are provided via the embedded features and functions of several onboard components including: the Integrated I/O Module (IIO) of the Intel® Xeon™ Phi™ processor product family, the Intel® C610 chipset, the Intel® Ethernet controller I320, and the I/O controllers embedded within the Emulex* Pilot-III Management Controller.

See the block diagram for an overview of the features and interconnects of each of the major subsystem components.

5.1 PCI Express* Support

The Integrated I/O (IIO) module of the Intel® Xeon™ Phi™ processor product family provides the PCI express interface for general purpose PCI Express* (PCIe) devices at up to Gen 3 speeds. The primary responsibility of the Knights Landing Processor Integrated I/O (IIO) module is to translate between the PCI Express protocol and the coherent Mesh. The IIO supports a wide variety of system configurations. The IIO module incorporates the sub-modules and provides the following capabilities.

The IIO module provides the following PCIe Capabilities and Features:

- The Intel® Xeon™ Phi™ processor implements 36 lanes of PCIe Gen 3 interface, comprising of 3 ports. The first and second ports are x16 PCIe that can be bifurcated to 2 (x8), 3 (two x4 and x8) or 4 (x4) Root Ports each, and the third port is a x4 PCIe that cannot be bifurcated. •x4 DMI2 interface
- All processors support a x4 DMI2 lane which can be connected to a PCH.

  IOxAPIC controller used to convert legacy interrupts from I/O devices into messages to the CPU's Local APIC
- I/O Virtualization Logic (Intel® VT-d2)
- PCI Express Features
  - Compliant with the PCI Express Base Specification Revision 3.0
  - 2.5 GT/s (Gen1) and 5 GT/s (Gen2) and 8 GT/s (Gen3)
  - x16 PCI Express Gen3 interface supports up to 4 x4 controllers and is configurable to 4x4 links, 2x8, 2x4 and 1x8, or 1x16
  - Full support for software-initiated PCI Express power management
  - Address Translation Services (ATS 1.0)
  - PCIe Atomic Operations Completer Capability
  - Autonomous Linkwidth
- Direct Media Interface 2 (DMI2) Features
  - One x4 DMI Gen2 link interface supporting 5 GT/s (PCIe physical layer) peak bandwidth
  - Support for Intel® C10 Series PCH
  - Autonomous Linkwidth
- Intel I/O Acceleration Technology (Intel ®I/OAT)
  - Gen3 I/O caching hints used for Intel ®I/OAT
  - ATR enhancements (SocketID returned on reads)
- I/O Virtualization Intel

- ®Virtualization Technology (Intel® VT) for Directed I/O
- (Intel®VT-d2) Features
    - Original Intel®
- VT-d features
    - Improved invalidation architecture
    - End point caching support (ATS)
    - Interrupt remapping
    - 4K/2M/1G super page support
- RAS Features
    - PECI to read and write Device registers within the IIO module
    - Supports PCI Express Base Specification, Revision 2.0 and Revision 3.0 CRC with link-level retry
    - Advanced Error Reporting (AER) capability for PCI Express link interfaces
- Power Management Support Features
    - PCI Express Link states (L0, L1, L3)
    - System states (S0, S5)
    - Coarse Grained clock gating
    - Autonomous Linkwidth
- Security Features
    - Trusted Platform Module (TPM) 2.0 and 1.2 for the Knights Landing Processor
    - Intel® VT-d for server security
- Features not supported in the IIO
    - PHOLD is not supported in the IIO
    - Posted Interrupts are not supported in the IIO
    - ROL is not supported in the IIO
    - IOSAV is not supported in the processor

5.1.1 PCIe Enumeration and Allocation

The BIOS assigns PCI bus numbers in a depth-first hierarchy, in accordance with the PCI Local Bus Specification, Revision 2.2. The bus number is incremented when the BIOS encounters a PCI-PCI bridge device.

Scanning continues on the secondary side of the bridge until all subordinate buses are assigned numbers. PCI bus number assignments may vary from boot to boot with varying presence of PCI devices with PCI-PCI bridges.

If a bridge device with a single bus behind it is inserted into a PCI bus, all subsequent PCI bus numbers below the current bus are increased by one. The bus assignments occur once, early in the BIOS boot process, and never change during the pre-boot phase.

The BIOS resource manager assigns the PIC-mode interrupt for the devices that are accessed by the legacy code. The BIOS ensures that the PCI BAR registers and the command registers for all devices are correctly set

up to match the behavior of the legacy BIOS after booting to a legacy OS. Legacy code cannot make any assumption about the scan order of devices or the order in which resources are allocated to them.

The BIOS automatically assigns IRQs to devices in the system for legacy compatibility. A method is not provided to manually configure the IRQs for devices.

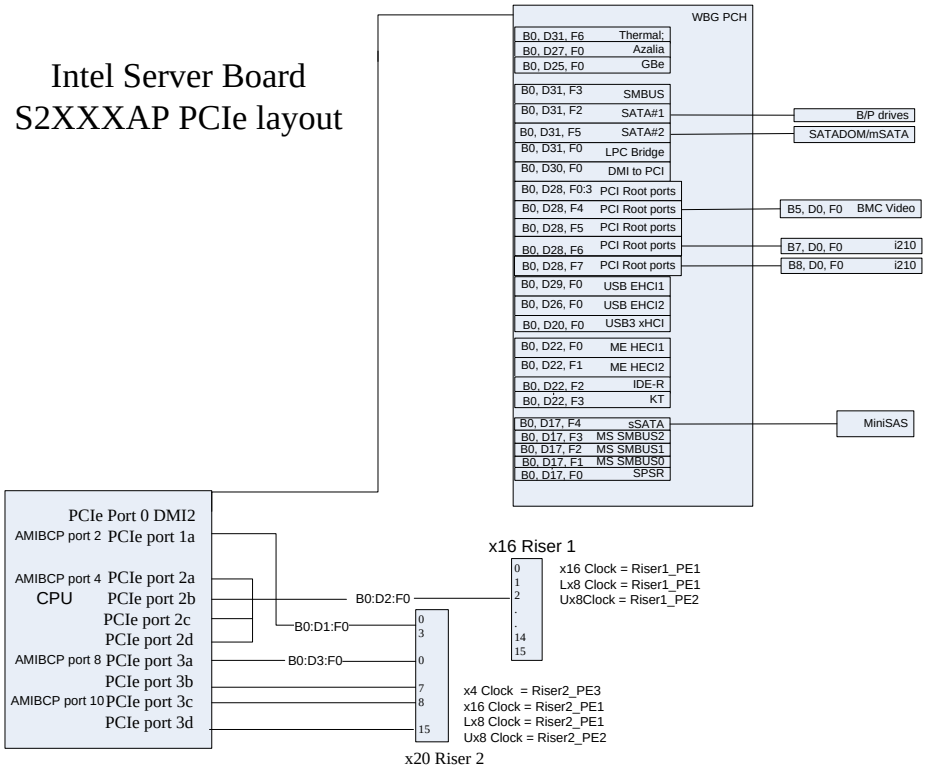


Figure 25. PCIe Bus/Device/Function Map

5.1.2 PCIe Non-Transparent Bridge (NTB)

PCI Express Non-Transparent Bridge (NTB) is unsupported on this platform.

5.2 Add-in Card Support

The following sub-sections describe the server board features that are directly supported by the processor IIO module. These include the Riser Card Slots, Network Interface, and optional boards.

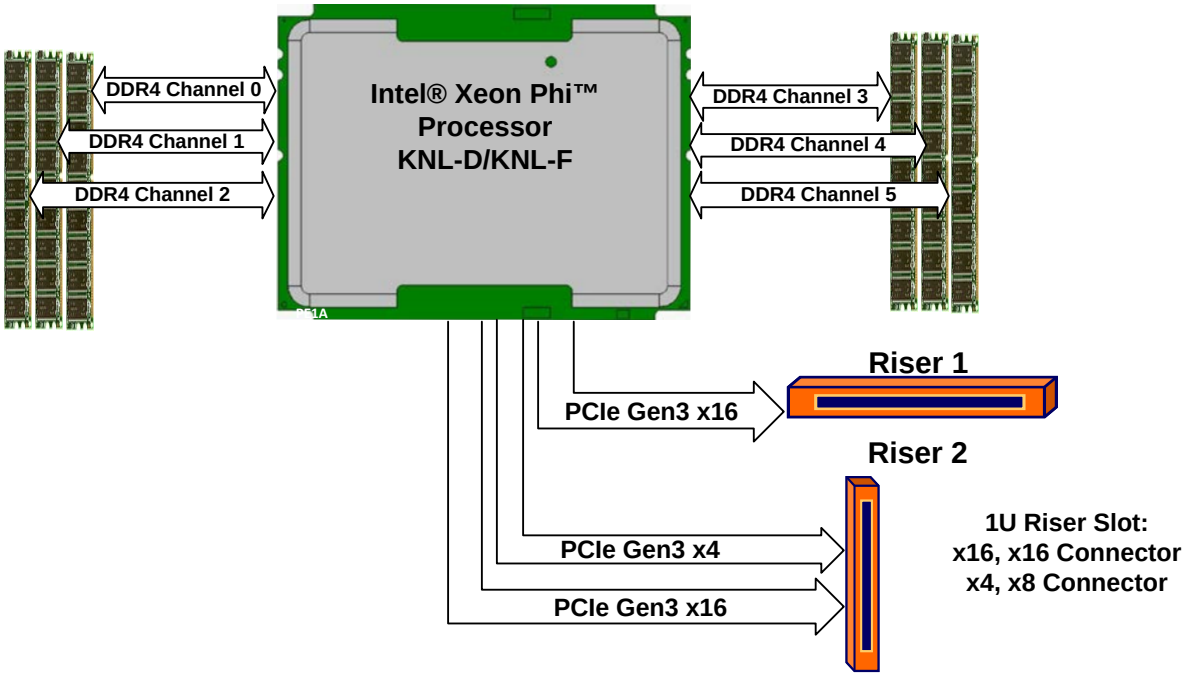


Figure 26. Add-in Card Support Block Diagram (S7200AP)

5.2.1 Riser Card Support for Add-in Cards

The server board includes features for concurrent support of several add-in card types including: PCIe add-in cards via 2 riser cards slots (RISER_SLOT_1 and RISER_SLOT_2). The distance between riser slots allows for simultaneous support of LP PCIe adapter cards. The following illustration identifies the location of the onboard connector features and general board placement for add-in modules and riser cards.

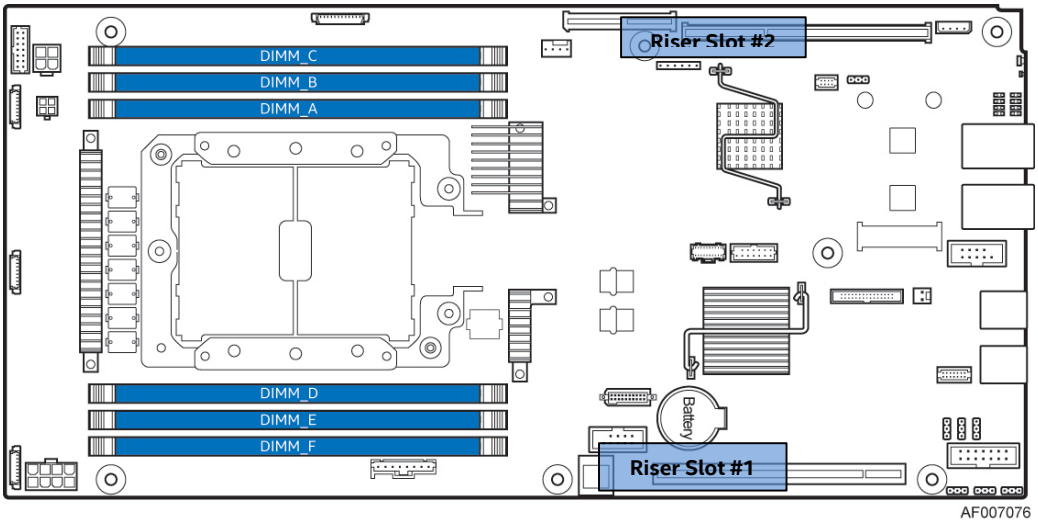


Figure 27. Server Board Riser Slots (S7200AP)

Below is the PCIe* port connection from the CPU to the Risers:

Table 13. PCIe Port connections

Device	Riser	Physical Conn	IOU#	PORT	Electrical Width
CPU	Slot1	X16	0	2A-2D	1x16, 2x8,
CPU	Slot2	X24	1	3A-3D	1x16, 2x8
CPU	Slot2	X24	2	1A	x4

5.2.2 Host Fabric Interface Support

The Intel® Server Board S7200AP supports the bootable Intel® Xeon™ Phi™ processor (KNL-F) with integrated Intel® Omni-Path Host Fabric Interface (HFI), to include associated add-in adapter Intel® Through Fabric (IFT) carrier. Intel® Omni-Path (IOP) Fabric enabled processors have two 25 GB/s fabric I/O ports, up to 50 GB/s bi-directional peak total bandwidth. The Two x16 PCIe Gen3 ports are replaced by two IOP fabric ports. Processors with integrated on-package fabric route x4 PCIe Gen3 lanes for use at the platform level. The CPUs have x32 lanes of Gen3 PCIe are routed from the CPU die to the on-package fabric chip to support the IOP fabric interface.

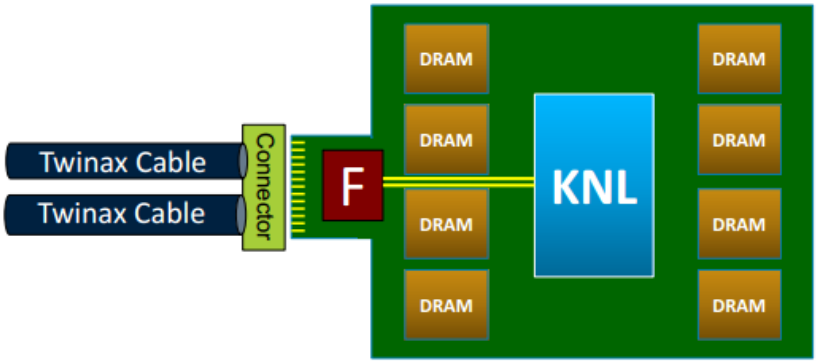


Figure 28. Intel® Xeon™ Phi™ processor (KNL-F) with integrated Intel® Omni-Path

5.2.3 Intel® Fabric Through (IFT) Carrier

The Intel® Fabric Through (IFT) Carrier board allows external access to the Intel® Xeon™ Phi™ processor (KNL-F) Host Fabric Interface (HFI) from the CPU socket. The external connection access is accomplished through use of two QSFP+28 style connections. LEDs are present to provide status indications. The IFT CARRIER is mounted in PCIe slot 1, however, no PCIe related signals are passed through as the slot. The slot is used only for mechanical mounting and provides power and SMBus signals. For external connection, the IFT CARRIER will be using two QSFP+28 style connectors along with their associated cages and heatsink. The signal definition of these QSFP+28 style connectors consists of the high speed diff pairs, miscellaneous side band signals, and 3.3V power. As shown in Figure 30, The IFT Carrier mounts into RISER 1 slot on the server baseboard. The riser slot has 3 pins assigned for 3.3V power delivery, which provide ups to 3A for the carrier card.

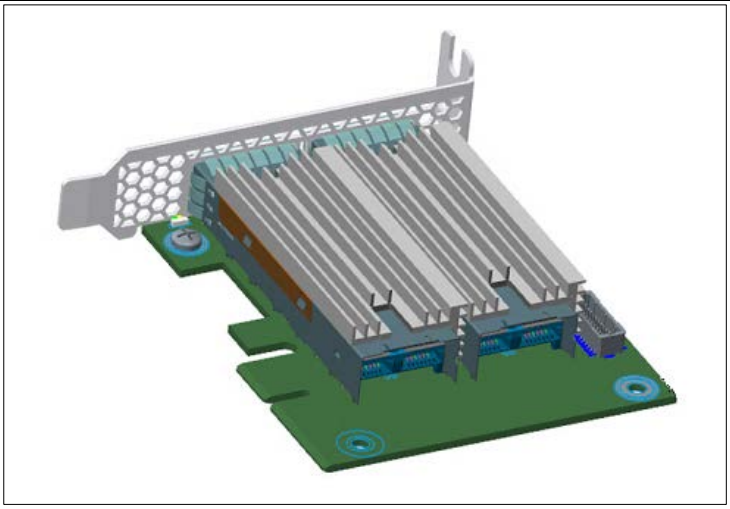


Figure 29. IFT Carrier Card

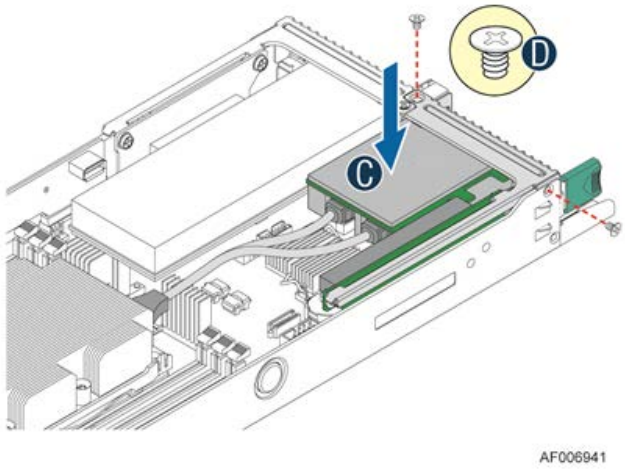


Figure 30. IFT Card Mounting

5.2.4 Intel® Fabric Passive (IFP) Cable

The Intel® Fabric Passive (IFP) Internal Cable Assembly enables high speed, low loss data connections between the Intel® Xeon™ Phi™ processor (KNL-F) and chassis connections to an external network interface via QSFP +28 style connectors. IFP internal cable assembly enables a direct connection from an Intel processor to an Intel Fabric network. It connects an Intel processor package using a 54-pin Linear Edge Connector (LEC54x), and connects to Intel Fabric network using a 28-pin plug that mates to the Intel® Fabric Through connector. See Figure 31 for an illustration of the IFP internal cable assembly. Connectivity between the LEC54-Pin connector and the 28-Pin IFP plug is made via twinax cable. IFP internal cable assemblies are designed to support 25 Gbps data transfer rates.

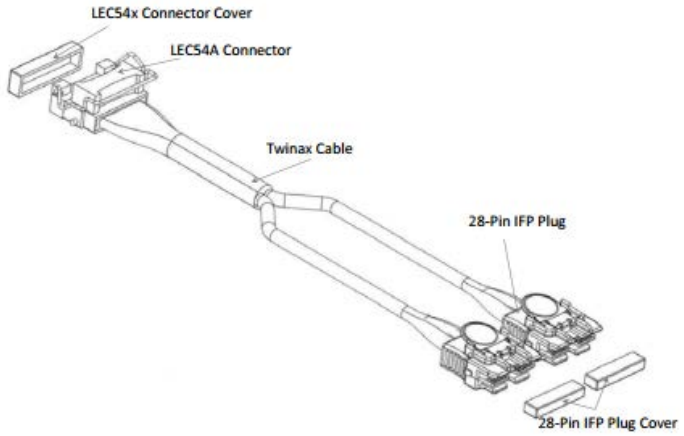


Figure 31. Intel® Fabric Passive (IFP) Internal Cable

The following image depicts the full system connection of the Intel® Xeon™ Phi™ processor (KNL-F), Intel® Fabric Passive (IFP) Internal Cable, and the Intel® Fabric Through (IFT) Carrier Assembly to the rear of the server system.

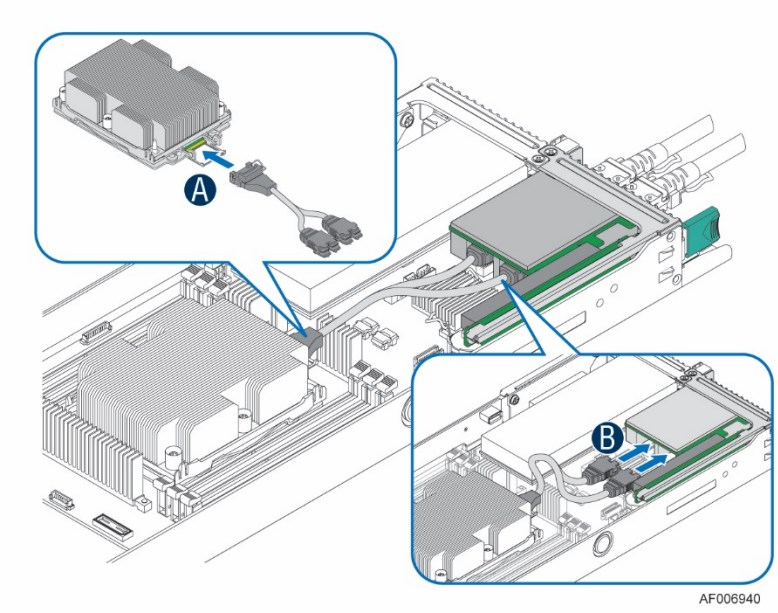


Figure 32. Intel® Xeon™ Phi™ processor (KNL-F) full connection

5.3 Serial ATA (SATA) Support

The server board utilizes two chipset embedded AHCI SATA controllers, identified as **SATA** and **sSATA** ("s" is for secondary), providing for up to ten 6 Gb/sec Serial ATA (SATA) ports.

The AHCI ***SATA*** controller provides support for up to six SATA ports on the server board:

- Four SATA ports to the bridge board connector and then to the backplane through the bridge board
- The SATA DOM connector on the bridge board has been depopulated in order to accommodate a x16 PCIe Riser.
- The AHCI ***sSATA*** controller provides four SATA ports on the server board via the mini-SAS HD connector.
- An 8mm single stacked mSATA connector is also supported on the server board.

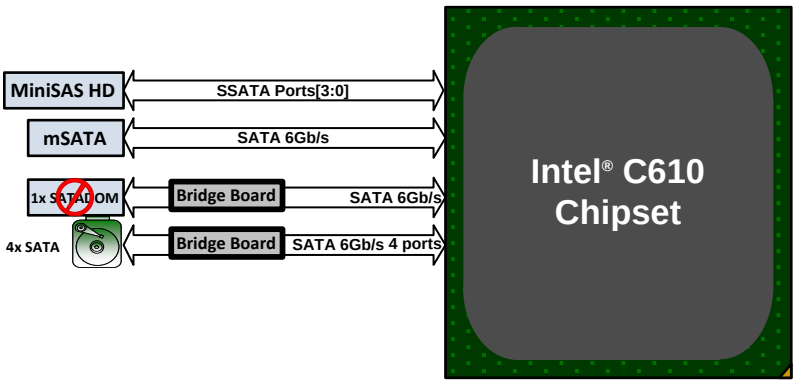


Figure 33. SATA Support

5.3.1 Bridge Board

To enable the board to be hot swapped, power control, PCH SATA disk and BMC management signals are routed from the baseboard through the bridge board and delivered to the chassis backplane.

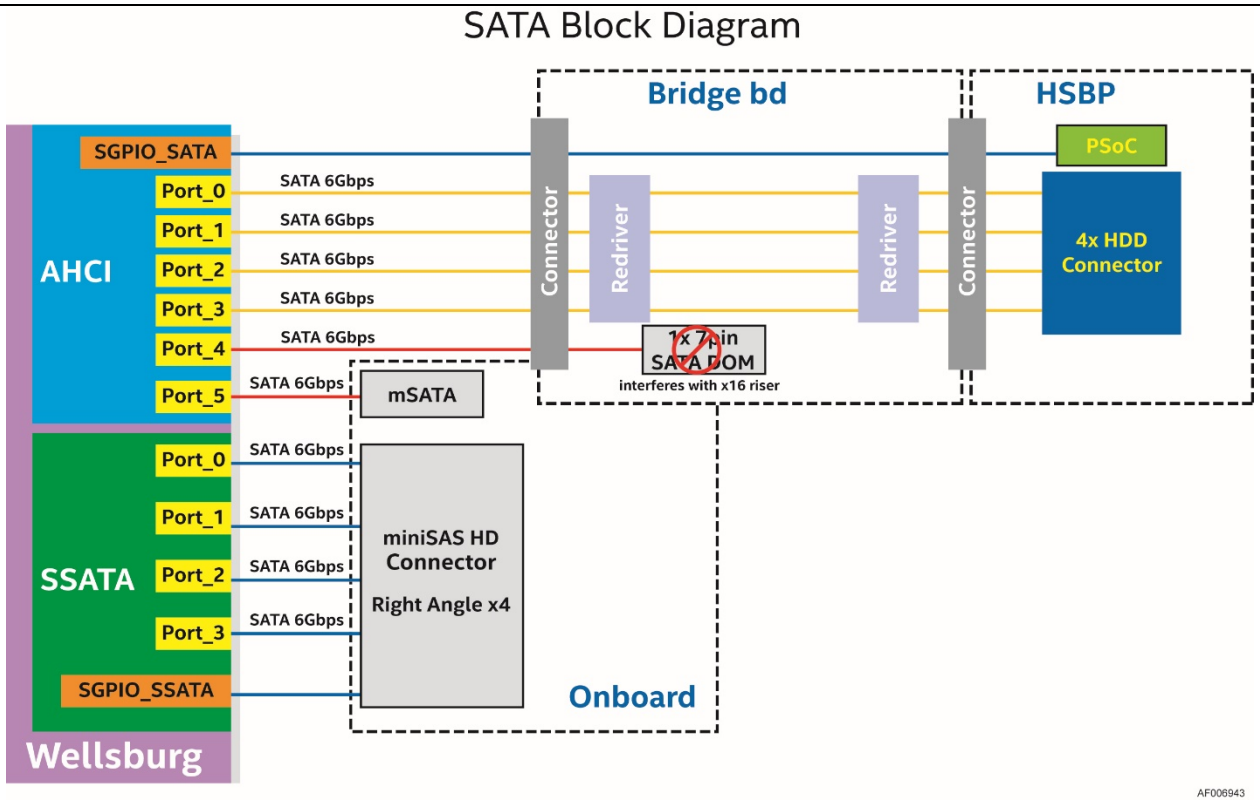


Figure 34. SATA Block Diagram

5.3.2 Bridge Features

The Bridge slot can pass all electrical connectivity through a 2x40pin card edge interconnect between the compute node and rest of the server. The bridge connector passes the follow features (per compute node) to the bridge board. Note that on the PCSD bridge board implementation, both ends of the Bridge board do not have identical signal assignments.

- 4x 6Gbs SATA ports from Wellsburg to HSBP drives
- Two x4 lane 6Gbs SATA re-drivers
- Four chassis ID signals to determine the physical location of the compute module
- One SGPIO SFF-8485 interface to the HSBP microcontroller
- 5V_AUX power generated on HSBP and provided to the compute module
- 3.3V power generated on HSBP and provided to bridge board to run SATA redrivers
- Global PMBus alert signal for CLST support
- Four SMB bus interfaces(Note: Signal re-drives may be needed due to capacitive load)
  - SMBUS R1 - For chassis temp sensor & chassis FRU EEPROM device
  - SMBUS R5 - Connectivity to up to two HSBP controllers & one shared 12V current monitoring device
  - SMBUS R7 - Connectivity to up to two common redundant power supply (CRPS) module PMBus
  - IPMB - For Internode BMC communication.

- Front panel buttons: Power, reset, ID
- Front panel LEDs: Power, fault, status, fabric activity, ID, HDD activity (NOTE: HDD activity signal not used on Intel® Server Systems is provided for OEM requirement)
- COM Serial bus transmit and receive signals for OEM requirement, not used on HSBP of Intel® Server Systems.
- One 7pin 6Gbs SATA port connector for DOM device docking to the bridge board
- USB2.0 interface to a 4pin type-A connector for flash device docking to bridge board
- 2Pin 5V_AUX power for the SATA DOM for case of cabling power
- Power for SATA DOM should include some type of over current protection for example a resettable fuse, this is to cover case where incorrect DOM device is plugged causing short of 5V_AUX to ground
- 3pin jumper header to allow customer to switch between DOM pin 7 connected to GND or to 5V_AUX
- 1.2V switcher VR (from 3.3V) for the SATA re-drivers
- Inductor on 5V_AUX to limit inrush current during compute module hot swap

---

Note: In order to accommodate a x16 PCIe connector on riser 2, the bridge board for server system has de-populated the SATADOM function.

The SATA controller (AHCI Capable Controller 1) and the sSATA controller (AHCI Capable Controller 2) can be independently enabled and disabled and configured through the <F2> BIOS Setup Utility under the “Mass Storage Controller Configuration” menu screen.

Table 14. SATA and sSATA Controller BIOS Utility Setup Options

SATA Controller	sSATA Controller	Supported
AHCI	AHCI	Yes
AHCI	Enhanced	Yes
AHCI	Disabled	Yes
AHCI	ESRT2	Microsoft* Windows Only
Enhanced	AHCI	Yes
Enhanced	Enhanced	Yes
Enhanced	Disabled	Yes
Enhanced	ESRT2	Yes
Disabled	AHCI	Yes
Disabled	Enhanced	Yes
Disabled	Disabled	Yes
Disabled	ESRT2	Yes
ESRT2	AHCI	Microsoft* Windows Only
ESRT2	Enhanced	Yes
ESRT2	Disabled	Yes
ESRT2	ESRT2	Yes

Table 15. SATA and sSATA Controller Feature Support

Feature	Description	AHCI / RAID Disabled	AHCI / RAID Enabled
Native Command Queuing (NCQ)	Allows the device to reorder commands for more efficient data transfers	N/A	Supported
Auto Activate for DMA	Collapses a DMA Setup then DMA Activate sequence into a DMA Setup only	N/A	Supported
Hot Plug Support	Allows for device detection without power being applied and ability to connect and disconnect devices without prior notification to the system	N/A	Supported
Asynchronous Signal Recovery	Provides a recovery from a loss of signal or establishing communication after hot plug	N/A	Supported
6 Gb/s Transfer Rate	Capable of data transfers up to 6 Gb/s	Supported	Supported
ATAPI Asynchronous Notification	A mechanism for a device to send a notification to the host that the device requires attention	N/A	Supported
Host & Link Initiated Power Management	Capability for the host controller or device to request Partial and Slumber interface power states	N/A	Supported
Staggered Spin-Up	Enables the host the ability to spin up hard drives sequentially to prevent power load problems on boot	Supported	Supported
Command Completion Coalescing	Reduces interrupt and completion overhead by allowing a specified number of commands to complete and then generating an interrupt to process the commands		N/A

5.3.3 Staggered Disk Spin-Up

Because of the high density of disk drives that can be attached to the onboard AHCI SATA controller and the sSATA controller, the combined startup power demand surge for all drives at once can be much higher than the normal running power requirements and could require a much larger power supply for startup than for normal operations.

In order to mitigate this and lessen the peak power demand during system startup, both the AHCI SATA controller and the sSATA controller implement a Staggered Spin-Up capability for the attached drives. This means that the drives are started up separately, with a certain delay between disk drives starting.

For the onboard SATA controller, Staggered Spin-Up is an option – AHCI HDD Staggered Spin-Up – in the Setup Mass Storage Controller Configuration screen found in the <F2> BIOS Setup Utility.

5.4 Embedded SATA RAID Support

The server board has the SATA RAID options below:

- Intel® Embedded Server RAID Technology 2 (ESRT2) based on LSI* MegaRAID technology
- Intel® Rapid Storage Technology (RSTe) 4.0 is not supported on this platform

Using the <F2> BIOS Setup Utility, accessed during system POST, options are available to enable/disable RAID, and select which embedded software RAID option to use.

Note: RAID partitions created using ESRT2 cannot span across the two embedded SATA controllers. Only drives attached to a common SATA controller can be included in a RAID partition.

5.4.1 Intel® Embedded Server RAID Technology 2 (ESRT2)

Features of ESRT2 include the following:

- Based on LSI* MegaRAID Software Stack
- Software RAID with system providing memory and CPU utilization
- **RAID Level 0** – Non-redundant striping of drive volumes with performance scaling up to six drives, enabling higher throughput for data intensive applications such as video editing.
- Data security is offered through **RAID Level 1**, which performs mirroring.
- **RAID Level 10** provides high levels of storage performance with data protection, combining the fault-tolerance of RAID Level 1 with the performance of RAID Level 0. By striping RAID Level 1 segments, high I/O rates can be achieved on systems that require both performance and fault-tolerance. RAID Level 10 requires four hard drives, and provides the capacity of two drives.

Note: RAID configurations cannot span across the two embedded AHCI SATA controllers.

5.5 Network Interface

On the back edge of the server board there are two RJ45 networking ports shown as “NIC 1”, and “NIC 2”.

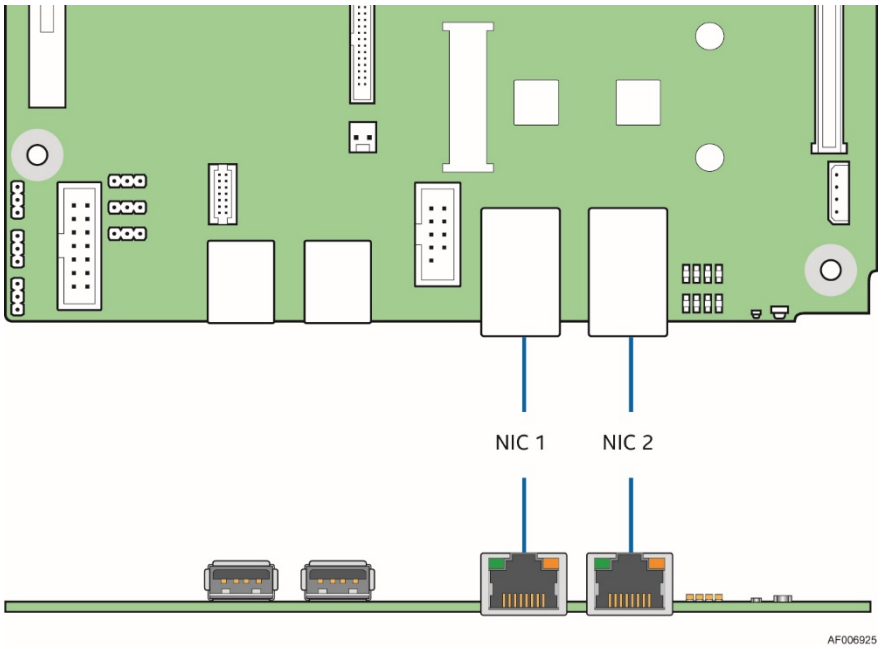


Figure 35. Network Interface Connectors

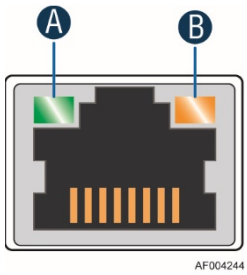
Network interface support is provided by two onboard Intel® Ethernet Controllers I210, which is a single-port, compact component with a fully integrated GbE Media Access Control (MAC) and Physical Layer (PHY) ports. The Intel® I210 NIC provides the server board with support for two LAN ports designed for 10/100/1000 Mbps operation. Refer to the I210 *Ethernet Controller Datasheet* for full details of the NIC feature set.

The NIC device provides a standard IEEE 802.3 Ethernet interface for 1000BASE-T, 100BASE-TX, and 10BASE-T applications (802.3, 802.3u, and 802.3ab) and is capable of transmitting and receiving data at rates of 1000 Mbps, 100 Mbps, or 10 Mbps.

The Intel® I210 NIC will be used in conjunction with the Emulex* PILOT III BMC for in band Management traffic. The BMC will communicate with Intel® i350 over a NC-SI interface (RMII physical). The NIC will be on standby power so that the BMC can send management traffic over the NC-SI interface to the network during sleep states S4 and S5.

The NIC supports the normal RJ-45 LINK/Activity speed LEDs as well as the Preset ID function. These LEDs are powered from a Standby voltage rail.

The link/activity LED (at the right of the connector) indicates network connection when on, and transmit/receive activity when blinking. The speed LED (at the left of the connector) indicates 1000-Mbps operation when green, 100-Mbps operation when amber, and 10-Mbps when off. The following table provides an overview of the LEDs.



LED Color	LED State	NIC State
Green/Amber (B)	Off	10 Mbps
	Amber	100 Mbps
	Green	1000 Mbps
Green (A)	On	Active Connection
	Blinking	Transmit/Receive activity

Figure 36. RJ45 NIC Port LED

5.5.1 MAC Address Definition

The Intel® Server Board S7200AP product family has the following five MAC addresses assigned to it at the Intel factory:

- NIC 1 MAC address (for OS usage)
- NIC 2 MAC address = NIC 1 MAC address + 1 (for OS usage)
- BMC LAN channel 1 MAC address = NIC1 MAC address + 2
- BMC LAN channel 2 (Dedicated Server Management NIC) MAC address = NIC1 MAC address + 3
- BMC LAN channel 3 (DMN) MAC address = NIC1 MAC address + 4

The Intel® Server Board S7200AP has a white MAC address sticker included with the board. The sticker displays the NIC 1 MAC address in both bar code and alphanumeric formats.

5.5.2 LAN Manageability

Port 1 of the Intel® I320 NIC will be used by the BMC firmware to send management traffic.

5.6 Video Support (Internal Header)

There is a video controller which is actually a functional block included in the Baseboard Management Controller integrated on the server board. The Onboard Video Controller can support the 2D video resolutions shown in the following table. The board supports an internal header so that a DB15 connector can be attached via a ribbon cable. Removing the connector from the back panel allows improved venting for the server system.

Table 16. Onboard Video Resolution and Refresh Rate (Hz)

2D Mode	2D Video Mode Support (Color Bit)			
Resolution	**8 bpp**	**16 bpp**	**24 bpp**	**32 bpp**
640x480	60, 72, 75, 85	60, 72, 75, 85	Not supported	60, 72, 75, 85
800x600	60, 72, 75, 85	60, 72, 75, 85	Not supported	60, 72, 75, 85
1024x768	60, 70, 75, 85	60, 70, 75, 85	Not supported	60, 70, 75, 85
1152x864	75	75	75	75
1280x800	60	60	60	60
1280x1024	60	60	60	60
1440x900	60	60	60	60
1600x1200	60	60	Not Supported	Not Supported
1680x1050	60	60	Not Supported	Not Supported
1920x1080	60	60	Not Supported	Not Supported
1920x1200	60	60	Not Supported	Not Supported

The user can use an add-in PCIe video adapter to either replace or complement the Onboard Video Controller.

There are enable/disable options in BIOS Setup screen for “Add-in Video Adapter” and “Onboard Video”.

- When Onboard Video is *Enabled*, and Add-in Video Adapter is also *Enabled*, then both video displays can be active. The onboard video is still the primary console and active during BIOS POST; the add-in video adapter would be active under an OS environment with the video driver support.
- When Onboard Video is *Enabled*, and Add-in Video Adapter is *Disabled*, then only the onboard video would be active.
- When Onboard Video is *Disabled*, and Add-in Video Adapter is *Enabled*, then only the add-in video adapter would be active.

5.7 Universal Serial Bus (USB) Ports

There are eight USB 2.0 ports and six USB 3.0 ports available from Intel® C610 chipset. All ports are high-speed, full-speed and low-speed capable. A total of 4 USB dedicated ports are used. The USB port distribution is as follows:

- Two external USB 3.0 ports on the rear side of server board.
- One internal USB 2.0 port for extension of front-panel USB port on server board.
- One internal USB 2.0 port on bridge board of the compute module.

5.8 Serial Port

The server board has support for one internal serial port - Serial Port A.

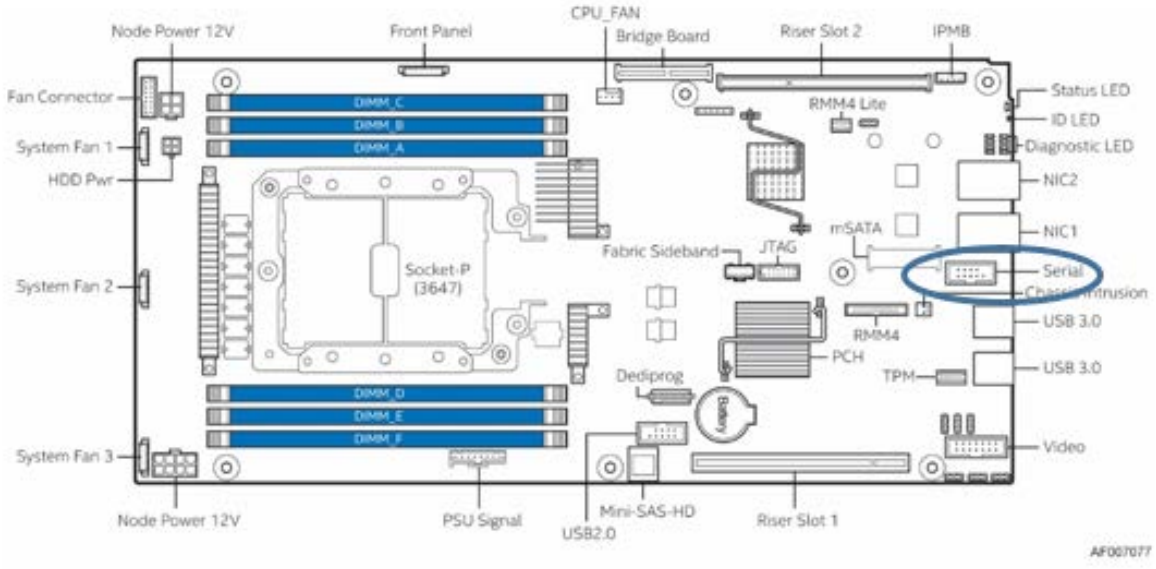


Figure 37. Serial Port A Location

Serial Port A is an internal 10-pin DH-10 connector labeled “Serial_A”.

.

6 Connector and Header

6.1 Power Connectors

6.1.1 Main Power Connector

The main power connector is implemented through use of two 2x4 vertical Minifit Jr* connectors.

Table 17. Main Input Power Supply Connector 8-pin 2x4 Connector

Pin	Signal Name	Pin	Signal Name
1	GND	5	+12V
2	GND	6	+12V
3	GND	7	+12V
4	GND	8	+12V

6.2 System Management Headers

6.2.1 Intel® Remote Management Module 4 (Intel® RMM4) Lite Connector

A 7-pin Intel® RMM4 Lite connector is included on the server board to support the optional Intel® Remote Management Module 4. There is no support for third-party management cards on this server board.

Note: This connector is not compatible with the Intel® Remote Management Module 3 (Intel® RMM3).

Table 18. Intel® RMM4 Lite Connector

Pin	Signal Description	Pin	Signal Description
1	P3V3_AUX	2	SPI_RMM4_LITE_DI
3	Key Pin	4	SPI_RMM4_LITE_CLK
5	SPI_RMM4_LITE_DO	6	GND
7	SPI_RMM4_LITE_CS_N	8	GND

6.2.2 IPMB Header

Table 19. IPMB Header

Pin	Signal Name	Description
1	SMB_IPMB_5VSB_DAT	BMC IPMB 5V standby data line
2	GND	Ground
3	SMB_IPMB_5VSB_CLK	BMC IPMB 5V standby clock line
4	P5V_STBY	+5V standby power

6.3 Bridge Board Connector

The bridge board delivers SATA/SAS signals, disk back plane management signals, BMC SMBus*es as well as control panel and miscellaneous compute module specific signals.

Table 20. Bridge Board Connector

Pin	Signal	Pin	Signal
80	SATA_SAS_SEL	79	GND
78	GND	77	GND
76	SAS0_RX_DP	75	SAS0_TX_DN
74	SAS0_RX_DN	73	SAS0_TX_DP
72	GND	71	GND
70	SAS1_TX_DP	69	SAS1_RX_DN
68	SAS1_TX_DN	67	SAS1_RX_DP
66	GND	65	GND
64	SAS2_RX_DP	63	SAS2_TX_DN
62	SAS2_RX_DN	61	SAS2_TX_DP
60	GND	59	GND
58	SAS3_TX_DP	57	SAS3_RX_DN
56	SAS3_TX_DN	55	SAS3_RX_DP
54	GND	53	GND
52	SGPIO CLK	51	SPKR_IN
50	IBMC_ID0	49	SGPIO_SAS1_LOAD
48	IBMC_ID1	47	SGPIO_SAS1_DATA_OUT
46	IBMC_ID2(reserved)	45	SGPIO_SAS1_DATA_IN
KEY			
44	IBMC_ID3(reserved)	43	PS_EN_PSU _N
42	SPA_SIN_N	41	IRQ_PMBUS Alert_N
40	SPA_SOUT_N	39	GND
38	FP NMI BTN_N	37	SMB_PMBUS_CLK
36	FP PWR BTN_N	35	SMBUS_PMBUS_DATA
34	FP RST BTN_N	33	GND
32	FP ID BTN_N	31	SMB_HSBP_3V3STBY_CLK
30	FP ID LED_N	29	SMB_HSBP_3V3STBY_DATA
28	FP PWR LED_N	27	GND
26	FP STS LED G_N	25	SMB_3V3STBY_CLK
24	FP STS LED A_N	23	SMB_3V3STBY_DATA
22	FP ACT LED_N	21	GND
20	FP HDD ACT LED_N	19	IPMB-5VSTBY_Clk
18	GND	17	IPMB-5VSTBY_Data
16	USB2_P0_DN	15	GND
14	USB2_P0_DP	13	SPARE
12	GND	11	ALL_NODE_OFF
10	SATA0_RX_DP	9	GND
8	SATA0_RX_DN	7	GND
6	GND	5	SATA0_TX_DP
4	USB_OC_FP	3	SATA0_TX_DN
2	5V Aux	1	5V Aux

Combined system BIOS and the Integrated BMC support provide the functionality of the various supported control panel buttons and LEDs. The following sections describe the supported functionality of each control panel feature.

6.3.1 Power Button

The BIOS supports a front control panel power button. Pressing the power button initiates a request that the Integrated BMC forwards to the ACPI power state machines in the chipset. It is monitored by the Integrated BMC and does not directly control power on the power supply.

- **Power Button — Off to On**
  The Integrated BMC monitors the power button and the wake-up event signals from the chipset. A transition from either source results in the Integrated BMC starting the power-up sequence. Since the processors are not executing, the BIOS does not participate in this sequence. The hardware receives the power good and reset signals from the Integrated BMC and then transitions to an ON state.
- **Power Button — On to Off (operating system absent)**
  The System Control Interrupt (SCI) is masked. The BIOS sets up the power button event to generate an SMI and checks the power button status bit in the ACPI hardware registers when an SMI occurs. If the status bit is set, the BIOS sets the ACPI power state of the machine in the chipset to the OFF state. The Integrated BMC monitors power state signals from the chipset and de-asserts PS_PWR_ON to the power supply. As a safety mechanism, if the BIOS fails to service the request, the Integrated BMC automatically powers off the system in four to five seconds.
- **Power Button — On to Off (operating system present)**
  If an ACPI operating system is running, pressing the power button switch generates a request through SCI to the operating system to shut down the system. The operating system retains control of the system and the operating system policy determines the sleep state into which the system transitions, if any. Otherwise, the BIOS turns off the system.

6.3.2 Reset Button

The platform supports a front control panel reset button. Pressing the reset button initiates a request forwarded by the Integrated BMC to the chipset. The BIOS does not affect the behavior of the reset button.

6.4 I/O Connectors

6.4.1 PCI Express* Connectors

The Intel® Server Board S7200AP uses two PCI Express* slots physically with different pint out definition. Each riser slot has dedicated usage and cannot be used for normal PCIe based add-in card.

- Riser slot 1: Provide PCIe x16 to Riser (Using standard 164-pin connector).
- Riser slot 2: Provide PCIe x24 to Riser (Using 200-pin HSEC8 Connector).

The pin-outs for the slots are shown in the following tables.

Table 21. PCI Express* x16 Riser Slot 1 Connector

Pin	Pin Name	Pin	Pin Name
B1	12V	A1	Present2
B2	12V	A2	12V
B3	12V	A3	12V
B4	GND	A4	GND
B5	SMBUS_R4 CLK	A5	3.3V
B6	SMBUS_R4 DAT	A6	Pull_Up
B7	GND	A7	Pull_Up
B8	3.3V	A8	3.3V
B9	Pull Down	A9	3.3V
B10	3.3VAUX	A10	3.3V
B11	WAKE#	A11	PERST#
KEY			
B12	RSVD	A12	GND
B13	GND	A13	REFCLK1+
B14	PETxP0	A14	REFCLK1-
B15	PETxN0	A15	GND
B16	GND	A16	PERxP0
B17	Present2	A17	PERxN0
B18	GND	A18	GND
B19	PETxP1	A19	3.3V
B20	PETxN1	A20	GND
B21	GND	A21	PERxP1
B22	GND	A22	PERxN1
B23	PETxP2	A23	GND
B24	PETxN2	A24	GND
B25	GND	A25	PERxP2
B26	GND	A26	PERxN2
B27	PETxP3	A27	GND
B28	PETxN3	A28	GND
B29	GND	A29	PERxP3
B30	3.3V	A30	PERxN3
B31	Present2	A31	GND
B32	GND	A32	REFCLK2+
B33	PETxP4	A33	REFCLK2-
B34	PETxN4	A34	GND
B35	GND	A35	PERxP4
B36	GND	A36	PERxN4
B37	PETxP5	A37	GND
B38	PETxN5	A38	GND
B39	GND	A39	PERxP5
B40	GND	A40	PERxN5
B41	PETxP6	A41	GND
B42	PETxN6	A42	GND
B43	GND	A43	PERxP6
B44	GND	A44	PERxN6

B45	PETxP7	A45	GND
B46	PETxN7	A46	GND
B47	GND	A47	PERxP7
B48	PW ID 0	A48	PERxN7
B49	GND	A49	GND
B50	PETxP8	A50	RSVD
B51	PETxN8	A51	GND
B52	GND	A52	PERxP8
B53	GND	A53	PERxN8
B54	PETxP9	A54	GND
B55	PETxN9	A55	GND
B56	GND	A56	PERxP9
B57	GND	A57	PERxN9
B58	PETxP10	A58	GND
B59	PETxN10	A59	GND
B60	GND	A60	PERxP10
B61	GND	A61	PERxN10
B62	PETxP11	A62	GND
B63	PETxN11	A63	GND
B64	GND	A64	PERxP11
B65	GND	A65	PERxN11
B66	PETxP12	A66	GND
B67	PETxN12	A67	GND
B68	GND	A68	PERxP12
B69	GND	A69	PERxN12
B70	PETxP13	A70	GND
B71	PETxN13	A71	GND
B72	GND	A72	PERxP13
B73	GND	A73	PERxN13
B74	PETxP14	A74	GND
B75	PETxN14	A75	GND
B76	GND	A76	PERxP14
B77	GND	A77	PERxN14
B78	PETxP15	A78	GND
B79	PETxN15	A79	GND
B80	GND	A80	PERxP15
B81	Present2	A81	PERxN15
B82	RSVD	A82	GND

Table 22. PCI Express* x24 Riser Slot 2 Connector

Pin	Pin Name	Pin	Pin Name
1	12V	2	12V
3	12V	4	12V
5	GND	6	GND
7	GND	8	GND
9	3.3V	10	3.3V
11	3.3V	12	3.3V
13	GND	14	GND
15	3.3VAUX	16	5V Aux
17	GND	18	GND
19	Spare	20	Spare
21	Spare	22	Spare
23	Spare	24	Spare
25	ID1	26	Spare
27	GND	28	Spare
29	LED_ACT#	30	Spare
31	Riser ID0	32	GND
33	GND	34	PERST#
35	SMCLK_R2M1	36	WAKE#
37	SMDATA_R2M1	38	GND
39	GND	40	PE1_CLK3+
41	PE1_R00-	42	PE1_CLK3-
43	PE1_R00+	44	GND
45	GND	46	PE1_T00+
47	PE1_R01-	48	PE1_T00-
49	PE1_R01+	50	GND
51	GND	52	PE1_T01+
53	PE1_R02-	54	PE1_T01-
55	PE1_R02+	56	GND
57	GND	58	PE1_T02+
59	PE1_R03-	60	PE1_T02-
61	PE1_R03+	62	GND
63	GND	64	GND
65	Spare	66	
67	Spare	68	GND
69	GND	70	PE1_T03+
71	PE1_R04-	72	PE1_T03-
73	PE1_R04+	74	GND
75	GND	76	PE1_T04+
77	PE1_R05-	78	PE1_T04-
79	PE1_R05+	80	GND
81	GND	82	PE1_T05+
83	PE1_R06-	84	PE1_T05-
85	PE1_R06+	86	GND
87	GND	88	PE1_T06+
89	PE1_R07-	90	PE1_T06-

Pin	Pin Name	Pin	Pin Name
91	PE1_R07+	92	GND
93	GND	94	PE1_T07+
95	CLK2+	96	PE1_T07-
97	CLK2-	98	GND
99	GND	100	GND
101	GND	102	CLK1+
103	R00-	104	CLK1-
105	R00+	106	GND
107	GND	108	T00-
109	R01-	110	T00+
111	R01+	112	GND
113	GND	114	T01-
115	R02-	116	T01+
117	R02+	118	GND
119	GND	120	T02-
121	R03-	122	T02+
123	R03+	124	GND
125	GND	126	T03-
127	R04-	128	T03+
129	R04+	130	GND
131	GND	132	T04-
133	R05-	134	T04+
135	R05+	136	GND
137	GND	138	T05-
139	R06-	140	T05+
141	R06+	142	GND
143	GND	144	T06-
145	R07-	146	T06+
147	R07+	148	GND
149	GND	150	T07-
151	R08-	152	T07+
153	R08+	154	GND
155	GND	156	T08-
157	R09-	158	T08+
159	R09+	160	GND
161	GND	162	T09-
163	R10-	164	T09+
165	R10+	166	GND
167	GND	168	T10-
169	R11-	170	T10+
171	R11+	172	GND
173	GND	174	T11-
175	R12-	176	T11+
177	R12+	178	GND
179	GND	180	T12-
181	R13-	182	T12+

Pin	Pin Name	Pin	Pin Name
183	R13+	184	GND
185	GND	186	T13-
187	R14-	188	T13+
189	R14+	190	GND
191	GND	192	T14-
193	R15-	194	T14+
195	R15+	196	GND
197	GND	198	T15-
199	Spare	200	T15+

6.4.2 VGA Connector

The following table details the pin-out definition of the external VGA connector.

Table 23. VGA Internal Video Connector

Pin	Signal Name	Description
1	V_IO_R_CONN	Red (analog color signal R)
2	GND	Ground
3	V_IO_G_CONN	Green (analog color signal G)
4	GND	Ground
5	V_IO_B_CONN	Blue (analog color signal B)
6	GND	Ground
7	V_IO_VSYNC_CONN	VSYNC (vertical sync)
8	GND	Ground
9	V_IO_HSYNC_CONN	HSYNC (horizontal sync)
10	GND	Ground
11	V_IO_DDCDAT	DDCDAT
12	NC	No connection
13	V_IO_DDCCLK	DDCCLK
14	NC	No connection

6.4.3 NIC Connectors

The server board provides two independent RJ-45 connectors on the back edge of the board. The pin-out for NIC connectors are identical and are defined in the following table.

Table 24. RJ-45 10/100/1000 NIC Connector

Pin	Signal Name
1	LED_NIC_LINK0_100_N
2	LED_NIC_LINK0_1000_R_N
3	NIC_0_0_DP
4	NIC_0_0_DN
5	NIC_0_1_DP
6	NIC_0_1_DN
7	NIC_CT1
8	NIC_CT2
9	NIC_0_2_DP

Pin	Signal Name
10	NIC_0_2_DN
11	NIC_0_3_DP
12	NIC_0_3_DN
13	LED_NIC_LINK0_LNKUP_N
14	LED_NIC_LINK0_ACT_R_N

6.4.4 mSATA Connector

The server board provides one mSATA port connector.

Table 25. mSATA Connector

Pin	Signal Name	Pin	Signal Name
P1	P2	P2	+3.3 V
P3	Reserved	P4	GND
P5	Reserved	P6	Reserved
P7	Reserved	P8	Reserved
P9	GND	P10	Reserved
P11	Reserved	P12	Reserved
P13	Reserved	P14	Reserved
P15	GND	P16	Reserved
P17	Reserved	P18	GND
P19	Reserved	P20	Reserved
P21	GND	P22	Reserved
P23	+B	P24	+3.3 V
P25	-B	P26	GND
P27	GND	P28	Reserved
P29	GND	P30	Reserved
P31	-A	P32	Reserved
P33	+A	P34	GND
P35	GND	P36	Reserved
P37	GND	P38	Reserved
P39	+3.3 V	P40	GND
P41	+3.3 V	P42	Reserved
P43	Reserved	P44	Reserved
P45	Vendor	P46	Reserved
P47	Vendor	P48	Reserved
P49	DA/DSS	P50	GND
P51	Presence Detection	P52	+3.3 V

6.4.5 Hard Drive Activity LED Header

Table 26. SATA HDD Activity LED Header

Pin	Description
1	LED_HD_ACTIVE_N
2	NC

6.4.6 Serial Port Connectors

The server board provides one internal 9-pin serial type-A header. The following tables define the pin-outs.

Table 27. Internal 9-pin Serial A

Pin	Signal Name	Pin	Signal Name
1	SPA_DCD	2	SPA_DSR
3	SPA_SIN_N	4	SPA_RTS
5	SPA_SOUT_N	6	SPA_CTS
7	SPA_DTR	8	SPA_RI
9	GND	10	KEY

6.4.7 USB Connectors

The following table details the pin-out of the external stacked USB port connectors found on the back edge of the server board.

Table 28. External USB port Connector

Pin	Signal Name	Description
1	+5V	USB Power
2	USB_N	Differential data line paired with DATAH0
3	USB_P	Differential date line paired with DATAL0
4	GND	Ground

One 2x5 connector on the server board provides an option to support an additional internal USB port. The pin-out is detailed in the following table.

Table 29. Internal USB Connector

Pin	Signal Name	Description	Pin	Signal Name	Description
1	USB2_PWR +5V	USB Power (Ports 0,1)	2	USB2_PWR	USB Power (Ports 0,1)
3	USB2_13_L	USB Port 13 Negative Signal	4	USB2_8_L	USB Port 8 Negative Signal
5	USB2_13	USB Port 13 Positive Signal	6	USB2_8	USB Port 8 Positive Signal
7	Ground	GND	8	Ground	GND
9	NC	Key	10	TP_USB2PIN_P10	Test Point

6.4.8 IFT Connector

Table 30. IFT Connector

Pin	Signal Name	Pin	Signal Name
1	GND	28	GND
3	HFIx_RX_DN[1]	27	HFIx_RX_DN[2]
5	HFIx_RX_DP[1]	26	HFIx_RX_DP[2]
7	GND	25	GND
9	HFIx_RX_DN[3]	24	HFIx_RX_DN[4]
6	HFIx_RX_DP[3]	23	HFIx_RX_DP[4]
7	GND	22	GND

Pin	Signal Name	Pin	Signal Name
8	GND	21	GND
9	HFIx_TX_DP[4]	20	HFIx_TX_DP[3]
10	HFIx_TX_DN[4]	19	HFIx_TX_DN[3]
11	GND	18	GND
12	HFIx_TX_DP[2]	17	HFIx_TX_DP[1]
13	HFIx_TX_DN[2]	16	HFIx_TX_DN[1]
14	GND	15	GND

6.5 Fan Headers

6.5.1 FAN Control Cable Connector

To facilitate the connection of 3 double rotor fans, a 14 pin header is provided; all fans will share a PWM. Both rotor tachs can be monitored.

Table 31. Baseboard Fan Connector

Pin	Signal Name	Pin	Signal Name
1	FAN_PWM_OUT	2	Key
3	FAN_BMC_TACH0	4	FAN_BMC_TACH1
5	FAN_BMC_TACH2	6	FAN_BMC_TACH3
7	FAN_BMC_TACH4	8	FAN_BMC_TACH5
9	PS_HOTSWAP_EN	10	GND
11	SMB_HOST_3V3_CLK	12	SMB_HOST_3V3_DATA
13	NODE_ADR0(GND)	14	PWRGD_PS_PWROK

The SMBus* is used to connect to the hot swap controller that provides inrush current protection and can measure the power being used by the compute module. The NODE_ON signal is used to turn on the hot swap controller. Note that the polarity is correct as the ADI1275 controller uses a high true enable signal. When the compute module is turned off, the fans will continue to rotate at a preset rate; this rate is selected by Intel® and preset by the Fan manufacturer. This is done to stop air recirculation between compute modules. When docking the board to a live 12V rail, the fans could spin up immediately; it may be required to phase their connection to power to minimize the inrush current. Bench testing of the fans should determine if this is necessary.

6.5.2 Discrete System FAN Connector

To support the 3rd party chassis, three discrete system fan connectors are provided on baseboard. They are used for connecting FANs with tach meters directly.

Table 32. Baseboard Fan Connector

Fan 1		Fan 2		Fan 3	
Pin	**Signal Description**	**Pin**	**Signal Description**	**Pin**	**Signal Description**
1	GND	1	GND	1	GND
2	P12V	2	P12V	2	P12V
3	BMC_TACH0_R	3	BMC_TACH2_R	3	BMC_TACH4_R
4	PWM0	4	PWM0	4	PWM0

5	GND	5	GND	5	GND
6	P12V	6	P12V	6	P12V
7	BMC_TACH1_R	7	BMC_TACH3_R	7	BMC_TACH5_R
8	PWM0	8	PWM0	8	PWM0

6.6 Node Power Docking Board Connectors

The table below lists the connector type and pin definition on the Node Power Docking Board.

Table 33. Main Power Input Connector

Pin	Signal Description	Pin	Signal Description
Lower Blade (Circuit 1)			
1	GND	2	GND
3	GND	4	GND
5	GND	6	GND
Upper Blade (Circuit 2)			
7	P12V	8	P12V
9	P12V	10	P12V
11	P12V	12	P12V

Table 34. Fan Control Signal Connector

Pin	Signal Description	Pin	Signal Description
1	PWM1	2	Reserved
3	Tach0	4	Tach1
5	Tach2	6	Tach3
7	Tach4	8	Tach5
9	NODE_ON	10	GND
11	SMBUS_R4 CLK	12	SMBUS_R4 DAT
13	NODE_ADR0	14	NODE_PWRGD

Table 35. Compute Module Fan Connector

Pin	Signal Description
1	GND
2	P12V
3	TACH1
4	PWM1
5	GND
6	P12V
7	TACH2
8	PWM1

Table 36. Main Power Output Connector

Pin	Signal Description	Pin	Signal Description
1	GND	7	P12V_HS
2	GND	8	P12V_HS
3	GND	9	P12V_HS
4	GND	10	P12V_HS
5	GND	11	P12V_HS
6	GND	12	P12V_HS

7 Configuration Jumpers

The following table provides a summary and description of configuration, test, and debug jumpers. The server board has several 3-pin jumper blocks that can be used. **Pin 1** on each jumper block can be identified by the following symbol on the silkscreen: □

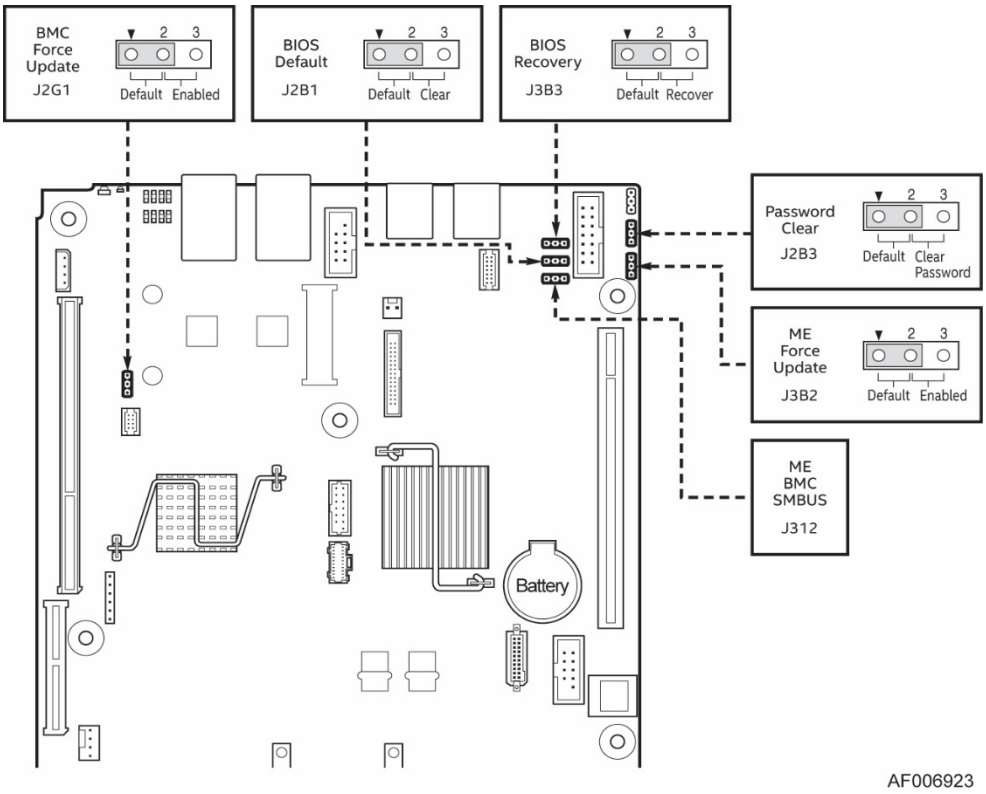


Figure 38. Jumper Location

Table 37. Jumper Modes Selection

Jumper Name	Jumper Position	Mode of Operation	Note
J2G1: BMC Force Update	1-2	Normal	Normal mode
	2-3	Update	BMC in force update mode
J2B1: BIOS Default	1-2	Normal	Normal mode
	2-3	Clear BIOS Settings	BIOS settings are reset to factory default
J3B3: BIOS Recovery Mode	1-2	Normal	Normal mode
	2-3	Recovery	BIOS in recovery mode
J2B3: Password Clear	1-2	Normal	Normal mode, password in protection
	2-3	Clear Password	BIOS password is cleared
J3B2: ME Force Update	1-2	Normal	Normal mode
	2-3	Update	ME in force update mode

7.1 BMC Force Update (J2G1)

When performing a standard BMC firmware update procedure, the update utility places the BMC into an update mode, allowing the firmware to load safely onto the flash device. In the unlikely event the BMC firmware update process fails due to the BMC not being in the proper update state, the server board provides a BMC Force Update jumper (J2G1) which will force the BMC into the proper update state. The following procedure should be followed in the event the standard BMC firmware update process fails.

Table 38. Force Integrated BMC Update Jumper (J2G1)

Jumper Position	Mode of Operation	Note
1-2	Normal	Normal operation
2-3	Update	BMC in force update mode

Steps to perform Force BMC Update:

1. Plug out the compute module.
2. Remove the air duct. See the *Service Guide* for instructions.
3. Move the jumper (J2G1) from the default operating position (covering pins1 and 2) to the enabled position (covering pins 2 and 3).
4. Restore the air duct to the compute module.
5. Insert the compute module back to the chassis.
6. Power on the compute module by pressing the power button on the front panel.
7. Perform the BMC firmware update procedure as documented in the *Release Note*s included in the given BMC firmware update package. After successful completion of the firmware update process, the firmware update utility may generate an error stating the BMC is still in update mode.
8. Power down and plug out the compute module.
9. Remove the air duct.
10. Move the jumper from the enabled position (covering pins 2 and 3) to the disabled position (covering pins 1 and 2).
11. Restore the air duct to the compute module.
12. Plug in the compute module back to the chassis and power up the server.

Note: Normal BMC functionality is disabled with the Force BMC Update jumper is set to the enabled position. You should never run the server with the BMC Force Update jumper set in this position. You should only use this jumper setting when the standard firmware update process fails. This jumper should remain in the default/disabled position when the server is running normally.

The server board has several 3-pin jumper blocks that can be used to configure, protect, or recover specific features of the server board.

7.2 ME Force Update (J3B2)

When this 3-pin jumper is set, it manually puts the ME firmware in update mode, which enables the user to

update ME firmware code when necessary.

Table 39. Force ME Update Jumper (J3B2)

Jumper Position	Mode of Operation	Note
1-2	Normal	Normal operation
2-3	Update	ME in force update mode

Note: Normal ME functionality is disabled with the ME Force Update jumper set to the enabled position. You should never run the server with the ME Force Update jumper set in this position. You should only use this jumper setting when the standard firmware update process fails. This jumper should remain in the default/disabled position when the server is running normally.

Steps to perform the Force ME Update:

1. Plug out the compute module from the chassis.
2. Remove the air duct. See the *Service Guide* for instructions.
3. Move the jumper (J3B2) from the default operating position (covering pins 1 and 2) to the enabled position (covering pins 2 and 3).
4. Restore the air duct back to the compute module.
5. Plug in the compute module back to the chassis.
6. Perform the ME firmware update procedure as documented in the Release Notes file that is included in the given system update package.
7. After update process is done, plug out the compute module out of the chassis.
8. Remove the air duct.
9. Move the jumper from the enabled position (covering pins 2 and 3) to the disabled position (covering pins 1 and 2).
10. Restore the compute module back to the chassis.

7.3 Password Clear (J2B3)

The user sets this 3-pin jumper to clear the password.

Table 40. Password Clear Jumper (J2B3)

Jumper Position	Mode of Operation	Note
1-2	Normal	Normal mode, password in protection
2-3	Clear Password	BIOS password is cleared

This jumper causes both the User password and the Administrator password to be cleared if they were set. The operator should be aware that this creates a security gap until passwords have been installed again.

Note: No method of resetting BIOS configuration settings to the default values will affect either the Administrator or User passwords.

This is the only method by which the Administrator and User passwords can be cleared unconditionally. Other than this jumper, passwords can only be set or cleared by changing them explicitly in BIOS Setup or by similar means.

The recommended steps to clear the User and Administrator passwords are:

1. Plug out the compute module and remove the air duct.
2. Move the jumper from pins 1-2 to pins 2-3. It is necessary to leave the jumper in place while rebooting the system in order to clear the passwords.
3. Installed the air duct and plug in and power up the compute module.
4. Boot into the BIOS Setup. Check the Error Manager tab for POST Error Codes:

   5221 Passwords cleared by jumper

   5224 Password clear jumper is set
5. Power down and plug out the compute module and remove the air duct again.
6. Restore the jumper from pins 2-3 to the normal setting of pins 1-2.
7. Installed the air duct and plug in and power up the compute module.
8. **Strongly recommended:** Boot into the BIOS Setup immediately, go to the Security tab and set the Administrator and User passwords if you intend to use BIOS password protection.

7.4 BIOS Recovery Mode (J3B3)

If a system is completely unable to boot successfully to an OS, hangs during POST, or even hangs and fails to start executing POST, it may be necessary to perform a BIOS Recovery procedure, which can replace a defective copy of the Primary BIOS.

The BIOS introduces three mechanisms to start the BIOS recovery process, which is called Recovery Mode:

- Recovery Mode Jumper – This jumper causes the BIOS to boot in Recovery Mode.
- The BootBlock detects partial BIOS update and automatically boots in Recovery Mode.
- The BMC asserts Recovery Mode GPIO in case of partial BIOS update and FRB2 time-out.

Table 41. BIOS Recovery Mode Jumper (J3B3)

Jumper Position	Mode of Operation	Note
1-2	Normal	Normal mode
2-3	Recovery	BIOS in recovery mode

The BIOS Recovery takes place without any external media or Mass Storage device as it utilizes the Backup BIOS inside the BIOS flash in Recovery Mode.

The Recovery procedure is included here for general reference. However, if in conflict, the instructions in the BIOS Release Notes are the definitive version.

When Recovery Mode Jumper is set, the BIOS begins with a “Recovery Start” event logged to the SEL, loads and boots with the Backup BIOS image inside the BIOS flash itself. This process takes place before any video or console is available. The system boots up into the Shell directly while a “Recovery Complete” SEL logged. An external media is required to store the BIOS update package and steps are the same as the normal BIOS

update procedures. After the update is complete, there will be a message displayed stating that the “BIOS has been updated successfully" indicating the BIOS update process is finished. The User should then switch the recovery jumper back to normal operation and restart the system by performing a power cycle.

If the BIOS detects partial BIOS update or the BMC asserts Recovery Mode GPIO, the BIOS will boot up with Recovery Mode. The difference is that the BIOS boots up to the Error Manager Page in the BIOS Setup utility. In the BIOS Setup utility, boot device, Shell or Linux for example, could be selected to perform the BIOS update procedure under Shell or OS environment.

Again, before starting to perform a Recovery Boot, be sure to check the BIOS Release Notes and verify the Recovery procedure shown in the Release Notes.

The following steps demonstrate this recovery process:

1. Plug out the compute module and remove the air duct.
2. Move the jumper (J3B3) from the default operating position (covering pins 1 and 2) to the BIOS Recovery position (covering pins 2 and 3).
3. Restore the air duct back to the compute module.
4. Plug in the compute module back to the chassis.
5. Power on the compute module.
6. The BIOS will load and boot with the backup BIOS image without any video or display.
7. When compute module boots into the EFI shell directly, the BIOS recovery is successful.
8. Power off the compute module.
9. Plug out the compute module from the chassis.
10. Remove the air duct and put the jumper (J3B3) back to the normal position (covering pin 1 and 2).
11. Restore the air duct and put the compute module back to the chassis.
12. A normal BIOS update can be performed if needed.

7.5 BIOS Default (J2B1)

Table 42. BIOS Default Jumper

Jumper Position	Mode of Operation	Note
1-2	Normal	Normal mode
2-3	Clear BIOS settings	BIOS settings are reset to factory default

This jumper causes the BIOS Setup settings to be reset to their default values. On previous generations of server boards, this jumper has been referred to as “Clear CMOS”, or “Clear NVRAM”. Setting this jumper according to the procedure below will clear all current contents of NVRAM variable storage, and then load the BIOS default settings.

Note that this jumper does not reset Administrator or User passwords. In order to reset passwords, the Password Clear jumper must be used.

The recommended steps to reset to the BIOS defaults are:

1. Plug out the compute module and remove the air duct.
2. Move the jumper from pins 1-2 to pins 2-3 momentarily. It is not necessary to leave the jumper in place while rebooting.
3. Restore the jumper from pins 2-3 to the normal setting of pins 1-2.
4. Installed the air duct and plug in the compute module, and power up.
5. Boot the system into Setup. Check the Error Manager tab, and you should see POST Error Codes:

   0012 System RTC date/time not set

   5220 BIOS Settings reset to default settings
6. Go to the Setup Main tab, and set the System Date and System Time to the correct current settings. Make any other changes that are required in Setup – for example, Boot Order.

8 Intel® Light-Guided Diagnostics

Intel® Server Board S7200AP has several onboard diagnostic LEDs to assist in troubleshooting board-level issues. This section provides a description of the location and function of each LED on the server board.

8.1 Status LED

Note: The status LED state shows the state for the current, most severe fault. For example, if there was a critical fault due to one source and a non-critical fault due to another source, the status LED state would be solid on (the critical fault state).

The status LED is a bicolor LED. Green (status) shows a normal operation state or a degraded operation. Amber (fault) shows the hardware state and overrides the green status.

The Integrated BMC-detected state and the state from the other controllers, such as the SCSI/SATA hot-swap controller state, are included in the LED state. For fault states monitored by the Integrated BMC sensors, the contribution to the LED state follows the associated sensor state, with the priority going to the most critical state currently asserted.

When the server is powered down (transitions to the DC-off state or S5), the Integrated BMC is still on standby power and retains the sensor and front panel status LED state established prior to the power-down event.

The following table maps the server state to the LED state.

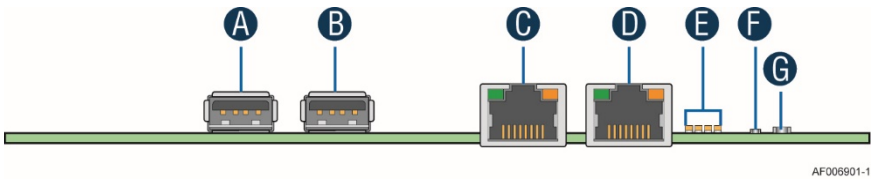


Figure 39. Status LED (G) and ID LED (F)

Table 43. Status LED State Definitions

Color	State	Criticality	Description
Off	System is not operating	Not ready	▪ System is powered off (AC and/or DC). ▪ System is in EuP Lot6 Off Mode. ▪ System is in S5 Soft-Off State.
Green	Solid on	Ok	Indicates that the System is running (in S0 State) and its status is 'Healthy'. The system is not exhibiting any errors. AC power is present and BMC has booted and manageability functionality is up and running. After a BMC reset, and in conjuction with the Chassis ID solid ON, the BMC is booting Linux*. Control has been passed from BMC uBoot to BMC Linux* itself. It will be in this state for ~10-~20 seconds

Color	State	Criticality	Description
Green	~1 Hz blink	Degraded - system is operating in a degraded state although still functional, *or* system is operating in a redundant state but with an impending failure warning	System degraded: ▪ Redundancy loss such as power-supply or fan. Applies only if the associated platform sub-system has redundancy capabilities. ▪ Fan warning or failure when the number of fully operational fans is less than minimum number needed to cool the system. ▪ Non-critical threshold crossed – Temperature (including HSBP temp), voltage, input power to power supply, output current for main power rail from power supply and Processor Thermal Control (Therm Ctrl) sensors. ▪ Power supply predictive failure occurred while redundant power supply configuration was present. ▪ Unable to use all of the installed memory (more than 1 DIMM installed). ▪ Correctable Errors over a threshold and migrating to a spare DIMM (memory sparing). This indicates that the system no longer has spared DIMMs (a redundancy lost condition). Corresponding DIMM LED lit. ▪ In mirrored configuration, when memory mirroring takes place and system loses memory redundancy. ▪ Battery failure. ▪ BMC executing in uBoot. (Indicated by Chassis ID blinking at 3Hz). System in degraded state (no manageability). BMC uBoot is running but has not transferred control to BMC Linux*. Server will be in this state 6-8 seconds after BMC reset while it pulls the Linux* image into flash. ▪ BMC Watchdog has reset the BMC. ▪ Power Unit sensor offset for configuration error is asserted. ▪ HDD HSC is off-line or degraded.
Amber	~1 Hz blink	Non-critical - System is operating in a degraded state with an impending failure warning, although still functioning	Non-fatal alarm – system is likely to fail: ▪ Critical threshold crossed – Voltage, temperature (including HSBP temp), input power to power supply, output current for main power rail from power supply and PROCHOT (Therm Ctrl) sensors. ▪ VRD Hot asserted. ▪ Minimum number of fans to cool the system not present or failed ▪ Hard drive fault ▪ Power Unit Redundancy sensor – Insufficient resources offset (indicates not enough power supplies present) ▪ In non-sparing and non-mirroring mode if the threshold of correctable errors is crossed within the window

Color	State	Criticality	Description
Amber	Solid on	Critical, non-recoverable – System is halted	Fatal alarm – system has failed or shutdown: ▪ CPU CATERR signal asserted ▪ MSID mismatch detected (CATERR also asserts for this case). ▪ CPU 1 is missing ▪ CPU Thermal Trip ▪ No power good – power fault ▪ DIMM failure when there is only 1 DIMM present and hence no good memory present. ▪ Runtime memory uncorrectable error in non-redundant mode. ▪ DIMM Thermal Trip or equivalent ▪ SSB Thermal Trip or equivalent ▪ CPU ERR2 signal asserted ▪ BMC/Video memory test failed. (Chassis ID shows blue/solid-on for this condition) ▪ Both uBoot BMC FW images are bad. (Chassis ID shows blue/solid-on for this condition) ▪ 240VA fault ▪ Fatal Error in processor initialization: o Processor family not identical o Processor model not identical o Processor core/thread counts not identical o Processor cache size not identical o Unable to synchronize processor frequency o Unable to synchronize QPI link frequency ▪ Uncorrectable memory error in a non-redundant mode

8.2 ID LED

The ID LED provides a visual indication of the server board or compute module being serviced. The state of the ID LED is affected by the following:

- Toggled by the ID button
- Controlled by the *Chassis Identify* command (IPMI)

Table 44. ID LED

State	LED State
Identify active through button	Solid on
Identify active through command	~1 Hz blink
Off	Off

There is no precedence or lock-out mechanism for the control sources. When a new request arrives, all previous requests are terminated. For example, if the ID LED is blinking and the chassis ID button is pressed, then the ID LED changes to solid on. If the button is pressed again with no intervening commands, the ID LED turns off.

8.3 BMC Boot/Reset Status LED Indicators

During the BMC boot or BMC reset process, the System Status LED and System ID LED are used to indicate BMC boot process transitions and states. A BMC boot will occur when AC power is first applied to the system. A BMC reset will occur after a BMC FW update, upon receiving a BMC cold reset command, and upon a BMC watchdog initiated reset. The following table defines the LED states during the BMC Boot/Reset process.

Table 45. BMC Boot/Reset Status LED Indicators

BMC Boot/Reset State	Chassis ID LED	Status LED	Comment
BMC/Video memory test failed	Solid Blue	Solid Amber	Non-recoverable condition. Contact your Intel representative for information on replacing this motherboard.
Both Universal Bootloader (u-Boot) images bad	Blink Blue 6 Hz	Solid Amber	Non-recoverable condition. Contact your Intel representative for information on replacing this motherboard.
BMC in u-Boot	Blink Blue 3 Hz	Blink Green 1Hz	Blinking green indicates degraded state (no manageability), blinking blue indicates u-Boot is running but has not transferred control to BMC Linux. Server will be in this state 6-8 seconds after BMC reset while it pulls the Linux image into flash.
BMC Booting Linux	Solid Blue	Solid Green	Solid green with solid blue after an AC cycle/BMC reset, indicates that the control has been passed from u-Boot to BMC Linux itself. It will be in this state for ~10-~20 seconds.
End of BMC boot/reset process. Normal system operation	Off	Solid Green	Indicates BMC Linux has booted and manageability functionality is up and running. Fault/Status LEDs operate as per usual.

8.4 POST Code Diagnostic LEDs

Eight POST code diagnostic LEDs are located on the back left edge of the server board, between the ID LED and NIC 2.

During the system boot process, the BIOS executes a number of platform configuration processes, each of which is assigned a specific hex POST code number. As each configuration routine is started, the BIOS displays the given POST code to the POST code diagnostic LEDs on the back edge of the server board. To assist in troubleshooting a system hang during the POST process, you can use the Diagnostic LEDs to identify the last POST process executed. For a complete description of how these LEDs are read and a list of all supported POST codes, refer to appendix.

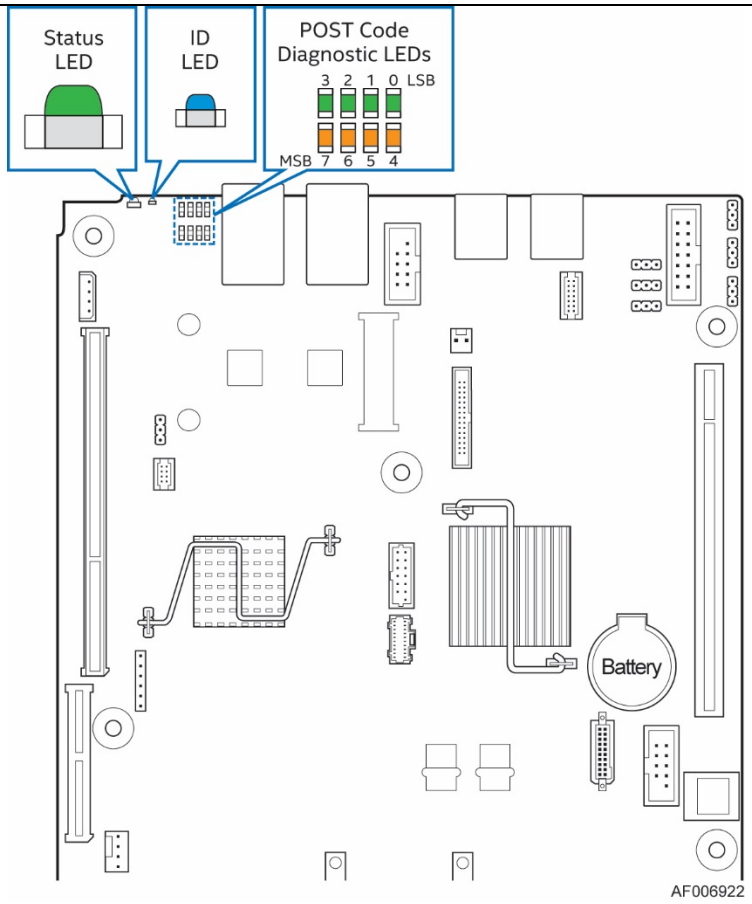


Figure 40. Rear Panel Diagnostic LEDs

9 Platform Management

Platform management is supported by several hardware and software components integrated on the server board that work together to support the following:

- Control systems functions – power system, ACPI, system reset control, system initialization, front panel interface, system event log
- Monitor various board and system sensors, regulate platform thermals and performance in order to maintain (when possible) server functionality in the event of component failure and/or environmentally stressed conditions
- Monitor and report system health
- Provide an interface for Server Management Software applications

This chapter provides a high level overview of the platform management features and functionality implemented on the server board.

The Intel® Server System *BMC Firmware External Product Specification (EPS)* and the Intel® Server System *BIOS External Product Specification (EPS)* for Intel® Server products based on the Intel® Xeon® processor E5 v3 product family should be referenced for more in-depth and design level platform management information.

9.1 Management Feature Set Overview

The following sections outline features that the integrated BMC firmware can support. Support and utilization for some features is dependent on the server platform in which the server board is integrated and any additional system level components and options that may be installed.

9.1.1 IPMI 2.0 Features Overview

- Baseboard management controller (BMC)
- IPMI Watchdog timer
- Messaging support, including command bridging and user/session support
- Chassis device functionality, including power/reset control and BIOS boot flags support
- Event receiver device: The BMC receives and processes events from other platform subsystems
- Field Replaceable Unit (FRU) inventory device functionality: The BMC supports access to system FRU devices using IPMI FRU commands
- System Event Log (SEL) device functionality: The BMC supports and provides access to a SEL including SEL Severity Tracking and the Extended SEL
- Sensor Data Record (SDR) repository device functionality: The BMC supports storage and access of system SDRs
- Sensor device and sensor scanning/monitoring: The BMC provides IPMI management of sensors. It polls sensors to monitor and report system health
- IPMI interfaces
    - Host interfaces include system management software (SMS) with receive message queue support, and server management mode (SMM)
    - IPMB interface

  - LAN interface that supports the IPMI-over-LAN protocol (RMCP, RMCP+)
- Serial-over-LAN (SOL)
- ACPI state synchronization: The BMC tracks ACPI state changes that are provided by the BIOS
- BMC self-test: The BMC performs initialization and run-time self-tests and makes results available to external entities

See also the *Intelligent Platform Management Interface Specification Second Generation v2.0*.

9.1.2 Non IPMI Features Overview

The BMC supports the following non-IPMI features.

- In-circuit BMC firmware update
- Fault resilient booting (FRB): FRB2 is supported by the watchdog timer functionality.
- Chassis intrusion detection (dependent on platform support)
- Basic fan control using Control version 2 SDRs
- Fan redundancy monitoring and support
- Enhancements to fan speed control
- Power supply redundancy monitoring and support
- Hot-swap fan support
- Acoustic management: Support for multiple fan profiles
- Signal testing support: The BMC provides test commands for setting and getting platform signal states
- The BMC generates diagnostic beep codes for fault conditions
- System GUID storage and retrieval
- Front panel management: The BMC controls the system status LED and chassis ID LED. It supports secure lockout of certain front panel functionality and monitors button presses. The chassis ID LED is turned on using a front panel button or a command.
- Power state retention
- Power fault analysis
- Intel® Light-Guided Diagnostics
- Power unit management: Support for power unit sensor. The BMC handles power-good dropout conditions.
- DIMM temperature monitoring: New sensors and improved acoustic management using closed-loop fan control algorithm taking into account DIMM temperature readings.
- Address Resolution Protocol (ARP): The BMC sends and responds to ARPs (supported on embedded NICs).
- Dynamic Host Configuration Protocol (DHCP): The BMC performs DHCP (supported on embedded NICs).
- Platform environment control interface (PECI) thermal management support
- E-mail alerting
- Support for embedded web server UI in Basic Manageability feature set.
- Enhancements to embedded web server

- Human-readable SEL
- Additional system configurability
- Additional system monitoring capability
- Enhanced on-line help

- Integrated KVM.
- Enhancements to KVM redirection
  - Support for higher resolution
- Integrated Remote Media Redirection
- Lightweight Directory Access Protocol (LDAP) support
- Intel® Intelligent Power Node Manager support
- Embedded platform debug feature which allows capture of detailed data for later analysis.
- Provisioning and inventory enhancements:
  - Inventory data/system information export (partial SMBIOS table)
- DCMI 1.5 compliance (product-specific).
- Management support for PMBus* rev1.2 compliant power supplies
- BMC Data Repository (Managed Data Region Feature)
- Support for an Intel® Local Control Display Panel
- System Airflow Monitoring
- Exit Air Temperature Monitoring
- Ethernet Controller Thermal Monitoring
- Global Aggregate Temperature Margin Sensor
- Memory Thermal Management
- Power Supply Fan Sensors
- Energy Star Server Support
- Smart Ride Through (SmaRT)/ Closed Loop System Throttling (CLST)
- Power Supply Cold Redundancy
- Power Supply FW Update
- Power Supply Compatibility Check
- BMC FW reliability enhancements:
  - Redundant BMC boot blocks to avoid possibility of a corrupted boot block resulting in a scenario that prevents a user from updating the BMC.
  - BMC System Management Health Monitoring.

9.2 Platform Management Features and Functions

9.2.1 Power Subsystem

The server board supports several power control sources which can initiate power-up or power-down activity.

Source	External Signal Name or Internal Subsystem	Capabilities
Power button	Front panel power button	Turns power on or off
BMC watchdog timer	Internal BMC timer	Turns power off, or power cycle
BMC chassis control Commands	Routed through command processor	Turns power on or off, or power cycle
Power state retention	Implemented by means of BMC internal logic	Turns power on when AC power returns
Chipset	Sleep S4/S5 signal (same as *POWER_ON*)	Turns power on or off
CPU Thermal	Processor Thermtrip	Turns power off
PCH Thermal	PCH Thermtrip	Turns power off

9.2.2 Advanced Configuration and Power Interface (ACPI)

The server board has support for the following ACPI states.

Table 46. ACPI Power States

State	Supported	Description
S0	Yes	Working. ▪ The front panel power LED is on (not controlled by the BMC). ▪ The fans spin at the normal speed, as determined by sensor inputs. ▪ Front panel buttons work normally.
S1	No	Not supported.
S2	No	Not supported.
S3	No	Supported only on Workstation platforms. See appropriate Platform Specific Information for more information.
S4	No	Not supported.
S5	Yes	Soft off. ▪ The front panel buttons are not locked. ▪ The fans are stopped. ▪ The power-up process goes through the normal boot process. ▪ The power, reset, and ID buttons are unlocked.

9.2.3 System Initialization

During system initialization, both the BIOS and the BMC initialize the following items.

9.2.3.1 Processor Tcontrol Setting

Processors used with this chipset implement a feature called Tcontrol, which provides a processor-specific value that can be used to adjust the fan-control behavior to achieve optimum cooling and acoustics. The BMC reads these from the CPU through PECI Proxy mechanism provided by Manageability Engine (ME). The BMC uses these values as part of the fan-speed-control algorithm.

9.2.3.2 Fault Resilient Booting (FRB)

Fault resilient booting (FRB) is a set of BIOS and BMC algorithms and hardware support that allow a multiprocessor system to boot even if the bootstrap processor (BSP) fails. Only FRB2 is supported using watchdog timer commands.

FRB2 refers to the FRB algorithm that detects system failures during POST. The BIOS uses the BMC watchdog timer to back up its operation during POST. The BIOS configures the watchdog timer to indicate that the BIOS is using the timer for the FRB2 phase of the boot operation.

After the BIOS has identified and saved the BSP information, it sets the FRB2 timer use bit and loads the watchdog timer with the new timeout interval.

If the watchdog timer expires while the watchdog use bit is set to FRB2, the BMC (if so configured) logs a watchdog expiration event showing the FRB2 timeout in the event data bytes. The BMC then hard resets the system, assuming the BIOS-selected reset as the watchdog timeout action.

The BIOS is responsible for disabling the FRB2 timeout before initiating the option ROM scan and before displaying a request for a boot password. If the processor fails and causes an FRB2 timeout, the BMC resets the system.

The BIOS gets the watchdog expiration status from the BMC. If the status shows an expired FRB2 timer, the BIOS enters the failure in the system event log (SEL). In the OEM bytes entry in the SEL, the last POST code generated during the previous boot attempt is written. FRB2 failure is not reflected in the processor status sensor value.

The FRB2 failure does not affect the front panel LEDs.

9.2.3.3 Post Code Display

The BMC, upon receiving standby power, initializes internal hardware to monitor port 80h (POST code) writes. Data written to port 80h is output to the system POST LEDs.

The BMC deactivates POST LEDs after POST had completed.

9.2.4 System Event Log (SEL)

The BMC implements the system event log as specified in the *Intelligent Platform Management Interface Specification,* Version 2.0. The SEL is accessible regardless of the system power state through the BMC's in-band and out-of-band interfaces.

The BMC allocates 95231bytes (approx. 93 KB) of non-volatile storage space to store system events. The SEL timestamps may not be in order. Up to 3,639 SEL records can be stored at a time. Because the SEL is circular, any command that results in an overflow of the SEL beyond the allocated space will overwrite the oldest entries in the SEL, while setting the overflow flag.

9.3 Sensor Monitoring

The BMC monitors system hardware and reports system health. The information gathered from physical sensors is translated into IPMI sensors as part of the “IPMI Sensor Model”. The BMC also reports various

system state changes by maintaining virtual sensors that are not specifically tied to physical hardware. This section describes general aspects of BMC sensor management as well as describing how specific sensor types are modeled. Unless otherwise specified, the term “sensor” refers to the IPMI sensor-model definition of a sensor.

9.3.1 Sensor Scanning

The value of many of the BMC’s sensors is derived by the BMC FW periodically polling physical sensors in the system to read temperature, voltages, and so on. Some of these physical sensors are built-in to the BMC component itself and some are physically separated from the BMC. Polling of physical sensors for support of IPMI sensor monitoring does not occur until the BMC’s operational code is running and the IPMI FW subsystem has completed initialization. IPMI sensor monitoring is not supported in the BMC boot code. Additionally, the BMC selectively polls physical sensors based on the current power and reset state of the system and the availability of the physical sensor when in that state. For example, non-standby voltages are not monitored when the system is in S4 or S5 power state.

9.3.2 Sensor Rearm Behavior

9.3.2.1 Manual versus Re-arm Sensors

Sensors can be either manual or automatic re-arm. An automatic re-arm sensor will "re-arm" (clear) the assertion event state for a threshold or offset it if that threshold or offset is de-asserted after having been asserted. This allows a subsequent assertion of the threshold or an offset to generate a new event and associated side-effect. An example side-effect would be boosting fans due to an upper critical threshold crossing of a temperature sensor. The event state and the input state (value) of the sensor track each other. Most sensors are auto-rearm.

A manual re-arm sensor does not clear the assertion state even when the threshold or offset becomes de-asserted. In this case, the event state and the input state (value) of the sensor do not track each other. The event assertion state is "sticky". The following methods can be used to re-arm a sensor:

- Automatic re-arm – Only applies to sensors that are designated as “auto-rearm”.
- IPMI command Re-arm Sensor Event
- BMC internal method – The BMC may re-arm certain sensors due to a trigger condition. For example, some sensors may be re-armed due to a system reset. A BMC reset will re-arm all sensors.
- System reset or DC power cycle will re-arm all system fan sensors.

9.3.2.2 Re-arm and Event Generation

All BMC-owned sensors that show an asserted event status generate a de-assertion SEL event when the sensor is re-armed, provided that the associated SDR is configured to enable a de-assertion event for that condition. This applies regardless of whether the sensor is a threshold/analog sensor or a discrete sensor.

To manually re-arm the sensors, the sequence is outlined below:

1. A failure condition occurs and the BMC logs an assertion event.
2. If this failure condition disappears, the BMC logs a de-assertion event (if so configured).
3. The sensor is re-armed by one of the methods described in the previous section.

4. The BMC clears the sensor status.
5. The sensor is put into "reading-state-unavailable" state until it is polled again or otherwise updated.
6. The sensor is updated and the “reading-state-unavailable” state is cleared. A new assertion event will be logged if the fault state is once again detected.

All auto-rearm sensors that show an asserted event status generate a de-assertion SEL event at the time the BMC detects that the condition causing the original assertion is no longer present; and the associated SDR is configured to enable a de-assertion event for that condition.

9.3.3 BIOS Event-Only Sensors

BIOS-owned discrete sensors are used for event generation only and are not accessible through IPMI sensor commands like the *Get Sensor Reading* command. Note that in this case the sensor owner designated in the SDR is not the BMC.

An example of this usage would be the SELs logged by the BIOS for uncorrectable memory errors. Such SEL entries would identify a BIOS-owned sensor ID.

9.3.4 Margin Sensors

There is sometimes a need for an IPMI sensor to report the difference (margin) from a non-zero reference offset. For the purposes of this document, these type sensors are referred to as margin sensors. For instance, for the case of a temperature margin sensor, if the reference value is 90 degrees and the actual temperature of the device being monitored is 85 degrees, the margin value would be -5.

9.3.5 IPMI Watchdog Sensor

The BMC supports a Watchdog Sensor as a means to log SEL events due to expirations of the IPMI 2.0 compliant Watchdog Timer.

9.3.6 BMC Watchdog Sensor

The BMC supports an IPMI sensor to report that a BMC reset has occurred due to action taken by the BMC Watchdog feature. A SEL event will be logged whenever either the BMC FW stack is reset or the BMC CPU itself is reset.

9.3.7 BMC System Management Health Monitoring

The BMC tracks the health of each of its IPMI sensors and reports failures by providing a “BMC FW Health” sensor of the IPMI 2.0 sensor type Management Subsystem Health with support for the Sensor Failure offset. Only assertions should be logged into the SEL for the Sensor Failure offset. The BMC Firmware Health sensor asserts for any sensor when 10 consecutive sensor errors are read. These are not standard sensor events (that is, threshold crossings or discrete assertions). These are BMC Hardware Access Layer (HAL) errors. If a successful sensor read is completed, the counter resets to zero.

9.3.8 VR Watchdog Timer

The BMC FW monitors that the power sequence for the board VR controllers is completed when a DC power-on is initiated. Incompletion of the sequence indicates a board problem, in which case the FW powers down the system.

The BMC FW supports a discrete IPMI sensor for reporting and logging this fault condition.

9.3.9 System Airflow Monitoring

The BMC provides an IPMI sensor to report the volumetric system airflow in CFM (cubic feet per minute). The air flow in CFM is calculated based on the system fan PWM values. The specific Pulse Width Modulation (PWM or PWMs) used to determine the CFM is SDR configurable. The relationship between PWM and CFM is based on a lookup table in an OEM SDR.

The airflow data is used in the calculation for exit air temperature monitoring. It is exposed as an IPMI sensor to allow a datacenter management application to access this data for use in rack-level thermal management.

9.3.10 Thermal Monitoring

The BMC provides monitoring of component and board temperature sensing devices. This monitoring capability is instantiated in the form of IPMI analog/threshold or discrete sensors, depending on the nature of the measurement.

For analog/threshold sensors, with the exception of *Processor Temperature* sensors, critical and non-critical thresholds (upper and lower) are set through SDRs and event generation enabled for both assertion and de-assertion events.

For discrete sensors, both assertion and de-assertion event generation are enabled.

Mandatory monitoring of platform thermal sensors includes:

- Inlet temperature (physical sensor is typically on system front panel or HDD back plane)
- Board ambient thermal sensors
- Processor temperature
- Memory (DIMM) temperature
- CPU VRD Hot monitoring
- Power supply (only supported for PMBus*-compliant PSUs)

Additionally, the BMC FW may create “virtual” sensors that are based on a combination of aggregation of multiple physical thermal sensors and application of a mathematical formula to thermal or power sensor readings.

9.3.10.1 Absolute Value versus Margin Sensors

Thermal monitoring sensors fall into three basic categories:

- Absolute temperature sensors – These are analog/threshold sensors that provide a value that corresponds to an absolute temperature value.

- Thermal margin sensors – These are analog/threshold sensors that provide a value that is relative to some reference value.
- Thermal fault indication sensors – These are discrete sensors that indicate a specific thermal fault condition.

9.3.10.2 Processor DTS-Spec Margin Sensor(s)

Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family incorporate a DTS based thermal spec. This allows a much more accurate control of the thermal solution and will enable lower fan speeds and lower fan power consumption. The main usage of this sensor is as an input to the BMC's fan control algorithms. The BMC implements this as a threshold sensor. There is one DTS sensor for each installed physical processor package. Thresholds are not set and alert generation is not enabled for these sensors.

9.3.10.3 Processor Thermal Margin Sensor(s)

Each processor supports a physical thermal margin sensor per core that is readable through the PECI interface. This provides a relative value representing a thermal margin from the core's throttling thermal trip point. Assuming that temp controlled throttling is enabled; the physical core temp sensor reads '0', which indicates the processor core is being throttled.

The BMC supports one IPMI processor (margin) temperature sensor per physical processor package. This sensor aggregates the readings of the individual core temperatures in a package to provide the hottest core temperature reading. When the sensor reads '0', it indicates that the hottest processor core is throttling.

Due to the fact that the readings are capped at the core's thermal throttling trip point (reading = 0), thresholds are not set and alert generation is not enabled for these sensors.

9.3.10.4 Processor Thermal Control Monitoring (Prochot)

The BMC FW monitors the percentage of time that a processor has been operationally constrained over a given time window (nominally six seconds) due to internal thermal management algorithms engaging to reduce the temperature of the device. When any processor core temperature reaches its maximum operating temperature, the processor package PROCHOT# (processor hot) signal is asserted and these management algorithms, known as the Thermal Control Circuit (TCC), engage to reduce the temperature, provided TCC is enabled. TCC is enabled by BIOS during system boot. This monitoring is instantiated as one IPMI analog/threshold sensor per processor package. The BMC implements this as a threshold sensor on a per-processor basis.

Under normal operation, this sensor is expected to read '0' indicating that no processor throttling has occurred.

The processor provides PECI-accessible counters, one for the total processor time elapsed and one for the total thermally constrained time, which are used to calculate the percentage assertion over the given time window.

9.3.10.5 Processor Voltage Regulator (VRD) Over-Temperature Sensor

The BMC monitors processor VRD_HOT# signals. The processor VRD_HOT# signals are routed to the respective processor PROCHOT# input in order to initiate throttling to reduce processor power draw, therefore indirectly lowering the VRD temperature.

There is one processor VRD_HOT# signal per CPU slot. The BMC instantiates one discrete IPMI sensor for each VRD_HOT# signal. This sensor monitors a digital signal that indicates whether a processor VRD is running in an over-temperature condition. When the BMC detects that this signal is asserted it will cause a sensor assertion which will result in an event being written into the sensor event log (SEL).

9.3.10.6 Inlet Temperature Sensor

Each platform supports a thermal sensor for monitoring the inlet temperature. The inlet temperature sensor is on the backplane of Intel® Server Chassis with address 21h. For third-party chassis, sensor 20h which is on the front edge of the baseboard can be used as inlet temperature sensor with several degrees of preheat from front end.

9.3.10.7 Baseboard Ambient Temperature Sensor(s)

The server baseboard provides one or more physical thermal sensors for monitoring the ambient temperature of a board location. This is typically to provide rudimentary thermal monitoring of components that lack internal thermal sensors.

9.3.10.8 Chpiset Thermal Monitoring

The BMC monitors the chipset temperature. This is instantiated as an analog (threshold) IPMI thermal sensor.

9.3.10.9 Exit Air Temperature Monitoring

This sensor is only available in Intel® Server Chassis. The BMC synthesizes a virtual sensor to approximate system exit air temperature for use in fan control. This is calculated based on the total power being consumed by the system and the total volumetric air flow provided by the system fans. Each system shall be characterized in tabular format to understand total volumetric flow versus fan speed. The BMC calculates an average exit air temperature based on the total system power, front panel temperature, the volumetric system air flow (cubic feet per meter or CFM), and altitude range.

The Exit Air temp sensor is only available when PMBus* power supplies are installed.

9.3.10.10 Ethernet Controller Thermal Monitoring

The Intel® Ethernet Controller I350 supports an on-die thermal sensor. For baseboard Ethernet controllers that use these devices, the BMC will monitor the sensors and use this data as input to the fan speed control. The BMC will instantiate an IPMI temperature sensor for each device on the baseboard.

9.3.10.11 Memory VRD-Hot Sensor(s)

The BMC monitors memory VRD_HOT# signals. The memory VRD_HOT# signals are routed to the respective processor MEMHOT# inputs in order to throttle the associated memory to effectively lower the temperature of the VRD feeding that memory.

For Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor family there are 2 memory VRD_HOT# signals per CPU slot. The BMC instantiates one discrete IPMI sensor for each memory VRD_HOT# signal.

9.3.10.12 Add-in Module Thermal Monitoring

Some boards have dedicated slots for an IO module and/or a SAS module. For boards that support these slots, the BMC will instantiate an IPMI temperature sensor for each slot. The modules themselves may or may not provide a physical thermal sensor (a TMP75 device). If the BMC detects that a module is installed, it will attempt to access the physical thermal sensor and, if found, enable the associated IPMI temperature sensor.

9.3.10.13 Processor ThermTrip

When a Processor ThermTrip occurs, the system hardware will automatically power down the server. If the BMC detects that a ThermTrip occurred, then it will set the ThermTrip offset for the applicable processor status sensor.

9.3.10.14 Server South Bridge (SSB) ThermTrip Monitoring

The BMC supports SSB ThermTrip monitoring that is instantiated as an IPMI discrete sensor. When a SSB ThermTrip occurs, the system hardware will automatically power down the server and the BMC will assert the sensor offset and log an event.

9.3.10.15 DIMM ThermTrip Monitoring

The BMC supports DIMM ThermTrip monitoring that is instantiated as one aggregate IPMI discrete sensor per CPU. When a DIMM ThermTrip occurs, the system hardware will automatically power down the server and the BMC will assert the sensor offset and log an event.

This is a manual re-arm sensor that is rearmed on system resets and power-on (AC or DC power on transitions).

9.3.11 Processor Sensors

The BMC provides IPMI sensors for processors and associated components, such as voltage regulators and fans. The sensors are implemented on a per-processor basis.

Table 47. Processor Sensors

Sensor Name	Per-Processor Socket	Description
Processor Status	Yes	Processor presence and fault state
Digital Thermal Sensor	Yes	Relative temperature reading by means of PECI
Processor VRD Over-Temperature Indication	Yes	Discrete sensor that indicates a processor VRD has crossed an upper operating temperature threshold

Sensor Name	Per-Processor Socket	Description
Processor Voltage	Yes	Threshold sensor that indicates a processor power-good state
Processor Thermal Control (Prochot)	Yes	Percentage of time a processor is throttling due to thermal conditions

9.3.11.1 Processor Status Sensors

The BMC provides an IPMI sensor of type processor for monitoring status information for each processor slot. If an event state (sensor offset) has been asserted, it remains asserted until one of the following happens:

1. A Rearm Sensor Events command is executed for the processor status sensor.
2. AC or DC power cycle, system reset, or system boot occurs.

The BMC provides system status indication to the front panel LEDs for processor fault conditions shown in Table 48.

CPU Presence status is not saved across AC power cycles and therefore will not generate a de-assertion after cycling AC power.

Table 48. Processor Status Sensor Implementation

Offset	Processor Status	Detected By
0	Internal error (IERR)	Not Supported
1	Thermal trip	BMC
2	FRB1/BIST failure	Not Supported
3	FRB2/Hang in POST failure	BIOS[1]
4	FRB3/Processor startup/initialization failure (CPU fails to start)	Not Supported
5	Configuration error (for DMI)	BIOS[1]
6	SM BIOS uncorrectable CPU-complex error	Not Supported
7	Processor presence detected	BMC
8	Processor disabled	Not Supported
9	Terminator presence detected	Not Supported

Note:

1. Fault is not reflected in the processor status sensor.

9.3.11.2 Processor Population Fault (CPU Missing) Sensor

The BMC supports a *Processor Population Fault* sensor. This is used to monitor for the condition in which processor slots are not populated as required by the platform hardware to allow power-on of the system.

At BMC startup, the BMC will check for the fault condition and set the sensor state accordingly. The BMC also checks for this fault condition at each attempt to DC power-on the system. At each DC power-on attempt, a beep code is generated if this fault is detected.

The following steps are used to correct the fault condition and clear the sensor state:

1. AC power down the server.

2. Install the missing processor into the correct slot.
3. AC power on the server.

9.3.11.3 ERR2 Timeout Monitoring

The BMC supports an ERR2 Timeout Sensor (1 per CPU) that asserts if a CPU's ERR2 signal has been asserted for longer than a fixed time period (> 90 seconds). ERR2 is a processor signal that indicates when the IIO (Integrated IO module in the processor) has a fatal error which could not be communicated to the core to trigger SMI. ERR2 events are fatal error conditions, where the BIOS and OS will attempt to gracefully handle error, but may not be always do so reliably. A continuously asserted ERR2 signal is an indication that the BIOS cannot service the condition that caused the error. This is usually because that condition prevents the BIOS from running.

When an ERR2 timeout occurs, the BMC asserts/de-asserts the ERR2 Timeout Sensor, and logs a SEL event for that sensor. The default behavior for BMC core firmware is to initiate a system reset upon detection of an ERR2 timeout. The BIOS setup utility provides an option to disable or enable system reset by the BMC for detection of this condition.

9.3.11.4 CATERR Sensor

The BMC supports a CATERR sensor for monitoring the system CATERR signal.

The CATERR signal is defined as having 3 states:

- high (no event)
- pulsed low (possibly fatal may be able to recover)
- low (fatal)

All processors in a system have their CATERR pins tied together. The pin is used as a communication path to signal a catastrophic system event to all CPUs. The BMC has direct access to this aggregate CATERR signal.

The BMC only monitors for the "CATERR held low" condition. A pulsed low condition is ignored by the BMC. If a CATERR-low condition is detected, the BMC logs an error message to the SEL against the CATERR sensor and the default action after logging the SEL entry is to reset the system. The BIOS setup utility provides an option to disable or enable system reset by the BMC for detection of this condition.

The sensor is rearmed on power-on (AC or DC power on transitions). It is not rearmed on system resets in order to avoid multiple SEL events that could occur due to a potential reset loop if the CATERR keeps recurring, which would be the case if the CATERR was due to an MSID mismatch condition.

When the BMC detects that this aggregate CATERR signal has asserted, it can then go through PECI to query each CPU to determine which one was the source of the error and write an OEM code identifying the CPU slot into an event data byte in the SEL entry. If PECI is non-functional (it isn't guaranteed in this situation), then the OEM code should indicate that the source is unknown.

Event data byte 2 and byte 3 for CATERR sensor SEL events

ED1 – 0xA1
ED2 - CATERR type.
0: Unknown

1: CATERR
2: CPU Core Error (not supported on Intel® Xeon™ Phi™ processor product family)
3: MSID Mismatch

ED3 - CPU bitmap that causes the system CATERR.
[0]: CPU1

When a CATERR Timeout event is determined to be a CPU 3-strike timeout, The BMC shall log the logical FRU information (e.g. bus/dev/func for a PCIe device, CPU, or DIMM) that identifies the FRU that caused the error in the extended SEL data bytes. In this case, Ext-ED0 will be set to 0x70 and the remaining ED1-ED7 will be set according to the device type and info available.

9.3.11.5 MSID Mismatch Sensor

The BMC supports a *MSID Mismatch* sensor for monitoring for the fault condition that will occur if there is a power rating incompatibility between a baseboard and a processor.

The sensor is rearmed on power-on (AC or DC power on transitions).

9.3.12 Voltage Monitoring

The BMC provides voltage monitoring capability for voltage sources on the main board and processors such that all major areas of the system are covered. This monitoring capability is instantiated in the form of IPMI analog/threshold sensors.

9.3.12.1 DIMM Voltage Sensors

Some systems support either LVDDR (Low Voltage DDR) memory or regular (non-LVDDR) memory. During POST, the system BIOS detects which type of memory is installed and configures the hardware to deliver the correct voltage.

Since the nominal voltage range is different, this necessitates the ability to set different thresholds for any associated IPMI voltage sensors. The BMC FW supports this by implementing separate sensors (that is, separate IPMI sensor numbers) for each nominal voltage range supported for a single physical sensor and it enables/disables the correct IPMI sensor based on which type memory is installed. The sensor data records for both these DIMM voltage sensor types have scanning disabled by default. Once the BIOS has completed its POST routine, it is responsible for communicating the DIMM voltage type to the BMC which will then enable sensor scanning of the correct DIMM voltage sensor.

9.3.13 Fan Monitoring

BMC fan monitoring support includes monitoring of fan speed (RPM) and fan presence.

9.3.13.1 Fan Tach Sensors

Fan Tach sensors are used for fan failure detection. The reported sensor reading is proportional to the fan's RPM. This monitoring capability is instantiated in the form of IPMI analog/threshold sensors.

Most fan implementations provide for a variable speed fan, so the variations in fan speed can be large. Therefore the threshold values must be set sufficiently low as to not result in inappropriate threshold crossings.

Fan tach sensors are implemented as manual re-arm sensors because a lower-critical threshold crossing can result in full boosting of the fans. This in turn may cause a failing fan's speed to rise above the threshold and can result in fan oscillations.

As a result, fan tach sensors do not auto-rearm when the fault condition goes away but rather are rearmed for either of the following occurrences:

1. The system is reset or power-cycled.
2. The fan is removed and either replaced with another fan or re-inserted. This applies to hot-swappable fans only. This re-arm is triggered by change in the state of the associated fan presence sensor.

After the sensor is rearmed, if the fan speed is detected to be in a normal range, the failure conditions shall be cleared and a de-assertion event shall be logged.

9.3.13.2 Fan Presence Sensors

Some chassis and server boards provide support for hot-swap fans. These fans can be removed and replaced while the system is powered on and operating normally. The BMC implements fan presence sensors for each hot swappable fan. These are instantiated as IPMI discrete sensors.

Events are only logged for fan presence upon changes in the presence state after AC power is applied (no events logged for initial state).

9.3.13.3 Fan Redundancy Sensor

The BMC supports redundant fan monitoring and implements fan redundancy sensors for products that have redundant fans. Support for redundant fans is chassis-specific.

A fan redundancy sensor generates events when its associated set of fans transits between redundant and non-redundant states, as determined by the number and health of the component fans. The definition of fan redundancy is configuration dependent. The BMC allows redundancy to be configured on a per fan-redundancy sensor basis through OEM SDR records.

There is a fan redundancy sensor implemented for each redundant group of fans in the system.

Assertion and de-assertion event generation is enabled for each redundancy state.

9.3.13.4 Power Supply Fan Sensors

Monitoring is implemented through IPMI discrete sensors, one for each power supply fan. The BMC polls each installed power supply using the PMBus* fan status commands to check for failure conditions for the power supply fans. The BMC asserts the "performance lags" offset of the IPMI sensor if a fan failure is detected.

Power supply fan sensors are implemented as manual re-arm sensors because a failure condition can result in boosting of the fans. This in turn may cause a failing fan's speed to rise above the "fault" threshold and can

result in fan oscillations. As a result, these sensors do not auto-rearm when the fault condition goes away but rather are rearmed only when the system is reset or power-cycled, or the PSU is removed and replaced with the same or another PSU.

After the sensor is rearmed, if the fan is no longer showing a failed state, the failure condition in the IPMI sensor shall be cleared and a de-assertion event shall be logged.

9.3.13.5 Monitoring for "Fans Off" Scenario

On Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family, it is likely that there will be situations where specific fans are turned off based on current system conditions. BMC Fan monitoring will comprehend this scenario and not log false failure events. The recommended method is for the BMC FW to halt updates to the value of the associated fan tach sensor and set that sensor's IPMI sensor state to "reading-state-unavailable" when this mode is active. Management software must comprehend this state for fan tach sensors and not report these as failure conditions.

The scenario for which this occurs is that the BMC Fan Speed Control (FSC) code turns off the fans by setting the PWM for the domain to 0. This is done when based on one or more global aggregate thermal margin sensor readings dropping below a specified threshold.

By default the fans-off feature will be disabled. There is a BMC command and BIOS setup option to enable/disable this feature.

The SmaRT/CLST system feature will also momentarily gate power to all the system fans to reduce overall system power consumption in response to a power supply event (for example, to ride out an AC power glitch). However, for this scenario, the fan power is gated by hardware for only 100ms, which should not be long enough to result in triggering a fan fault SEL event

9.3.14 Standard Fan Management

The BMC controls and monitors the system fans. Each fan is associated with a fan speed sensor that detects fan failure and may also be associated with a fan presence sensor for hot-swap support. For redundant fan configurations, the fan failure and presence status determines the fan redundancy sensor state.

The system fans are divided into fan domains, each of which has a separate fan speed control signal and a separate configurable fan control policy. A fan domain can have a set of temperature and fan sensors associated with it. These are used to determine the current fan domain state.

A fan domain has three states:

- The sleep and boost states have fixed (but configurable through OEM SDRs) fan speeds associated with them.
- The nominal state has a variable speed determined by the fan domain policy. An OEM SDR record is used to configure the fan domain policy.

The fan domain state is controlled by several factors. They are listed below in order of precedence, high to low:

- Boost

- Associated fan is in a critical state or missing. The SDR describes which fan domains are boosted in response to a fan failure or removal in each domain. If a fan is removed when the system is in 'Fans-off' mode it will not be detected and there will not be any fan boost till system comes out of 'Fans-off; mode.
- Any associated temperature sensor is in a critical state. The SDR describes which temperature-threshold violations cause fan boost for each fan domain.
- The BMC is in firmware update mode, or the operational firmware is corrupted.
- If any of the above conditions apply, the fans are set to a fixed boost state speed.

- Nominal
  - A fan domain's nominal fan speed can be configured as static (fixed value) or controlled by the state of one or more associated temperature sensors.
  - See section 9.3.14.3 for more details.

9.3.14.1 Fan Redundancy Detection

The BMC supports redundant fan monitoring and implements a fan redundancy sensor. A fan redundancy sensor generates events when its associated set of fans transits between redundant and non-redundant states, as determined by the number and health of the fans. The definition of fan redundancy is configuration dependent. The BMC allows redundancy to be configured on a per fan redundancy sensor basis through OEM SDR records.

A fan failure or removal of hot-swap fans up to the number of redundant fans specified in the SDR in a fan configuration is a non-critical failure and is reflected in the front panel status. A fan failure or removal that exceeds the number of redundant fans is a non-fatal, insufficient-resources condition and is reflected in the front panel status as a non-fatal error.

Redundancy is checked only when the system is in the DC-on state. Fan redundancy changes that occur when the system is DC-off or when AC is removed will not be logged until the system is turned on.

9.3.14.2 Fan Domains

System fan speeds are controlled through pulse width modulation (PWM) signals, which are driven separately for each domain by integrated PWM hardware. Fan speed is changed by adjusting the duty cycle, which is the percentage of time the signal is driven high in each pulse.

The BMC controls the average duty cycle of each PWM signal through direct manipulation of the integrated PWM control registers.

The same device may drive multiple PWM signals.

9.3.14.3 Nominal Fan Speed

A fan domain's nominal fan speed can be configured as static (fixed value) or controlled by the state of one or more associated temperature sensors.

OEM SDR records are used to configure which temperature sensors are associated with which fan control domains and the algorithmic relationship between the temperature and fan speed. Multiple OEM SDRs can reference or control the same fan control domain; and multiple OEM SDRs can reference the same temperature sensors.

The PWM duty-cycle value for a domain is computed as a percentage using one or more instances of a stepwise linear algorithm and a clamp algorithm. The transition from one computed nominal fan speed (PWM value) to another is ramped over time to minimize audible transitions. The ramp rate is configurable by means of the OEM SDR.

Multiple stepwise linear and clamp controls can be defined for each fan domain and used simultaneously. For each domain, the BMC uses the maximum of the domain's stepwise linear control contributions and the sum of the domain's clamp control contributions to compute the domain's PWM value, except that a stepwise linear instance can be configured to provide the domain maximum.

Hysteresis can be specified to minimize fan speed oscillation and to smooth fan speed transitions. If a Tcontrol SDR record does not contain a hysteresis definition, for example, an SDR adhering to a legacy format, the BMC assumes a hysteresis value of zero.

9.3.14.4 Thermal and Acoustic Management

This feature refers to enhanced fan management to keep the system optimally cooled while reducing the amount of noise generated by the system fans. Aggressive acoustics standards might require a trade-off between fan speed and system performance parameters that contribute to the cooling requirements, primarily memory bandwidth. The BIOS, BMC, and SDRs work together to provide control over how this trade-off is determined.

This capability requires the BMC to access temperature sensors on the individual memory DIMMs. Additionally, closed-loop thermal throttling is only supported with buffered DIMMs.

9.3.14.5 Thermal Sensor Input to Fan Speed Control

The BMC uses various IPMI sensors as input to the fan speed control. Some of the sensors are IPMI models of actual physical sensors whereas some are "virtual" sensors whose values are derived from physical sensors using calculations and/or tabular information.

The following IPMI thermal sensors are used as input to the fan speed control:

- Front panel temperature sensor
- Baseboard temperature sensors
- CPU DTS-Spec margin sensors
- DIMM thermal margin sensors
- Exit air temperature sensor
- Global aggregate thermal margin sensors
- SSB (Intel® C610 Series Chipset) temperature sensor
- On-board Ethernet controller temperature sensors (support for this is specific to the Ethernet controller being used)
- Add-in Intel® SAS/IO module temperature sensor(s) (if present)
- Power supply thermal sensors (only available on PMBus*-compliant power supplies)

A simple model is shown in the following figure which gives a high level graphic of the fan speed control structure creates the resulting fan speeds.

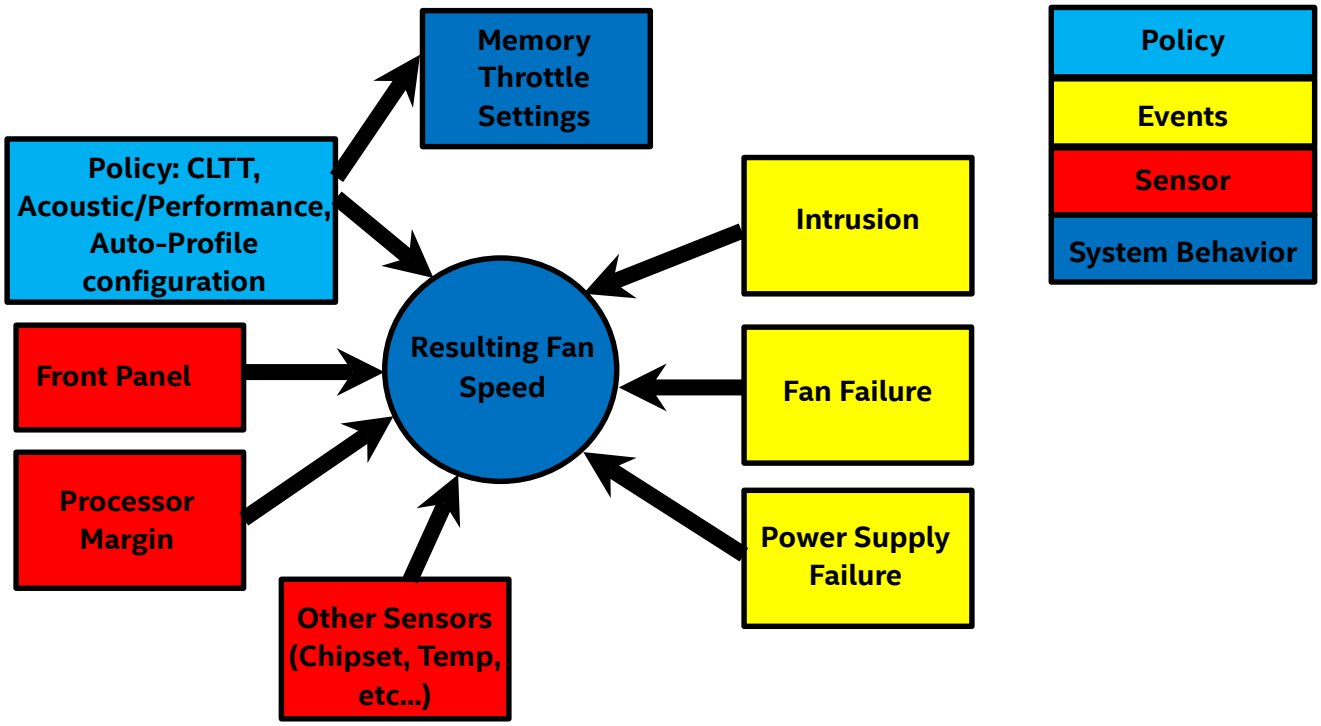


Figure 41. High-level Fan Speed Control Process

9.3.14.5.1 Processor Thermal Management

Processor thermal management utilizes clamp algorithms for which the Processor DTS-Spec margin sensor is a controlling input. This replaces the use of the (legacy) raw DTS sensor reading that was utilized on previous generation platforms. The legacy DTS sensor is retained only for monitoring purposes and is not used as an input to the fan speed control.

9.3.14.5.2 Memory Thermal Management

The system memory is the most complex subsystem to thermally manage as it requires substantial interactions between the BMC, BIOS, and the embedded memory controller. This section provides an overview of this management capability from a BMC perspective.

9.3.14.5.2.1 Memory Thermal Throttling

The system only supports thermal management through closed loop throttling (CLTT) on systems that are installed with DDR4 memory with temperature sensors. Throttling levels are changed dynamically to cap throttling based on memory and system thermal conditions as determined by the system and DIMM power and thermal parameters. The BMC fan speed control functionality is related to the memory throttling mechanism used.

The following terminology is used for the various memory throttling options:

- **Static Closed Loop Thermal Throttling (Static-CLTT):** CLTT control registers are configured by BIOS MRC during POST. The memory throttling is run as a closed-loop system with the DIMM temperature

sensors as the control input. Otherwise, the system does not change any of the throttling control registers in the embedded memory controller during runtime.

- **Dynamic Closed Loop Thermal Throttling (Dynamic-CLTT):** CLTT control registers are configured by BIOS MRC during POST. The memory throttling is run as a closed-loop system with the DIMM temperature sensors as the control input. Adjustments are made to the throttling during runtime based on changes in system cooling (fan speed).

Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family introduce a new type of CLTT which is referred to as Hybrid CLTT for which the Integrated Memory Controller estimates the DRAM temperature in between actual reads of the TSODs. Hybrid CLTTT shall be used on all Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family that have DIMMs with thermal sensors. Therefore, the terms Dynamic-CLTT and Static-CLTT are really referring to this 'hybrid' mode. Note that if the IMC's polling of the TSODs is interrupted, the temperature readings that the BMC gets from the IMC shall be these estimated values.

9.3.14.5.3 DIMM Temperature Sensor Input to Fan Speed Control

A clamp algorithm is used for controlling fan speed based on DIMM temperatures. Aggregate DIMM temperature margin sensors are used as the control input to the algorithm.

9.3.14.5.4 Dynamic (Hybrid) CLTT

The system will support dynamic (memory) CLTT for which the BMC FW dynamically modifies thermal offset registers in the IMC during runtime based on changes in system cooling (fan speed). For static CLTT, a fixed offset value is applied to the TSOD reading to get the die temperature; however this does not provide as accurate results as when the offset takes into account the current airflow over the DIMM, as is done with dynamic CLTT.

In order to support this feature, the BMC FW will derive the air velocity for each fan domain based on the PWM value being driven for the domain. Since this relationship is dependent on the chassis configuration, a method must be used which support this dependency (for example, through OEM SDR) that establishes a lookup table providing this relationship.

BIOS will have an embedded lookup table that provides thermal offset values for each DIMM type, altitude setting, and air velocity range (3 ranges of air velocity are supported). During system boot BIOS will provide 3 offset values (corresponding to the 3 air velocity ranges) to the BMC for each enabled DIMM. Using this data the BMC FW constructs a table that maps the offset value corresponding to a given air velocity range for each DIMM. During runtime the BMC applies an averaging algorithm to determine the target offset value corresponding to the current air velocity and then the BMC writes this new offset value into the IMC thermal offset register for the DIMM.

9.3.14.5.5 Auto-profile

The auto-profile feature is to improve upon previous platform configuration-dependent FSC and maintain competitive acoustics. This feature is not available for third-party customization.

The BIOS and BMC will handshake to automatically understand configuration details and automatically select the optimal fan speed control profile in the BMC.

Users will only select a performance or an acoustic profile selection from the BIOS menu for use with Intel® Server Chassis and the fan speed control will be optimal for the configuration loaded.

Users can still choose performance or acoustic profile in BIOS setting. Default is acoustic. Performance option is recommended if any other high power add-in cards (higher than 75W) are installed.

9.3.14.5.6 ASHRAE Compliance

System requirements for ASHRAE compliance is defined in the *Common Fan Speed Control & Thermal Management Platform Architecture Specification*. Altitude-related changes in fan speed control are handled through profiles for different altitude ranges.

9.3.14.6 Power Supply Fan Speed Control

This section describes the system level control of the fans internal to the power supply over the PMBus*. Some, but not all Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family will require that the power supplies be included in the system level fan speed control. For any system that requires either of these capabilities, the power supply must be PMBus*-compliant.

9.3.14.6.1 System Control of Power Supply Fans

Some products require that the BMC control the speed of the power supply fans, as is done with normal system (chassis) fans, except that the BMC cannot reduce the power supply fan any lower than the internal power supply control is driving it. For these products the BMC FW must have the ability to control and monitor the power supply fans through PMBus* commands. The power supply fans are treated as a system fan domain for which fan control policies are mapped, just as for chassis system fans, with system thermal sensors (rather than internal power supply thermal sensors) used as the input to a clamp algorithm for the power supply fan control. This domain has both piecewise clipping curves and clamped sensors mapped into the power supply fan domain. All the power supplies can be defined as a single fan domain.

9.3.14.6.2 Use of Power Supply Thermal Sensors as Input to System (Chassis) Fan Control

Some products require that the power supply internal thermal sensors are used as control inputs to the system (chassis) fans in the same manner as other system thermal sensors are used. The power supply thermal sensors are included as clamped sensors into one or more system fan domains, which may include the power supply fan domain.

9.3.14.7 Fan Boosting due to Fan Failures

Intel® Server Systems supporting the Intel® Xeon™ Phi™ processor product family introduce additional capabilities for handling fan failure or removal as described in this section.

Each fan failure shall be able to define a unique response from all other fan domains. An OEM SDR table defines the response of each fan domain based on a failure of any fan, including both system and power supply fans (for PMBus*-compliant power supplies only). This means that if a system has six fans, then there will be six different fan fail reactions.

9.3.14.8 Programmable Fan PWM Offset

The system provides a BIOS Setup option to boost the system fan speed by a programmable positive offset or a “Max” setting. Setting the programmable offset causes the BMC to add the offset to the fan speeds that it would otherwise be driving the fans to. The Max setting causes the BMC to replace the domain minimum speed with alternate domain minimums that also are programmable through SDRs.

This capability is offered to provide system administrators the option to manually configure fans speeds in instances where the fan speed optimized for a given platform may not be sufficient when a high end add-in is configured into the system. This enables easier usage of the fan speed control to support Intel® as well as third party chassis and better support of ambient temperatures higher than 35C.

9.3.15 Power Management Bus (PMBus*)

The Power Management Bus (“PMBus*”) is an open standard protocol that is built upon the SMBus* 2.0 transport. It defines a means of communicating with power conversion and other devices using SMBus*-based commands. A system must have PMBus*-compliant power supplies installed in order for the BMC or ME to monitor them for status and/or power metering purposes.

For more information on PMBus*, please see the System Management Interface Forum Web site http://www.powersig.org/.

9.3.16 Power Supply Dynamic Redundancy Sensor

The BMC supports redundant power subsystems and implements a Power Unit Redundancy sensor per platform. A Power Unit Redundancy sensor is of sensor type Power Unit (09h) and reading type Availability Status (0Bh). This sensor generates events when a power subsystem transitions between redundant and non-redundant states, as determined by the number and health of the power subsystem’s component power supplies. The BMC implements Dynamic Power Supply Redundancy status based upon current system load requirements as well as total Power Supply capacity. This status is independent of the Cold Redundancy status. This prevents the BMC from reporting Fully Redundant Power supplies when the load required by the system exceeds half the power capability of all power supplies installed and operational. Dynamic Redundancy detects this condition and generates the appropriate SEL event to notify the user of the condition. Power supplies of different power ratings may be swapped in and out to adjust the power capacity and the BMC will adjust the Redundancy status accordingly. The definition of redundancy is power subsystem dependent and sometimes even configuration dependent. See the appropriate Platform Specific Information for power unit
redundancy support.

This sensor is configured as manual-rearm sensor in order to avoid the possibility of extraneous SEL events that could occur under certain system configuration and workload conditions. The sensor shall rearm for the following conditions:

- PSU hot-add
- System reset
- AC power cycle
- DC power cycle

System AC power is applied but on standby – Power unit redundancy is based on OEM SDR power unit record and number of PSU present.

System is (DC) powered on - The BMC calculates Dynamic Power Supply Redundancy status based upon current system load requirements as well as total Power Supply capacity.

The BMC allows redundancy to be configured on a per power-unit-redundancy sensor basis by means of the OEM SDR records.

9.3.17 Component Fault LED Control

Several sets of component fault LEDs are supported on the server board (see Figure 7). Some LEDs are owned by the BMC and some by the BIOS.

The BMC owns control of the following FRU/fault LED:

- **Hard Disk Drive Status LEDs** – The HSBP PSoC* owns the hardware control for these LEDs and detection of the fault/status conditions that the LEDs reflect.

Table 49. Component Fault LEDs

Component	Owner	Color	State	Description
HDD Fault LED	HSBP PSoC*	Amber	On	HDD Fault
		Amber	Blink	Predictive failure, rebuild, identify
		Amber	Off	Ok (no errors)

9.3.18 CMOS Battery Monitoring

The BMC monitors the voltage level from the CMOS battery; which provides battery backup to the chipset RTC. This is monitored as an auto-rearm threshold sensor.

Unlike monitoring of other voltage sources for which the Emulex* Pilot III component continuously cycles through each input, the voltage channel used for the battery monitoring provides a software enable bit to allow the BMC FW to poll the battery voltage at a relatively slow rate in order to conserve battery power.

9.4 Intel® Intelligent Power Node Manager (NM)

Power management deals with requirements to manage processor power consumption and manage power at the platform level to meet critical business needs. Node Manager (NM) is a platform resident technology that enforces power capping and thermal-triggered power capping policies for the platform. These policies are applied by exploiting subsystem settings (such as processor P and T states) that can be used to control power consumption. NM enables data center power management by exposing an external interface to management software through which platform policies can be specified. It also implements specific data center power management usage models such as power limiting, and thermal monitoring.

The NM feature is implemented by a complementary architecture utilizing the ME, BMC, BIOS, and an ACPI-compliant OS. The ME provides the NM policy engine and power control/limiting functions (referred to as Node Manager or NM) while the BMC provides the external LAN link by which external management software can interact with the feature. The BIOS provides system power information utilized by the NM algorithms and also exports ACPI Source Language (ASL) code used by OS-Directed Power Management (OSPM) for

negotiating processor P and T state changes for power limiting. PMBus*-compliant power supplies provide the capability to monitor input power consumption, which is necessary to support NM.

The NM architecture applicable to this generation of servers is defined by the *NPTM Architecture Specification v2.0*. NPTM is an evolving technology that is expected to continue to add new capabilities that will be defined in subsequent versions of the specification. The ME NM implements the NPTM policy engine and control/monitoring algorithms defined in the Node Power and Thermal Manager (NPTM) specification.

9.4.1 Hardware Requirements

NM is supported only on platforms that have the NM FW functionality loaded and enabled on the Management Engine (ME) in the SSB and that have a BMC present to support the external LAN interface to the ME. NM power limiting features requires a means for the ME to monitor input power consumption for the platform. This capability is generally provided by means of PMBus*-compliant power supplies although an alternative model using a simpler SMBus* power monitoring device is possible (there is potential loss in accuracy and responsiveness using non-PMBus* devices). The NM SmaRT/CLST feature does specifically require PMBus*-compliant power supplies as well as additional hardware on the server board.

9.4.2 Features

NM provides feature support for policy management, monitoring and querying, alerts and notifications, and an external interface protocol. The policy management features implement specific IT goals that can be specified as policy directives for NM. Monitoring and querying features enable tracking of power consumption. Alerts and notifications provide the foundation for automation of power management in the data center management stack. The external interface specifies the protocols that must be supported in this version of NM.

9.4.3 ME System Management Bus (SMBus*) Interface

- The ME uses the SMLink0 on the SSB in multi-master mode as a dedicated bus for communication with the BMC using the IPMB protocol. The BMC FW considers this a secondary IPMB bus and runs at 400 kHz.
- The ME uses the SMLink1 on the SSB in multi-master mode bus for communication with PMBus* devices in the power supplies for support of various NM-related features. This bus is shared with the BMC, which polls these PMBus* power supplies for sensor monitoring purposes (for example, power supply status, input power, and so on). This bus runs at 100 KHz.
- The Management Engine has access to the “Host SMBus*”.

9.4.4 PECI 3.0

- The BMC owns the PECI bus for all Intel server implementations and acts as a proxy for the ME when necessary.

9.4.5 NM “Discovery” OEM SDR

An NM “discovery” OEM SDR must be loaded into the BMC’s SDR repository if and only if the NM feature is supported on that product. This OEM SDR is used by management software to detect if NM is supported and to understand how to communicate with it.

Since PMBus* compliant power supplies are required in order to support NM, the system should be probed when the SDRs are loaded into the BMC's SDR repository in order to determine whether or not the installed power supplies do in fact support PMBus*. *If the installed power supplies are not PMBus* compliant then the NM "discovery" OEM SDR should not be loaded.*

Please refer to the Intel® Intelligent Power Node Manager 2.0 External Architecture Specification using IPMI for details of this interface.

9.4.6 SmaRT/CLST

The power supply optimization provided by SmaRT/CLST relies on a platform HW capability as well as ME FW support. When a PMBus*-compliant power supply detects insufficient input voltage, an overcurrent condition, or an over-temperature condition, it will assert the SMBAlert# signal on the power supply SMBus* (such as, the PMBus*). Through the use of external gates, this results in a momentary assertion of the PROCHOT# and MEMHOT# signals to the processors, thereby throttling the processors and memory. The ME FW also sees the SMBAlert# assertion, queries the power supplies to determine the condition causing the assertion, and applies an algorithm to either release or prolong the throttling, based on the situation.

System power control modes include:

1. SmaRT: Low AC input voltage event; results in a one-time momentary throttle for each event to the maximum throttle state.
2. Electrical Protection CLST: High output energy event; results in a throttling hiccup mode with fixed maximum throttle time and a fix throttle release ramp time.
3. Thermal Protection CLST: High power supply thermal event; results in a throttling hiccup mode with fixed maximum throttle time and a fix throttle release ramp time.

When the SMBAlert# signal is asserted, the fans will be gated by HW for a short period (~100ms) to reduce overall power consumption. It is expected that the interruption to the fans will be of short enough duration to avoid false lower threshold crossings for the fan tach sensors; however, this may need to be comprehended by the fan monitoring FW if it does have this side-effect.

ME FW will log an event into the SEL to indicate when the system has been throttled by the SmaRT/CLST power management feature. This is dependent on ME FW support for this sensor. Please refer ME FW EPS for SEL log details.

9.4.6.1.1 Dependencies on PMBus*-compliant Power Supply Support

The SmaRT/CLST system feature depends on functionality present in the ME NM SKU. This feature requires power supplies that are compliant with the *PMBus.*

Note: For additional information on Intel® Intelligent Power Node Manager usage and support, please visit the following Intel Website:
http://www.intel.com/content/www/us/en/data-center/data-center-management/node-manager-general.html?wapkw=node+manager

---

9.5 Basic and Advanced Server Management Features

The integrated BMC has support for basic and advanced server management features. Basic management features are available by default. Advanced management features are enabled with the addition of an optionally installed Remote Management Module 4 Lite (RMM4 Lite) key.

Table 50. Intel® Remote Management Module 4 (RMM4) Options

Intel Product Code	Description	Kit Contents	Benefits
AXXRMM4LITE	Intel® Remote Management Module 4 Lite	RMM4 Lite Activation Key	Enables KVM & media redirection

When the BMC FW initializes, it attempts to access the Intel® RMM4 lite. If the attempt to access Intel® RMM4 lite is successful, then the BMC activates the a*dvanced* features.

The following table identifies both basic and advanced server management features.

Table 51. Basic and Advanced Server Management Features Overview

Feature	Basic	Advanced
IPMI 2.0 Feature Support	X	X
In-circuit BMC Firmware Update	X	X
FRB 2	X	X
Chassis Intrusion Detection	X	X
Fan Redundancy Monitoring	X	X
Hot-Swap Fan Support	X	X
Acoustic Management	X	X
Diagnostic Beep Code Support	X	X
Power State Retention	X	X
ARP/DHCP Support	X	X
PECI Thermal Management Support	X	X
E-mail Alerting	X	X
Embedded Web Server	X	X
SSH Support	X	X
Integrated KVM		X
Integrated Remote Media Redirection		X
Lightweight Directory Access Protocol (LDAP)	X	X
Intel® Intelligent Power Node Manager Support	X	X
SMASH CLP	X	X

9.5.1 Dedicated Management Port

The server board includes a dedicated 1GbE RJ45 Management Port. The management port is active with or without the RMM4 Lite key installed.

9.5.2 Embedded Web Server

BMC Base manageability provides an embedded web server and an OEM-customizable web GUI which exposes the manageability features of the BMC base feature set. It is supported over all on-board NICs that have management connectivity to the BMC as well as an optional dedicated add-in management NIC. At least

two concurrent web sessions from up to two different users is supported. The embedded web user interface shall support the following client web browsers:

- Microsoft Internet Explorer 9.0*
- Microsoft Internet Explorer 10.0*
- Mozilla Firefox 24*
- Mozilla Firefox 25*

The embedded web user interface supports strong security (authentication, encryption, and firewall support) since it enables remote server configuration and control. The user interface presented by the embedded web user interface, shall authenticate the user before allowing a web session to be initiated. Encryption using 128-bit SSL is supported. User authentication is based on user id and password.

The GUI presented by the embedded web server authenticates the user before allowing a web session to be initiated. It presents all functions to all users but grays-out those functions that the user does not have privilege to execute. For example, if a user does not have privilege to power control, then the item shall be displayed in grey-out font in that user's UI display. The web GUI also provides a launch point for some of the advanced features, such as KVM and media redirection. These features are grayed out in the GUI unless the system has been updated to support these advanced features. The embedded web server only displays US English or Chinese language output.

Additional features supported by the web GUI includes:

- Presents all the Basic features to the users
- Power on/off/reset the server and view current power state
- Display BIOS, BMC, ME and SDR version information
- Display overall system health.
- Configuration of various IPMI over LAN parameters for both IPV4 and IPV6
- Configuration of alerting (SNMP and SMTP)
- Display system asset information for the product, board, and chassis.
- Display of BMC-owned sensors (name, status, current reading, enabled thresholds), including color-code status of sensors.
- Provide ability to filter sensors based on sensor type (Voltage, Temperature, Fan and Power supply related)
- Automatic refresh of sensor data with a configurable refresh rate
- On-line help
- Display/clear SEL (display is in easily understandable human readable format)
- Supports major industry-standard browsers (Microsoft Internet Explorer* and Mozilla Firefox*)
- The GUI session automatically times-out after a user-configurable inactivity period. By default, this inactivity period is 30 minutes.
- Embedded Platform Debug feature - Allow the user to initiate a "debug dump" to a file that can be sent to Intel® for debug purposes.
- Virtual Front Panel. The Virtual Front Panel provides the same functionality as the local front panel. The displayed LEDs match the current state of the local panel LEDs. The displayed buttons (for example, power button) can be used in the same manner as the local buttons.

- Display of ME sensor data. Only sensors that have associated SDRs loaded will be displayed.
- Ability to save the SEL to a file
- Ability to force HTTPS connectivity for greater security. This is provided through a configuration option in the UI.
- Display of processor and memory information as is available over IPMI over LAN.
- Ability to get and set Node Manager (NM) power policies
- Display of power consumed by the server
- Ability to view and configure VLAN settings
- Warn user the reconfiguration of IP address will cause disconnect.
- Capability to block logins for a period of time after several consecutive failed login attempts. The lock-out period and the number of failed logins that initiates the lock-out period are configurable by the user.
- Server Power Control - Ability to force into Setup on a reset
- System POST results – The web server provides the system's Power-On Self-Test (POST) sequence for the previous two boot cycles, including timestamps. The timestamps may be viewed in relative to the start of POST or the previous POST code.
- Customizable ports - The web server provides the ability to customize the port numbers used for SMASH, http, https, KVM, secure KVM, remote media, and secure remote media.

For additional information, reference the Intel® Remote Management Module 4 and Integrated BMC Web Console Users Guide.

9.5.3 Advanced Management Feature Support (RMM4 Lite)

The integrated baseboard management controller has support for advanced management features which are enabled when an optional Intel® Remote Management Module 4 Lite (RMM4 Lite) is installed. The Intel RMM4 add-on offers convenient, remote KVM access and control through LAN and internet. It captures, digitizes, and compresses video and transmits it with keyboard and mouse signals to and from a remote computer. Remote access and control software runs in the integrated baseboard management controller, utilizing expanded capabilities enabled by the Intel RMM4 hardware.

Key Features of the RMM4 add-on are:

- KVM redirection from either the dedicated management NIC or the server board NICs used for management traffic; up to two KVM sessions
- Media Redirection – The media redirection feature is intended to allow system administrators or users to mount a remote IDE or USB CDROM, floppy drive, or a USB flash disk as a remote device to the server. Once mounted, the remote device appears just like a local device to the server allowing system administrators or users to install software (including operating systems), copy files, update BIOS, or boot the server from this device.
- KVM – Automatically senses video resolution for best possible screen capture, high performance mouse tracking and synchronization. It allows remote viewing and configuration in pre-boot POST and BIOS setup.

9.5.3.1 Keyboard, Video, Mouse (KVM) Redirection

The BMC firmware supports keyboard, video, and mouse redirection (KVM) over LAN. This feature is available remotely from the embedded web server as a Java applet. This feature is only enabled when the Intel® RMM4 lite is present. The client system must have a Java Runtime Environment (JRE) version 6.0 or later to run the KVM or media redirection applets.

The BMC supports an embedded KVM application (*Remote Console*) that can be launched from the embedded web server from a remote console. USB1.1 or USB 2.0 based mouse and keyboard redirection are supported. It is also possible to use the KVM-redirection (KVM-r) session concurrently with media-redirection (media-r). This feature allows a user to interactively use the keyboard, video, and mouse (KVM) functions of the remote server as if the user were physically at the managed server. KVM redirection console supports the following keyboard layouts: English, Dutch, French, German, Italian, Russian, and Spanish.

KVM redirection includes a “soft keyboard” function. The “soft keyboard” is used to simulate an entire keyboard that is connected to the remote system. The “soft keyboard” functionality supports the following layouts: English, Dutch, French, German, Italian, Russian, and Spanish.

The KVM-redirection feature automatically senses video resolution for best possible screen capture and provides high-performance mouse tracking and synchronization. It allows remote viewing and configuration in pre-boot POST and BIOS setup, once BIOS has initialized video.

Other attributes of this feature include:

- Encryption of the redirected screen, keyboard, and mouse
- Compression of the redirected screen
- Ability to select a mouse configuration based on the OS type
- Supports user definable keyboard macros

KVM redirection feature supports the following resolutions and refresh rates:

- 640x480 at 60Hz, 72Hz, 75Hz, 85Hz, 100Hz
- 800x600 at 60Hz, 72Hz, 75Hz, 85Hz
- 1024x768 at 60Hx, 72Hz, 75Hz, 85Hz
- 1280x960 at 60Hz
- 1280x1024 at 60Hz
- 1600x1200 at 60Hz
- 1920x1080 (1080p)
- 1920x1200 (WUXGA)
- 1650x1080 (WSXGA+)

9.5.3.2 Remote Console

The Remote Console is the redirected screen, keyboard and mouse of the remote host system. To use the Remote Console window of your managed host system, the browser must include a Java* Runtime Environment plug-in. If the browser has no Java support, such as with a small handheld device, the user can maintain the remote host system using the administration forms displayed by the browser.

The Remote Console window is a Java Applet that establishes TCP connections to the BMC. The protocol that is run over these connections is a unique KVM protocol and not HTTP or HTTPS. This protocol uses ports #7578 for KVM, #5120 for CDROM media redirection, and #5123 for Floppy/USB media redirection. When encryption is enabled, the protocol uses ports #7582 for KVM, #5124 for CDROM media redirection, and #5127 for Floppy/USB media redirection. The local network environment must permit these connections to be made, that is, the firewall and, in case of a private internal network, the NAT (Network Address Translation) settings have to be configured accordingly.

9.5.3.3 Performance

The remote display accurately represents the local display. The feature adapts to changes to the video resolution of the local display and continues to work smoothly when the system transitions from graphics to text or vice-versa. The responsiveness may be slightly delayed depending on the bandwidth and latency of the network.

Enabling KVM and/or media encryption will degrade performance. Enabling video compression provides the fastest response while disabling compression provides better video quality.

For the best possible KVM performance, a 2Mb/sec link or higher is recommended.

The redirection of KVM over IP is performed in parallel with the local KVM without affecting the local KVM operation.

9.5.3.4 Security

The KVM redirection feature supports multiple encryption algorithms, including RC4 and AES. The actual algorithm that is used is negotiated with the client based on the client's capabilities.

9.5.3.5 Availability

The remote KVM session is available even when the server is powered-off (in stand-by mode). No re-start of the remote KVM session shall be required during a server reset or power on/off. A BMC reset (for example, due to a BMC Watchdog initiated reset or BMC reset after BMC FW update) will require the session to be re-established.

KVM sessions persist across system reset, but not across an AC power loss.

9.5.3.6 Usage

As the server is powered up, the remote KVM session displays the complete BIOS boot process. The user is able to interact with BIOS setup, change and save settings as well as enter and interact with option ROM configuration screens.

At least two concurrent remote KVM sessions are supported. It is possible for at least two different users to connect to same server and start remote KVM sessions.

9.5.3.7 Force-enter BIOS Setup

KVM redirection can present an option to force-enter BIOS Setup. This enables the system to enter F2 setup while booting which is often missed by the time the remote console redirects the video.

9.5.3.8 Media Redirection

The embedded web server provides a Java applet to enable remote media redirection. This may be used in conjunction with the remote KVM feature, or as a standalone applet.

The media redirection feature is intended to allow system administrators or users to mount a remote IDE or USB CD-ROM, floppy drive, or a USB flash disk as a remote device to the server. Once mounted, the remote device appears just like a local device to the server, allowing system administrators or users to install software (including operating systems), copy files, update BIOS, and so on, or boot the server from this device.

The following capabilities are supported:

- The operation of remotely mounted devices is independent of the local devices on the server. Both remote and local devices are useable in parallel.
- Either IDE (CD-ROM, floppy) or USB devices can be mounted as a remote device to the server.
- It is possible to boot all supported operating systems from the remotely mounted device and to boot from disk IMAGE (*.IMG) and CD-ROM or DVD-ROM ISO files. See the Tested/supported Operating System List for more information.
- Media redirection supports redirection for both a virtual CD device and a virtual Floppy/USB device concurrently. The CD device may be either a local CD drive or else an ISO image file; the Floppy/USB device may be either a local Floppy drive, a local USB device, or else a disk image file.
- The media redirection feature supports multiple encryption algorithms, including RC4 and AES. The actual algorithm that is used is negotiated with the client based on the client's capabilities.
- A remote media session is maintained even when the server is powered-off (in standby mode). No restart of the remote media session is required during a server reset or power on/off. A BMC reset (for example, due to a BMC reset after BMC FW update) will require the session to be re-established.
- The mounted device is visible to (and useable by) managed system's OS and BIOS in both pre-boot and post-boot states.
- The mounted device shows up in the BIOS boot order and it is possible to change the BIOS boot order to boot from this remote device.
- It is possible to install an operating system on a bare metal server (no OS present) using the remotely mounted device. This may also require the use of KVM-r to configure the OS during install.

USB storage devices will appear as floppy disks over media redirection. This allows for the installation of device drivers during OS installation.

If either a virtual IDE or virtual floppy device is remotely attached during system boot, both the virtual IDE and virtual floppy are presented as bootable devices. It is not possible to present only a single-mounted device type to the system BIOS.

9.5.3.8.1 Availability

The default inactivity timeout is 30 minutes and is not user-configurable. Media redirection sessions persist across system reset but not across an AC power loss or BMC reset.

9.5.3.8.2 Network Port Usage

The KVM and media redirection features use the following ports:

- 5120 – CD Redirection
- 5123 – FD Redirection
- 5124 – CD Redirection (Secure)
- 5127 – FD Redirection (Secure)
- 7578 – Video Redirection
- 7582 – Video Redirection (Secure)

For additional information, reference the *Intel® Remote Management Module 4 and Integrated BMC Web Console Users Guide*.

10 Thermal Management

The compute module is designed to operate at external ambient temperatures of between 10ºC and 35ºC with limited excursion based operation up to 45ºC. Working with integrated platform management, several features within the compute module are designed to move air in a front-to-back direction, through the compute module and over critical components to prevent them from overheating and allow the system to operate with best performance.

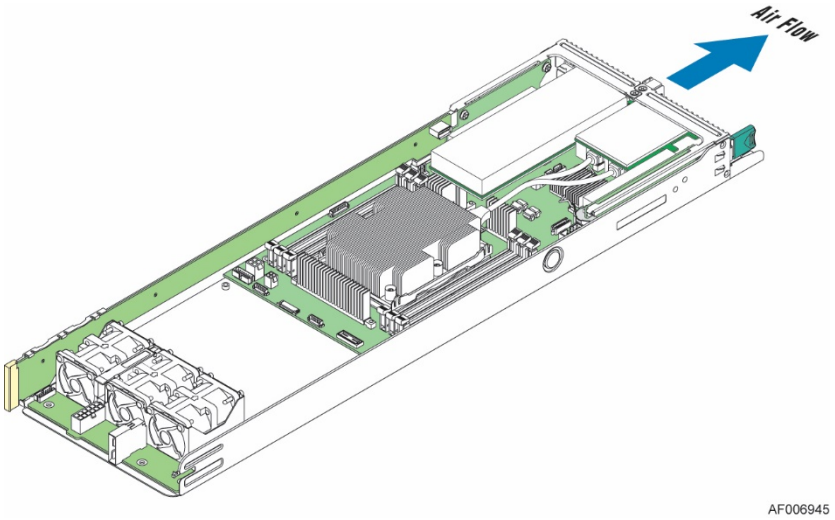


Figure 42. Air Flow and Fan Identification

The following table provides air flow data associated with the different product models within this product family, and is provided for reference purposes only. The data was derived from actual wind tunnel test methods and measurements using fully configured system configurations. Lesser system configurations may produce slightly different data results. As such, the CFM data provided using server management utilities that utilize platform sensor data, may vary from the data listed in the table.

Table 52. Air Flow

	Single Compute Module Airflow
With Intel® Server Chassis H2312XXKR2	4 ~43CFM
With Intel® Server Chassis H2216XXKR2	5 ~ 62CFM

The compute module supports short-term, excursion-based, operation up to 45°C (ASHRAE A4) with limited performance impact. The configuration requirements and limitations are described in the configuration matrix found in the *Intel® Server Board S7200AP Product Family Power Budget and Thermal Configuration Tool,* to be available as a download online at http://www.intel.com/support.

The installation and functionality of several components are used to maintain compute module thermals. They include three compute module fans, air duct, and installed CPU heatsinks.

To keep the compute module operating within supported maximum thermal limits, the compute module must meet the following operating and configuration guidelines:

- The compute module operating ambient is designed for sustained operation up to 35ºC (ASHRAE Class A2) with short-term excursion-based operation up to 45ºC (ASHRAE Class A4).
  - The compute module can operate up to 40ºC (ASHRAE Class A3) for up to 900 hours per year.

- The compute module can operate up to 45ºC (ASHRAE Class A4) for up to 90 hours per year.
- The compute module performance may be impacted when operating within the extended operating temperature range.
- There is no long-term system reliability impact when operating at the extended temperature range within the approved limits.

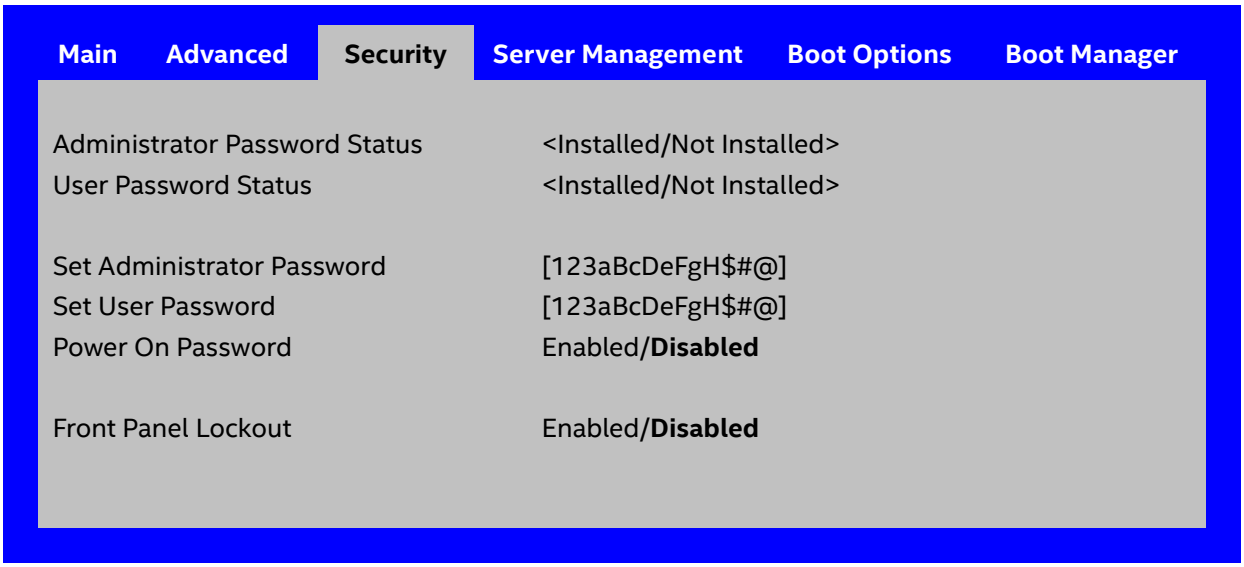
- Specific configuration requirements and limitations are documented in the configuration matrix found in the *Intel® Server Board S7200AP product family Power Budget and Thermal Configuration Tool*, available as a download online at http://www.intel.com/supoprt.
- CPU heatsink must be installed first.

11 System Security

The server board supports a variety of system security options designed to prevent unauthorized system access or tampering of server settings. System security options supported include:

- Password Protection
- Front Panel Lockout

The <F2> BIOS Setup Utility, accessed during POST, includes a Security tab where options to configure passwords, and front panel lockout can be found.



11.1 Password Setup

The BIOS uses passwords to prevent unauthorized access to the server. Passwords can restrict entry to the BIOS Setup utility, restrict use of the Boot Device popup menu during POST, suppress automatic USB device re-ordering, and prevent unauthorized system power on. It is strongly recommended that an Administrator Password be set. A system with no Administrator password set allows anyone who has access to the server to change BIOS settings.

An Administrator password must be set in order to set the User password.

The maximum length of a password is 14 characters and can be made up of a combination of alphanumeric (a-z, A-Z, 0-9) characters and any of the following special characters:

! @ # $ % ^ & * ( ) - _ + = ?

Passwords are case sensitive.

The Administrator and User passwords must be different from each other. An error message will be displayed and a different password must be entered if there is an attempt to enter the same password for both. The use of “Strong Passwords” is encouraged, but not required. In order to meet the criteria for a strong password, the password entered must be at least 8 characters in length, and must include at least one

each of alphabetic, numeric, and special characters. If a weak password is entered, a warning message will be displayed, and the weak password will be accepted.

Once set, a password can be cleared by changing it to a null string. This requires the Administrator password, and must be done through BIOS Setup or other explicit means of changing the passwords. Clearing the Administrator password will also clear the User password. Passwords can also be cleared by using the Password Clear jumper on the server board. See Chapter 10 – Reset and Recovery Jumpers.

Resetting the BIOS configuration settings to default values (by any method) has no effect on the Administrator and User passwords.

As a security measure, if a User or Administrator enters an incorrect password three times in a row during the boot sequence, the system is placed into a halt state. A system reset is required to exit out of the halt state. This feature makes it more difficult to guess or break a password.

In addition, on the next successful reboot, the Error Manager displays a Major Error code 0048, which also logs a SEL event to alert the authorized user or administrator that a password access failure has occurred.

11.1.1 System Administrator Password Rights

When the correct Administrator password is entered when prompted, the user has the ability to perform the following:

- Access the <F2> BIOS Setup Utility.
- Has the ability to configure all BIOS setup options in the <F2> BIOS Setup Utility.
- Has the ability to clear both the Administrator and User passwords.
- Access the <F6> Boot Menu during POST.
- If the Power On Password function is enabled in BIOS Setup, the BIOS will halt early in POST to request a password (Administrator or User) before continuing POST.

11.1.2 Authorized System User Password Rights and Restrictions

When the correct User password is entered, the user has the ability to perform the following:

- Access the <F2> BIOS Setup Utility.
- View, but not change any BIOS Setup options in the <F2> BIOS Setup Utility.
- Modify System Time and Date in the BIOS Setup Utility.
- If the Power On Password function is enabled in BIOS Setup, the BIOS will halt early in POST to request a password (Administrator or User) before continuing POST.

In addition to restricting access to most Setup fields to viewing only when a User password is entered, defining a User password imposes restrictions on booting the system. In order to simply boot in the defined boot order, no password is required. However, the F6 Boot popup menu prompts for a password, and can only be used with the Administrator password. Also, when a User password is defined, it suppresses the USB Reordering that occurs, if enabled, when a new USB boot device is attached to the system. A User is restricted from booting in anything other than the Boot Order defined in the Setup by an Administrator.

11.2 Front Panel Lockout

If enabled in BIOS setup, this option disables the following front panel features:

- The OFF function of the Power button
- System Reset button

If [Enabled] is selected, system power off and reset must be controlled via a system management interface.

12 Environmental Limits Specification

Operation of the server board at conditions beyond those shown in the following table may cause permanent damage to the system. Exposure to absolute maximum rating conditions for extended periods may affect long term system reliability.

Note: The Energy Star compliance is at systems level, but not board level. Use of Intel® boards alone does not guarantee Energy Star compliance.

Table 53. Server Board Design Specifications

Parameter	Limits
Operating Temperature	+10°C to +35°C with the maximum rate of change not to exceed 10°C per hour.
Non-Operating Temperature	-40°C to +70°C
Non-Operating Humidity	90%, non-condensing at 35°C
Acoustic noise	Sound power: 7.0BA with hard disk drive stress only at room ambient temperature (23 +/-2°C)
Shock, operating	Half sine, 2g peak, 11 mSec
Shock, unpackaged	Trapezoidal, 25g, velocity change 205 inches/second (80 lbs. to < 100 lbs.)
Vibration, unpackaged	5 Hz to 500 Hz, 2.20 g RMS random
Shock and vibration, packaged	ISTA (International Safe Transit Association) Test Procedure 3A
ESD	+/-12 KV except I/O port +/- 8 KV per Intel® Environmental Test Specification
System Cooling Requirement in BTU/Hr.	2130 Watt Max – 7263 BTU/hour

Disclaimer Note: *Intel ensures the unpackaged server board and system meet the shock requirement mentioned above through its own chassis development and system configuration. It is the responsibility of the system integrator to determine the proper shock level of the board and system if the system integrator chooses different system configuration or different chassis. Intel Corporation cannot be held responsible if components fail or the server board does not operate correctly when used outside any of its published operating or non-operating limits.*

13 Power Supply Specification Guidelines

This section provides power supply specification guidelines recommended for providing the specified server platform with stable operating power requirements.

Note: The power supply data provided in this section is for reference purposes only. It reflects Intel's own DC power out requirements for a 2130W power supply as used in an Intel designed 2U server platform. The intent of this section is to provide customers with a guide to assist in defining and/or selecting a power supply for custom server platform designs that utilize the server boards detailed in this document.

13.1 Mechanical Overview

The physical size of the power supply enclosure is 39/40mm x 73.5mm x 265mm. The power supply contains a single 40mm fan. The power supply has a card edge output that interfaces with a 2x25 card edge connector in the system. The AC plugs directly into the external face of the power supply. Refer to the following **Error! Reference source not found.**. All dimension are nominal

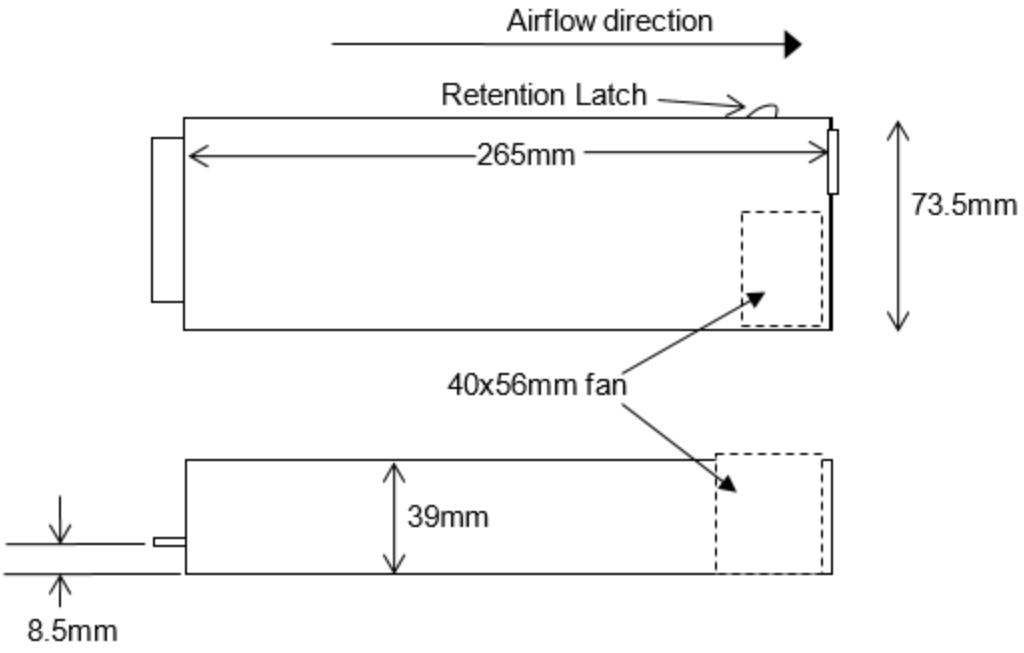


Figure 43. Mechanical Dimensions

13.2 LED Indicator States

The table below shows the LED indicator states:

Table 54. LED Indicator States

Power Supply Condition	LED State
Output ON and OK	Solid GREEN
No AC power to all power supplies	OFF
AC present / Only 12VSB on (PS off) or PS in Cold redundant state	1Hz Blink GREEN
AC cord unplugged; with a second power supply in parallel still with AC input power.	Solid AMBER
Power supply warning events where the power supply continues to operate; high temp, high power, high current, slow fan.	1Hz Blink AMBER
Power supply critical event causing a shutdown; failure, OCP, OVP, Fan Fail	Solid AMBER
Power supply in FW upload mode	2Hz Blink GREEN

13.3 Server Board DC Output Connector

The server board includes two main power Minifit Jr* connectors allowing for power supplies to attach directly to the server board. The connectors are two sets of 2x3 pin and can be used to deliver 12amps per pin or 60+Amps total. Note that no over-voltage protective circuits will exist on the board.

Table 55. Power Supply DC Power Input Connector Pin-out

Pin	Signal Name	Pin	Signal Name
1	+12V	4	GND
2	+12V	5	GND
3	+12V	6	GND

13.4 Power Supply DC Output Connector

Table 56. Power Supply DC Output Connector

Pin	Name	Pin	Name
A1	GND	B1	GND
A2	GND	B2	GND
A3	GND	B3	GND
A4	GND	B4	GND
A5	GND	B5	GND
A6	GND	B6	GND
A7	GND	B7	GND
A8	GND	B8	GND
A9	GND	B9	GND
A10	+12V	B10	+12V

Pin	Name	Pin	Name
A11	+12V	B11	+12V
A12	+12V	B12	+12V
A13	+12V	B13	+12V
A14	+12V	B14	+12V
A15	+12V	B15	+12V
A16	+12V	B16	+12V
A17	+12V	B17	+12V
A18	+12V	B18	+12V
A19	PMBus SDA [1]	B19	A0 (SMBus address) [1]
A20	PMBus SCL [1]	B20	A1 (SMBus address) [1]
A21	PSON	B21	12V stby
A22	SMBAlert#	B22	Cold Redundancy Bus [1]
A23	Return Sense	B23	12V load share bus
A24	+12V remote Sense	B24	No Connect
A25	PWOK	B25	Compatibility Bus [1]

13.5 AC Input Requirement

13.5.1 Power Factor

The power supply must meet the power factor requirements stated in the Energy Star® Program Requirements for Computer Servers. These requirements are stated below.

Output power	10% load	20% load	50% load	100% load
Power factor	> 0.80	> 0.90	> 0.90	> 0.95

Tested at 230Vac, 50Hz and 60Hz and 115VAC, 60Hz
Tested according to Generalized Internal Power Supply Efficiency Testing Protocol Rev 6.6. This is posted at: http://www.plugloadsolutions.com/docs/collatrl/print/Generalized_Internal_Power_Supply_Efficiency_Test_Protocol_R6.6.pdf

13.5.2 AC Inlet Connector

The AC input connector shall be an *IEC 320 C14* power inlet.

13.5.3 AC Input Voltage Specification

The power supply must operate within all specified limits over the following input voltage range. Harmonic distortion of up to 10% of the rated line voltage must not cause the power supply to go out of specified limits. Application of an input voltage below 85VAC or between 127VAC and 180VAC shall not cause damage to the

power supply, including a blown fuse.

Table 57. Input Voltage Range

PARAMETER	MIN	VAC RATED	IAC RATED	V_{MAX}	Start up VAC	Power Off VAC
Voltage (low line)	90 V_{rms}	100 V_{rms} 110/115 V_{rms} 120/127 V_{rms}	12 A_{rms}	140 V_{rms}	85VAC +/- 4VAC	75VAC +/- 5VAC
Voltage (high line)	180 V_{rms}	200/208 V_{rms} 220 V_{rms} 230 V_{rms} 240 V_{rms}	10 A_{rms}	264 V_{rms}		
Frequency	47 Hz	50/60		63 Hz		

Notes:

1 Maximum input current at low line voltage range shall be measured at 90VAC, 110VAC, and 120VAC at max loads.
2 Maximum input current at high input voltage range shall be measured at 180VAC, 220VAC, 230VAC, and 240VAC at max load.
3 This requirement is not to be used for determining agency input current markings.

13.5.4 AC Line Isolation Requirements

The power supply shall meet all safety agency requirements for dielectric strength. Additionally, power supply vendor must provide Intel with written confirmation of dielectric withstand test which includes: voltage level, duration of test and identification detailing how each power supply is marked to indicate dielectric withstand test had been completed successfully. Transformers' isolation between primary and secondary windings must comply with the 3000Vac (4242Vdc) dielectric strength criteria. If the working voltage between primary and secondary dictates a higher dielectric strength test voltage the highest test voltage should be used. In addition the insulation system must comply with reinforced insulation per safety standard IEC 950. Separation between the primary and secondary circuits, and primary to ground circuits, must comply with the IEC 950 spacing requirement.

13.5.5 AC Line Dropout/Holdup

An AC line dropout is defined to be when the AC input drops to 0VAC at any phase of the AC line for any length of time. During an AC dropout the power supply must meet dynamic voltage regulation requirements. An AC line dropout of any duration shall not cause tripping of control signals or protection circuits. If the AC dropout lasts longer than the holdup time the power supply should recover and meet all turn on requirements. The power supply shall meet the AC dropout requirement over rated AC voltages and frequencies. A dropout of the AC line for any duration shall not cause damage to the power supply.

Table 58. AC Line Dropout/Holdup

Loading	Holdup time
75%	10msec

13.5.6 AC 12VSB Holdup

The 12VSB output voltage should stay in regulation under its full load (static or dynamic) during an AC dropout of **70ms min** (=12VSB holdup time) whether the power supply is in ON or OFF state (PSON asserted or de-asserted).

13.5.7 AC Line Fuse

The power supply shall have one line fused in the **single line fuse** on the line (Hot) wire of the AC input. The line fusing shall be acceptable for all safety agency requirements. The input fuse shall be a slow blow type. AC inrush current shall not cause the AC line fuse to blow under any conditions. All protection circuits in the power supply shall not cause the AC fuse to blow unless a component in the power supply has failed. This includes DC output load short conditions.

13.5.8 AC Inrush

AC line inrush current shall not exceed **65A peak**, for up to one-quarter of the AC cycle, after which, the input current should be no more than the specified maximum input current. The peak inrush current shall be less than the ratings of its critical components (including input fuse, bulk rectifiers, and surge limiting device).
The power supply must meet the inrush requirements for any rated AC voltage, during turn on at any phase of AC voltage, during a single cycle AC dropout condition as well as upon recovery after AC dropout of any duration, and over the specified temperature range ($T_{op1 \text{ and } Top2}$).

13.6 Power Supply DC Output Specification

13.6.1 Output Power/Currents

The following tables define the minimum power and current ratings. The power supply must meet both static and dynamic voltage regulation requirements for all conditions. The power supplies may be used in a 1+1 redundant configuration.

Table 59. Output Load Ratings and Peak Loading for a single power supply

Parameter	VAC Rating	Power Rating (W)	Min Current (A)	Current Rating (A)	20sec Peak Current (A) [2, 3]	25msec Peak Current (A)
12V main (90-110VAC)	100	1120	0.0	93	140	180
12V main (110-120VAC)	110/115	1250	0.0	104	140	180
12V main (120-140VAC)	120/127	1370	0.0	114	140	180
12V main (180-208VAC)	200	1780	0.0	148	210	250
12V main (208-220VAC)	208	1850	0.0	154	210	250
12V main (220-240VAC)	220/230	1960	0.0	163	210	250
12V main (240-264VAC)	240	2130	0.0	178	210	250
12Vstby [1]			0.0	3.5	4.0	

Notes:

1. 12Vstby must be able to provide 4.0A peak load with single power supply. The power supply fan is allowed to run in standby mode for loads > 1.5A.
2. Length of time peak power can be supported is based on thermal sensor and assertion of the SMBAlert# signal. Minimum peak power duration shall be 20 seconds without asserting the SMBAlert# signal. The peak load requirement should apply to

full operating temperature range.

3. The setting of $I_{Peak} < I_{OCW} < I_{OCP}$ needs to be followed to make the CLST work reasonably.
4. Power supply must protect itself in case system doesn't take any action to reduce load based on SMBAlert# signal asserting.
5. The power supply shall support 25msec peak power at 20% duty cycle step loading for an average current at the current rating.
6. With two power supplies in parallel; the power supplies must support 2130W at any AC voltage range.

13.6.2 Standby Output

The 12VSB output shall be present when an AC input greater than the power supply turn on voltage is applied.

13.6.3 Voltage Regulation

The power supply output voltages must stay within the following voltage limits when operating at steady state and dynamic loading conditions. These limits include the peak-peak ripple/noise. These shall be measured at the output connectors.

Table 60. Voltage Regulation Limits

Parameter	Tolerance	Min	Nom	Max	Units
+12V	- 5%/+5%	+11.40	+12.00	+12.60	V_{rms}
+12V stby	- 5% / +5%	+11.40	+12.00	+12.60	V_{rms}

13.6.4 Dynamic Loading

The output voltages shall remain within limits specified for the step loading and capacitive loading specified in the table below. The load transient repetition rate shall be tested between 50Hz and 5kHz at duty cycles ranging from 10%-90%. The load transient repetition rate is only a test specification. The Δ step load may occur anywhere within the MIN load to the MAX load conditions.

Table 61. Transient Load Requirements

Output	Δ Step Load Size	Load Slew Rate	Test capacitive Load
+12VSB	1.0A	0.25 A/μsec	20 μF
+12V	60% of max load	0.25 A/μsec	2000 μF

Note: For dynamic condition +12V min loading is 1A.

13.6.5 Capacitive Loading

The power supply shall be stable and meet all requirements except transient loading with the following capacitive loading ranges.

Table 62. Capacitive Loading Conditions

Output	MIN	MIN (Cold Redundancy)	MAX	Units
+12VSB	0	20	3100	μF
+12V	0	1,000	70,000	μF

13.6.6 Grounding

The output ground of the pins of the power supply provides the output power return path. The output connector ground pins shall be connected to the safety ground (power supply enclosure). This grounding

should be well designed to ensure passing the max allowed Common Mode Noise levels.

The power supply shall be provided with a reliable protective earth ground. All secondary circuits shall be connected to protective earth ground. Resistance of the ground returns to chassis shall not exceed 1.0 mΩ. This path may be used to carry DC current.

13.6.7 Closed-loop Stability

The power supply shall be unconditionally stable under all line/load/transient load conditions including specified capacitive load ranges. A minimum of **45 degrees phase margin** and **10dB-gain margin** is required. Closed-loop stability must be ensured at the maximum and minimum loads as applicable.

13.6.8 Residual Voltage Immunity in Standby Mode

The power supply should be immune to any residual voltage placed on its outputs (Typically a leakage voltage through the system from standby output) up to **500mV.** There shall be no additional heat generated, nor stressing of any internal components with this voltage applied to any individual or all outputs simultaneously. It also should not trip the protection circuits during turn on.

The residual voltage at the power supply outputs for no load condition shall not exceed **100mV** when AC voltage is applied and the PSON# signal is de-asserted.

13.6.9 Common Mode Noise

The Common Mode noise on any output shall not exceed **350mV pk-pk** over the frequency band of 10Hz to 20MHz. The measurement shall be made across a 100Ω resistor between each of DC outputs, including ground at the DC power connector and chassis ground (power subsystem enclosure).

13.6.10 Soft Starting

The Power Supply shall contain control circuit which provides monotonic soft start for its outputs without overstress of the AC line or any power supply components at any specified AC line or load conditions.

13.6.11 Zero Load Stability Requirements

When the power subsystem operates in a no load condition, it does not need to meet the output regulation specification, but it must operate without any tripping of over-voltage or other fault circuitry. When the power subsystem is subsequently loaded, it must begin to regulate and source current without fault.

13.6.12 Hot Swap Requirements

Hot swapping a power supply is the process of inserting and extracting a power supply from an operating power system. During this process the output voltages shall remain within the limits with the capacitive load specified. The hot swap test must be conducted when the system is operating under static, dynamic, and zero loading conditions.

13.6.13 Forced Load Sharing

The +12V output will have active load sharing. The output will share within 10% at full load. The failure of a power supply should not affect the load sharing or output voltages of the other supplies still operating. The supplies must be able to load share in parallel and operate in a hot-swap/redundant 1+1 configurations. The 12VSB output is not required to actively share current between power supplies (passive sharing). The 12VSB output of the power supplies are connected together in the system so that a failure or hot swap of a redundant power supply does not cause these outputs to go out of regulation in the system.

13.6.14 Ripple/Noise

The maximum allowed ripple/noise output of the power supply is defined in the following table. This is measured over a bandwidth of 10Hz to 20MHz at the power supply output connectors. A 10μF tantalum capacitor in parallel with a 0.1μF ceramic capacitor is placed at the point of measurement.

Table 63. Ripples and Noise

+12V main	+12VSB
120mVp-p	120mVp-p

13.6.15 Timing Requirement

These are the timing requirements for the power supply operation. The output voltages must rise from 10% to within regulation limits (T_{vout_rise}) within 5 to 70ms. For 12VSB, it is allowed to rise from 1.0 between 25ms. **All outputs must rise monotonically**. Table below shows the timing requirements for the power supply being turned on and off two ways; 1) via the AC input with PSON held low; 2) via the PSON signal with the AC input applied. The PSU needs to remain off for 1 second minimum after POK is de-asserted.

Table 64. Timing Requirements

Item	Description	Min	Max	Units
T_{vout_rise}	Output voltage rise time	5.0 *	70 *	ms
Tsb_on_delay	Delay from AC being applied to 12VSBbeing within regulation.		1500	ms
T ac_on_delay	Delay from AC being applied to all output voltages being within regulation.		3000	ms
Tvout_holdup	Time 12V output voltage stay within regulation after loss of AC.	11		ms
Tpwok_holdup	Delay from loss of AC to de-assertion of PWOK	10		ms
Tpson_on_delay	Delay from PSON# active to output voltages within regulation limits.	5	400	ms
T pson_pwok	Delay from PSON# deactivate to PWOK being de-asserted.		5	ms
Tpwok_on	Delay from output voltages within regulation limits to PWOK asserted at turn on.	100	500	ms
T pwok_off	Delay from PWOK de-asserted to output voltages dropping out of regulation limits.	1		ms
Tpwok_low	Duration of PWOK being in the de-asserted state during an off/on cycle using AC or the PSON signal.	100		ms
Tsb_vout	Delay from 12VSBbeing in regulation to O/Ps being in regulation at AC turn on.	50	1000	ms

Item	Description	Min	Max	Units
T12VSB_holdup	Time the 12VSBoutput voltage stays within regulation after loss of AC.	70		ms

Note:
* The 12VSB output voltage rise time shall be from 1.0ms to 25ms.

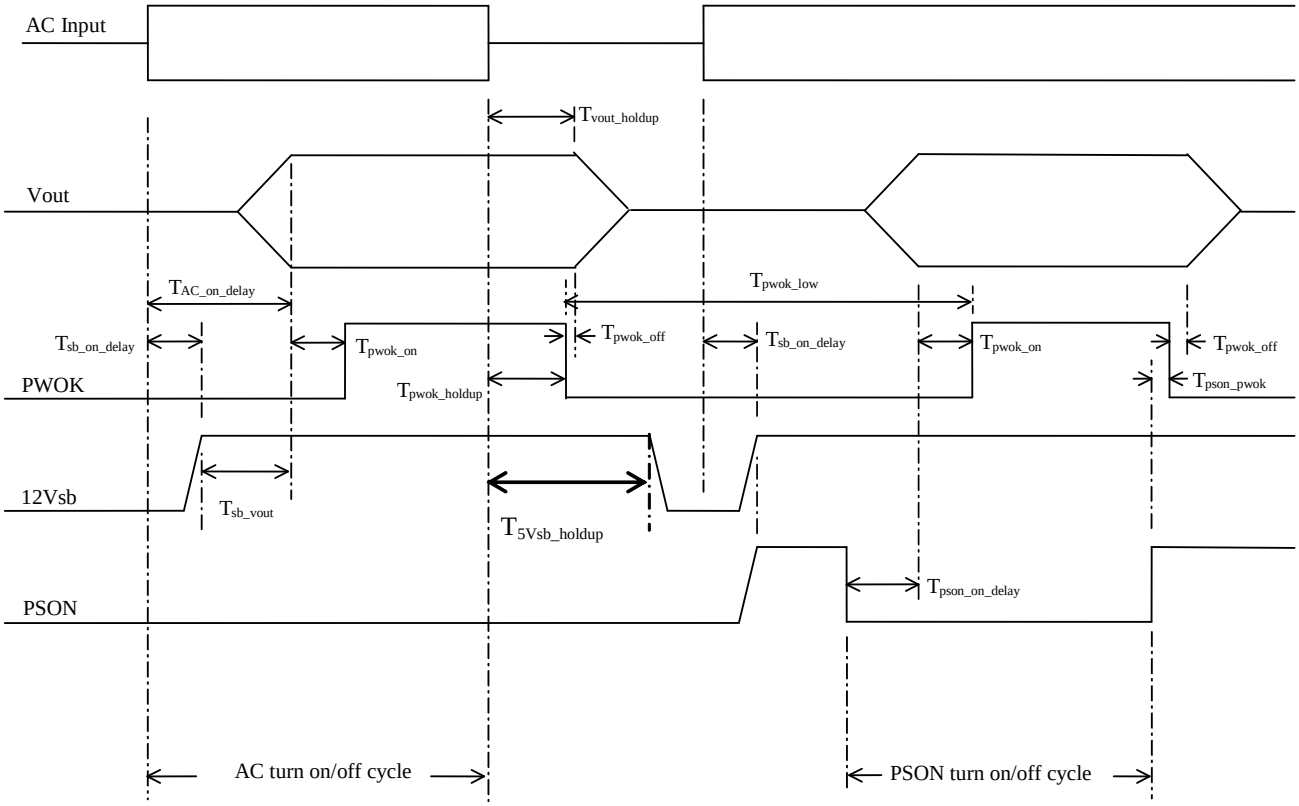


Figure 44. Turn On/Off Timing (Power Supply Signals)

13.7 Power Supply DC Output Specification

Protection circuits inside the power supply shall cause only the power supply's main outputs to shut down. If the power supply latches off due to a protection circuit tripping, an AC cycle OFF for 15sec and a PSON# cycle HIGH for 1sec shall be able to reset the power supply.

13.7.1 Current Limit & Power Protection (OCP & OPP)

The power supply shall have current limit to prevent the outputs from exceeding the values shown in table below. If the current limits are exceeded the power supply shall shutdown and latch off. The latch will be cleared only by an AC power interruption. The power supply shall not be damaged from repeated power cycling in this condition. 12VSB will be auto-recovered after removing OCP limit.

Table 65. Over current protection (OCP) and warning

Name	Description	Current Threshold		Trip timing		Testing range	Comments
		MIN	MAX	MIN	MAX		
OCP1	Fast over current protection (shutdown, latch)	280A	300A	10µsec	100usec	OCP1 to Short Circuit	
OPP	Over power protection (voltage foldback)	265A	280A	NA	NA	OPP to Vfoldback to 8V	

Name	Description	Current Threshold		Trip timing		Testing range	Comments
OCW1	Fast over current warning (SMBAlert#)	250A	265A	5µsec	20µsec	OCW1 to OCP1	Latch and hold for 50-150msec
OCP2	Slow over current protection (shutdown, latch)	210A	230A	50msec	100msec	OCP2 to OCP1	
OCW2	Slow over current warning (SMBAlert#)	210A	230A	25msec	50msec	OCW2 to OCW1	
OCPstby	Stby over current protection (shutdown, hiccup mode)	4.5A	5.5A				10msec minimum delay

13.7.2 Fast Output Current Sharing

Fast output is supported with a circuit in the power supply to quickly assert the SMBAlert signal when the output current exceeds the Ithrottle threshold. A current sense resistor on the output side of the PSUs output capacitors shall be used to quickly sense current exceeding the Ithrottle threshold. The SMBAlert# signal shall assert within Tfast_smbalert time. The PSU shall hold the SMBAlert# signal asserted for Tsmbalert_latch duration then release it.

Table 66. Fast output OCP and warning

Name	Description	Current Threshold		Trip timing		Testing range	Comments
		MIN	MAX	MIN	MAX		
OCP1	Fast over current protection (shutdown, latch)	280A	300A	10µsec	100usec	OCP1 to Short Circuit	
OPP	Over power protection (voltage foldback)	265A	280A	NA	NA	OPP to Vfoldback to 8V	
OCW1	Fast over current warning (SMBAlert#)	250A	265A	5µsec	20µsec	OCW1 to OCP1	Latch and hold for 50-150msec
OCP2	Slow over current protection (shutdown, latch)	210A	230A	50msec	100msec	OCP2 to OCP1	
OCW2	Slow over current warning (SMBAlert#)	210A	230A	25msec	50msec	OCW2 to OCW1	
OCPstby	Stby over current protection (shutdown, hiccup mode)	4.5A	5.5A				10msec minimum delay

13.7.3 Over Voltage Protection

The power supply over voltage protection shall be locally sensed. The power supply shall shutdown and latch off after an over voltage condition occurs. This latch shall be cleared by toggling the PSON# signal or by an AC power interruption. The values are measured at the output of the power supply's connectors. The voltage shall never exceed the maximum levels when measured at the power connectors of the power supply connector during any single point of fail. The voltage shall never trip any lower than the minimum levels when measured at the power connector. 12VSB will be auto-recovered after removing OVP limit.

Table 67. Over Voltage Protection Limits

Output Voltage	MIN (V)	MAX (V)
+12V	**13.3**	**14.5**
+12VSB	**13.3**	**14.5**

13.7.4 Over Temperature Protection

The power supply will be protected against over temperature conditions caused by loss of fan cooling or excessive ambient temperature. In an OTP condition the PSU will shut down. OT warning SMBAlert assertion (see Common Hardware & Firmware Requirements for CRPS Power Supplies must always precede the OTP shutdown. When the power supply temperature drops to within specified limits, the power supply shall restore power automatically, while the 12VSB remains always on. The OTP circuit must have built in margin such that the power supply will not oscillate on and off due to temperature recovering condition. The OTP trip temperature level shall be at least 5°C higher than SMBAlert over temperature warning threshold level.

Appendix A. Integration and Usage Tips

- When adding or removing components or peripherals from the server board, AC power must be removed. With AC power plugged into the server board, 5V standby is still present even though the server board is powered off.
- This server board supports the Intel® Xeon™ Phi™ processor product family with a Thermal Design Power (TDP) of up to and including 215 Watts. Previous generations of the Intel® Xeon™ Phi™ processors are not supported.
- On the back edge of the server board are eight diagnostic LEDs that display a sequence of amber POST codes during the boot process. If the server board hangs during POST, the LEDs display the last POST event run before the hang.
- Normal Integrated BMC functionality is disabled with the BMC Force Update jumper set to the “enabled” position (pins 2-3). The server should never be run with the BMC Force Update jumper set in this position and should only be used when the standard firmware update process fails. This jumper should remain in the default (disabled) position (pins 1-2) when the server is running normally.
- When performing a normal BIOS update procedure, the BIOS recovery jumper must be set to its default position (pins 1-2).

Appendix B. Initial Usage

- Refer to the Service Guide when available for initial installation
- Follow the guidelines to install memory, the CPU, PCI-e add-in devices, and RMM4 management module if required.
- For the Intel® Xeon™ Phi Processor X200 Product Family CPUs, please download and install the Intel® Many Integrated Core (MIC) Architecture Software Package to enable all functionality of the Intel® Xeon™ Phi Processor X200 Processor. Refer to the user guide that is provided in the package. This software can be obtained via the Intel® Premier Support tool via download. Search for “Knights Landing” and then look for “MPSP” software.
- For installation of the node into the chassis and chassis related information, please refer to the Intel® Server Chassis H2000G Product Family Technical Product Specification. “Adams Pass” specific information will be updated in this document during the Beta timeframe. However, compute node installation, drive bay configuration and scheme are essentially the same.

Appendix C. Integrated BMC Sensor Tables

This section will be updated in the Beta time frame. This appendix lists the sensor identification numbers and information about the sensor type, name, supported thresholds, assertion and de-assertion information, and a brief description of the sensor purpose. See the *Intelligent Platform Management Interface Specification, Version 2.0*, for sensor and event/reading-type table information.

- **Sensor Type**
  The sensor type references the values in the Sensor Type Codes table in the *Intelligent Platform Management Interface Specification Second Generation v2.0*. It provides a context to interpret the sensor.
- **Event/Reading Type**
  The event/reading type references values from the Event/Reading Type Code Ranges and the Generic Event/Reading Type Code tables in the *Intelligent Platform Management Interface Specification Second Generation v2.0*. Digital sensors are specific type of discrete sensors that only have two states.
- **Event Thresholds/Triggers**
  The following event thresholds are supported for threshold type sensors:

  [u,l][nr,c,nc] upper non-recoverable, upper critical, upper non-critical, lower non-recoverable, lower critical, lower non-critical uc, lc upper critical, lower critical

  Event triggers are supported event-generating offsets for discrete type sensors. The offsets can be found in the *Generic Event/Reading Type Code* or *Sensor Type Code* tables in the *Intelligent Platform Management Interface Specification Second Generation v2.0*, depending on whether the sensor event/reading type is generic or a sensor-specific response.
- **Assertion/Deassertion**
  Assertion and de-assertion indicators reveal the type of events this sensor generates:

  As: Assertion

  De: De-assertion
- **Readable Value/Offsets**
  Readable value indicates the type of value returned for threshold and other non-discrete type sensors.

  Readable offsets indicate the offsets for discrete sensors that are readable by means of the *Get Sensor Reading* command. Unless otherwise indicated, event triggers are readable. Readable offsets consist of the reading type offsets that do not generate events.
- **Event Data**
  Event data is the data that is included in an event message generated by the associated sensor. For threshold-based sensors, these abbreviations are used:

  R: Reading value

  T: Threshold value
- **Rearm Sensors**
  The rearm is a request for the event status for a sensor to be rechecked and updated upon a transition between good and bad states. Rearming the sensors can be done manually or automatically. This column indicates the type supported by the sensor. The following abbreviations are used in the comment column to describe a sensor:

  A: Auto-rearm

  M: Manual rearm

  I: Rearm by init agent

- **Default Hysteresis**
  The hysteresis setting applies to all thresholds of the sensor. This column provides the count of hysteresis for the sensor, which can be 1 or 2 (positive or negative hysteresis).
- **Criticality**
  Criticality is a classification of the severity and nature of the condition. It also controls the behavior of the front panel status LED.
- **Standby**
  Some sensors operate on standby power. These sensors may be accessed and/or generate events when the main (system) power is off, but AC power is present.

Note: All sensors listed below may not be present on all platforms. Please reference the BMC EPS for platform applicability. Redundancy sensors will only be present on systems with appropriate hardware to support redundancy (for instance, fan or power supply).

Table 68. BMC Sensor Table

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
Power Unit Status *(Pwr Unit Status)*	01h	All	Power Unit 09h	Sensor Specific 6Fh	00 - Power down	OK	As and De	–	Trig Offset	A	X
					02 - 240 VA power down	Fatal					
					04 - A/C lost	OK					
					05 - Soft power control failure	Fatal					
					06 - Power unit failure						
Power Unit Redundancy[1] *(Pwr Unit Redund)*	02h	Chassis-specific	Power Unit 09h	Generic 0Bh	00 - Fully Redundant	OK	As	–	Trig Offset	M	X
					01 - Redundancy lost	Degraded					
					02 - Redundancy degraded	Degraded					
					03 - Non-redundant: sufficient resources. Transition from full redundant state.	Degraded					
					04 – Non-redundant: sufficient resources. Transition from insufficient state.	Degraded					
					05 - Non-redundant: insufficient resources	Fatal					

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
					06 – Redundant: degraded from fully redundant state.	Degraded					
					07 – Redundant: Transition from non-redundant state.	Degraded					
IPMI Watchdog *(IPMI Watchdog)*	03h	All	Watchdog 2 23h	Sensor Specific 6Fh	00 - Timer expired, status only	OK	As	–	Trig Offset	A	X
					01 - Hard reset						
					02 - Power down						
					03 - Power cycle						
					08 - Timer interrupt						
Physical Security *(Physical Scrty)*	04h	Chassis Intrusion is chassis-specific	Physical Security 05h	Sensor Specific 6Fh	00 - Chassis intrusion	Degraded OK	As and De	–	Trig Offset	A	X
					04 - LAN leash lost						
FP Interrupt *(FP NMI Diag Int)*	05h	Chassis - specific	Critical Interrupt 13h	Sensor Specific 6Fh	00 - Front panel NMI/diagnostic interrupt	OK	As	–	Trig Offset	A	–
QPI Correctable Event (QPI Corr Sensor)	06h	All	Critical Event 13h	72h							
QPI Uncorrectable Event (QPI Fatl Sensor)	07h	All	Critical Event 13h	73h							
SMI Timeout *(SMI Timeout)*	06h	All	SMI Timeout F3h	Digital Discrete 03h	01 – State asserted	Fatal	As and De	–	Trig Offset	A	–
System Event Log *(System Event Log)*	07h	All	Event Logging Disabled 10h	Sensor Specific 6Fh	02 - Log area reset/cleared	OK	As	–	Trig Offset	A	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
System Event *(System Event)*	08h	All	System Event 12h	Sensor Specific 6Fh	04 – PEF action	OK	As	-	Trig Offset	A	X
Button Sensor *(Button)*	09h	All	Button/Switch 14h	Sensor Specific 6Fh	00 – Power Button 02 – Reset Button	OK	AS	–	Trig Offset	A	X
BMC Watchdog	0Ah	All	Mgmt System Health 28h	Digital Discrete 03h	01 – State Asserted	Degraded	As	–	Trig Offset	A	X
Voltage Regulator Watchdog (VR Watchdog)	0Bh	All	Voltage 02h	Digital Discrete 03h	01 – State Asserted	Fatal	As and De	–	Trig Offset	M	X
Fan Redundancy[1] *(Fan Redundancy)*	0Ch	Chassis-specific	Fan 04h	Generic 0Bh	00 - Fully redundant	OK	As and De	–	Trig Offset	A	–
					01 - Redundancy lost	Degraded					
					02 - Redundancy degraded	Degraded					
					03 - Non-redundant: Sufficient resources. Transition from redundant	Degraded					
					04 - Non-redundant: Sufficient resources. Transition from insufficient.	Degraded					

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
					05 - Non-redundant: insufficient resources.	Non-Fatal					
					06 – Non-Redundant: degraded from fully redundant.	Degraded					
					07 - Redundant degraded from non-redundant	Degraded					
SSB Thermal Trip *(SSB Therm Trip)*	0Dh	All	Temperature 01h	Digital Discrete 03h	01 – State Asserted	Fatal	As and De	–	Trig Offset	M	X
BMC Firmware Health (BMC FW Health)	10h	All	Mgmt Health 28h	Sensor Specific 6Fh	04 – Sensor Failure	Degraded	As	-	Trig Offset	A	X
System Airflow (System Airflow)	11h	All	Other Units 0Bh	Threshold 01h	–	–	–	Analog	–	–	–
FW Update Status	12h	All	Version Change 2Bh	OEM defined 70h	00h – Update started 01h – Update completed successfully. 02h – Update failure	OK	As	–	Trig Offset	A	–
Baseboard Temperature 5 *(*Platform Specific*)*	14h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Baseboard Temperature 6 *(*Platform Specific*)*	15h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
NM Health *(NM Health)*	19h	Platform-specific	OEM DCh	OEM defined 73h	-	-	-	-	-	-	–
NM Capabilities *(NM Capabilities)*	1Ah	Platform-specific	OEM DCh	OEM defined 74h	-	-	-	-	-	-	–
Baseboard Temperature 1 *(*Platform Specific*)*	20h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Front Panel Temperature *(Front Panel Temp)*	21h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc] UNR	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
SSB Temperature *(SSB Temp)*	22h	All	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	–
Baseboard Temperature 2 *(*Platform Specific*)*	23h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Baseboard Temperature 3 *(*Platform Specific*)*	24h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Baseboard Temperature 4 *(*Platform Specific*)*	25h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
IO Module Temperature (I/O Mod Temp)	26h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
PCI Riser 1 Temperature (PCI Riser 1 Temp)	27h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
IO Riser Temperature (IO Riser Temp)	28h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Hot-swap Backplane 1 Temperature (HSBP 1 Temp)	29h	Chassis-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	-
Hot-swap Backplane 2 Temperature (HSBP 2 Temp)	2Ah	Chassis-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	-
Hot-swap Backplane 3 Temperature (HSBP 3 Temp)	2Bh	Chassis-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	-
PCI Riser 2 Temperature (PCI Riser 2 Temp)	2Ch	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
SAS Module Temperature (SAS Mod Temp)	2Dh	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Exit Air Temperature (Exit Air Temp)	2Eh	Chassis and Platform Specific	Temperature 01h	Threshold 01h	This sensor does not generate any events.	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	-
Network Interface Controller Temperature (LAN NIC Temp)	2Fh	All	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Fan Tachometer Sensors[2] (Chassis specific sensor names)	30h–3Fh	Chassis and Platform Specific	Fan 04h	Threshold 01h	[l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	M	-
Fan Present Sensors *(Fan x Present)*	40h–4Fh	Chassis and Platform Specific	Fan 04h	Generic 08h	01 - Device inserted	OK	As and De	-	Triggered Offset	Auto	-
Power Supply 1 Status[3] *(PS1 Status)*	50h	Chassis-specific	Power Supply 08h	Sensor Specific 6Fh	00 - Presence	OK	As and De	–	Trig Offset	A	X
					01 - Failure	Degraded					
					02 – Predictive Failure	Degraded					
					03 - A/C lost	Degraded					
					06 – Configuration error	OK					
Power Supply 2 Status[3] *(PS2 Status)*	51h	Chassis-specific	Power Supply 08h	Sensor Specific 6Fh	00 - Presence	OK	As and De	–	Trig Offset	A	X
					01 - Failure	Degraded					

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
					02 – Predictive Failure	Degraded					
					03 - A/C lost	Degraded					
					06 – Configuration error	OK					
Power Supply 1 AC Power Input *(PS1 Power In)*	54h	Chassis-specific	Other Units 0Bh	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Power Supply 2 AC Power Input *(PS2 Power In)*	55h	Chassis-specific	Other Units 0Bh	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Power Supply 1 +12V % of Maximum Current Output *(PS1 Curr Out %)*	58h	Chassis-specific	Current 03h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Power Supply 2 +12V % of Maximum Current Output *(PS2 Curr Out %)*	59h	Chassis-specific	Current 03h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Power Supply 1 Temperature *(PS1 Temperature)*	5Ch	Chassis-specific	Temperature 01h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Power Supply 2 Temperature *(PS2 Temperature)*	5Dh	Chassis-specific	Temperature 01h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
Hard Disk Drive 15 - 23 Status *(HDD 15 - 23 Status)*	60h – 68h	Chassis-specific	Drive Slot 0Dh	Sensor Specific 6Fh	00 - Drive Presence	OK	As and De	–	Trig Offset	A	–
					01- Drive Fault	Degraded					
					07 - Rebuild/Remap in progress	Degraded					
Processor 1 Status *(P1 Status)*	70h	All	Processor 07h	Sensor Specific 6Fh	01 - Thermal trip/ FIVR	Fatal	As and De	–	Trig Offset	M	X
					07 - Presence	OK					
					07 - Presence	OK					
Processor 1 Thermal Margin *(P1 Therm Margin)*	74h	All	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Processor 1 Thermal Control % *(P1 Therm Ctrl %)*	78h	All	Temperature 01h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	Trig Offset	A	–
Processor ERR2 Timeout *(CPU ERR2)*	7Ch	All	Processor 07h	Digital Discrete 03h	01 – State Asserted	Fatal	As and De	–	Trig Offset	A	–
IERR recovery dump info (IERR Rec Info)	7Dh	All	OEM sensor type D1h	OEM defined 70h	00h – Dump successfully 01h – Dump failure	OK	As	–	Trig Offset	A	–
InternalCatastrophic Error *(IERR \ CATERR)*	80h	All	Processor 07h	Digital Discrete 03h	01 – State Asserted	Fatal	As and De	–	Trig Offset	M	–
MTM Level Change *(MTM Lvl Change)*	81h	All	Mgmt Health 28h	Digital Discrete 03h	01 – State Asserted	-	As and De	–	Trig Offset	A	–
Processor Population Fault *(CPU Missing)*	82h	All	Processor 07h	Digital Discrete 03h	01 – State Asserted	Fatal	As and De	–	Trig Offset	M	X

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
Processor 1 DTS Thermal Margin *(P1 DTS Therm Mgn)*	83h	All	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Auto Config Status *(AutoCfg Status)*	87h	All	Mgmt Health 28h	Digital Discrete 03h	01 – State Asserted	-	As and De	–	Trig Offset	A	–
VRD Over Temperature (*VRD Hot*)	90h	All	Temperature 01h	Digital Discrete 05h	01 - Limit exceeded	Non-fatal	As and De	–	Trig Offset	A	–
Power Supply 1 Fan Fail 1[3] *(PS1 Fan Fail 1)*	A0h	Chassis-specific	Fan 04h	Generic – digital discrete 03h	01 – State Asserted	Non-fatal	As and De	-	Trig Offset	M	X-
Power Supply 1 Fan Fail 2[3] *(PS1 Fan Fail 2)*	A1h	Chassis-specific	Fan 04h	Generic – digital discrete 03h	01 – State Asserted	Non-fatal	As and De	-	Trig Offset	M	X-
PHI 1 Status *(GPGPU1 Status)*	A2h	Platform Specific	Status C0h	OEM Defined 70h	-	-	-	-	-	-	-
PHI 2 Status *(GPGPU2 Status)*	A3h	Platform Specific	Status C0h	OEM Defined 70h	-	-	-	-	-	-	-
Power Supply 2 Fan Fail 1[3] *(PS2 Fan Fail 1)*	A4h	Chassis-specific	Fan 04h	Generic – digital discrete 03h	01 – State Asserted	Non-fatal	As and De	-	Trig Offset	M	X-
Power Supply 2 Fan Fail 2[3] *(PS2 Fan Fail 2)*	A5h	Chassis-specific	Fan 04h	Generic – digital discrete 03h	01 – State Asserted	Non-fatal	As and De	-	Trig Offset	M	X
PHI 3 Status *(GPGPU3 Status)*	A6h	Platform Specific	Status C0h	OEM Defined 70h	-	-	-	-	-	-	-

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
PHI 4 Status *(GPGPU4 Status)*	A7h	Platform Specific	Status C0h	OEM Defined 70h	-	-	-	-	-	-	-
PHI 1 Avg Pwr	AAh	Platform Specific	Power 03h	Threshold 01h	-	-	-	Analog	-	-	-
PHI 2 Avg Pwr	ABh	Platform Specific	Power 03h	Threshold 01h	-	-	-	Analog	-	-	-
Processor 1 DIMM Aggregate Thermal Margin 1 *(P1 DIMM Thrm Mrgn1)*	B0h	All	Temperature 01h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	–
Processor 1 DIMM Aggregate Thermal Margin 2 *(P1 DIMM Thrm Mrgn2)*	B1h	All	Temperature 01h	Threshold 01h	[u] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	–
Node Auto-Shutdown Sensor *(Auto Shutdown)*	B8h	Multi-Node Specific	Power Unit 09h	Generic – digital discrete 03h	01 – State Asserted	Non-fatal	As and De	-	Trig Offset	A	X
Fan Tachometer Sensors (Chassis specific sensor names)	BAh–BFh	Chassis and Platform Specific	Fan 04h	Threshold 01h	[l] [c,nc]	nc = Degraded c = Non-fatal[2]	As and De	Analog	R, T	M	-
Processor 1 DIMM Thermal Trip *(P1 Mem Thrm Trip)*	C0h	All	Memory 0Ch	Sensor Specific 6Fh	0A- Critical overtemperature	Fatal	As and De	–	Trig Offset	M	X
PHI 1 Temp (GPGPU1 Core Temp)	C4h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	-	-	-	-
PHI 2 Temp (GPGPU2 Core Temp)	C5h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	-	-	-	-
PHI 3 Temp (GPGPU3 Core Temp)	C6h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	-	-	-	-

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
PHI 4 Temp (GPGPU4 Core Temp)	C7h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	-	-	-	-
Global Aggregate Temperature Margin 1 (Agg Therm Mrgn 1)	C8h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 2 (Agg Therm Mrgn 2)	C9h	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 3 (Agg Therm Mrgn 3)	CAh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 4 (Agg Therm Mrgn 4)	CBh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 5 (Agg Therm Mrgn 5)	CCh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 6 (Agg Therm Mrgn 6)	CDh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 7 (Agg Therm Mrgn 7)	CEh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–
Global Aggregate Temperature Margin 8 (Agg Therm Mrgn 8)	CFh	Platform Specific	Temperature 01h	Threshold 01h	-	-	-	Analog	R, T	A	–

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
Baseboard +12V *(BB +12.0V)*	D0h	All	Voltage 02h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	–
Baseboard Temperature 5 *(MEM VRM Temp)*	D5h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Baseboard Temperature 6 *(MEM EFVRD Temp)*	D6h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Voltage Fault *(Voltage Fault)*	D1h	All	Voltage 02h	Discrete 03h	01 – Asserted	Degraded	-	-	-	A	-
Baseboard Temperature 5 *(Platform Specific)*	D5h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Baseboard Temperature 6 *(Platform Specific)*	D6h	Platform-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Bad User PWD	D7h	All	Session Audit 2Ah	Sensor Specific 6Fh	02h - Invalid Username or Password 03h – Invalid password disable	OK	As for 02h. As and De for 03h	–	Trig Offset	A	–

Full Sensor Name (Sensor name in SDR)	Sensor #	Platform Applicability	Sensor Type	Event/ Reading Type	Event Offset Triggers	Contrib. To System Status	Assert/ De-assert	Readable Value/ Offsets	Event Data	Rearm	Stand -by
Baseboard CMOS Battery *(BB +3.3V Vbat)*	DEh	All	Voltage 02h	Threshold 01h	[l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	–
Hot-swap Backplane 4 Temperature (HSBP 4 Temp)	E0h	Chassis-specific	Temperature 01h	Threshold 01h	[u,l] [c,nc]	nc = Degraded c = Non-fatal	As and De	Analog	R, T	A	X
Rear Hard Disk Drive 0 -1 Status (Rear HDD 0 - 1 Stat)	E2h - E3h	Chassis-specific	Drive Slot 0Dh	Sensor Specific 6Fh	00 - Drive Presence	OK	As and De	–	Trig Offset	A	–
					01- Drive Fault	Degraded					
					07 - Rebuild/Remap in progress	Degraded					
Hard Disk Drive 0 -14 Status (HDD 0 - 14 Status)	F0h - FEh	Chassis-specific	Drive Slot 0Dh	Sensor Specific 6Fh	00 - Drive Presence	OK	As and De	–	Trig Offset	A	–
					01- Drive Fault	Degraded					
					07 - Rebuild/Remap in progress	Degraded					

Notes:

1. Redundancy sensors will be only present on systems with appropriate hardware to support redundancy (for instance, fan or power supply). Note that power supply redundancy may be lost even when both supplies are operational if the system is loaded beyond the capacity of a single power supply.
2. This is only applicable when the system doesn't support redundant fans. When fan redundancy is supported, then the contribution to system state is driven by the fan redundancy sensor, not individual sensors. On a system with fan redundancy, the individual sensor severities will read the same as the fan redundancy sensor's severity.
3. This is only applicable when the system doesn't support redundant power supplies. When redundancy is supported, then the contribution to system state is driven by the power unit redundancy sensor. On a system with power supply redundancy, the individual sensor severities will read the same as the power unit redundancy sensor's severity.

Appendix D. BIOS Sensors and SEL Data

BIOS owns a set of IPMI-compliant Sensors. These are actually divided in ownership between BIOS POST (GID = 01) and BIOS SMI Handler (GID = 33). The SMI Handler Sensors are typically for logging runtime error events, but they are active during POST and may log errors such as Correctable Memory ECC Errors if they occur.

It is important to remember that a Sensor is uniquely identified by the combination of Sensor Owner plus Sensor Number. There are cases where the same Sensor Number is used with different Sensor Owners – this is not a conflict. For example, in the BIOS Sensors list there is a Sensor Number 83h for Sensor Owner 01h (BIOS POST) as well as for Sensor Owner 33h (SMI Handler), but these are two distinct sensors reporting the same type of event from different sources (Generator IDs 01h and 33h).

On the other hand, each distinct Sensor (GID + Sensor Number) is defined by one specific Sensor Type describing the kind of data being reported, and one specific Event Type describing the type of event and the format of the data being reported.

Table 69. BIOS Sensor and SEL Data

The BIOS Sensors list below includes all BIOS Sensors used with BIOS IPMI System Event Logging, with the exception of additional specialized sensors defined for Compute Module boards. These sensors are described in the *PCSD Blade BIOS Extension External Product Specification*.

Sensor Name	Sensor Number	Sensor Owner (GID)	Sensor Type	Event/Reading Type Offset Values	Event Data 2 Event Data 3
Mirroring Redundancy State	01h	33h (SMI Handler)	0Ch (Memory)	0Bh (Discrete, Redundancy State) *0h = Fully Redundant* *2h = Redundancy Degraded*	ED2 = [7:4] = Mirroring Domain 0-1 = Channel Pair for Socket [3:2] = Reserved [1:0] = Rank on DIMM 0-3 = Rank Number ED3 = [7:5]= Socket ID 0-3 = CPU1-4 [4:3] = Channel 0-3 = Channel A-D for Socket [2:0] = DIMM 0-2 = DIMM 1-3 on Channel

Sensor Name	Sensor Number	Sensor Owner (GID)	Sensor Type	Event/Reading Type Offset Values	Event Data 2 Event Data 3
Memory RAS Configuration Status	02h	01h (BIOS POST)	0Ch (Memory)	09h (Digital Discrete) *0h = RAS Configuration Disabled* *1h = RAS Configuration Enabled*	ED2 = [7:4] = Reserved [3:0] Config Err 0 = None 3 = Invalid DIMM Config for RAS Mode ED3 = [7:4] = Reserved [3:0] = RAS Mode 0 = None 1 = Mirroring 2 = Lockstep 4 = Rank Sparing
Memory ECC Error	02h	33h (SMI Handler)	0Ch (Memory)	6Fh (Sensor Specific Offset) *0h = Correctable Error* *1h = Uncorrectable Error*	ED2 = [7:2] = Reserved [1:0] = Rank on DIMM 0-3 = Rank Number ED3 = [7:5]= Socket ID 0-3 = CPU1-4 [4:3] = Channel 0-3 = Channel A-D for Socket [2:0] = DIMM 0-2 = DIMM 1-3 on Channel
Legacy PCI Error	03h	33h (SMI Handler)	13h (Critical Interrupt)	6Fh (Sensor Specific Offset) *4h = PCI PERR* *5h = PCI SERR*	ED2 = [7:0] = Bus Number ED3 = [7:3] = Device Number [2:0] = Function Number

Sensor Name	Sensor Number	Sensor Owner (GID)	Sensor Type	Event/Reading Type Offset Values	Event Data 2 Event Data 3
PCIe Fatal Error (Standard AER Errors) (see Sensor 14h for continuation)	04h	33h (SMI Handler)	13h (Critical Interrupt)	70h (OEM Discrete) *0h = Data Link Layer Protocol Error* *1h = Surprise Link Down Error* *2h = Completer Abort* *3h = Unsupported Request* *4h = Poisoned TLP* *5h = Flow Control Protocol* *6h = Completion Timeout* *7h = Receiver Buffer Overflow* *8h = ACS Violation* *9h = Malformed TLP* *Ah = ECRC Error* *Bh = Received Fatal Message From Downstream* *Ch = Unexpected Completion* *Dh = Received ERR_NONFATAL Message* *Eh = Uncorrectable Internal* *Fh = MC Blocked TLP*	ED2 = [7:0] = Bus Number ED3 = [7:3] = Device Number [2:0] = Function Number
PCIe Correctable Error (Standard AER Errors)	05h	33h (SMI Handler)	13h (Critical Interrupt)	71h (OEM Discrete) *0h = Receiver Error* *1h = Bad DLLP* *2h = Bad TLP* *3h = Replay Num Rollover* *4h = Replay Timer timeout* *5h = Advisory Non-fatal* *6h = Link BW Changed* *7h = Correctable Internal* *8h = Header Log Overflow*	ED2 = [7:0] = Bus Number ED3 = [7:3] = Device Number [2:0] = Function Number
BIOS POST Error	06h	01h (BIOS POST)	0Fh (System Firmware Progress)	6Fh (Sensor Specific Offset) *0h = System Firmware Error (POST Error Code)*	ED2 = [7:0] = LSB of POST Error Code ED3 = [7:0] MSB of POST Error Code
Memory Error Extension *(reserved for Validation)*	*10h*	*33h (SMI Handler)*	0Ch (Memory)	7Fh (OEM Discrete) *Offset Reserved*	ED2 = Reserved ED3 = Reserved

Sensor Name	Sensor Number	Sensor Owner (GID)	Sensor Type	Event/Reading Type Offset Values	Event Data 2 Event Data 3
Sparing Redundancy State	11h	33h (SMI Handler)	0Ch (Memory)	0Bh (Discrete, Redundancy State) *0h = Fully Redundant* *2h = Redundancy Degraded*	ED2 = [7:4] = Sparing Domain 0-3 = Channel A-D for Socket [3:2] = Reserved [1:0] = Rank on DIMM 0-3 = Rank Number ED3 = [7:5]= Socket ID 0-3 = CPU1-4 [4:3] = Channel 0-3 = Channel A-D for Socket [2:0] = DIMM 0-2 = DIMM 1-3 on Channel
Memory RAS Mode Select	12h	01h (BIOS POST)	0Ch (Memory)	09h (Digital Discrete) *0h = RAS Configuration Disabled* *1h = RAS Configuration Enabled*	ED2 = Prior Mode [7:4] = Reserved [3:0] = RAS Mode 0 = None 1 = Mirroring 2 = Lockstep 4 = Rank Sparing ED3 = Selected Mode [7:4] = Reserved [3:0] = RAS Mode 0 = None 1 = Mirroring 2 = Lockstep 4 = Rank Sparing

Sensor Name	Sensor Number	Sensor Owner (GID)	Sensor Type	Event/Reading Type Offset Values	Event Data 2 Event Data 3
Memory Parity Error	13h	33h (SMI Handler)	0Ch (Memory)	6Fh (Sensor Specific Offset) *3h = Command and Address Parity Error*	ED2 = Validity [7:5] = Reserved [4] = Channel Validity Check 0 = ED3 Chan # Not Valid 1 = ED3 Chan # Is Valid [3] = DIMM Validity Check 0 = ED3 DIMM # Not Valid 1 = ED3 DIMM # Is Valid [2:0] = Error Type 0 = Not Known 3 = Command and Address Parity Error ED3 = Location [7:5]= Socket ID 0-3 = CPU1-4 [4:2] = Channel 0-3 = Channel A-D for Socket [1:0] = DIMM 0-2 = DIMM 1-3 on Channel
PCIe Fatal Error#2 (Standard AER Errors) (continuation of Sensor 04h)	14h	33h (SMI Handler)	13h (Critical Interrupt)	76h (OEM Discrete) *0h = Atomic Egress Blocked* *1h = TLP Prefix Blocked* *Fh = Unspecified Non-AER Fatal Error*	ED2 = [7:0] = Bus Number ED3 = [7:3] = Device Number [2:0] = Function Number
BIOS Recovery Start	15h	33h (SMI Handler)	0Fh (System Firmware Progress)	70h (OEM Discrete) *1h = BIOS Recovery Start*	ED2 = Reserved ED3 = Reserved
BIOS Recovery Completion	15h	33h (SMI Handler)	0Fh (System Firmware Progress)	F0h (OEM Discrete) 1h = BIOS Recovery Complete	ED2 = Reserved ED3 = Reserved

Appendix E. POST Code Diagnostic LED Decoder

During the system POST process, Diagnostic LED Codes are used extensively as a mechanism to indicate progress and Fatal Halt conditions independently of the video display. If the system hangs or halts, the Diagnostic LED display can help determine the reason even when video is not available.

These Diagnostic LEDs are equivalent to the Legacy “Port 80 POST Codes”, and a Legacy I/O Port 80 output will be displayed as a Diagnostic LED code.

The Diagnostic LEDs are a set of LEDs found on the back edge of the server board. There are 8 Diagnostic LEDs which form a 2 hex digit (8 bit) code read left-to-right as facing the rear of the server.

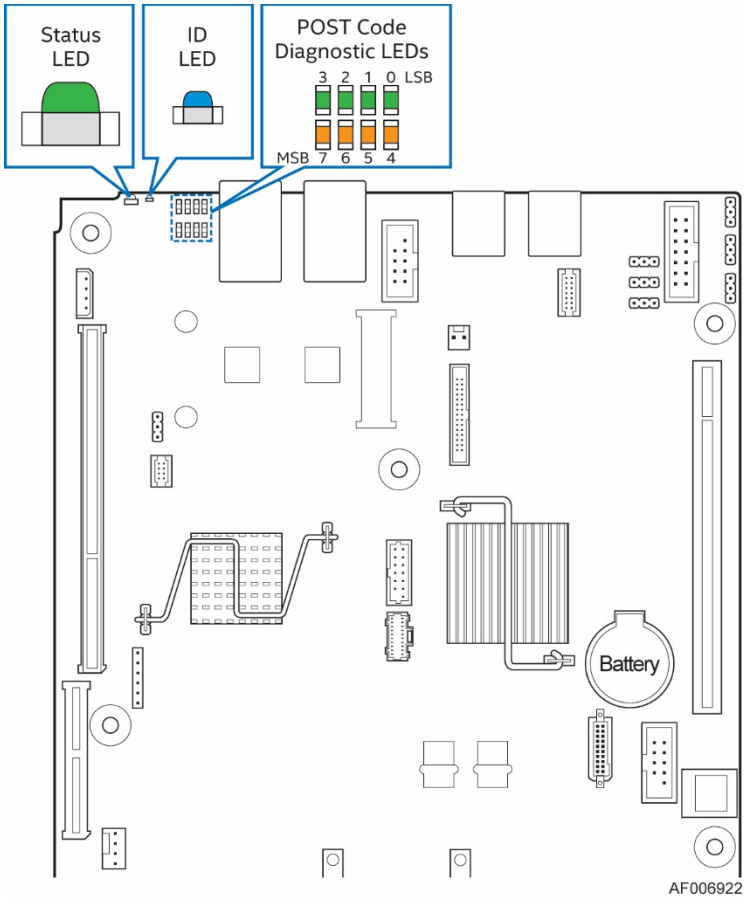


Figure 45. Diagnostic LED Placement Diagram

An LED which is ON represents a 1 bit value, and an LED which is OFF represents a 0 bit value. The LED bit values are read as Most Significant Bit to the left, Least Significant Bit to the right.

In the following example, the BIOS sends a value of ACh to the diagnostic LED decoder. The LEDs are decoded as follows:

Table 70. POST Code LED Example

LEDs	Upper Nibble AMBER LEDs				Lower Nibble GREEN LEDs			
	MSB							LSB
	LED #7	LED #6	LED #5	LED #4	LED #3	LED #2	LED #1	LED #0
	8h	4h	2h	1h	8h	4h	2h	1h
Status	ON	OFF	ON	OFF	ON	ON	OFF	OFF

Results	1	0	1	0	1	1	0	0
	Ah				Ch			

- Upper nibble bits = 1010b = Ah; Lower nibble bits = 1100b = Ch; the two are concatenated as ACh.

POST Memory Initialization MRC Diagnostic Codes

This is a brief list of the Diagnostic LED codes displayed during memory initialization by the Memory Reference Code (MRC), the BIOS component responsible for it. There are two types of POST Diagnostic Codes used by the MRC, Fatal Error Codes and Progress Codes.

MRC Fatal Error Codes are necessary because if the Memory Initialization fails badly for some reason – like no usable memory installed – the system would not have the resources to give any other error indication. So in the case of a major failure during Memory Initialization, the system outputs a Fatal Error Code to Port 80 (the Diagnostic LEDs) and executes a Halt. These Fatal Error Halts do not change the Status LED, and they do not get logged as SEL Events.

The MRC Progress Codes are displays to the Diagnostic LEDs that show the execution point in the MRC operational path at each step. The intent is that if the system hangs during execution of the MRC, the LED display will tell at what point in the code the system was executing.

Be aware that these are Diagnostic LED display codes used in early POST by the MRC. Later in POST these same Diagnostic LED display codes are used for other BIOS Progress Codes.

Also, be aware that the MRC Fatal Error Codes and MRC Progress Codes are ***not controlled by the BIOS*** and are subject to change at the discretion of the Memory Reference Code teams.

Table 71. MRC Fatal Error Codes

Error Code	Fatal Error Code Explanation (with MRC Internal Minor Code)
0xE8	No Usable Memory Error: 01h = No memory was detected via SPD read, or invalid config that causes no operable memory. 02h = Memory DIMMs on all channels of all sockets are disabled due to hardware memtest error. 03h = No memory installed. All channels are disabled.
0xE9	Memory is locked by Intel® Trusted Execution Technology and is inaccessible.
0xEA	DDR4 Channel Training Error: 01h = Error on read DQ/DQS (Data/Data Strobe) init 02h = Error on Receive Enable 03h = Error on Write Leveling 04h = Error on write DQ/DQS (Data/Data Strobe)
0xEB	Memory Test Failure: 01h = Software memtest failure. 02h = Hardware memtest failed. 03h = Hardware Memtest failure in Lockstep Channel mode requiring a channel to be disabled. *This is a fatal error which requires a reset and calling MRC with a different RAS mode to retry.*

Error Code	Fatal Error Code Explanation (with MRC Internal Minor Code)
0xED	DIMM Configuration/Population Error: 01h = Different DIMM types (RDIMM, LRDIMM) are detected installed in the system. 02h = Violation of DIMM population rules. 03h = The third DIMM slot cannot be populated when QR DIMMs are installed. 04h = UDIMMs are not supported. 05h = Unsupported DIMM Voltage.
0xEF	Indicates a CLTT table structure error.

Table 72. MRC Progress Codes

Progress Code	Main Sequence	Subsequences/Subfunctions
0xB0	Detect DIMM population	*—n/a—*
0xB1	Set DDR4 frequency	*—n/a—*
0xB2	Gather remaining SPD data	*—n/a—*
0xB3	Program registers on the memory controller level	*—n/a—*
0xB4	Evaluate RAS modes and save rank information	*—n/a—*
0xB5	Program registers on the channel level	*—n/a—*
0xB6	Perform the JEDEC defined initialization sequence	*—n/a—*
0xB7	Train DDR4 ranks	*—n/a—*
0x01	⇩	Read DQ/DQS training
0x02	⇩	Receive Enable training
0x03	⇩	Write Leveling training
0x04	⇩	Write DQ/DQS training
0x05	⇩	DDR channel training done
0xB8	Initialize CLTT/OLTT	*—n/a—*
0xB9	Hardware memory test and init	*—n/a—*
0xBA	Execute software memory init	*—n/a—*
0xBB	Program memory map and interleaving	*—n/a—*
0xBC	Program RAS configuration	*—n/a—*
0xBF	MRC is done	*—n/a—*

POST Progress Code Checkpoints

During the system boot process, the BIOS executes a number of platform configuration processes, each of which is assigned a specific POST Progress Code, a 2-digit hexadecimal number. As each configuration routine is started, the BIOS displays the POST Progress Code on the Diagnostic LEDs found on the back edge of the server board.

To assist in troubleshooting a system hang during the POST process, the POST Progress Code displayed in the Diagnostic LEDs can be used to identify the last POST process to begin execution.

These POST Progress Codes are the same for all board series using this BIOS and included in this BIOS EPS.

Table 73. POST Progress Codes

Progress Code	Description
SEC Phase	
0x01	First POST code after CPU reset
0x02	Microcode load begin
0x03	CRAM initialization begin
0x04	Pei Cache When Disabled
0x05	SEC Core At Power On Begin.
0x06	Early CPU initialization during Sec Phase.
0x07	Early SB initialization during Sec Phase.
0x08	Early NB initialization during Sec Phase.
0x09	End Of Sec Phase.
0x0E	Microcode Not Found.
0x0F	Microcode Not Loaded.
PEI Phase	
0x10	PEI Core
0x11	CPU PEIM
0x15	NB PEIM
0x19	SB PEIM
MRC Progress Codes *At this point the MRC Progress Code sequence is executed See Table 72.*	
0x31	Memory Installed
0x32	CPU PEIM (CPU Init)
0x33	CPU PEIM (Cache Init)
0x4F	Dxe IPL started
DXE Phase	
0x60	DXE Core started
0x61	DXE NVRAM Init
0x62	DXE Setup Init
0x63	DXE CPU Init
0x65	DXE CPU BSP Select
0x66	DXE CPU AP Init
0x68	DXE PCI Host Bridge Init

Progress Code	Description
0x69	DXE NB Init
0x6A	DXE NB SMM Init
0x70	DXE SB Init
0x71	DXE SB SMM Init
0x72	DXE SB devices Init
0x78	DXE ACPI Init
0x79	DXE CSM Init
0x80	DXE BDS Started
0x81	DXE BDS connect drivers
0x82	DXE PCI Bus begin
0x83	DXE PCI Bus HPC Init
0x84	DXE PCI Bus enumeration
0x85	DXE PCI Bus resource requested
0x86	DXE PCI Bus assign resource
0x87	DXE CON_OUT connect
0x88	DXE CON_IN connect
0x89	DXE SIO Init
0x8A	DXE USB start
0x8B	DXE USB reset
0x8C	DXE USB detect
0x8D	DXE USB enable
0x91	DXE IDE begin
0x92	DXE IDE reset
0x93	DXE IDE detect
0x94	DXE IDE enable
0x95	DXE SCSI begin
0x96	DXE SCSI reset
0x97	DXE SCSI detect
0x98	DXE SCSI enable
0x99	DXE verifying SETUP password
0x9B	DXE SETUP start
0x9C	DXE SETUP input wait
0x9D	DXE Ready to Boot
0x9E	DXE Legacy Boot

Progress Code	Description
0x9F	DXE Exit Boot Services
0xC0	RT Set Virtual Address Map Begin
0xC2	DXE Legacy Option ROM init
0xC3	DXE Reset system
0xC4	DXE USB Hot plug
0xC5	DXE PCI BUS Hot plug
0xC6	DXE NVRAM cleanup
0xC7	DXE ACPI Enable
0x00	Clear POST Code
S3 Resume	
0x40	S3 Resume PEIM (S3 started)
0x41	S3 Resume PEIM (S3 boot script)
0x42	S3 Resume PEIM (S3 Video Repost)
0x43	S3 Resume PEIM (S3 OS wake)
BIOS Recovery	
0x46	PEIM which detected forced Recovery condition
0x47	PEIM which detected User Recovery condition
0x48	Recovery PEIM (Recovery started)
0x49	Recovery PEIM (Capsule found)
0x4A	Recovery PEIM (Capsule loaded)

Appendix F. POST Code Errors

The table below lists the supported POST Error Codes, with a descriptive Error Message text for each. There is also a Response listed, which classifies the error as Minor, Major, or Fatal depending on how serious the error is and what action the system should take.

The Response section in the following table indicates one of these actions:

- **Minor:** The message is displayed on the screen or on the Error Manager screen, and an error is logged to the SEL. The system continues booting in a degraded state. The user may want to replace the erroneous unit. The POST Error Pause option setting in the BIOS setup does not have any effect on this error.
- **Major:** The message is displayed on the Error Manager screen, and an error is logged to the SEL. The POST Error Pause option setting in the BIOS setup determines whether the system pauses to the Error Manager for this type of error so the user can take immediate corrective action or the system continues booting.

  Note that for 0048 "Password check failed", the system halts, and then after the next reset/reboot will display the error code on the Error Manager screen.
- **Fatal:** The system halts during post at a blank screen with the text **"Unrecoverable fatal error found. System will not boot until the error is resolved"** and **"Press <F2> to enter setup."** The POST Error Pause option setting in the BIOS setup does not have any effect on this class of error.

  When the operator presses the F2 key on the keyboard, the error message is displayed on the Error Manager screen, and an error is logged to the SEL with the error code. The system cannot boot unless the error is resolved. The user needs to replace the faulty part and restart the system.

Table 74. POST Error Codes and Messages

Error Code	Error Message	Response
0012	System RTC date/time not set	Major
0048	Password check failed	Major
0140	PCI component encountered a PERR error	Major
0141	PCI resource conflict	Major
0146	PCI out of resources error	Major
5220	BIOS Settings reset to default settings	Major
5221	Passwords cleared by jumper	Major
5224	Password clear jumper is Set	Major
8130	Processor 01 disabled	Major
8160	Processor 01 unable to apply microcode update	Major
8170	Processor 01 failed Self-Test (BIST)	Major
8180	Processor 01 microcode update not found	Minor
8190	Watchdog timer failed on last boot	Major
8198	OS boot watchdog timer failure	Major
8300	Baseboard management controller failed self-test	Major
8305	Hot Swap Controller failure	Major
83A0	Management Engine (ME) failed self-test	Major
83A1	Management Engine (ME) Failed to respond.	Major
84F2	Baseboard management controller failed to respond	Major
84F3	Baseboard management controller in update mode	Major

Error Code	Error Message	Response
84F4	Sensor data record empty	Major
84FF	System event log full	Minor
8500	Memory component could not be configured in the selected RAS mode	Major
8501	DIMM Population Error	Major
8520	DIMM_A1 failed test/initialization	Major
8523	DIMM_B1 failed test/initialization	Major
8526	DIMM_C1 failed test/initialization	Major
8529	DIMM_D1 failed test/initialization	Major
852C	DIMM_E1 failed test/initialization	Major
852F	DIMM_F1 failed test/initialization	Major
8540	DIMM_A1 disabled	Major
8543	DIMM_B1 disabled	Major
8546	DIMM_C1 disabled	Major
8549	DIMM_D1 disabled	Major
854C	DIMM_E1 disabled	Major
854F	DIMM_F1 disabled	Major
8560	DIMM_A1 encountered a Serial Presence Detection (SPD) failure	Major
8563	DIMM_B1 encountered a Serial Presence Detection (SPD) failure	Major
8566	DIMM_C1 encountered a Serial Presence Detection (SPD) failure	Major
8569	DIMM_D1 encountered a Serial Presence Detection (SPD) failure	Major
856C	DIMM_E1 encountered a Serial Presence Detection (SPD) failure	Major
856F	DIMM_F1 encountered a Serial Presence Detection (SPD) failure	Major
8604	POST Reclaim of non-critical NVRAM variables	Minor
8605	BIOS Settings are corrupted	Major
8606	NVRAM variable space was corrupted and has been reinitialized	Major
8607	Recovery boot has been initiated. Note: The Primary BIOS image may be corrupted or the system may hang during POST. A BIOS update is required.	Fatal
92A3	Serial port component was not detected	Major
92A9	Serial port component encountered a resource conflict error	Major
A000	TPM device not detected.	Minor
A001	TPM device missing or not responding.	Minor
A002	TPM device failure.	Minor
A003	TPM device failed self-test.	Minor
A100	BIOS ACM Error	Major
A421	PCI component encountered a SERR error	Fatal
A5A0	PCI Express component encountered a PERR error	Minor
A5A1	PCI Express component encountered an SERR error	Fatal
A6A0	DXE Boot Services driver: Not enough memory available to shadow a Legacy Option ROM.	Minor

POST Error Beep Codes

The following table lists POST Error Beep Codes. Prior to system video initialization, the BIOS uses these beep codes to inform users of error conditions. The beep code is followed by a user visible code displayed on the Diagnostic LEDs.

Table 75. POST Error Beep Codes

Beeps	Error Message	POST Progress Code	Description
1	USB device action	N/A	Short beep sounded whenever USB device is discovered in POST, or inserted or removed during runtime.
1 long	Intel® TXT security violation	0xAE, 0xAF	System halted because Intel® Trusted Execution Technology detected a potential violation of system security.
3	Memory error	Multiple	System halted because a fatal error related to the memory was detected.
3 long and 1	CPU mismatch error	0xE5, 0xE6	System halted because a fatal error related to the CPU family/core/cache mismatch was detected.
The following Beep Codes are sounded during BIOS Recovery.			
2	Recovery started	N/A	Recovery boot has been initiated.
4	Recovery failed	N/A	Recovery has failed. This typically happens so quickly after recovery is initiated that it sounds like a 2-4 beep code.
The following Beep Codes are from the BMC.			
1-5-2-1	CPU socket population error	N/A	CPU1 socket is empty, or sockets are populated incorrectly – CPU1 must be populated before CPU2.
1-5-2-4	MSID Mismatch	N/A	MSID mismatch occurs if a processor is installed into a system board that has incompatible power capabilities.
1-5-4-2	Power fault	N/A	DC power unexpectedly lost (power good dropout) – Power unit sensors report power unit failure offset.
1-5-4-4	Power control fault	N/A	Power good assertion timeout – Power unit sensors report soft power control failure offset.
1-5-1-2	VR Watchdog Timer	N/A	VR controller DC power on sequence not completed in time.
1-5-1-4	Power Supply Status	N/A	The system does not power on or unexpectedly powers off and a Power Supply Unit (PSU) is present that is an incompatible model with one or more other PSUs in the system.

Appendix G. Statement of Volatility

This Appendix describes the volatile and non-volatile components on the Intel® Server Board S7200AP Product Family. It is not the intention of this document to include any components not directly on the listed Intel server boards, such as the chassis components, processors, memory, hard drives, or add-in cards.

Server Board Components

Intel® servers contain several components that can be used to store data. A list of components for the Intel® Server Board S7200AP is included in the table below. The sections below the table provide additional information about the fields in this table.

Component Type	Size	Board Location	User Data	Name
Non-Volatile	16 MB	U4E2	No (FW)	Firmware Flash
Non-Volatile	16 MB	U4E1	No (BIOS)	BIOS Flash
Non-Volatile	2 KB	U1N1	No	IFT Carrier Board Flash
Non-Volatile	2 MB	U2F5 and U2F1	No	NIC EEPROM
Volatile	2 GB	U4G1	No	FW SDRAM

Component Type

Three types of components are on an Intel® server board. These types are:

- Non-volatile: Non-volatile memory is persistent, and is not cleared when power is removed from the system. Non-Volatile memory must be erased to clear data. The exact method of clearing these areas varies by the specific component. Some areas are required for normal operation of the server, and clearing these areas may render the server board inoperable
- Volatile: Volatile memory is cleared automatically when power is removed from the system.
- Battery powered RAM: Battery powered RAM is similar to volatile memory, but is powered by a battery on the server board. Data in Battery powered Ram is persistent until the battery is removed from the server board.

Size

The size of each component includes sizes in bits, Kbits, bytes, kilobytes (KB) or megabytes (MB).

Board Location

The physical location of each component is specified in the Board Location column. The board location information corresponds to information on the server board silkscreen.

User Data

The flash components on the server boards do not store user data from the operating system. No operating system level data is retained in any listed components after AC power is removed. The persistence of information written to each component is determined by its type as described in the table.

Each component stores data specific to its function. Some components may contain passwords that provide access to that device's configuration or functionality. These passwords are specific to the device and are unique and unrelated to operating system passwords. The specific components that may contain password data are:

- BIOS: The server board BIOS provides the capability to prevent unauthorized users from configuring BIOS settings when a BIOS password is set. This password is stored in BIOS flash, and is only used to set BIOS configuration access restrictions.

- BMC: The server boards support an Intelligent Platform Management Interface (IPMI) 2.0 conformant baseboard management controller (BMC). The BMC provides health monitoring, alerting and remote power control capabilities for the Intel® server board. The BMC does not have access to operating system level data.

  The BMC supports the capability for remote software to connect over the network and perform health monitoring and power control. This access can be configured to require authentication by a password. If configured, the BMC will maintain user passwords to control this access. These passwords are stored in the BMC flash.

Appendix H. Glossary

This glossary contains important terms used in the preceding chapters. For ease of use, numeric entries are listed first (for example, 82460GX) with alpha entries following (for example, AGP 4x). Acronyms are then entered in their respective place, with non-acronyms following.

Table 76. Glossary

Term	Definition
ACPI	Advanced Configuration and Power Interface
AP	Application Processor
ARP	Address Resolution Protocal
BIOS	Basic Input/Output System
BIST	Built-In Self Test
BMC	Baseboard Management Controller
Bridge	Circuitry connecting one computer bus to another, allowing an agent on one to access the other
BSP	Bootstrap Processor
Byte	8-bit quantity.
CATERR	On a catastrophic hardware event the core signals CATERR to the uncore. The core enters a halted state that can only be exited by a reset.
CMOS	In terms of this specification, this describes the PC-AT compatible region of battery-backed 128 bytes of memory, which normally resides on the server board.
DCMI	Data Center Management Interface
DHCP	Dynamic Host Configuration Protocal
DIMM	Dual In-line Memory Module
EEPROM	Electrically Erasable Programmable Read-Only Memory
EPS	External Product Specification
FRB	Fault Resilient Booting
FRU	Field Replaceable Unit
GB	1024 MB
GPIO	General Purpose I/O
HSC	Hot-Swap Controller
Hz	Hertz (1 cycle/second)
I^2C	Inter-Integrated Circuit Bus
IA	Intel® Architecture
ILM	Independent Loading Mechanism
IMC	Integrated Memory Controller
IP	Internet Protocol
IPMB	Intelligent Platform Management Bus
IPMI	Intelligent Platform Management Interface
KB	1024 bytes
LAN	Local Area Network
LED	Light Emitting Diode
LSB	Least Significant Bit
LUN	Logical Unit Number
MAC	Media Access Control
MB	1024KB

Term	Definition
ME	Management Engine
MCDRAM	Multi Channel Dynamic Random Acess Memory
ms	Milliseconds
MSB	Most Significant Bit
NIC	Network Interface Controller
NMI	Nonmaskable Interrupt
NTB	Non-Transparent Bridge
OEM	Original Equipment Manufacturer
PECI	Platform Environment Control Interface
PEF	Platform Event Filtering
POST	Power-On Self Test
PWM	Pulse-Width Modulation
QPI	QuickPath Interconnect
QSFP+	Quad Small Form-factor Pluggable Plus
RAM	Random Access Memory
ROM	Read Only Memory
RTC	Real-Time Clock (Component of ICH peripheral chip on the server board)
RMM4	Remote Management Module 4
SDR	Sensor Data Record
SEEPROM	Serial Electrically Erasable Programmable Read-Only Memory
SEL	System Event Log
SIO	Server Input/Output
SMBus*	System Management BUS
SMI	Server Management Interrupt (SMI is the highest priority nonmaskable interrupt)
SMM	Server Management Mode
SMS	Server Management Software
SNMP	Simple Network Management Protocol
TDP	Thermal Design Power
TIM	Thermal Interface Material
UART	Universal Asynchronous Receiver/Transmitter
VLSI	Very Large Scale Integration
VMM	Virtual Machine Monitor
VT	Intel® Virtualization Technology
VT-d	Intel® Virtualization Technology for Directed I/O
VT-x	Intel® Virtualization Technology for Intel® 64 and IA-32 Intel® Architecture
WFM	Wired For Management
WHEA	Windows Hardware Error Architecture
WHQL	Microsoft Windows* Hardware Quality Labs
WOL	Wake on LAN – PCI/PCIe capability for installed LAN adapter to “wake” system from sleep or Power Off state on reception of a “magic packet” from the LAN.

Appendix I. Reference Documents

- *Intel® Server System BIOS External Product Specification for Intel® Server Systems supporting the Intel® Xeon® Phi™ processor product family*
- *Intel® Server System BMC Firmware External Product Specification for Intel® Servers Systems supporting the Intel® Xeon® processor E5 V3 product family*
- *Intel® Remote Management Module 4 Technical Product Specification*
- *Intel® Remote Management Module 4 and Integrated BMC Web Console Users Guide*
- *Intel® Chipset C610 product family External Design Specification*
- *SmaRT & CLST Architecture on Intel Systems and Power Supplies Specification*
- *Advanced Configuration and Power Interface Specification, Revision 3.0, http://www.acpi.info/.*
- *Intelligent Platform Management Bus Communications Protocol Specification, Version 1.0. 1998. Intel Corporation, Hewlett-Packard Company, NEC Corporation, Dell Computer Corporation.*
- *Intelligent Platform Management Interface Specification, Version 2.0. 2004. Intel Corporation, Hewlett-Packard Company, NEC Corporation, Dell Computer Corporation.*
- *Platform Support for Serial-over-LAN (SOL), TMode, and Terminal Mode External Architecture Specification, Version 1.1, 02/01/02, Intel Corporation.*
- *Alert Standard Format (ASF) Specification, Version 2.0, 23 April 2003, ©2000-2003, Distributed Management Task Force, Inc., http://www.dmtf.org.*
- *Knights Landing Processor - External Design Specification (EDS) - Volume One: Architecture, Doc ID # 547141*
- *Knights Landing Processor - External Design Specification (EDS) - Volume Two: Registers, Part A, Doc ID # 546888*
- *Knights Landing Processor - External Design Specification (EDS) - Volume Two: Registers, Part B, Doc ID # 552048*
- *Knights Landing Processor - External Design Specification (EDS) - Volume Three: Electrical Specification, Doc ID # 540624*
- *Knights Landing Processor Mechanical Retention Assembly Models, Doc ID # 542094*
- *Knights Landing and Knights Landing with Fabric Processor Thermal Models' Doc ID # 542093*
- *Knights Landing with Fabric Thermal Test Vehicle Users Guide, Doc ID # 548231*
- *Intel®IA64 IA32 Architecture Software Developers Manual' Doc ID # 325462*
- *Intel® Fabric Passive (IFP) Internal Cable Assembly Design Specification, Doc ID # 548740*
- *Intel®Fabric Through (IFT) Connector Assembly Design Specification, Doc ID # 549649*