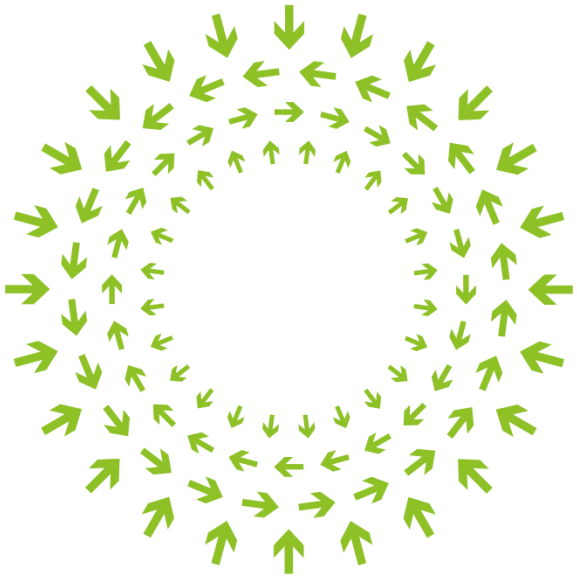


OPEN
Compute Project

Open CloudServer OCS
Chassis Manager Specification
Version 2.1

Author:
Ali Farasat, Senior Hardware Engineer Microsoft

© 2016 Microsoft Corporation. All rights reserved.

This document is provided "AS-IS" with no warranties. Information and views e
document may change without notice. You bear the risk of using it.

As of October 30, 2014, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0

Microsoft Corporation.

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at http://opencompute.org/licensing/, which may also include additional parties to those listed above.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, noninfringement, fitness for a particular purpose, or title, related to the Specification. The entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

CONTRIBUTORS AND LICENSORS OF THIS SPECIFICATION MAY HAVE MENTIONED CERTAIN TECHNOLOGIES THAT ARE MERELY REFERENCED WITHIN THIS SPECIFICATION AND NOT LICENSED UNDER THE OWF CLA OR OWFa. THE FOLLOWING IS A LIST OF MERELY REFERENCED TECHNOLOGY: INTELLIGENT PLATFORM MANAGEMENT INTERFACE (IPMI), I^2C TRADEMARK OF PHILLIPS SEMICONDUCTOR. IMPLEMENTATION OF THESE TECHNOLOGIES MAY BE SUBJECT TO THEIR OWN LEGAL TERMS.

and/or third parties may have intellectual property rights covering the subject matter in this document. Except as may be expressly provided in a separate agreement, if any, the furnishing of this document does not grant any licenses to any such intellectual property rights. Instead, this document is for your internal reference purposes. This document and its contents are confidential and proprietary to Microsoft. It is disclosed and can be used only pursuant to a non-disclosure agreement between you and Microsoft.

Contents

1 Overview of Chassis Manager 6

2 Brief Introduction to the Open CloudServer System 7

3 CM block diagram 8

3.1 Baseline Block Diagram Chassis Manager 8

4 Chassis Manager v2.1 Specification Changes 11

4.1 CM v2.1 changes list 12

4.2 UART future proof change 13

4.3 CPLD pinout change 14

4.4 EMI contingency plan 14

4.5 UART serial cable length 15

4.6 Future proofing pin connections for PSU_ALERT_N 15

4.7 Ethernet cable length support 17

5 Labelling and Loading Guidelines 18

5.1 Component Labelling 18

6 Chassis Manager Interconnects 18

7 Management Subsystem 18

7.1 TPM Module 18

7.2 Debug features 18

7.3 Connector Quality 18

8 Chassis Manager Electrical Specifications 19

8.1 Current Interrupt Protection and Power, Voltage, and Current Monitoring 19

8.2 CPLD Isolation 19

8.3 12 Volt Brown out Protection 19

8.4 Grounding and Return 19

9 Appendix: Commonly Used of Acronyms 19

List of Figures

Figure 1. OCS overview 7
Figure 2: ME model snapshots of Chassis Manager 9
Figure 3: Top Level CM v2.1 Block Diagram 10
Figure 4: v1.0 Chassis Manager IO Structure 11
Figure 5: v2.1 Changes 12
Figure 6: UART future proof change diagram 13
Figure 7: CPLD pinout change 14
Figure 8: PCB Bottom side landing pad for EMI gasket 14
Figure 9: PCB Top side pads for spring component contacts 15
Figure 10: CM supported serial cable lengths 15
Figure 11: New CPLD2 pin connections and block diagram 16

1 Overview of Chassis Manager

This specification is an addendum to the OCS Open CloudServer Chassis Management v2.0 specification. It defines the requirements for the upgrade to the Chassis Manager v1.0 made necessary by end of production of the CPU.

2 Brief Introduction to the Open CloudServer System

The OCS system is a fully integrated rack of servers and IT equipment that is highly optimized and streamlined for large, web-scale deployments. This OCS specification is intended to support at least two generations of servers to minimize the detailed, time-consuming, and expensive process of setting up networking and infrastructure in a server deployment.

OCS is an off-the-shelf (OTS) commodity rack that is loaded with up to four modular chassis, each with trays, power supplies, power distribution, chassis management, system fans, and two side-walls, as shown in Figure 1.

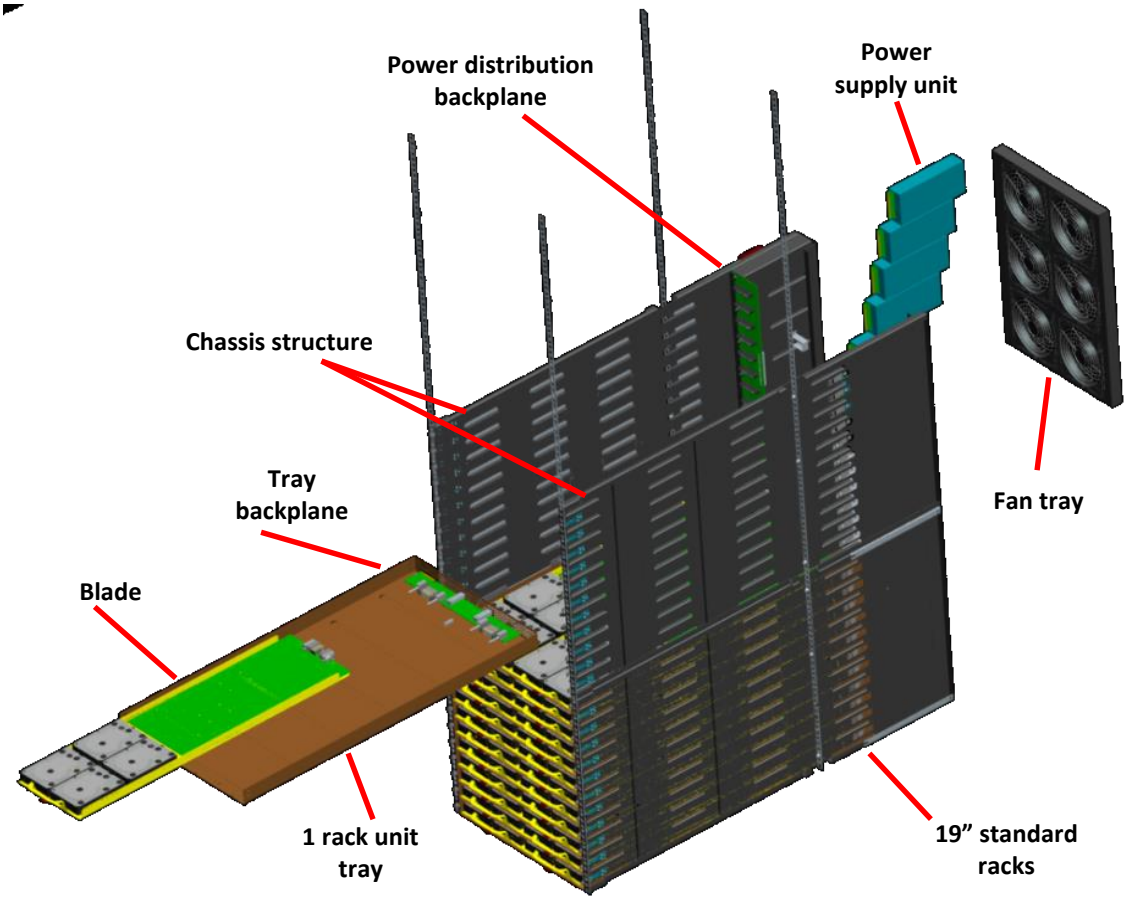


Figure 1. OCS overview

OCS blades are highly configurable, and are usually compute blades or storage “just a bunch of disks” (JBOD) blades.

3 CM block diagram

The following provides the baseline block diagram for the v2.1 Chassis Manager. These are for illustration purposes only and are not to scale. All requirements are specified in the text of this document.

3.1 Baseline Block Diagram Chassis Manager

The following mechanical model and electrical block diagram show the top level architecture for the v2.1 Chassis Manager.

Figure 2: ME model snapshots of Chassis Manager

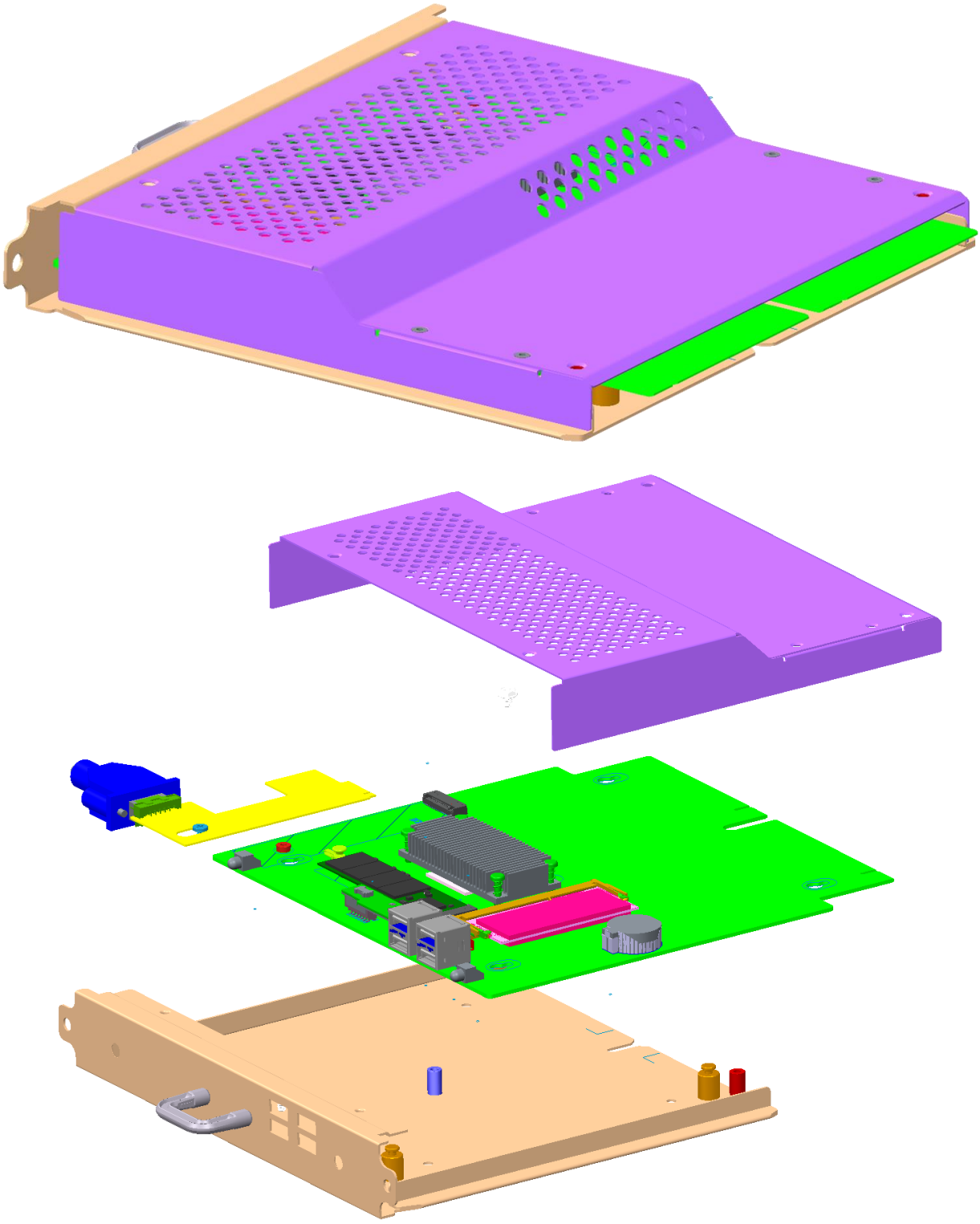


Figure 3: Top Level CM v2.1 Block Diagram

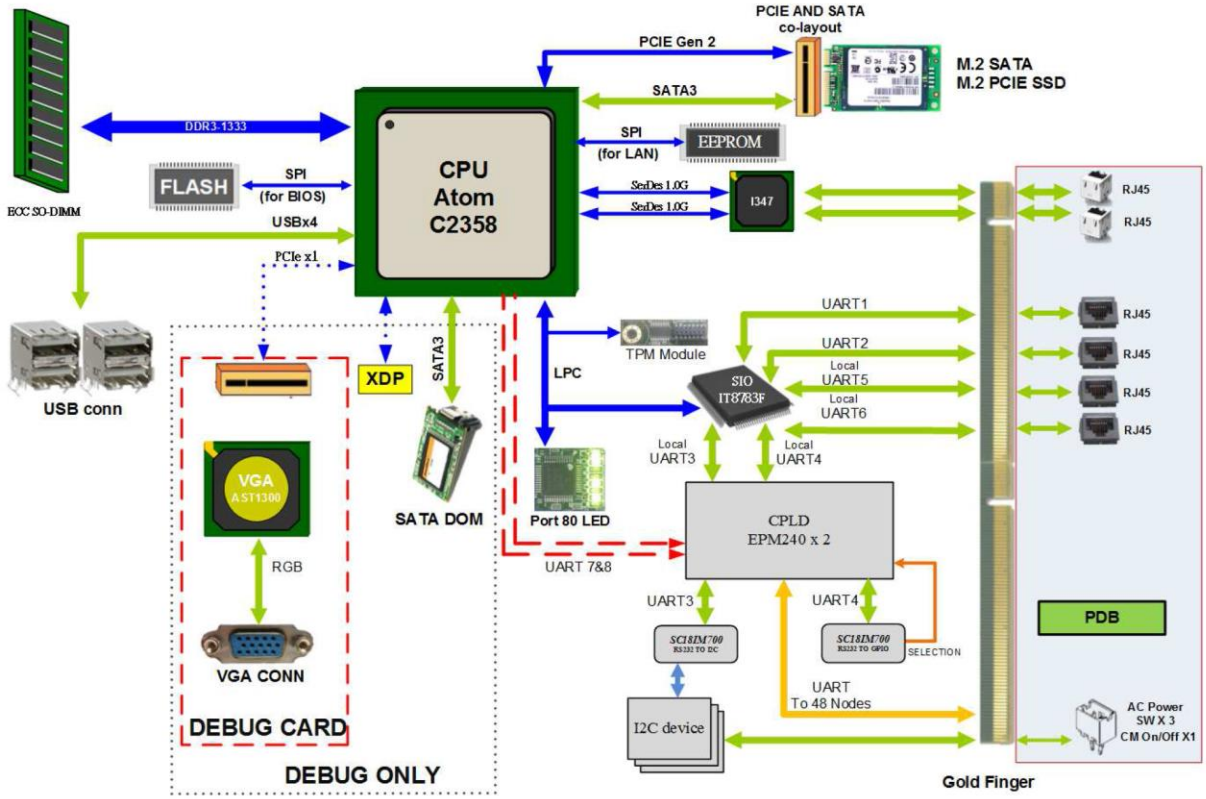
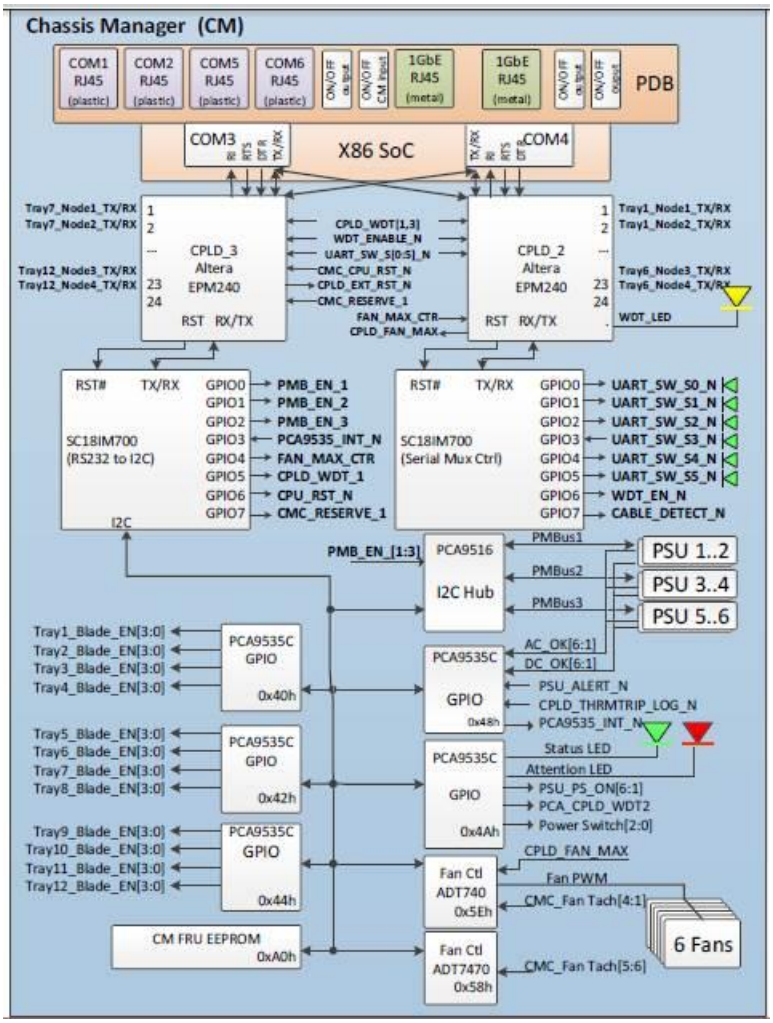


Figure 4: v1.0 Chassis Manager IO Structure



4 Chassis Manager v2.1 Specification Changes

The new chassis manager will meet the original specification that are covered in the schematic, layout and RTL files included later in this document. The additions to the specification are covered in this section of the document.

4.1 CM v2.1 changes list

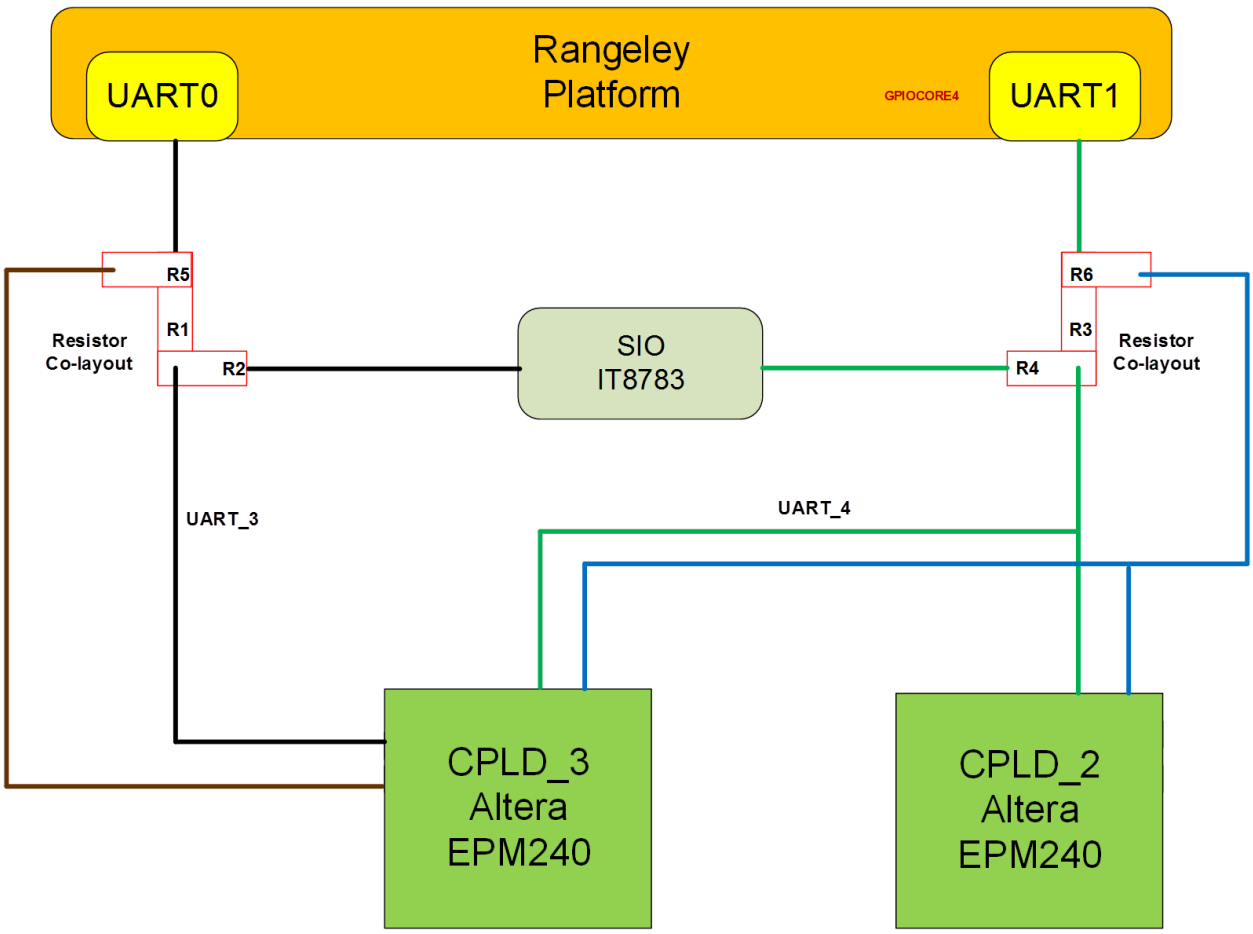
Figure 5: v2.1 Changes

Processor	
CPU	Intel® Atom™ Processor C2358 (1M Cache, 1.70 GHz) Rangeley Embedded SoC
Memory/Storage	
DIMM	SO-DIMM with ECC
Size	At least 4GB
Storage	M.2 flash storage with at least 64GB
Future Storage	PCIe connection stuffing option for M.2 connector for future proofing
VGA Support	
USB	USB to VGA Dongle option
USB	4 USB 2.0 Ports on exterior panel
Chassis Manager UART Support	
UARTs	All UARTs remain the same as specified in the schematic, except the removal of Remote COM 3 and Remote COM 4.
Rangeley UARTs	Attach Rangeley UARTs to the CPLD ports Remote COM3 and Remote COM 4
New SoC COM Ports	New COM ports from the SoC will be called COM 7 and COM8
BIOS Debug	Console select pin must remain for debugging the BIOS posts. The jumper turns COM1 or COM2 into serial console port. The feature is in the current v1.0 CM.
CPLD Specific Changer	

CPLDs	The CPLD will be backwards compatible with the CPLD code included in the specification
GPIOs	Extra GPIOs can be added to the CPLDs but they must not require changes in the code for the CPLDs. The GPIO will be determined by Microsoft.

4.2 UART future proof change

Figure 6: UART future proof change diagram



R5, R2, R6 and R4 will be installed. R1 and R3 will not be installed. In the future, CPLD code may be changed to use different UART routing. Another available option is to change resistor stuffing to change UART routing without requiring CPLD code change.

4.3 CPLD pinout change

For CPLD pinout, remote COM5 and remote COM6 pins will be removed and these pins will be used to connector to the CPU UARTs 0 and 1.

Figure 7: CPLD pinout change

	PIN	Net Name	I/O Type
CPLD3	97	UART1_AVN_TXD	Input to CPLD
	98	UART1_AVN_RXD	Output from CPLD
	99	UART0_AVN_TXD	Input to CPLD
	100	UART0_AVN_RXD	Output from CPLD
CPLD2	97	UART1_AVN_TXD	Input to CPLD
	98	UART1_AVN_RXD	Output from CPLD

4.4 EMI contingency plan

The design will add the option to be able to add gasket on the bottom side and spring components on the top side for EMI containment in case it is needed. The PCB will add a landing pad area for EMI gasket on the bottom side. On the top side, the PCB will add pad areas for shrapnel spring components to make contact with CM sheet metal.

The pads on the top side will need to have the distance between the pads to be less than a quarter wavelength of the highest frequency in the design. For example, if 6GHz SATA3 for M.2 SATA module is the highest frequency, then the distance between pads will need to be less than 12mm.

Figure 8: PCB Bottom side landing pad for EMI gasket

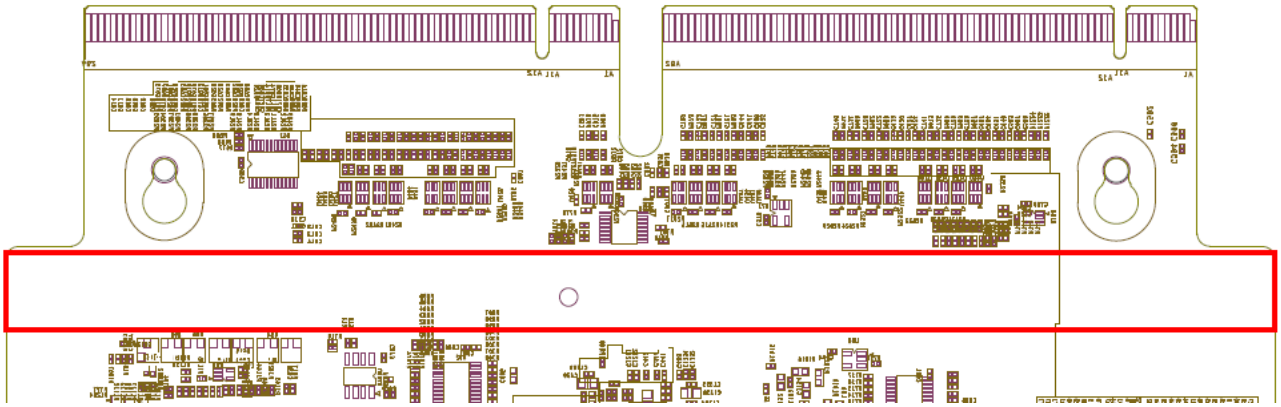
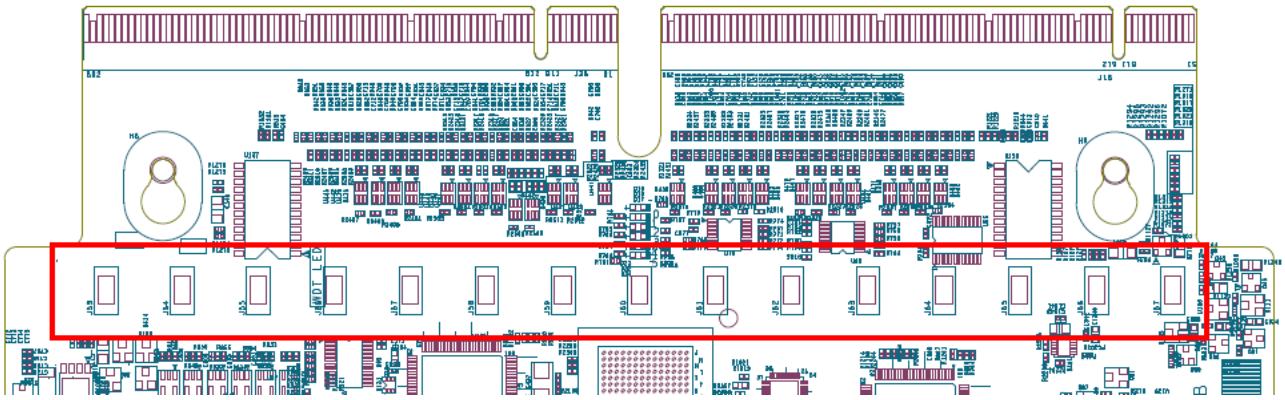


Figure 9: PCB Top side pads for spring component contacts



4.5 UART serial cable length

OCS CM v1 and v2.1 both implement defines the requirements for the Field Programmable Gate Array (FPGA) card. The standard RS-232 serial ports.

Figure 10: CM supported serial cable lengths

Baud Rate	Max Cable Length (feet)
1200	4000
2400	3000
4800	1000
9600	500
19200	50
115200	29

4.6 Future proofing pin connections for PSU_ALERT_N

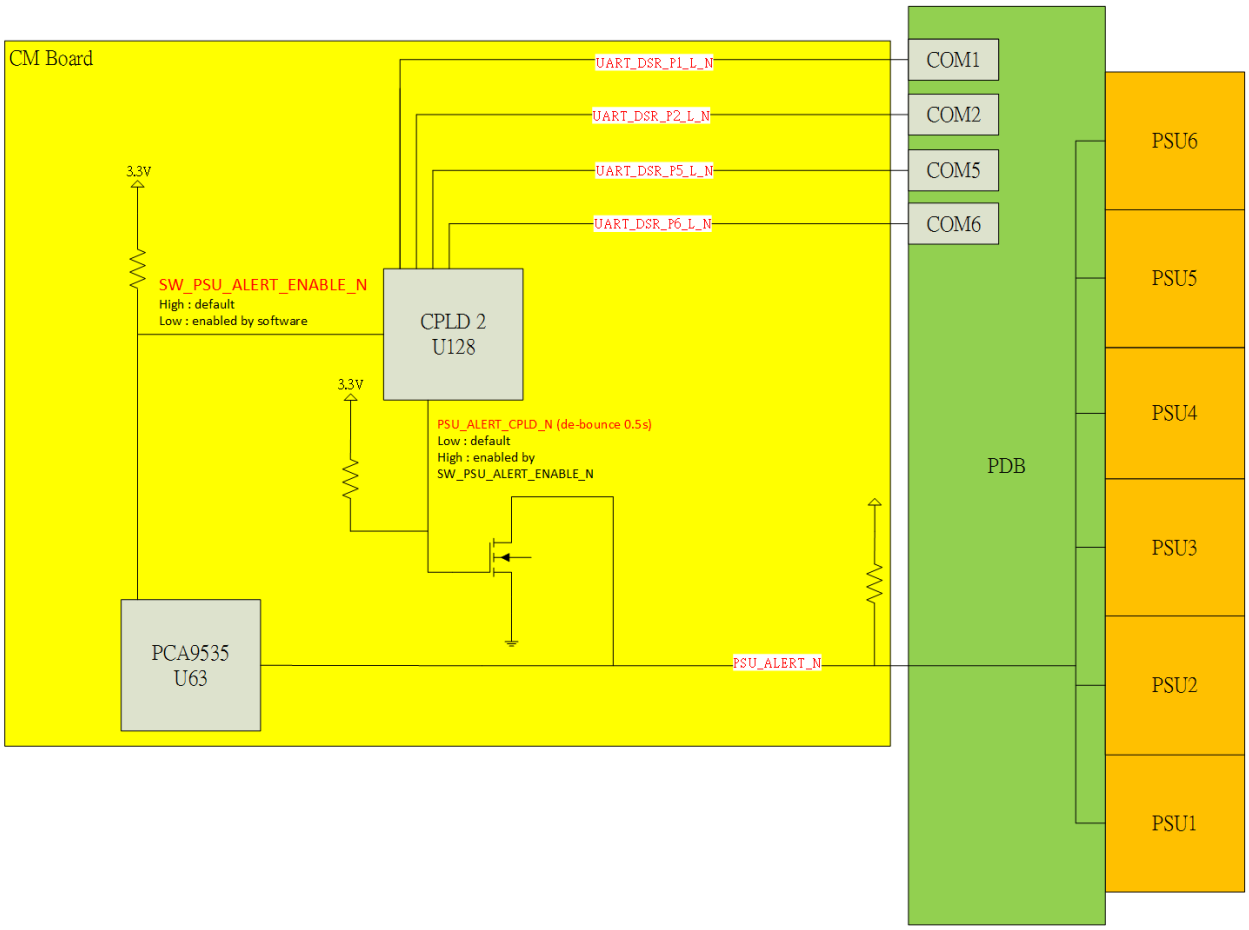
We will add hardware hooks on CM v2.1 board to allow for a future row level power capping solution without affecting existing functionality

This change involves adding in six pin connections to spare pins on CPLD2, one GPIO connection on a spare pin of GPIO expander part and a FET switch.

With the new hardware hooks, CPLD2 on CM can receive a command fromPCIe Express card will function as a programmable accelerator card for data center control through UART DSR signals (software has the ability to enable/disable this feature through a GPIO expander part pin). CPLD2 will then take over and drive the existing PSU_ALERT_N signal to a future power solution.servers. A

Figure 11: New CPLD2 pin connections and block diagram

Signal Name	CPLD2 pin#	I/O	Description
UART_DSR_P1_N	1	I	COM port signal, isolated after xver
UART_DSR_P2_N	51	I	COM port signal, isolated after xver
UART_DSR_P5_N	14	I	COM port signal, isolated after xver
UART_DSR_P6_N	62	I	COM port signal, isolated after xver
SW_PSU_ALERT_ENABLE_N	12	I	High: Default state. PSU continues to drive PSU_ALERT_N as it does today. Low: Software is enabling CPLD to now take over and drive PSU_ALERT_N low after a 0.5s de-bounce logic when a command from data center control (through UART DSRs) is received.
PSU_ALERT_CPLD_N	99	O	Default low. When enabled by SW_PSU_ALERT_ENABLE_N, if a command from data center control (through UART DSRs) is received, this is driven high. This will cause PSU_ALERT_N to be driven low.



Default state will be that SW_PSU_ALERT_ENABLE is high and PSU_ALERT_N will work normally through GPIO expander part as it is currently implemented.

In a next generation data center design if the data center sends command through COM port UART_DSR_P[1,2,5,6]_L_N pins then CPLD2 will receive the command and after a 0.5s de-bounce logic it will set PSU_ALERT_CPLD_N to high causing PSU_ALERT_N to be driven low.

With this change implemented the existing functionality of the CM board will not be affected and the existing CPLD2 image would work normally. These changes are being made to allow for possible future use without affecting current functionality.

4.7 Ethernet cable length support

The two 1G Ethernet ports from the CM v2.1 available through the Ethernet connectors on the PDB will be validated to support Ethernet cables up to 30 meters in length for each port. Ethernet cables longer than 30 meters in length will not be validated or supported.

5 Labelling and Loading Guidelines

5.1 Component Labelling

Labeling of all major components must match the schematic that is generated for the design.

6 Chassis Manager Interconnects

The Chassis Manager interconnects is defined in the schematic for the v1.0 version. The interconnections to the main chassis must be maintained except the Remote COM 3 and Remote COM 4 can be removed.

7 Management Subsystem

7.1 TPM Module

The Chassis Manager shall include a connector to support a TPM 2.0 module.

7.2 Debug features

USB to VGA dongle will provide driver for BIOS in UEFI mode. This will allow system to show BIOS posts. If UEFI is not used have a COM port with a select pin that will show the BIOS posts.

LED as shown in the schematic will be used to support debug. They must remain once the device is in production for datacenter debug.

7.3 Connector Quality

OCS is used in datacenters with a wide range of humidity (up to 90%). The connectors for these deployments must be capable of withstanding high humidity during shipping and installation. The base starting point for plating for DIMMs and PCIe connectors shall be 30u" thickness gold. Connectors can be made from different materials and thicknesses. The plating specifications for all connectors within the Blade Assembly must be reviewed with Microsoft and approved by Microsoft with extra focus on plating that is less than 30u" gold. DIMM connectors must also include lubricant/sealant applied by the connector manufacturer which can remain intact after soldering and other manufacturing processes. The sealant is required to displace any voids in the connector gold plating.

8 Chassis Manager Electrical Specifications

Follow the specification in the schematic for the v1.0 Chassis Manager. v2.1 must fit within the limits of the v1.0 device.

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

Match the specification for the hotswap controller in the schematic from the v1.0 Chassis Manager.

8.2 CPLD Isolation

The Isolation of the CPLD's UARTs from the blades need to be implemented using the same method as the I2C interface. The purpose is to stop current leaking from the blades to the Chassis Manager when the Chassis Manager is off.

8.3 12 Volt Brown out Protection

In no case the removal of the 12V power can cause the system to boot hang the Chassis Manager. The 12V supply voltage may drop or cut in and out and in these cases the Chassis Manager must be able to recover.

8.4 Grounding and Return

Follow the pin mapping included in the schematic from the v1.0 Chassis Manager.

9 Appendix: Commonly Used of Acronyms

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

ACPI – advanced configuration and power interface

AHCI – advanced host controller interface

AHJ – authority having jurisdiction

ANSI – American National Standards Institute

API – application programming interface

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

ASIC – application-specific integrated circuit

BCD – binary-coded decimal

BIOS – basic input/output system

BMC – baseboard management controller

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

CM – Chassis Manager

CMOS – complementary metal–oxide–semiconductor

COLO – co-location

CTS – clear to send

DDR3 – double data rate type 3

DHCP – dynamic host configuration protocol

DIMM – dual inline memory module

DPC - DIMMs per memory channel

DRAM –

DSR – data set ready

DTR – data terminal ready

ECC – error-correcting code

EEPROM - electrically erasable programmable read-only memory

EIA – Electronic Industries Alliance

EMC – electromagnetic compatibility

EMI – electromagnetic interference

FRU – field replaceable unit

FTP – file transfer protocol

GPIO – general purpose input output

GUID – globally unique identifier

HBI – high business intelligence

HCK – Windows Hardware Certification Kit

HMD – hardware monitoring device

HT – hyperthreading

I^2C – inter-integrated circuit

IBC – international building code

IDE – integrated development environment

IEC - International Electrotechnical Commission

IOC – I/O controller

IPMI – intelligent platform management interface

IPsec – IP security

ITPAC – IT pre-assembled components

JBOD – “just a bunch of disks”

KCS – keyboard controller style

L2 – layer 2

LAN – local area network

LFF – large form factor’

LPC – low pin count

LS – least significant

LUN – logical unit number

MAC – media access control

MDC – modular data center containers

MLC – multi-level call

MTBF – mean time between failures

MUX - multiplexer

NUMA – non-uniform memory access

OOB – out of band

OSHA - Occupational Safety & Health Administration

OTS – off the shelf

PCB – printed circuit board

PCIe – peripheral component interconnect express

PCH – platform control hub

PDB – power distribution backplane

PDU – power distribution unit

PECI – Platform Environment Control Interface

Ph-ph – phase to phase

Ph-N – phase to neutral

PNP – plug and play

POST – power-on self-test

PSU – power supply unit

PWM – pulse-width modulation

PXE – preboot execution environment

QDR – quad data rate

QFN – quad flat package no-lead

QPI – Intel QuickPath Interconnect

QSFP – Quad small form-factor pluggable

RAID – redundant array of independent disks

REST - representational state transfer

RM – Rack Manager

RMA – remote management agent

ROC – RAID-on-chip controller

RSS – receive-side scaling

RTS – request to send

RU – rack unit

RxD – received data

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

SATA – serial AT attachment

SCK – serial clock

SCSI – small computer system interface

SDA – serial data signal

SDR – sensor data record

SFF – small form factor

SFP - small form-factor pluggable

SMBUS – systems management bus

SMBIOS – systems management BIOS

SOL – serial over LAN

SPD – Serial Presence Detect

SPI – serial peripheral interface

SSD – solid-state drive

TBP – Tray Backplane

TDP – thermal design power

TMC – Tray Mezzanine Card

TOR – top of rack

TPM – trusted platform module

TxD – transmit data

v1.0 – Original Version of the Chassis Manager

v2.1 – Second Generation of the Chassis Manager

U – rack unit

UART – universal asynchronous receiver/transmitter

UEFI – unified extensible firmware interface

UL – Underwriters Laboratories

UPS – uninterrupted power supply

Vpp – voltage peak to peak

WMI –Windows Management Interface

CONFIDENTIAL