Datacenter-ready Secure Control Module (DC-SCM) and Interface (DC-SCI) for Modular Building Block Architecture (MBA, The Catalyst)

Rajeev Sharma
Director, Software & Data Center Solutions
Oct 23, 2019
Work-in-Progress as of 10/23

In preparation for an OCP specification, this slide deck is a progress report based on continued feedback received on **DC-SCM, DC-SCI, and PCIe Slot Cable Assembly**. It is subject to change without notice.
Feedback from

Lenovo
Wiwynn
Intel
AMD
Dell
Inspur
Quanta
Inventec
Dell
Sanmina
Supermicro

...
Outline

• Motivation, Background, and Review

• Update on MBA and on DC-SCM & DC-SCI

• Received Feedback

• Open-source Activities
Motivation

• Open-source Modular approach for faster TTM

• Modeled after well-known interfaces such as PCIe

• Standardizing Common Blocks and Interfaces

• Target interoperability with ease!
  • High-speed Interconnect (PCIe Gen-4 and Gen-5)
  • Datacenter-ready Security, Control, and Management
For a successful Modular Building Block Architecture, we need:

- Compute Modules (CPU/Memory/IO) (**CMIO**)
- IO & Accelerator Add-in Card Modules (**AIC**)
- Security, Control, and Management (**SCM**)
- Data-plane Control
- A suitable Interconnect

*Consume. Collaborate. Contribute.*
Modular Building Block Architecture (MBA)

- Is based on small building blocks
  - to allow flexible and agile system integration

- Clearly defines input/output ports
  - for interoperability with CPU boards from various suppliers

- Riser-based & Cable-based IO Slots
  - offer flexibility of choice-- ready for PCIe Gen-4 and Gen-5
MBA is a **Catalyst** for interoperable *Innovation*!

**DC-SCM** Facilitates MBA

A standards-ready secure control module, **DC-SCM**, enables the design and deployment of CPU/Memory Complexes and Expansion Chassis to become simply a *routine exercise* based on guidelines from CPU and SoC suppliers!
Examples of Modular Building Block Architecture (MBA)
CPU/Mem/IO Module

- Just the essential Central Compute Elements

- High-speed Memory and

- IO Connectors Close to the SoC

- Get ready for PCIe Gen-5!
DC-SCM

- Everything Else!

- Security, Control, Management
Interconnect

DC-SCM

CPU/Memory/IO

Add-in Card (AIC) Attachment
IO Slot to CPU Board Cable Harness
Ready for High-speed!

Interconnect
AIC Attachment…Cont’d
IO Slot to CPU Board Cable Harness

Gen-Z 4C Connector for attachment to CPU/Memory Module
SFF-TA-1002 4C Scalable Connector

PCIe Slot Connector for an Add-in Card in PCIe CEM form factor

CPU/Mem/IO + DC-SCM

Realized the concept

Single socket processor and SCM as separate elements
PCle Slot to Gen-Z Pin Map
(out for review—not final yet!)

<table>
<thead>
<tr>
<th>Ground pin</th>
<th>Zero volt reference, all tied together</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power pin</td>
<td>Supplies power to the card</td>
</tr>
<tr>
<td>High speed</td>
<td>High speed signals</td>
</tr>
<tr>
<td>Detect</td>
<td>Sense Pin</td>
</tr>
<tr>
<td>Other aux</td>
<td>May be pulled low or sensed by multiple cards</td>
</tr>
<tr>
<td>Reserved</td>
<td>Reserved for future use and no connect</td>
</tr>
</tbody>
</table>
3M™ Twin Ax Assembly

Cable compresses down to 7 mm without compromising performance.
Insertion loss is not significantly impacted by folding.
Adding Capacitors to the PCB

- Added high-frequency by-pass capacitors to power rails
  - to improve Power Integrity and
  - to reduce Simultaneous Switching Noise (SSN) effects
Two Assemblies in 1U Height (44.5 mm)

- Same PCB assembly for both assemblies to minimize number of SKUs
- Two PCIe Slots in 1U chassis
- Cable assembly folded onto itself
Enabled Options
Add-in Card (Riser Attached)
Enabled Options
Add-in Card (Cable Attached) *Ready for High-speed!*

**Closer approximation of PCIe Gen-5**

*Consume. Collaborate. Contribute.*
Realized Both Options (for Add-in Card attachment)
• Cable connected PCIe cards
• Riser connected PCIe cards
The Datacenter-Ready Secure Control Module (*DC-SCM*)

&

The Datacenter-Ready Secure Control Interface (*DC-SCI*)
DC-SCM (in a nutshell)

- DC-SCM is “the heart of the motherboard” when we extract CPU(PCH), Memory, and IO Slots

- Given a traditional 1S, 2S, 4S, … Motherboard, extract CPU/PCH, DIMM Slots, IO Slots, and the associated VRs, Clock Drivers, and Reset Circuitry, and move them to a new Module

- The residual is the DC-SCM which will include everything else such as BMC, RoT, Flash, and PSU control along with optional Boot SSD and connectors for Fan control
DC-SCM (Motivation)

- Don’t reinvent the wheel with each new server design
- Unify the solution to support multiple architectures

“Same as before” with F/W, S/W, & Services--maintaining the established tools and solutions experience with the same management, power sequencing, reset, FRU ID, VPD, ...

A vehicle to drive a common Boot, Monitoring, Control, and Remote Debug procedures for Xeon, EPYC, ARM64, and Power Servers with the same firmware, diagnostic tools, manufacturing tools
Software Standardization

Collaborating with CPU suppliers, Open Computing Project community (OCP), Linux Foundation, and Open System Firmware (OSF) to standardize the hardware and software for OpenBMC with RedFish interface and for the system BIOS/UEFI based on EDK-II
An example of DC-SCM
DC-SCM

- Receives Power
- Remote Control at Cloud Scale
  - CPU/Memory/IO Module (Xeon, EPYC, ARM64)
  - Expansion Chassis (JBOD, JBOG)
  - Fans, PSUs

Includes

- BMC and Rack Management Interface
- Flash Devices (all Firmware)
- RoT and TPM for Security
- Optional Boot SSD
- Remote, at-scale Debug

Another Example of DC-SCM (OCP NIC3 Form Factor)

SFF-TA-1002 4C+ 168-pin, Scalable Connector

With a mechanical key to avoid plugging in the NIC 3.0 connector & vice-versa

<table>
<thead>
<tr>
<th>Form Factor</th>
<th>Width</th>
<th>Depth</th>
<th>Primary Connector</th>
<th>Secondary Connector</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFF</td>
<td>W1 = 76 mm</td>
<td>L = 115 mm</td>
<td>“4C+” 168 pins</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Realized, Wiwynn has made
An Example of SCM Expander Connectors

<table>
<thead>
<tr>
<th>Item</th>
<th>Function</th>
<th>Qty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M2 socket</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>TPM connector</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>SPI socket</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>RoT connector</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>VGA cable connector</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>NCSI cable connector</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Front panel cable connector</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>FAN cable connector</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>BMC debug UART Pin header</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>HOST UART Pin header</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Auxiliary UART Pin header</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Reserved UART Pin header</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Pin header</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>PSU cable connector</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>JATG cable connector</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Reserved USB cable header</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>ID LED header</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>Battery</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>I2C Header</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>Golden Finger</td>
<td>1</td>
</tr>
</tbody>
</table>

SFF-TA-1002 4C+ 168-pin, Scalable Connector
DC-SCM to CPU/Mem Module Interface (DC-SCI)

Pinout and definition
Pin Reduction via SGPIO

Dedicated Signals through DC-SCI
Signals conditioned through the SCM CPLD

rsGPI: Signals Multiplexed via CPU_CPLD, presented in Registers at SCM_CPLD, READ by BMC
rsGPO: BMC WRITEs into SCM_CPLD Registers, signals Multiplexed via sGPO, CPU_CPLD de-Muxes them

sGPI: Signals Multiplexed via CPU_CPLD, de-Muxed by SCM_CPLD for BMC to use directly
sGPO: Signals Multiplexed via BMC_CPLD, de-Muxed by CPU_CPLD to various places
Win-win!

DC-SCM accelerates deploying servers from various suppliers into the datacenter
Standardizing DC-SCI for ease of integration into various datacenters
Flexibility to use BMC and RoT chips of choice on any platform

From a Datacenter point of view:
- with one DC-SCM, a datacenter may support multiple variants of servers (AMD-, Xeon-, ARM64-, Power-based 1S, 2S, 4S, ...) and expansion chassis, JBODs, JBOG, JBOFs, ...

From OEM/ODMs’ point of view:
- A product will fit datacenters of various CSPs or Hyperscalers

If we are smart, one DC-SCM may enable supplier products into different DC types; otherwise, each Datacenter Provider may have its own version of DC-SCM.
Thanks!
DC-SCM (Ingredients)

- Most MBA building blocks are stateless
- The secure control module (DC-SCM) includes all system related components (other than CPU/Mem/IO) that are normally present on Motherboards
- Baseboard Management Controller (BMC), Realtime Clock (RTC), FAN/PSU Control, Root of Trust Chip (RoT: Cerberus/Other and the associated circuitry), BIOS & BMC Flash, and the Boot Device
- SCM holds control bits secured (no firmware on CPU/Mem Module)
Call to Action

Design your Servers, Expansion Chassis, JBODs, JBOGs, JBOFs, multi-Server Chassis, etc. with DC-SCI connector in mind.

Make your solution Datacenter-Ready!

Join the effort to enhance DC-SCM and DC-SCI

https://www.opencompute.org/projects/server
DC-SCI (pinout implemented in PoC)

Connector Type:
SFF-TA-1002 4C+ 168-pin, Scalable Connector

DC-SCM Connector:
Gold-finger

CPU/Memory/IO Module Connector:
Right-angled or Vertical Receptacle

<table>
<thead>
<tr>
<th>Function</th>
<th>168 Pin</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOS SPI</td>
<td>6</td>
</tr>
<tr>
<td>BMC_PCIE</td>
<td>7</td>
</tr>
<tr>
<td>CPLD_SGPIO</td>
<td>9</td>
</tr>
<tr>
<td>Critical GPIO</td>
<td>30</td>
</tr>
<tr>
<td>GND</td>
<td>23</td>
</tr>
<tr>
<td>I2C ALERT</td>
<td>13</td>
</tr>
<tr>
<td>JTAG</td>
<td>5</td>
</tr>
<tr>
<td>LPC &amp; ESPI</td>
<td>12</td>
</tr>
<tr>
<td>M2_CLK</td>
<td>2</td>
</tr>
<tr>
<td>M2_PCIE</td>
<td>4</td>
</tr>
<tr>
<td>PECI</td>
<td>2</td>
</tr>
<tr>
<td>POWER</td>
<td>5</td>
</tr>
<tr>
<td>Sequence</td>
<td>3</td>
</tr>
<tr>
<td>I2C</td>
<td>26</td>
</tr>
<tr>
<td>SPI</td>
<td>0</td>
</tr>
<tr>
<td>UART</td>
<td>2</td>
</tr>
<tr>
<td>USB</td>
<td>4</td>
</tr>
<tr>
<td>Remote Debug</td>
<td>4</td>
</tr>
<tr>
<td>TPM</td>
<td>2</td>
</tr>
<tr>
<td>FPGA SPI</td>
<td>4</td>
</tr>
<tr>
<td>RSVRD</td>
<td>5</td>
</tr>
<tr>
<td>PSU Connector</td>
<td>0</td>
</tr>
</tbody>
</table>
DC-SCM (Form Factors)

• The SCM is small enough to fit anywhere in the Chassis

• Flexible as development vehicle or for Expansion Chassis:
  1. Cabled to Chassis Edge for external connections such as RJ45, Serial Console, and cabled internally for Fans, PSUs, ...

• IO connectors at the Edge of the Chassis; DC-SCI for interfacing to CPU/Mem Module:
  2. A plug-in module like OCP NIC-3: co-planar to CPU/Mem Module
  3. A plug-in module like low-profile PCIe cards: plugs vertically into the CPU/Mem module

CPU/Mem/IO
Ready for High-speed!

Started with concepts
Add-in Card (AIC) Attachment
IO Slot to CPU Board Cable Harness
Examples of Modular Designs

Pin Reduction Techniques

Various techniques for reducing required BMC and DC-SCI pins

CPU Module CPLD serially shifts GPIOs to/from SCM CPLD; SCM CPLD replicates them on pins connected to BMC GPIOs. (for latency-insensitive signals that need “exact” replication at BMC for firmware compatibility)

CPU Module CPLD serially shifts GPIs to SCM CPLD; In response to an Event/Interrupt, BMC READs GPIOs via SPI (or I2C) (for latency-insensitive signals and signals that don’t need “exact” replication at BMC for firmware compatibility)

BMC WRITEs GPOs via SPI (or I2C) into SCM CPLD. SCM CPLD serially shifts them onto CPU Module CPLD. CPU Module CPLD replicates them as parallel signals to go to various places on the CPU/Memory Module. (for latency-insensitive signals and signals that don’t need “exact” replication at BMC for firmware compatibility)