Open Compute Carrier-grade Edge Reference Architecture (OCCERA)

MCN-2600 Compute Node Specification (Short Form)

Rev. 0.2
May 9, 2017

LRI Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000
Leading EDGE COMPUTING

Revision History

Version	Date	Author	Changes
Rev. 0.1	2016-04-18	ADLINK Technology	Preliminary draft
Rev. 0.2	2017-05-09	ADLINK Technology	Update architecture name to OCCERA

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

Contents

Revision History 2
Figures 5
Tables 6
1 Scope 7
2 Overview 8
3 MCN-2600 Features 9
3.1 Block Diagram 9
3.2 Placement and Form Factor 10
3.3 Processor and Memory 11
3.4 Platform Controller Hub 11
4 Basic Input Output System (BIOS) 12
4.1 SPI Flash 12
4.2 BIOS Setup Menu 13
4.3 BIOS Boot Option 13
4.4 OS Support 13
5 Baseboard Management Controller (BMC) 14
5.1 LAN Interface 14
5.2 IPMB Interface 14
5.3 Serial Over LAN (SOL) 14
5.4 Remote Power Control 14
5.5 Power Monitoring and Power Limiting 15
5.6 Fan Monitoring and Control 15
5.7 System Event Log (SEL) 15
6 I/O System 16
6.1 PCIe Interfaces 16
6.2 LAN Bypass 16
6.3 Network 17
6.4 USB 17
6.5 SATA 17
6.6 VGA Interface 17
6.7 Serial Console 17
6.8 Power and Reset Button 18
6.9 Power LED 18
6.10 OOS LED 18
6.11 Health LED 18
6.12 Application LED 18
7 Power and Thermal Systems 19

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

7.1 Power Supply Input 19
7.2 Thermal Specifications 19
8 Mechanical 21
8.1 Layout Top View 21
8.2 Layout Bottom View 22
9 Environmental Specification 23
9.1 Vibration and Shock 23
Appendix MCN-2600T Sled 24

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

Figures

Figure 1: Functional block diagram of MCN-2600 9
Figure 2: Placement of the motherboard MCN-2600 10
Figure 3: High speed I/O map of the PCH 11
Figure 4: SPI flash selection 12
Figure 5: Main power dissipation components on motherboard 19
Figure 6: Motherboard layout top view 21
Figure 7: Motherboard layout top view 22
Figure 8: 3D view of MCN-2600T compute node sled 24

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

Tables

Table 1: Configuration options for LAN bypass setting based on host power status 16
Table 2: Thermal specifications of the key components on the motherboard 20
Table 3: Vibration and shock capability 23
Table 4: MCN-2600T specifications 24

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

1 Scope

This document outlines the technical specifications for the MCN-2600 compute node motherboard based on the Open Compute Carrier-grade Edge Reference Architecture (OCCERA).

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

2 Overview

With the recent technical advancements of the mobile Internet, data centers and big data processing, cloud computing is becoming the mainstream service delivery model of the future. A fundamental requirement of cloud computing is "resources on demand", meaning that computing, storage and IO resources can be combined flexibly to meet an application's requirements. This requires the computing, storage and IO to be decoupled into different functional modules in the early design phase. Then, based on the requirements of a concrete application, these independent functional modules are chosen in the necessary proportions, and then combined into a specific hardware entity. The Open Compute Carrier-grade Edge Reference Architecture (OCCERA) was created to realize this concept.

The three key characteristics of OCCERA are implied in its name: modular design, industrial features, and virtualization support building a next generation cloud computing infrastructure. One of the key components of OCCERA is a modularized compute node motherboard. This document describes the MCN-2600 compute node motherboard, which has integrated the latest technologies from Intel architecture processors, particularly the latest advancements in NFV. As a result, the MCN-2600 can bring high performance to the cloud while meeting the industrial computing requirements for flexible expansion, hardware acceleration, and high availability.

In this document, the MCN-2600's features will first be introduced (Section 3), and then BIOS and BMC features implemented on the MCN-2600 are reviewed (Sections 4, 5). Section 6 concentrates on the expansion connectors provided by the MCN-2600. Section 7 introduces the I/O options that can be achieved based on the MCN-2600. Section 8 introduces the switches and LEDs provided by this MCN-2600. Section 9 introduces the MCN-2600's power and thermal specifications. Sections 10 and 11 present an overview of the MCN-2600's mechanical and environmental specifications, respectively. The Appendix presents an overview of the MCN-2600T, which is a compute node sled built on the MCN-2600 motherboard.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

3 MCN-2600 Features

3.1 Block Diagram

The figure below illustrates the functional block diagram of the MCN-2600 motherboard. The motherboard features dual Intel® Xeon® processor E5 and its external IOs are routed to both edges of the motherboard for flexible expansion.

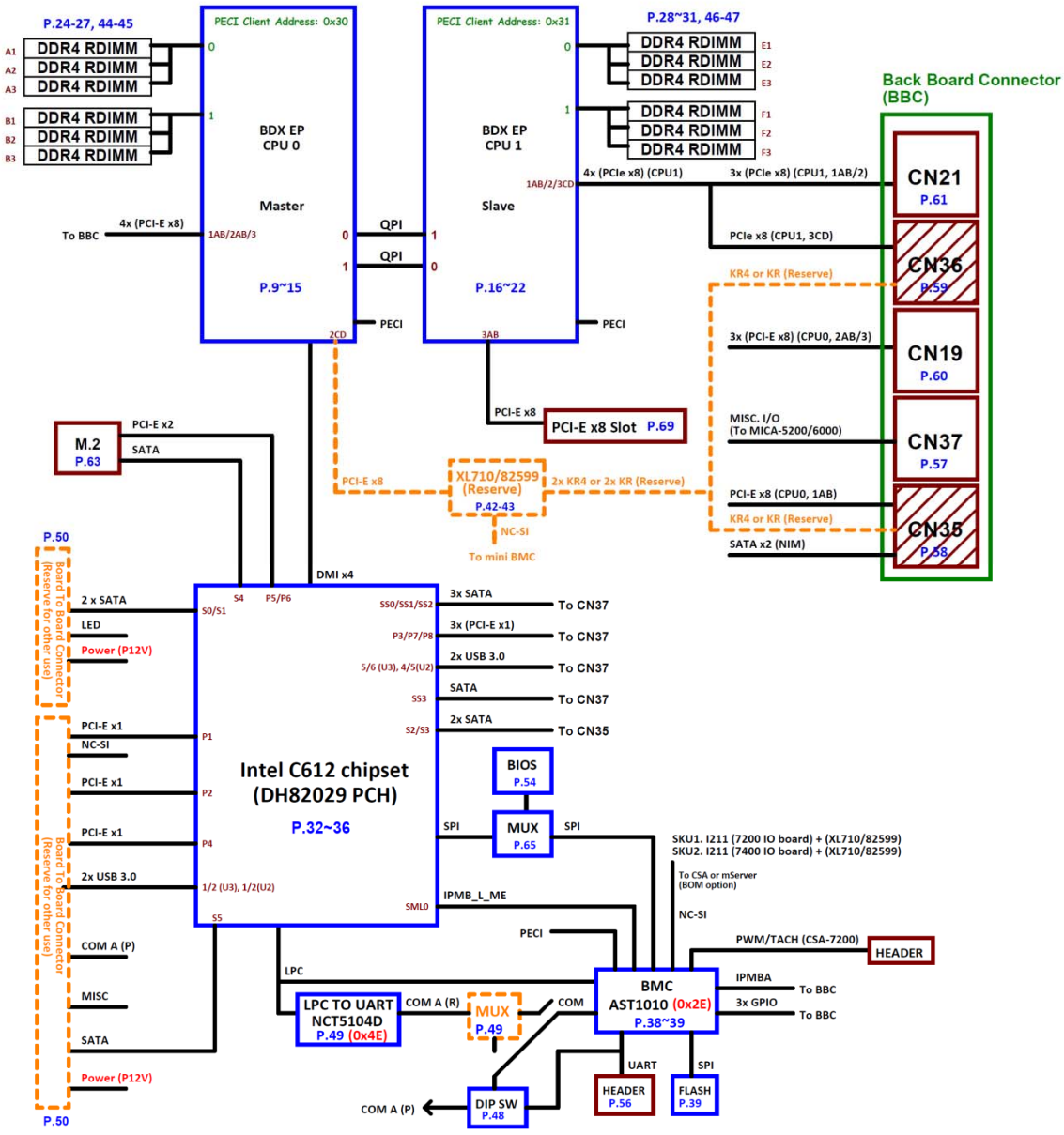
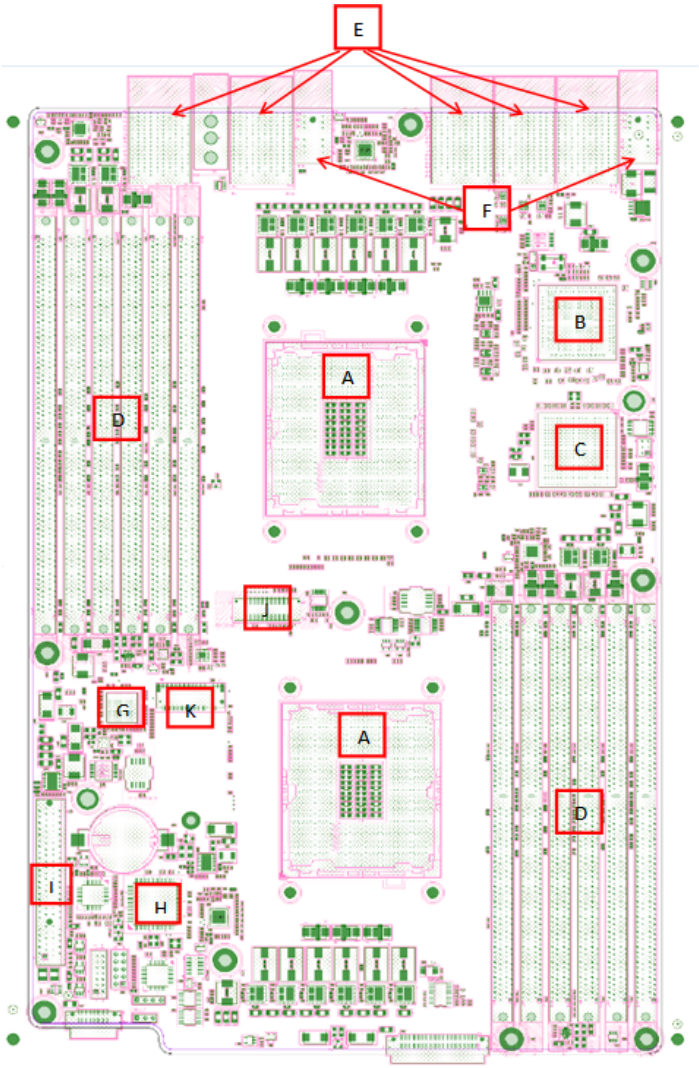


Figure 1: Functional block diagram of MCN-2600

LRI Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

3.2 Placement and Form Factor

The motherboard's form factor is 320mm x 204mm. Figure 2 illustrates the board placement. The placement shows the relative positions of key components. Exact dimensions and positioning information can be found in the mechanical drawing.



A	LGA-2011 R3 socket	**B**	Intel C612 chipset
C	Intel 82599 (optional)	**D**	DDR4 socket
E	Backplane connector	**F**	Power connector
G	BMC	**H**	CPLD
I	PCIe x8 slot	**J**	XDP connector
K	M.2 connector		

Figure 2: Placement of the motherboard MCN-2600

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

3.3 Processor and Memory

The motherboard uses Intel “Haswell/Broadwell” processors, which are 64-bit, multi-core processors built on 22/14-nanometer technology (Haswell/Broadwell respectively). The processors are designed to be paired with the C612 chipset on the Intel “Grantley” platform.

The processor supports the following technologies

- On-package cache memory integrated within Haswell/Broadwell-EP multi-core processor
- Intel® Turbo Boost Technology
- 2 socket Intel Haswell/Broadwell processors
- Single socket installation
- Direct attached DDR4 memory
- Six DDR4-RDIMM per processor
- Up to DDR4-2400 MHz for Broadwell-EP series

3.4 Platform Controller Hub

The motherboard uses the new generation Intel Platform Controller Hub (PCH), which provides extensive I/O support. The PCH supports flexible IO, a new architecture to allow some high speed IO signals to be configured as PCIe, SATA or USB 3.0, as shown in Figure 3.



Figure 3: High speed I/O map of the PCH

The main functions and capabilities of the PCH include:

- PCI Express Base Specification, Revision 2.0 support for up to eight ports with transfers up to 5 GT/s
- Integrated Serial ATA host controllers with independent DMA operation on up to ten ports
- USB controller provides support for up to 14 USB ports, of which six can be configured as SuperSpeed USB 3.0 ports
- Two legacy EHCI USB controllers
- Integrated 10/100/1000 Gigabit Ethernet MAC with System Defense
- System Management Bus (SMBus) Specification, Version 2.0 with additional support for I2C devices
- Supports Intel® Active Management Technology 9.x (Intel® AMT 9.x)
- Supports Intel® Rapid Storage Technology (Intel® RST)
- Integrated Clock Controller

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

4 Basic Input Output System (BIOS)

The BIOS firmware is based on the American Megatrends Inc. (AMI) EFI solution, and it can be customized to support different motherboard IO extensions. The BIOS is also customized to improve the power efficiency. The measures include:

- CPU Power Technology set to "Energy Efficient"
- CPU EIST enabled

4.1 SPI Flash

The BIOS firmware and configurations are stored in a SPI flash chip. The BIOS SPI flash controller supports the Serial Flash Discoverable Parameter (SFDP) JEDEC standard, which provides a consistent way of describing the functional and feature capabilities of serial flash devices in a standard set of internal parameter tables. The SPI flash device is up to 16 MB (128 Mb) in capacity.

There are two ways to update the BIOS firmware: via PCH or BMC. This is implemented by using an external bus switch, which can connect the SPI flash to the PCH or BMC, as shown in Figure 4.

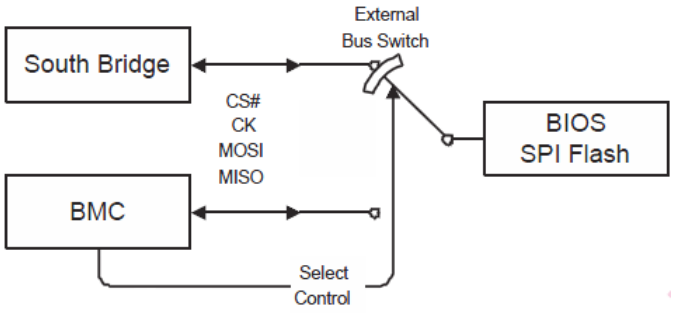


Figure 4: SPI flash selection

When updating the BIOS from the PCH, the AMI AFU tool is used, and it has different versions to support DOS, Windows, Linux, and UEFI environment. When updating the BIOS from the BMC, the AMI BMC Yafuflash tool is used, and it can perform the update using the BMC LAN or mailbox interface.

Update BIOS via BMC LAN interface:

Yafuflash -nw -ip <ip addr> -u root -p superuser -bios <BIOS image>

Update BIOS via BMC mailbox interface: (must be root user)

Yafuflash -mbox -bios <BIOS image>

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

4.2 BIOS Setup Menu

The BIOS setup menu provides the following configuration options.

- **Main BIOS Setup Screen**: BIOS Vendor, Core Version, Complied URFA Specification Version, Motherboard Name, BIOS Build Time, and settings for changing System Language, System Date, and System Time.
- **Advanced BIOS Setup Screen:** Board Automatic Power-up Configuration, ACPI Configuration, Super IO Configuration, Serial Port Console Redirection, iSCSI Configuration, PCIe SR-IOV Support, PXE boots Configuration, Compatibility Support Module Configuration, BMC Watchdog Configuration, and LAN Bypass Configuration.
- **Intel RefCode Setup Screen**: Processor Configuration Advanced Power Management Configuration, Common RefCode Configuration, Memory Configuration, IIO Configuration, PCH Configuration.
- **Server Management Setup Screen:** BMC Firmware Revision, and BMC Watchdog Configuration.
- **Security Setup Screen:** Administered Password, User Password, and HDD Security Configuration.
- **Boot Setup Screen:** Setup Prompt Timeout, Bootup NumLock State, Quiet Boot Configuration, and Boot Option Priorities.

4.3 BIOS Boot Option

The BIOS supports Legacy/UEFI PXE boot, and UEFI PXE supports both IPv4 and IPv6 environment. The BIOS supports UEFI iSCSI boot.

The default BIOS boot option priorities are

- 1st: UEFI storage devices(HDD/USB)
- 2nd: UEFI network devices
- 3rd: Legacy storage devices(HDD/USB)
- 4th: Legacy network devices
- 5th: UEFI: Built-in Shell

4.4 OS Support

The BIOS supports to run the following operation systems on the host.

- CentOS 6.3, 7.x
- Fedora™ latest version
- Ubuntu latest version
- Microsoft® Windows® 7 Ultimate bit version

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

5 Baseboard Management Controller (BMC)

The motherboard uses an AST1010 chip from Aspeed as the BMC. The BMC works as a standalone system, providing out-of-band IPMI management service that is not dependent on the health status of the host.

The BMC firmware can be configured to work in a single-board system or work together with a Chassis Management Module (CMM) in a multiple-board system. In a single-board system, the BMC is responsible for monitoring the thermal status of the board and control the fan speed adaptively. In a multiple-board system, the BMC is responsible for reporting the thermal status of the board to a CMM via the IPMB bus, and the CMM is responsible for controlling the fan speed.

5.1 LAN Interface

The BMC provides LAN accessing via the NC-SI interface built on the host network cards. The host network cards are be powered by standby power and are able to function even if the host is powered down (main power is turned off).

The motherboard has two network cards, one for control plane usage and soldered onboard by default, and the other for data plane usage and soldered onboard by request. If both network cards are soldered, the user can select which network card is used to implement IPMI over LAN by sending an OEM command to BMC.

5.2 IPMB Interface

When the MCN-2600 is used in a multiple-board system, an IPMB must be implemented in the chassis backplane. The IPMB is an I^2C-based serial bus that is routed between BMCs and CMM. The IPMB enables the BMC to accept IPMI requests and send IPMI responses from/to a CMM.

5.3 Serial Over LAN (SOL)

Because a GUI does not scale, the BMC does not support the graphic-based iKVM feature. Instead, the BMC provides a text-based Serial Over LAN (SOL) feature. With SOL redirection, system administrators can remotely view the text-based host console from anywhere and perform any task that doesn't require a GUI.

5.4 Remote Power Control

The BMC supports power on/off/cycle and cold/warm reset of the host system through out-of-band IPMI commands. Users can execute a host recovery action by triggering the power control regardless of the host operating status.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

5.5 Power Monitoring and Power Limiting

The BMC supports power monitoring through the hot-swap controller, and a current sensor is implemented to provide real-time power consumption of the motherboard. A series of voltage sensors are also implemented to monitor the key power rails on the motherboard.

The BMC also enables the Intel® Node Manager power limiting function. The user can enforce a power limit policy by bridging Intel® Node Manager commands to the Intel® Management Engine (ME) via BMC.

5.6 Fan Monitoring and Control

The BMC implements a series of temperature sensors to monitor the thermal status of key components of the motherboard, including inlet air temperature, CPU, DIMM and PCH.

When the BMC works in a single board system, based on the aggregated thermal situation, the BMC will run the fan based on an adaptive speeding algorithm

The BMC also implements tachometers to monitor the fan speed.

5.7 System Event Log (SEL)

The BMC supports a Sensor Event Log (SEL) with a size of 14 KB. The BMC uses a circular SEL type, and the SEL repository acts like a ring buffer. When the repository becomes full, the oldest logs are overwritten by new logs.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

6 I/O System

This section outlines the I/O options that can be implemented on the motherboard through an extension IO board.

6.1 PCIe Interfaces

The motherboard can provide nine groups of PCIe x8 Gen 3 interfaces for customization usage. Eight of them are routed to the midplane signal connectors CN21, CN35 and CN36, and one is provided via the onboard PCIe x8 expansion slot. The PCIe x8 interfaces can be used to expand the system with more network interface cards.

The motherboard also provides six groups of PCIe x1 interfaces for IO extension which are available on edge midplane signal connector CN19 (x3) and edge IO board connector CNX2 (x3). The PCIe x1 interfaces can be used to expand the system with 1G network cards or a VGA interface.

6.2 LAN Bypass

For the network interface cards expanded through the midplane PCIe x8 interfaces, a LAN bypass function is supported to allow uninterrupted network traffic for replaceable network interface cards in the case of power disruption or system failure. Under the following conditions, the BMC controls GPIO signals to trigger the LAN bypass mechanism:

- System crash occurs
- System is powered on but not running OS
- After booting into OS
- Power turned off or disconnected

The motherboard supports the configuration of the LAN bypass setting based on host power status in a BIOS setup screen, as shown in Table 1. The motherboard also allows changing current LAN bypass status via an IPMI interface.

Table 1: Configuration options for LAN bypass setting based on host power status

Host Power Status / **Configuration**		**Power On**	**Power Off**
Bypass Globally Disabled		Bypass Off	Bypass Off
Bypass Globally Enabled	**Bypass ON**	On boot, BIOS sets Bypass On	Bypass On
	Bypass OFF	On boot, BIOS sets Bypass Off	Bypass Off

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

6.3 Network

The motherboard provides several options to expand network interfaces. For control plane interfaces that only need 1G bandwidth, the six groups of PCIe x1 interfaces can be used. For data plane interfaces that need a more rich set of network interfaces, the eight groups of PCIe x8 interfaces can be used to expand with at most eight network interface cards with built-in LAN bypass support.

The motherboard also reserves placement for one Intel 82599 or XL710 chip, which can provide 2x 10G or 2x 40G interfaces on the midplane. When the Intel 82599 or XL710 chip is installed, the two KR4/KR interfaces are routed to the midplane connectors CN35 and CN36, respectively.

6.4 USB

The motherboard supports four USB3.0 interfaces that are compatible with USB 2.0. Two USB interfaces are routed to the edge midplane signal connector CN37, and the other two USB interfaces are routed to the edge IO board connector CNX2.

6.5 SATA

The motherboard supports ten SATA 6Gb/s (Gen 3) ports. One SATA port is connected to the M.2 connector for onboard storage expansion. On the midplane edge, four SATA ports are routed to the midplane connector CN37, and two SATA ports are routed to the midplane connector CN35. On the other board edge, two SATA ports are routed to the SATA connector, and the remaining SATA port is routed to the IO board connector. For a concrete application scenario, the user can select which SATA ports are implemented based on need.

6.6 VGA Interface

The motherboard does not have a built-in VGA card, but a VGA interface can be implemented using the PCIe lanes either on the edge midplane connector or the edge IO board connector.

6.7 Serial Console

The motherboard supports one serial console port, which is shared between the host and the BMC. Selection pins are available on both CN37 and CNX2 that can be used to implement a DIP switch, controlling serial console usage.

When the serial console is directed to the host, it provides a text based console port to the host processor. When the serial console is directed to the BMC, the user can log into the BMC shell, execute IPMI shell commands, and check IPMI debugging logs.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

6.8 Power and Reset Button

The GPIOs for implementing power and reset buttons are reserved on edge connectors CN37 and CNX2. The power and reset buttons can be implemented on either the midplane side or opposite side.

6.9 Power LED

The power LED is controlled by hardware directly, available on edge connectors CN37 and CNX2. The LED is on when main power is enabled, and the LED is off when main power is disabled.

6.10 OOS LED

The Out Of Service (OOS) LED is controlled by the BMC and available on edge connectors CN37 and CNX2. By default, the OOS LED indicates BIOS POST status, the OOS LED is turned on once the BMC detects that payload power is enabled. The OOS LED is turned off only when the BMC receives a message telling it that BIOS POST is completed. The BMC also implements an OEM command to allow the user to control the OOS LED via IPMI command.

6.11 Health LED

The health LED is controlled by the BMC and available on edge connectors CN37 and CNX2. By default the health LED indicates system status after BIOS POST. The health LED is turned on once BIOS POST is completed. BMC watchdog timeout will turn off the LED. If the BMC watchdog is not enabled, the green LED is always on. The BMC also implements an OEM command to allow the user to control the health LED via IPMI command.

6.12 Application LED

The application LED is controlled by the BMC and available on edge connectors CN37 and CNX2. By default the application indicates the LAN bypass status. There are at most 8 network cards that can be controlled via the LAN bypass function. If LAN bypass is enabled on any network card, the LED is on. When LAN bypass is disabled on all network cards, the LED is off. The BMC also implements an OEM command to allow the user to control the application LED via IPMI command.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

7 Power and Thermal Systems

The motherboard takes power from two edge midplane power connectors CN25 and CN29.

7.1 Power Supply Input

The nominal input voltage delivered by the power supply is 12 volts DC. The required current supply is based on the total power consumption of the motherboard plus extra extension IO boards.

7.2 Thermal Specifications

The following figure illustrates main power dissipation components on the motherboard.

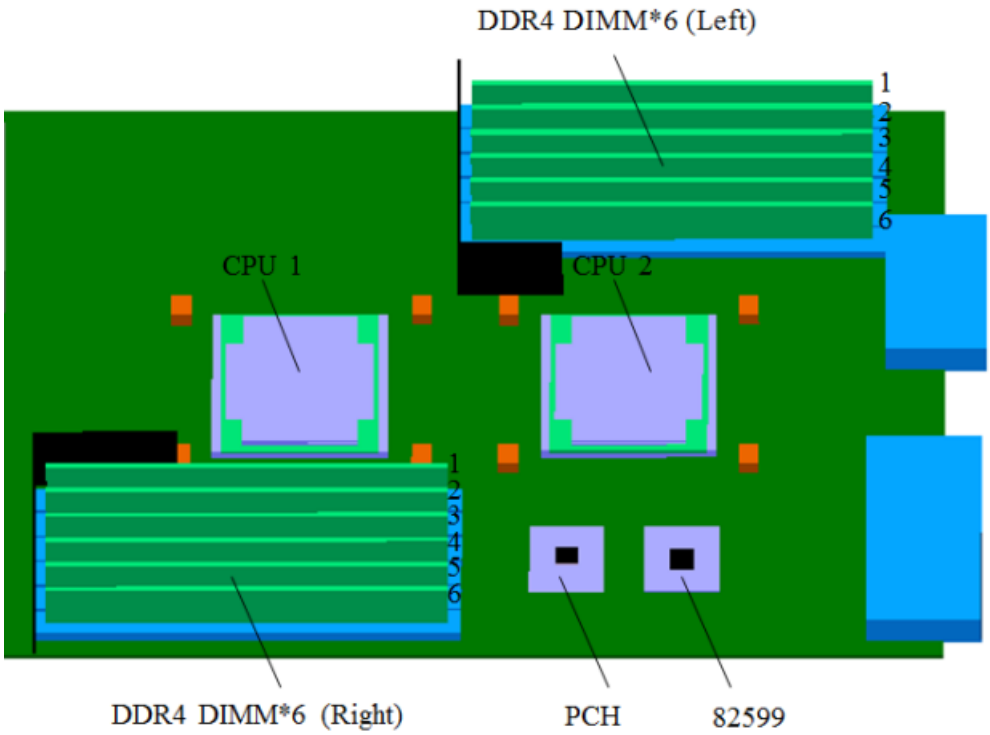


Figure 5: Main power dissipation components on motherboard

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

The thermal specifications of the key components on the motherboard, including the CPU, PCH, 82599 and DDR4 RDIMM, are listed in Table 2.

Table 2: Thermal specifications of the key components on the motherboard

Item	Power (W)	Quantity	Total (W)	Spec (°C)
CPU	120	2	240	84.5
PCH	4.0	1	4.0	100
82599	7.8	1	7.8	119
DDR4 RDIMM	3.5	12	42	85
Total			293.8	

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

8 Mechanical

The motherboard PCB thickness is 2.4mm ±5%, and PCB dimension is 320mm x204mm.

8.1 Layout Top View

The following figure depicts the top view of the motherboard layout.

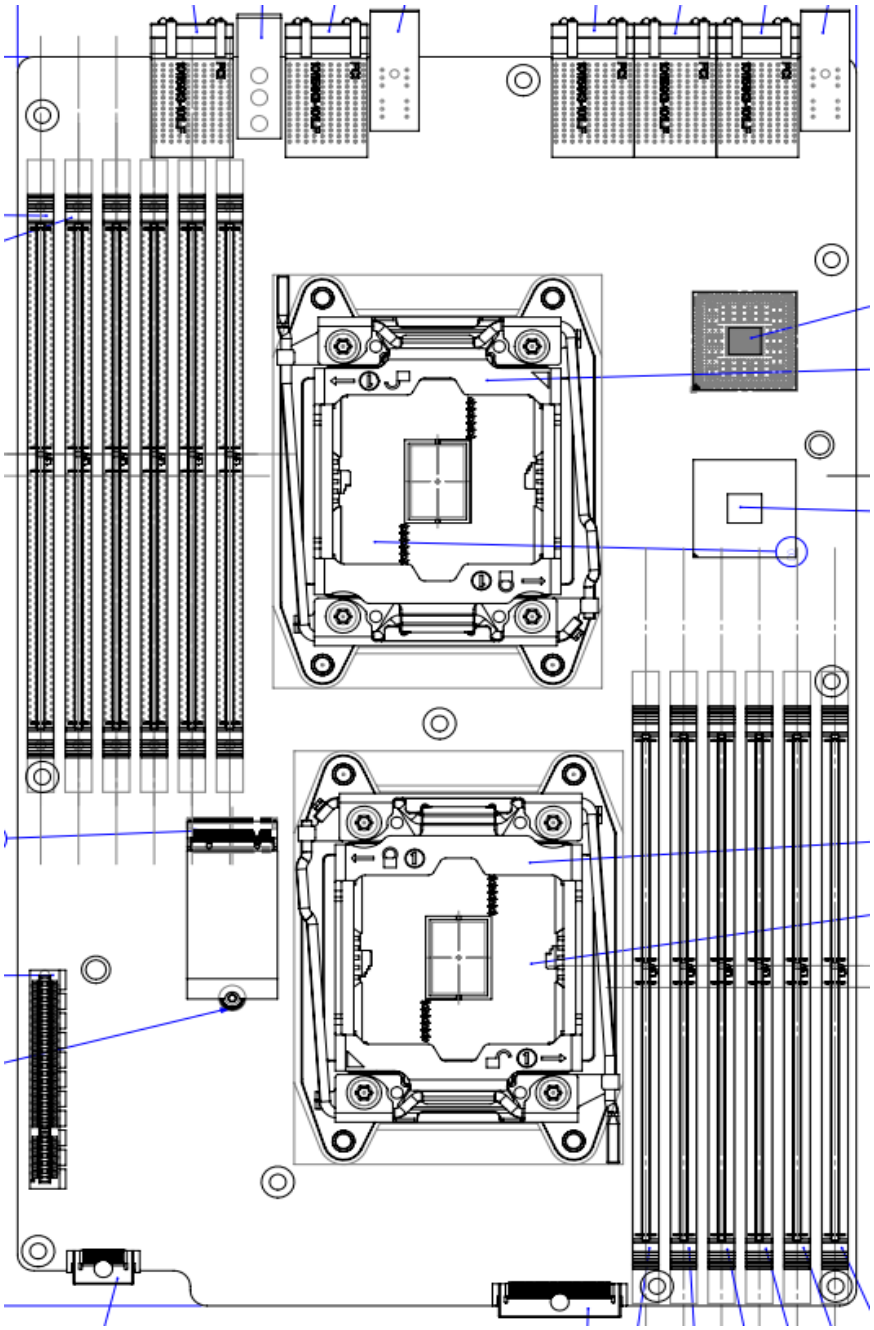


Figure 6: Motherboard layout top view

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

8.2 Layout Bottom View

The following figure depicts the bottom view of the motherboard layout.

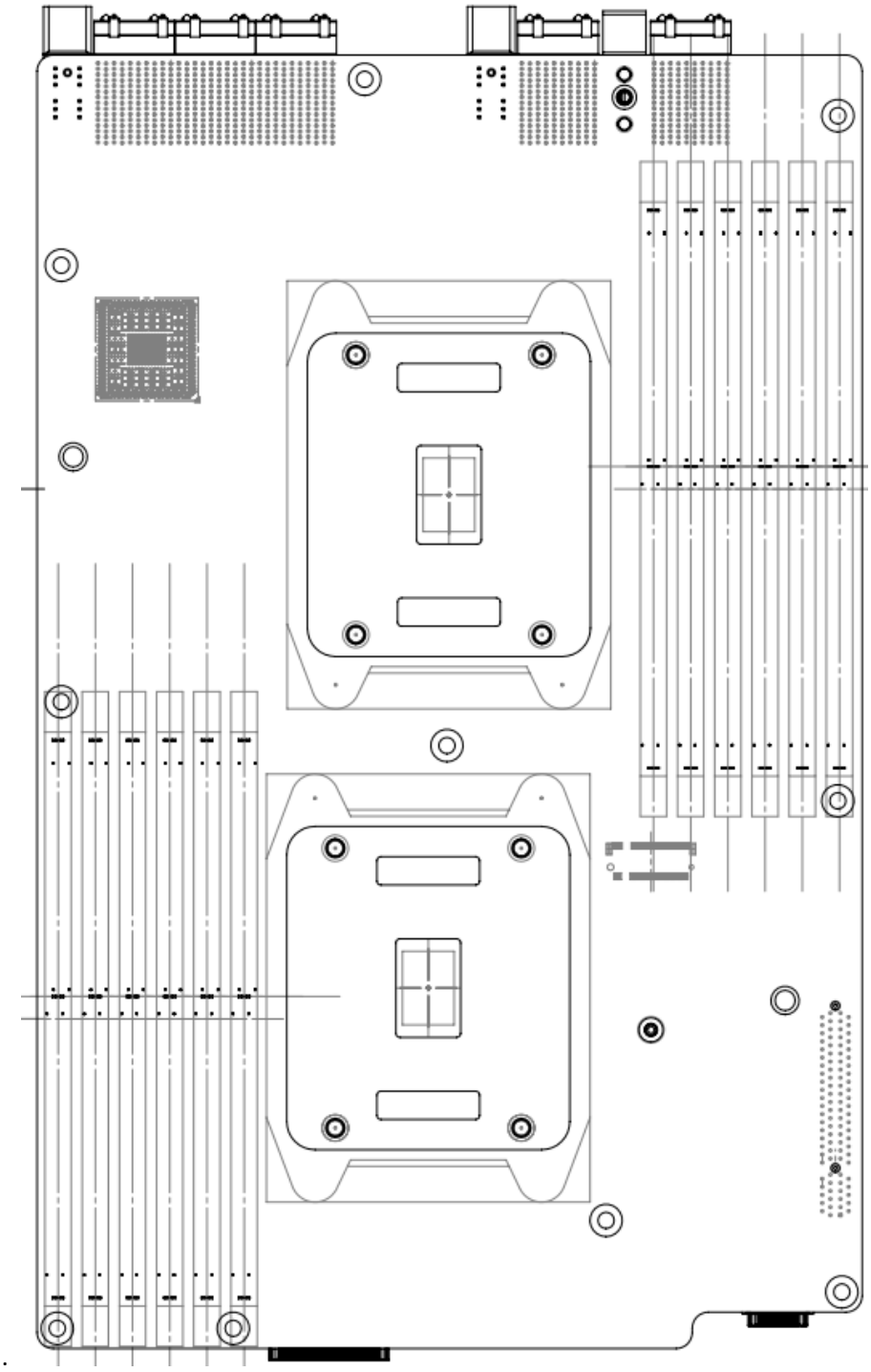


Figure 7: Motherboard layout top view

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

9 Environmental Specification

The motherboard is designed to meet the following environmental requirements:

- Operating temperature: 0 °C to +40 °C,
- Short term operating: -5°C to +45°C, 5%-90% relative humidity
- Storage temperature: -40 °C to +70 °C
- Operating humidity: 20% to 90%RH@40°C, non-condensing
- Non-operating: 5%-90% relative humidity
- EMI: NEBS Level-3 compliance, FCC class A, CE emission, CCC, UL, CB and RoHS compliant
- MTBF: 50,000 hours.

9.1 Vibration and Shock

The motherboard is designed to meet the shock and vibration requirements outlined in the following table.

Table 3: Vibration and shock capability

	Operating	**Non-operating**
Vibration	N/A	2.2Grms, 10 minutes per axis on all three axes
Shock	Half-sine 2G, 11ms pulse, 100 pulses in each direction, on each of the three axes	Trapezoidal, 25G, 170 inches/sec delta V, three drops in each direction, on each of the three axes.

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

Appendix MCN-2600T Sled

The MCN-2600T (Figure 8) is a compute node sled built on the MCN-2600 motherboard. The compute node sled is used in ADLINK OCCERA products MCS-2080 and CSA-7400.

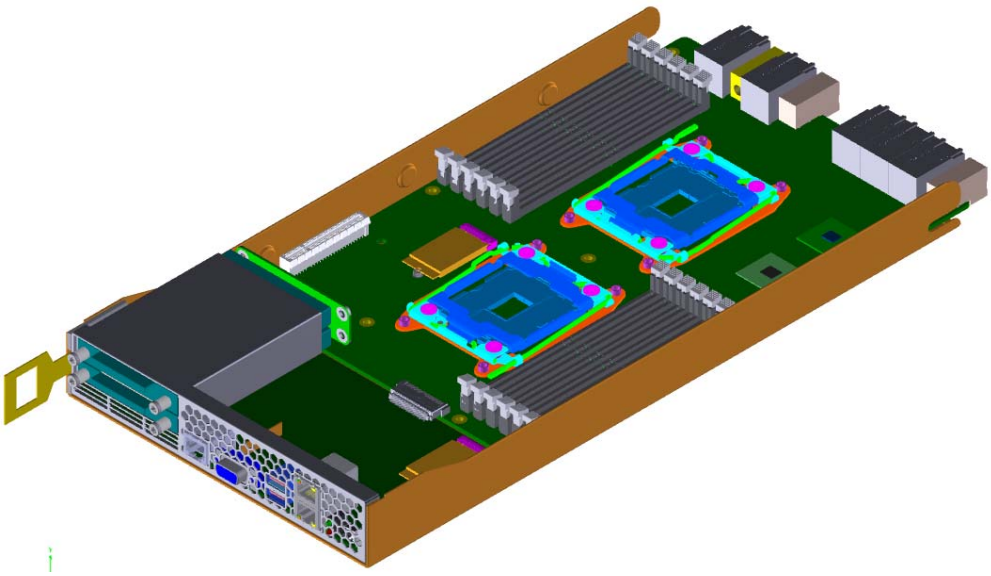


Figure 8: 3D view of MCN-2600T compute node sled

The specifications of the MCN-2600T are listed in Table 4

Table 4: MCN-2600T specifications

CPU/Chipset/Memory	**CPU**	Intel Xeon E5-2699 v3(18C/36T) Intel Xeon E5-2680 v3(12C/24T) Intel Xeon E5-2630 v3(8C/16T) Intel Xeon E5-2620 v3(6C/12T)
	L2 Cache	15MB/20MB/25MB/30MB depending on CPU
	Chipset	Intel C610 Chipset
	Memory	12x DDR4-1600/1866/2133 DIMM socket, ECC, up to 192GB
I/O Interface	**Ethernet**	2x RJ-45 10/100/1000BASE-T Ethernet ports
	Console	1x RJ-45 serial port
	USB	2x USB 3.0
	Graphics	1x VGA socket onboard
	HDD/SSD	2x 2.5'' Hot-swappable SATA drive bays & 2x M.2 SSD
	Backplane	2x PCIe G3 x 8 to Switch nodes
Control Buttons	**Power**	1x Power button(front)
	Reset	1x Reset Button(front)

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000

Other	**LEDs**	Power, Status, HDD
	Operating System	Window 7 64bit, Linux Kernel 2.6 or above
Mechanical	**Form Factor**	213.9mmx 40.44mm x 432.1mm (W x H x D)

Distribuída no Brasil pela LRI | vendas@lri.com.br | 51 2129-3000