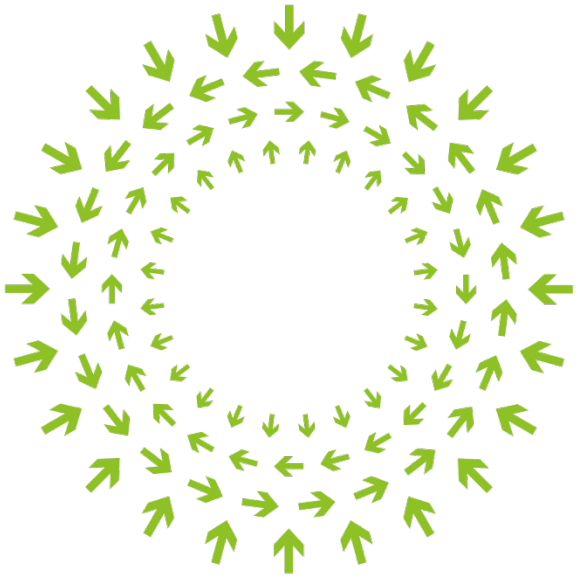

OPEN

Compute Project

Micro-Server Card Hardware v1.0 MB-draco-etamin-0.5

Author: Chris Petersen, Hardware Systems Engineer, Facebook

1 Scope

This specification provides a common form factor for emerging micro-server and SOC (System-On-Chip) server designs. For the purposes of this specification, the term "card" is used to describe a PCIe-like card that hosts the SOC, dynamic memory for the SOC, and a storage device.

2 Contents

3 Overview

When data center design and hardware design move in concert, they can improve efficiency and reduce power consumption. To this end, the Open Compute Project is a set of technologies that reduces energy consumption and cost, increases reliability and choice in the marketplace, and simplifies operations and maintenance. One key objective is openness—the project is starting with the opening of the specifications and mechanical designs for the major components of a data center, and the efficiency results achieved at facilities using Open Compute technologies.

One component of this project is a micro-server or SOC (System-On-Chip) server design. The micro-server is a PCIe-like card that hosts the SOC, dynamic memory for the SOC, and a storage device. This micro-server can be installed in slots on a baseboard. The baseboard provides power distribution and control, BMC management capabilities, and network distribution. The baseboard may or may not be plugged into a midplane.

3.1 License

4 Mechanical

4.1 Card Mechanical Outline

The x16 version of the card provides significantly more connections and increases the power capacity to address both high connectivity solutions and high power SOCs. The cards adhere to the mechanical dimensions provided in Sections 4.1.1 and 4.1.2. The dimensions are defined to accommodate multiple card lengths, but only on one end of the card.

4.1.1 X16 Card Design

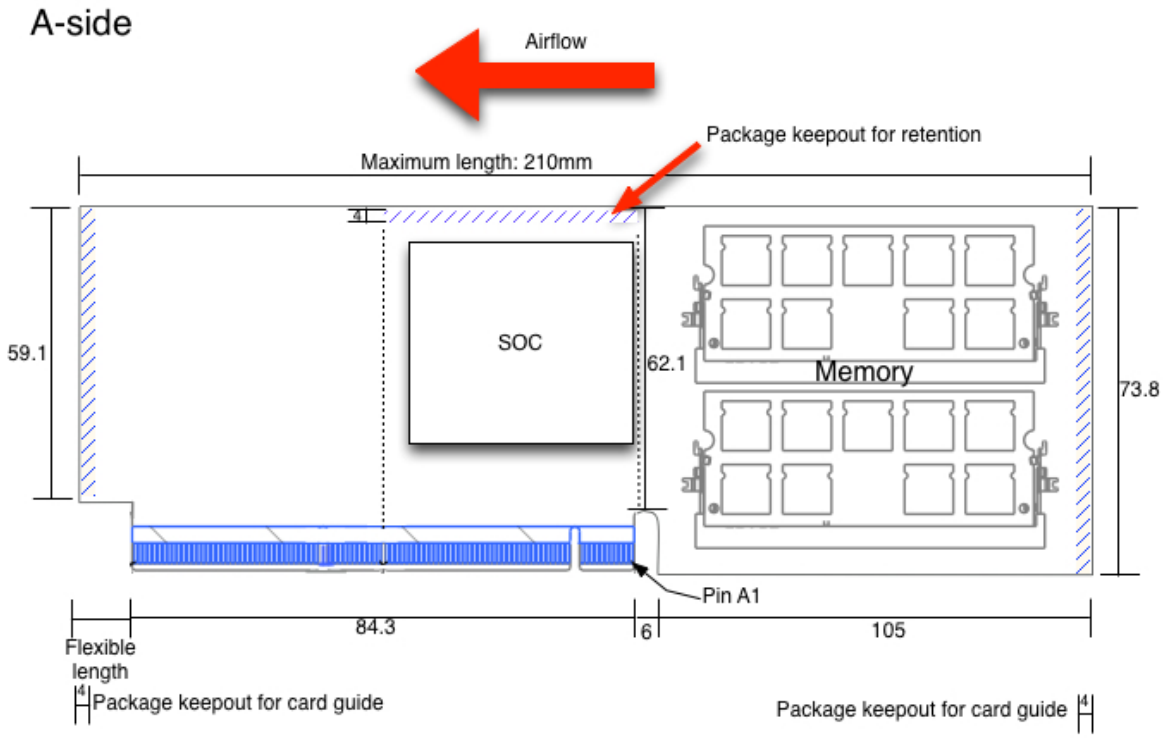

Figure 4-1 Mechanical Outline of the "A" Side of the x16 Card

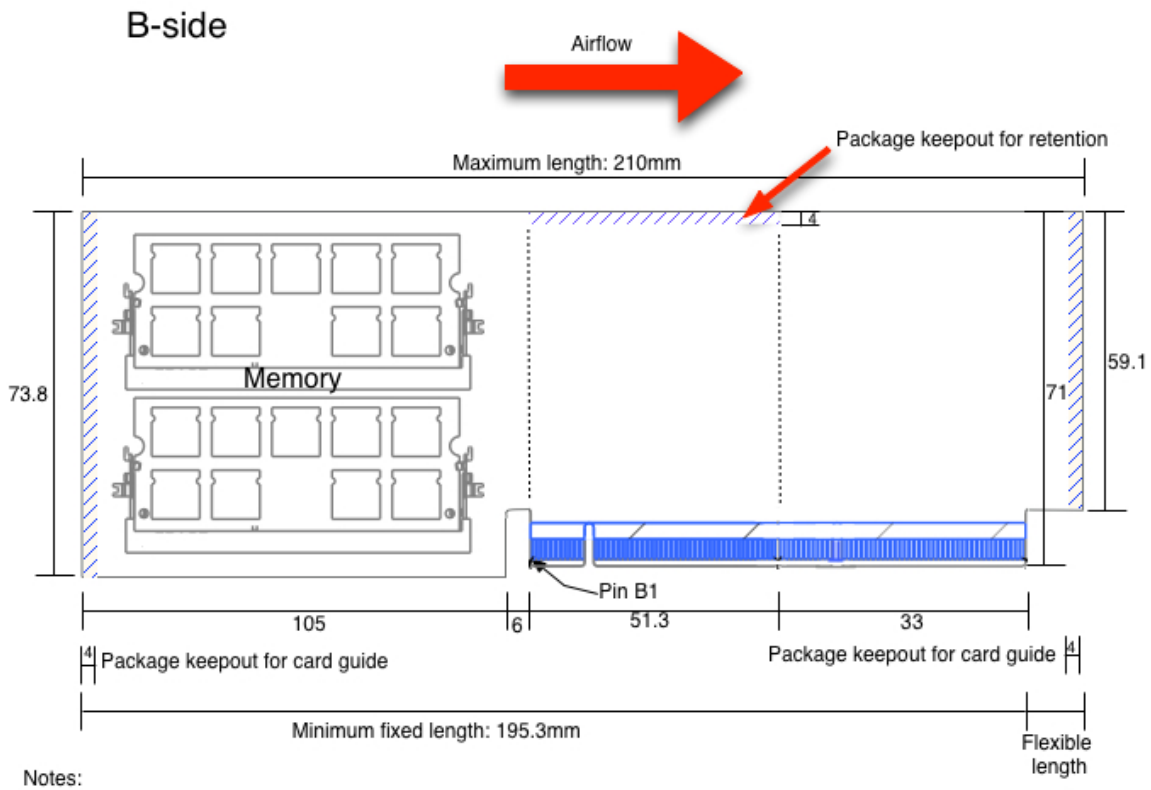

Figure 4-2 Mechanical Outline of the "B" Side of the x16 Card

4.2 Card Mechanical cross-section

The x16 cards must not exceed the single-width dimensions (see Figure 4-5) to enable high-density designs. It is strongly recommended that heatsinks are as tall as possible to minimize air bypass between card slots.

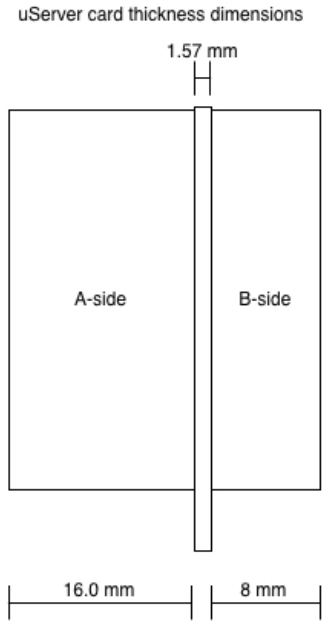

Figure 4-3 Mechanical Dimensions Component Heights Including Tolerance for Single-width cards

4.3 Card retention and removal

In addition to the 4mm keep-outs along the edges of the card, a 51.3m x 4mm keep-out is defined directly above the PCIe edge connector. This keep-out also includes two small holes that can be used to attach tool to ease insertion and extraction of the cards. See Figure 4-7 for more details.

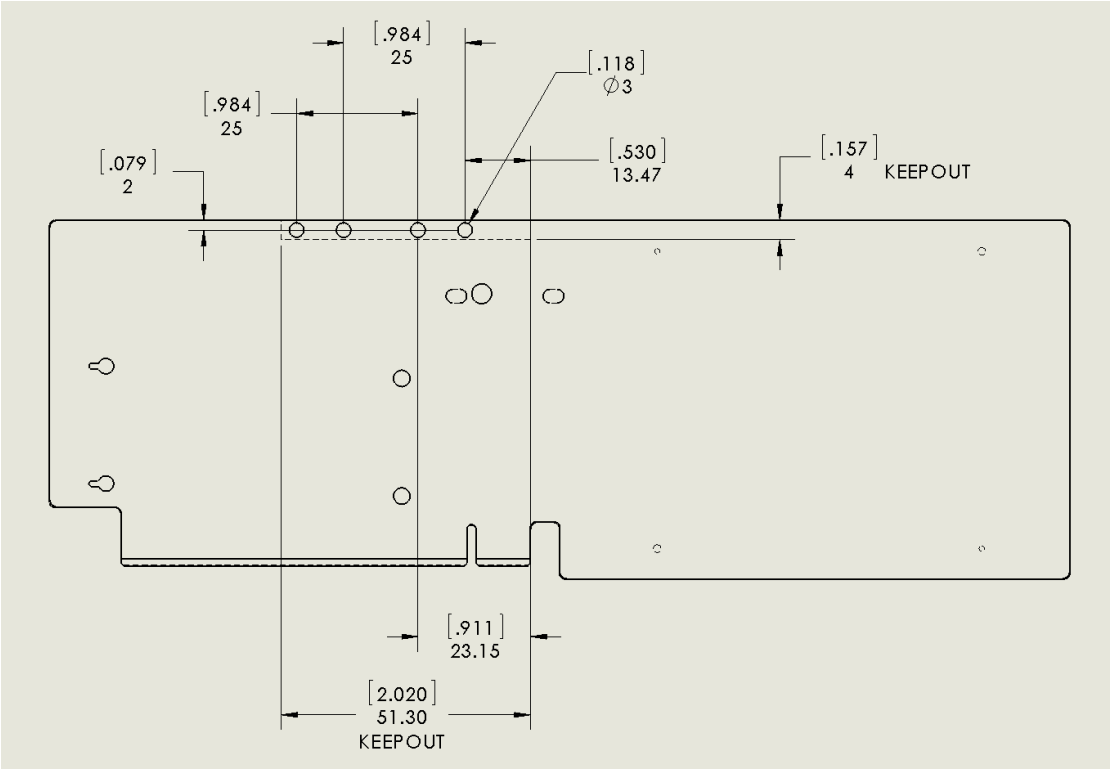

Figure 4-4: Hole definition for card insertion and removal

4.4 PCIe Edge Connector

The key dimensions, edge chamfer, pad layout (including a shorter pad for PRSNT# signal), placement, and dimensions of the card edge connector match the PCI Express Card Electromechanical Specification.

The GND planes underneath the pads for the edge connector on the card must be recessed according to the PCI Express Card Electromechanical Specification to improve signal integrity.

4.5 Baseboard Design

One possible implementation of the baseboard provides 12x micro-server card slots. The slots are situated as two rows of 6x slots. The slots in each row are spaced at a pitch of 27.5mm (center to center).

Future designs may accommodate thicker card designs to support higher-power versions.

5 Thermal

5.1 Data Center Environmental Conditions

5.1.1 Location of Data Center/Altitude

The data center may be located at a maximum of 1000m above sea level. Any variation of air properties or environmental differences due to the high altitude needs to be included in the thermal design.

5.1.2 Cold-Aisle Temperature

The data center will maintain the cold aisle temperature between 18°C and 30°C (65°F to 85°F). The mean temperature in the cold aisle is 24°C with 3°C standard deviation. The cold aisle temperature in the data center may fluctuate depending on the outside air temperature. Every component on the card must be cooled and maintained below its maximum spec temperature across the full cold aisle temperature range.

5.1.3 Cold-Aisle Pressurization

The data center will maintain the cold aisle pressure between 0 "H2O and 0.05 "H2O. The thermal solution of the system accommodates the worst-case operational pressurization in the data centers which is 0 "H2O with no fan failures and 0.01 "H2O with a single fan (or rotor) failure.

5.1.4 Relative Humidity

The data center will maintain the relative humidity between 20% and 85%. In the thermal design, the environmental condition changes due to the high altitude may not be considered when the thermal design can meet the requirement with maximum relative humidity (85%).

5.2 Server Operational Conditions

5.2.1 System Airflow or Volumetric Flow

The unit of airflow (or volumetric flow) used for this spec is CFM (cubic feet per minute). The maximum allowable airflow per watt in the system must be 0.16. The desired airflow per watt is 0.1 or lower in the system at the mean temperature (plus or minus standard deviation). See section 5.1.2 for the temperature definitions.

5.2.2 Delta T

The delta T is the air temperature difference across the system or the temperature difference between outlet air temperature of system and inlet air temperature of system. The delta T must be greater than 11.7°C (21°F). The desired delta T is 20°C (36°F) when the inlet air temperature to the system is lower than 30°C.

5.2.3 Thermal Margin

The thermal margin is the difference between the maximum theoretical safe temperature and the actual temperature. The board design operates at an inlet temperature of 35°C (95°F) outside of the system with a minimum 2% thermal margin for every component on the card. Otherwise, the thermal margin for every component in the system is at least 7% for temperatures up to 30°C.

5.2.4 Thermal Testing

Thermal testing must be performed up to 35°C (95°F) inlet temperature to guarantee high temperature reliability.

5.3 Heat Sink Requirements

The heat sink must be a thermally optimized design at the lowest cost. There must be no more than three heat pipes in the heat sink. Heat sink installation is simple and uncomplicated. Heat sinks must not block debug headers or connectors.

5.4 Temperature Sensors

Each card must provide temperature sensors for the SOC, the SO-DIMM(s) (if they are used), and one ambient temperature sensor. All temperature readings for each sensor must be readable via the management sideband interface to the baseboard. Additionally, over-temperature thresholds are configurable and an alert mechanism is provided to enable thermal shutdown and/or an increase in airflow. The sensors are accurate to +/-3C.

The ambient temperatures sensor is placed along the rear edge of the card on the A-side. Please see the blue, circled area in Figure 4 for more information.

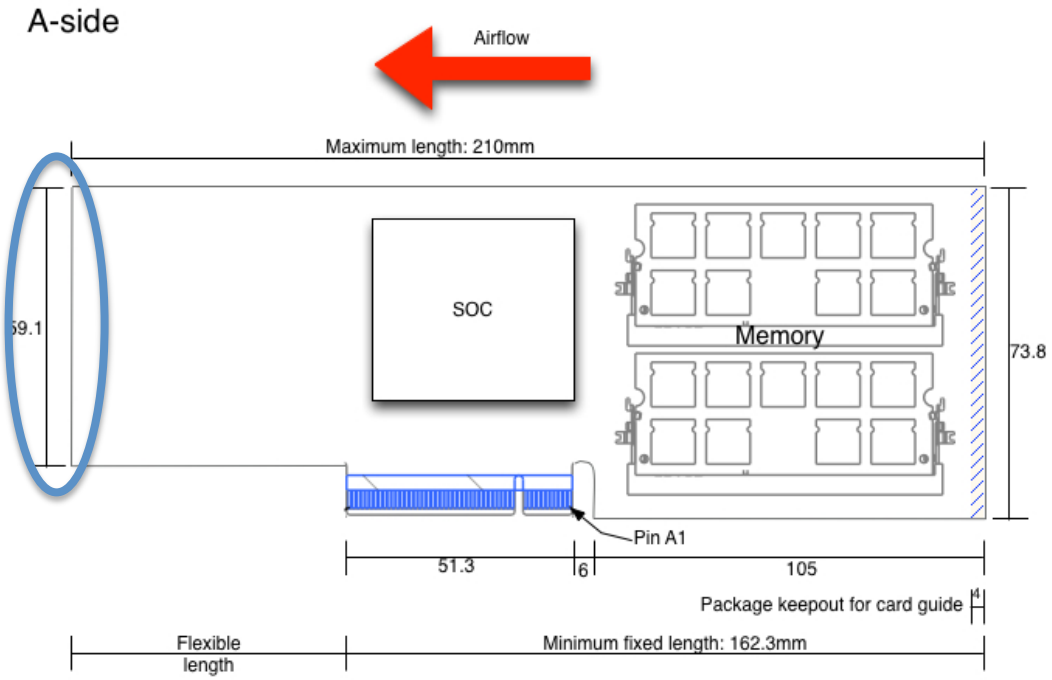

Figure 5-1 Ambient Temperature Sensor Location (Circled Area in Blue)

6 Electrical

6.1 X16 Edge Connector

The card must plug into an x16 PCIe edge connector. Figure 5 shows the default pin-out,. Other variations are possible (see below for more details).

Default Pin-out			
Pin Name	**Side B**	**Side A**	**Pin Name**
P12V	1	1	PRSNT#
P12V	2	2	P12V
P12V	3	3	P12V
GND	4	4	GND
I2C_SCL	5	5	SVR_ID0/GPIO0
I2C_DATA	6	6	SVR_ID1/GPIO1
GND	7	7	COM_TX
PWR_BTN#	8	8	COM_RX
USB_P	9	9	SVR_ID2/GPIO2
USB_N	10	10	SVR_ID3/GPIO3
SYS_RESET#	11	11	PCIE0_RESET#
I2C_ALERT#	12	12	GND
GND	13	13	PCIE0_REFCLK_P
GND	14	14	PCIE0_REFCLK_N
PCIE0_TX0_P	15	15	GND
PCIE0_TX0_N	16	16	GND
GND	17	17	PCIE0_RX0_P
GND	18	18	PCIE0_RX0_N
PCIE0_TX1_P	19	19	GND
PCIE0_TX1_N	20	20	GND
GND	21	21	PCIE0_RX1_P
GND	22	22	PCIE0_RX1_N
PCIE0_TX2_P	23	23	GND
PCIE0_TX2_N	24	24	GND
GND	25	25	PCIE0_RX2_P
GND	26	26	PCIE0_RX2_N

PCIE0_TX3_P	27	27	GND
PCIE0_TX3_N	28	28	GND
GND	29	29	PCIE0_RX3_P
GND	30	30	PCIE0_RX3_N
SATA0_TX_P	31	31	GND
SATA0_TX_N	32	32	GND
GND	33	33	SATA0_RX_P
GND	34	34	SATA0_RX_N
PCIE1_REFCLK_P	35	35	GND
PCIE1_REFCLK_N	36	36	GND
GND	37	37	PCIE2_REFCLK_P
GND	38	38	PCIE2_REFCLK_N
PCIE1_RESET#	39	39	GND
PCIE2_RESET#	40	40	GND
GND	41	41	RFU
GND	42	42	NIC_SMBUS_ALERT#
NIC_SMBUS_SCL	43	43	GND
NIC_SMBUS_SDA	44	44	GND
GND	45	45	GE0_RX_P
GND	46	46	GE0_RX_N
GE0_TX_P	47	47	GND
GE0_TX_N	48	48	GND
GND	49	49	PCIE0_RX4_P
GND	50	50	PCIE0_RX4_N
PCIE0_TX4_P	51	51	GND
PCIE0_TX4_N	52	52	GND
GND	53	53	PCIE0_RX5_P
GND	54	54	PCIE0_RX5_N
PCIE0_TX5_P	55	55	GND
PCIE0_TX5_N	56	56	GND
GND	57	57	PCIE0_RX6_P
GND	58	58	PCIE0_RX6_N
PCIE0_TX6_P	59	59	GND

PCIE0_TX6_N	60	60	GND
GND	61	61	PCIE0_RX7_P
GND	62	62	PCIE0_RX7_N
PCIE0_TX7_P	63	63	GND
PCIE0_TX7_N	64	64	GND
GND	65	65	PCIE1_RX0_P
GND	66	66	PCIE1_RX0_N
PCIE1_TX0_P	67	67	GND
PCIE1_TX0_N	68	68	GND
GND	69	69	PCIE1_RX1_P
GND	70	70	PCIE1_RX1_N
PCIE1_TX1_P	71	71	GND
PCIE1_TX1_N	72	72	GND
GND	73	73	PCIE1_RX2_P
GND	74	74	PCIE1_RX2_N
PCIE1_TX2_P	75	75	GND
PCIE1_TX2_N	76	76	GND
GND	77	77	PCIE1_RX3_P
GND	78	78	PCIE1_RX3_N
PCIE1_TX3_P	79	79	GND
PCIE1_TX3_N	80	80	GND
GND	81	81	P12V
GND	82	82	P12V

Required connection
Configurable connection
Reserved for future use

Figure 6-1 Pin Assignment Color-Coding Definitions

6.2 Pin Definitions

Figure 6-2 provides a detailed explanation of the pins. The direction of the signals is always defined from the perspective of the micro-server module.

Pin	Direction	Required/ Configurable	Pin Definition
P12V	Input	Required	12VAUX power
I2C_SCL	Input/Output	Required	I2C clock signal. I2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the baseboard.
I2C_SDA	Input/Output	Required	I2C data signal. I2C is the primary sideband interface for server management functionality. 3.3VAUX signal. Pull-up is provided on the baseboard.
I2C_ALERT#	Output	Required	I2C alert signal. Alerts the BMC that event has occurred that needs to be processed. 3.3VAUX signal. Pull-up is provided on the baseboard.
NIC_SMBUS_SCL	Input/Output	Required	Dedicated SMBus clock signal for network traffic between the BMC and the NIC. 3.3VAUX signal. Pull-up is provided on the baseboard.
NIC_SMBUS_SDA	Input/Output	Required	Dedicated SMBus data signal for network traffic between the BMC and the NIC. 3.3VAUX signal. Pull-up is provided on the baseboard.
NIC_SMBUS_ALERT#	Output	Required	Dedicated SMBus alert signal for network traffic between the BMC and the NIC. 3.3VAUX signal. Pull-up is provided on the baseboard.
PWR_BTN#	Input	Required	Power on signal. When driven low, it indicates that the server will begin its power-on sequence. 3.3VAUX signal. Pull-up is provided on the baseboard. If PWR_BTN# is held low for < 4 seconds, then this indicates a soft (graceful) power off. Otherwise, a hard shutdown is initiated.
SYS_RESET#	Input	Required	System reset signal. When driven low, it indicates that the server will begin its warm reboot process. 3.3VAUX signal. Pull-up is provided on the baseboard.
PRSNT#	Output	Required	Present signal. This is pulled low on the card to indicate that a card is installed. 3.3VAUX signal. Pull-up is provided on the baseboard.
COM_TX	Output	Required	Serial transmit signal. Data is sent from the micro-server module to the BMC. 3.3VAUX signal.
COM_RX	Input	Required	Serial receive signal. Data is sent from the BMC to the micro-server

			module. 3.3VAUX signal.
SVR_ID0/1/2/3 (GPIO0/1/2/3)	Input	Required	Slot ID bits or open-drain GPIOs. If a system contains more than one slot, each slot will be assigned a unique ID using pull-down resistors for 0 and open for a 1. Pull-ups should be provided on the card. The server can use this slot ID to identify its location in the system. Secondly, these pins can also be use as GPIOs when they are not needed as slot IDs. 3.3VAUX signal.
GE0_TX_P/N	Output	Required	Primary Ethernet transmit signal. Data is sent from the micro-server module to the baseboard.
GE0_RX_P/N	Input	Required	Primary Ethernet receive signal. Data is sent from the baseboard to the micro-server module.
PCIE0_RESET#	Output	Required	PCIe reset signal. If a PCIe bus is connected, this signal provides the reset signal indicating the card VRs and clocks are stable when driven high to 3.3V.
PCIE0_TX0/1/2/3_P/N	Output	Configurable	PCIe x4 bus transmit signals. Data is sent from the micro-server module to the baseboard. These signals may or may not be connected on the baseboard.
PCIE0_RX0/1/2/3_P/N	Input	Configurable	PCIe x4 bus receive signals. Data is sent from the baseboard to the micro-server module. These signals may or may not be connected on the baseboard.
PCIE0_REFCLK_P/_N	Output	Configurable	PCIe reference clock. This signal may or may not be connected on the baseboard.
PCIE1/2_RESET#	Output	Configurable	PCIe reset signals for to 2x additional PCIe buses. If a PCIe bus is connected, this signal provides the reset signal indicating the card VRs and clocks are stable when driven high to 3.3V.
PCIE1_TX0/1/2/3_P/N	Output	Configurable	Second set of PCIe x4 bus transmit signals. Data is sent from the micro-server module to the baseboard. These signals may or may not be connected on the baseboard.
PCIE1_RX0/1/2/3_P/N	Input	Configurable	Second set of PCIe x4 bus receive signals. Data is sent from the baseboard to the micro-server module. These signals may or may

			not be connected on the baseboard.
PCIE1/2_REFCLK_P/_N	Output	Configurable	Two additional PCIe reference clocks. These signals may or may not be connected on the baseboard.
SATA0_TX_P/N	Output	Configurable	SATA 2.0 or 3.0 transmit signals (two ports). Data is sent from the micro-server module to the baseboard. These signals may or may not be connected on the baseboard.
SATA0_RX_P/N	Input	Configurable	SATA 2.0 or 3.0 receive signals (two ports). Data is sent from the baseboard to the micro-server module. These signals may or may not be connected on the baseboard.
USB_P/N	Input/Output	Configurable	USB 2.0 differential pair.
RSVD	Input/Output	Configurable	Available differential pairs that could be configured as PCIe, SATA, SAS, Ethernet, or other high-speed interfaces.
RFU	Input/Output	Not connected	These pins are reserved for future use and are not used.

Figure 6-2 Detailed Pin Definitions

6.2.1 Required vs. Configurable Connections

The card pin-out is defined to provide basic functionality, but maintain flexibility for compatible alternate/future designs. All signals marked as "Required" must be connected on the card and will be connected on the baseboard. All signals marked "Configurable" may be used in some applications, but not in others. They may or may not be connected on the baseboard or the card. It is possible to create multiple card designs that support alternate pin assignments for the reserved signals.

All RX signals shall be placed on the A-side of the connector and all TX signals shall be placed on the B-side of the connector. This ensures the lowest amount of potential cross-talk between adjacent differential pairs.

6.2.2 Configurable Pin Assignment Algorithm

The baseboard will provide a table that defines the usage of each pair of reserved pins. This table will be accessible over I2C. During the power-on sequence, the SOC will query the baseboard and retrieve the contents of the table. If the connection type of a pair of reserved pins on the card matches to the connection type of the reserved pins on the baseboard, those signals are enabled and the initialization sequence can begin. If a match is not found, the table cannot be retrieved, or the connections are mismatched, the SOC signals are disabled and/or powered-off. The tables below provide the reserved pin mapping and encoding values for different functions.

Byte #	Field Definition
0	8 bit encoded value for pins A13 and A14
1	8 bit encoded value for pins A17 and A18
2	8 bit encoded value for pins A21 and A22

3	8 bit encoded value for pins A25 and A26
4	8 bit encoded value for pins A29 and A30
5	8 bit encoded value for pins A33 and A34
6	8 bit encoded value for pins A37 and A38
7	8 bit encoded value for pins A41 and A42
8	8 bit encoded value for pins A45 and A46
9	8 bit encoded value for pins B15 and B16
10	8 bit encoded value for pins B19 and B20
11	8 bit encoded value for pins B23 and B24
12	8 bit encoded value for pins B27 and B28
13	8 bit encoded value for pins B31 and B32
14	8 bit encoded value for pins B35 and B36
15	8 bit encoded value for pins B39 and B40
16	8 bit encoded value for pins B43 and B44
17	8 bit encoded value for pins B47 and B48

Figure 6-3 Byte to Configurable Pin Pair Mapping

Encoding	Signal Type
0x00	Off
0x01	PCIe gen 2
0x02	PCIe gen 3
0x03	PCIe clock
0x04	SATA 2.0
0x05	SATA 3.0
0x06	1000BASE-KX
0x07	10GBASE-KR
0x08	SGMII
0x09-0xff	RFU

Figure 6-4 Hex Value Encoding for Different Signal Types

6.3 Ethernet

At least one Ethernet connection (GE0_TX/RX) is required on the card. To enable maximum compatibility and a variety of potential topologies, this Ethernet port is a PHY layer device and must be capable of the following:

Mode	Standard	Encoding
1000BASE-X	IEEE Clause 36, 37	8b/10b
1000BASE-KX	IEEE 802.3ap	8b/10b

10GBase-KR	IEEE 802.3ap	64b/66b

Ideally, the Ethernet port should be a 10Gb port, but a first generation design may support 1GbE or 2.5GbE. Auto-negotiation may need to be disabled to work with some baseboard designs.

6.3.1 Routing guidelines

To support 10Gb speeds, it is critical that the differential pairs for transmit and receive adhere to the following rules:

- Each differential pair must be no longer than 2" including SOC breakout and the connection to the edge connector pad
- A minimum spacing of 120mils is maintained at all times (with the exception of the SOC breakout) between the transmit or receive differential pairs, and any adjacent signals.

6.4 SATA

The SATA connections are a minimum of SATA2.0 (3Gb/s) and may be SATA3.0 (6Gb/s).

6.5 PCIe

The PCIe connection is a minimum of PCIe 2.0 and may be PCIe 3.0.

6.6 I2C

The single I2C connection is the primary server management interface. It supports a minimum speed of 400kHz. The card does not contain a BMC as the primary BMC will be located on the baseboard. The BMC on the baseboard will act as the master on the bus during normal operation. During POST, the SOC may also act as a master. It is acceptable to have a small amount of glue logic on the card to support translating and/or buffering SOC signals. The I2C alert is signal is required and is used as an interrupt for the BMC.

Both the SOC and the BMC on the baseboard may be masters on the bus so a multi-master environment must be supported.

7 Power

7.1 Input

Power for the card is provided via five 12V pins on the PCIe connector. Each pin supports a maximum of up to 1.1A of current.

7.1.1 X8 card

The card is designed to only support a maximum of 60W (up to 5.5A). The nominal voltage is 12.0V, but may vary between 10.8V and 13.2V.

7.1.2 X16 card

The card is designed to only support a maximum of 80W (up to 7.7A). The nominal voltage is 12.0V, but may vary between 10.8V and 13.2V.

7.2 Hot-Plug Support

Hot-plug circuitry is not required on the card.

7.3 VR Efficiency

All VRs providing over 15W on the card are at least 91% efficient when loaded between 30% and 90%.

7.4 Input Capacitance

The capacitance provided on the input 12V rail shall not exceed 400uF.

8 Functional

8.1 Memory

Each card must support a minimum of 8GB of ECC DDR3L low-voltage (1.35V) SDRAM memory and two memory channels. Memory may be provided as one or more modules, or soldered directly to the card.

8.2 NIC

The NIC in the SOC must support at least 1Gb, IPv4 and IPv6, and iSCSI boot. iSCSI and TCP/IP offload capabilities are highly desirable.

8.3 Storage

Each card must support a minimum of 128GB of low-cost, MLC flash to be used as a boot device. If the flash is SATA-based, it must be connected to SATA port 0. Additional SATA connections are connected to port 1 and higher. The flash may be provided as a mSATA or NGFF card, or be soldered directly to the card. The solution must be designed so that all storage on the card can be de-populated.

8.4 EEPROM

Each card must include an I2C-accessible EEPROM. The EEPROM must be accessible from the I2C connection going to the baseboard and be at least 128Kbits. The I2C address for the EEPROM should be set to 0xA2. The EEPROM will contain the FRUID information and any additional configuration information that may be required. The FRUID is formatted in accordance with the IPMI Platform Management FRU Information Storage Definition document.

The EEPROM must contain the following:

- Device Description
- Board Mfg
- Board Product
- Board Serial
- Board Part Number
- Product Manufacturer
- Product Name
- Product Part Number

- Product Serial Number
- Product Asset Tag
- Product Build: e.g. EVT, DVT, PVT, MP
- Product Version: e.g. C1
- Product Comments: e.g.
- Manufacturing date code:
- Manufacturing lot code: (preferred, but optional)
- Manufacturing work order: (preferred, but optional)
- PCB revision
- SOC model name/number
- SOC revision
- SOC Tj (junction temperature) max

This information will be available via IPMI commands from the BMC.

8.5 BIOS

The card supplier is responsible for supplying and customizing the BIOS for the SOC. The requirements are outlined in this section.

8.5.1 UEFI

The BIOS shall be a UEFI compatible BIOS.

8.5.2 Configuration and Features

The BIOS is tuned to minimize card power consumption. It has the following features:

- Unused devices are disabled including PCIe* lanes, USB ports, SATA/SAS ports, etc
- BIOS setup menu
- SOC settings are provided to allow tuning to achieve the optimal combination of performance and power consumption.

8.5.3 BIOS Settings Tools

The card supplier shall provide a tool to make BIOS setting changes without requiring a BIOS re-flash. The BIOS settings update tool must also support success/failure codes so that updates can easily be scripted.

8.5.4 PXE Boot

The BIOS supports PXE boot and provide the ability to modify the boot sequence. When PXE booting, the card first attempts to boot from the first Ethernet device (etho). If this fails, PXE boot is attempted on the next Ethernet device.

The default boot device priority is:

1. Network (search all configured network interfaces)
2. HDD, SSD, or flash device (local or remote)
3. CD-ROM
4. Removable Device

This process loops indefinitely and requires no user-intervention.

8.5.5 iSCSI Boot

The BIOS shall be capable of iSCSI network boot.

8.5.6 Other Boot Options

The BIOS also supports booting from SATA/SAS and USB interfaces. The BIOS provides the capability to select boot options.

8.5.7 Remote BIOS Update

The BIOS can be updated remotely under these scenarios:

- Scenario 1: Sample/Audit BIOS settings
    - Return current BIOS settings, or
    - Save/export BIOS settings in a human-readable form that can be restored/imported (as in scenario 2)
- Scenario 2: Update BIOS with pre-configured set of BIOS settings
    - Update/change multiple BIOS settings
    - Reboot
- Scenario 3: BIOS/firmware update with a new revision
    - Load new BIOS/firmware on machine and update, retaining current BIOS settings
    - Reboot

Additionally, the update tools have the following capabilities:

- Update from the operating system over the LAN – the OS standard is CentOS v5.2
- Can complete BIOS update or setup change with a single reboot (no PXE boot, no multiple reboots)
- No user interaction (like prompts)
- BIOS updates and option changes do not take longer than five minutes to complete
- Can be scripted and propagated to multiple machines

8.5.8 Event Log

An event log needs to be implemented to capture informational and error messages. Per SMBIOS specification Rev 2.6, the BIOS implements SMBIOS type 15 for an event log; the assigned area is large enough to hold more than 500 event records (assuming the maximum event record length is 24 bytes, then the size will be larger than 12KB), and follow the SMBIOS event log organization format for the event log.

A system access interface and application software must be provided to retrieve and clear the event log from the BIOS, including, at minimum, a Linux application for the CentOS operating system and driver as needed. The event log must be retrieved and stored as a readable text file that is easy to handle by a scripting language under Linux. Each event record includes enhanced information identifying the error source device's vendor ID, card slot ID, and device ID.

8.5.9 Logged Errors

The following list of errors are logged by the BIOS and must include date, time, and location information which can easily be used to identify the failing component.

- CPU/Memory errors: Both correctable ECC and uncorrectable ECC errors are logged into the event log. Error categories include DRAM, Link, and L3 cache.
- PCIe* errors: Any errors that have a status register are logged into the event log, including root complex, endpoint device, and any switch upstream/downstream ports if available. Link disable on errors are also be logged. Fatal, non-fatal, or correctable classification follows the chipset vendor's recommendation.
- POST errors: All POST errors detected by the BIOS during POST are logged into the event log.
- SATA or SAS errors: All correctable and uncorrectable errors are logged.
- System reboot events
- Sensor values exceeding warning or critical thresholds

8.5.10 Error Thresholds

An error threshold setting must be enabled for both correctable and uncorrectable errors. Once the programmed threshold is reached, an event is triggered and logged.

- Memory Correctable ECC: The threshold value is 1000. When the threshold is reached, the BIOS logs the event and includes the physical DIMM location.

8.5.11 POST Codes

The BIOS outputs a set of POST codes identifying the current initialization step and any errors encountered along the way. The output is provided on the serial console and errors are logged.

During the boot sequence the BIOS shall initialize and test each DIMM module. If a module fails initialization or does not pass the BIOS test the following post codes should flash on the debug card to indicate which DIMM has failed. The first hex character indicates which CPU interfaces the DIMM module; the second hex character indicates the number of the DIMM module. POST code will also display error major code and minor code from Intel memory reference code. The display sequence will be “00”, DIMM location, Major code and Minor code with 1 second delay for every code displayed. The BIOS shall repeat the display sequence indefinitely to allow time for a technician to service the system. DIMM location code table is as Table 8-1. DIMM number count starts from the furthest DIMM from CPU.

Table 8-1 DIMM Error Code Table

Code	Result
A0	Channel 0 DIMM 0 Failure
A1	Channel 0 DIMM 1 Failure
A2	Channel 1 DIMM 0 Failure
A3	Channel 1 DIMM 1 Failure

8.6 Management

The primary server management functions will be provided using a BMC on the baseboard. The BMC on the baseboard will use the I2C connection as the sideband

interface. This section identifies the required information that must be accessible from the BMC.

8.6.1 BMC feature support

The BMC will support DCMI 1.5 plus an extended set of commands to enable support for the multi-node environment. The following is a list of features that the BMC must support:

- All SEL commands
- All sensor commands
- All SDR commands
- LAN print/set commands
- Power on/off/reset/soft/cycle commands
- Chassis identify force/off
- MC reset cold/warm
- MC info
- SOL activate/de-activate
- FRU list
- Chassis boot parameters

For a more detailed description and specification, please see the TBD baseboard specification.

8.6.2 I2C addressing

The SOC and BMC will communicate using IPMI 2.0 commands transmitted over the I2C connection. While each I2C connection will be an independent channel between the SOC and the BMC, each slot will need to be uniquely addressed to ensure the commands are routed properly by the BMC. To support this, each SOC must have a unique I2C address that is determined using the slot ID. The upper 3 bits of the I2C address are defined as 0b111 while the lower 4 bits are composed of the slot ID bits. For example, slot 0x4 on the baseboard, results in the SOC having an I2C address of 0b1110110.

8.6.3 IPMI commands

The following list of IPMI commands must be supported by the SOC.

TBD

8.6.4 Management interface and block diagram

The management interface to the card consists of a supplier-agnostic interface that combines a simple register interface with dual KCS buffers to abstract the card-specific details. This interface can be implemented in hardware, software, or a combination of the two.

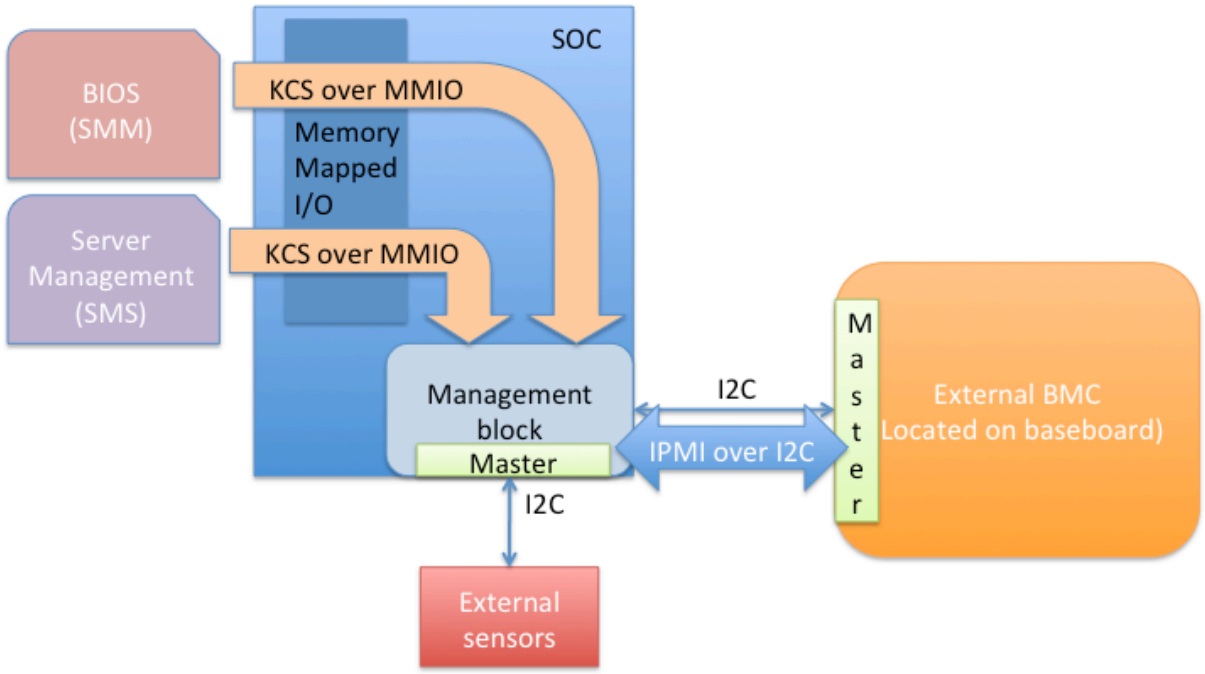

Figure 8-1 Block Diagram of IPMI interface

8.6.5 Micro-server register interface

The register interface and implementation details are covered in the Micro-server management specification version TBD.

8.6.6 Serial Console

The SOC provides a serial UART that is connected directly to the card edge. This connection will be used as the BIOS or OS serial console and will also be available as a Serial Over LAN (SOL) connection via the BMC. The BIOS menus must be fully accessible and text-based. Any hot keys that are required must be transmittable through a serial console session.

The BIOS should default to 57600bps/8N1.

8.6.7 Power Control

The BMC controls power on, off, and reset directly via the signals defined in the pin-out. If 12V to the card is lost and returns ("AC Lost"), the BIOS must be configurable to enable immediate or delayed power-on, or the last power state prior to the event.

8.6.8 Thermal Alerts

The SOC provides a mechanism to provide thermal alerts and overtemp notifications. The BMC must be able to receive these alerts in a timely fashion to allow it take action quickly. The I2C alert signal must be used. In some cases, an overtemp condition may occur which forces the SOC to power-off immediately. This condition must be logged.

8.6.9 Sensors

The following list of analog and discrete sensors are provided and are reported by the BMC.

Analog sensors include:

- CPU temp/thermal margin
- DIMM temp
- Ambient temp

Discrete sensors include:

- CPU therm trip
- Power threshold event
- SEL status
- DCMI watchdog
- POST error
- ProcHot assertion
- MemHot assertion
- Machine check error
- PCIe error
- Other IO error
- Memory error

8.7 Battery

A battery on the card is not required. The BIOS pulls its settings from non-volatile memory.

8.8 LEDs

Each card contains a power LED that illuminates when the power-on sequence on the card has completed successfully. The LED is blue in color and placed on the leading edge of the card (cold-aisle). Flashing this LED may also be used to identify a card.

9 Environmental Requirements

The full system with the server card installed meets the following environmental requirements:

- Gaseous Contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C (long-term storage)
- Transportation temperature range: -55°C to +85°C (short-term storage)
- Operating altitude with no de-rating to 1000m (3300 feet)

Operating temperatures and relative humidity can be found in section 5.2.

9.1 Vibration and Shock

The motherboard meets shock and vibration requirements according to the following IEC specifications: IEC78-2-(*) and IEC721-3-(*) Standard & Levels.

	Operating	Non-Operating
Vibration	0.5g acceleration, 1.5mm amplitude, 5 to 500 Hz, 10 sweeps at 1 octave / minute for each of the three axes (one sweep is 5 to 500 to 5	1g acceleration, 3mm amplitude, 5 to 500 Hz, 10 sweeps at 1 octave / minute for each of the three axes

	Hz)	(one sweep is 5 to 500 to 5 Hz)
Shock	6g, half-sine 11mS, 5 shocks for each of the three axes	12g, half-sine 11mS, 10 shocks for each of the three axes

Figure 9-1 Vibration and Shock Requirements

10 Prescribed Materials

10.1 Disallowed Components

The following components are not used in the design of the motherboard:

- Components disallowed by the European Union's Restriction of Hazardous Substances Directive (RoHS 6)
- Trimmers and/or potentiometers
- Dip switches

10.2 Capacitors and Inductors

The following limitations apply to the use of capacitors:

- Only aluminum organic polymer capacitors made by high quality manufacturers are used; they must be rated 105°C
- All capacitors have a predicted life of at least 50,000 hours at 45°C inlet air temperature, under worst conditions
- Tantalum capacitors using manganese dioxide cathodes are forbidden
- SMT ceramic capacitors with case size > 1206 are forbidden (size 1206 are still allowed when installed far from the PCB edge and with a correct orientation that minimizes risks of cracks)
- Ceramic material for SMT capacitors must be X7R or better material (COG or NP0 type are used in critical portions of the design) Only SMT inductors may be used. The use of through hole inductors is disallowed. 13.3

10.3 Component De-rating

For inductors, capacitors, and FETs, de-rating analysis is based on at least 20% de-rating.

11 Revision History

Author	Description	Revision	Date
Chris Petersen	▪ Initial draft	0.5	1/20/2013
Chris Petersen	▪ Add x16 form factor and connector pin assignments. Lots of additional details added.	0.66	11/07/2013
Chris Petersen	Removed x8 variant	1.0	02/05/2015


OPEN
Compute Project