



Data Center-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



Platform Teams Present

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



SERVER

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



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!*

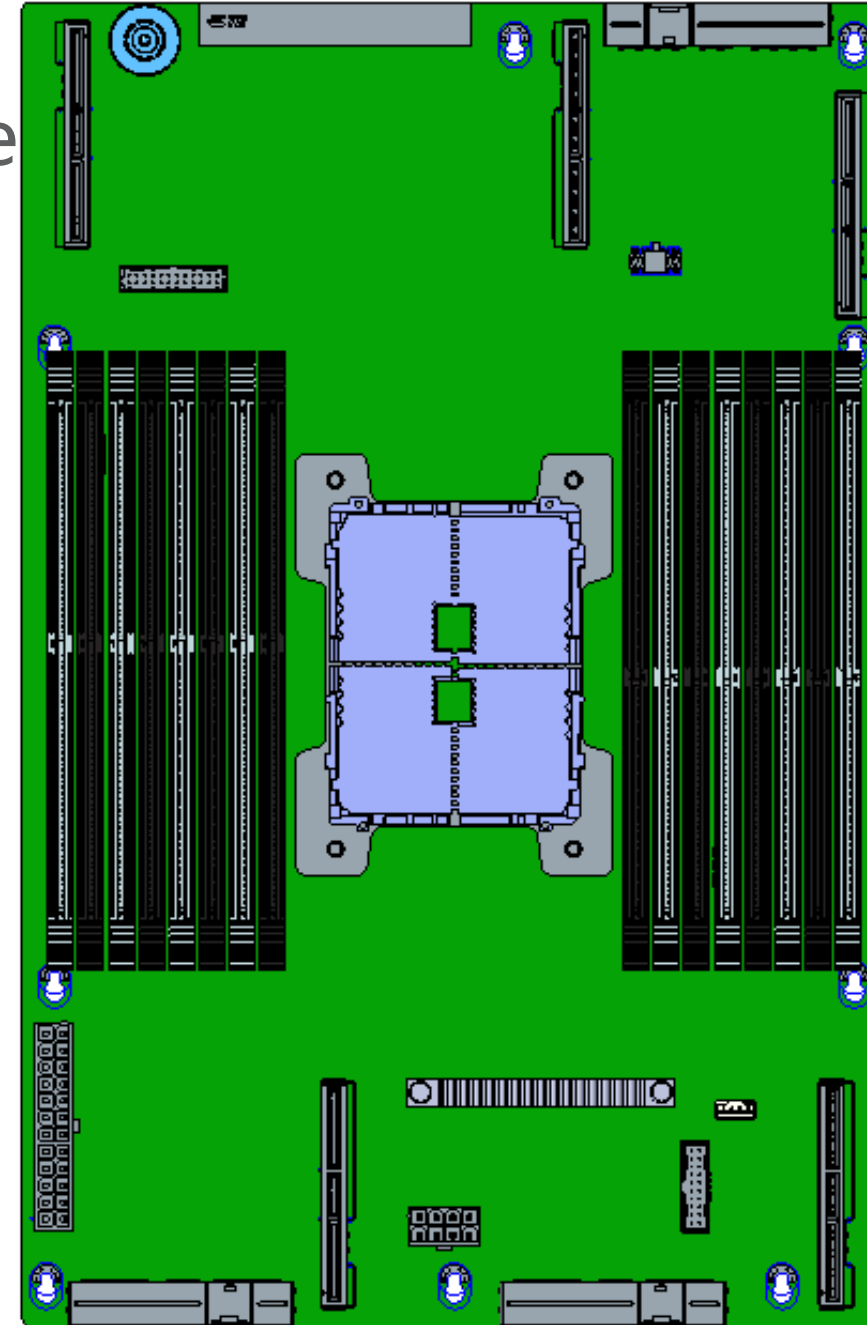
DC-SCM Facilitates MBA

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)

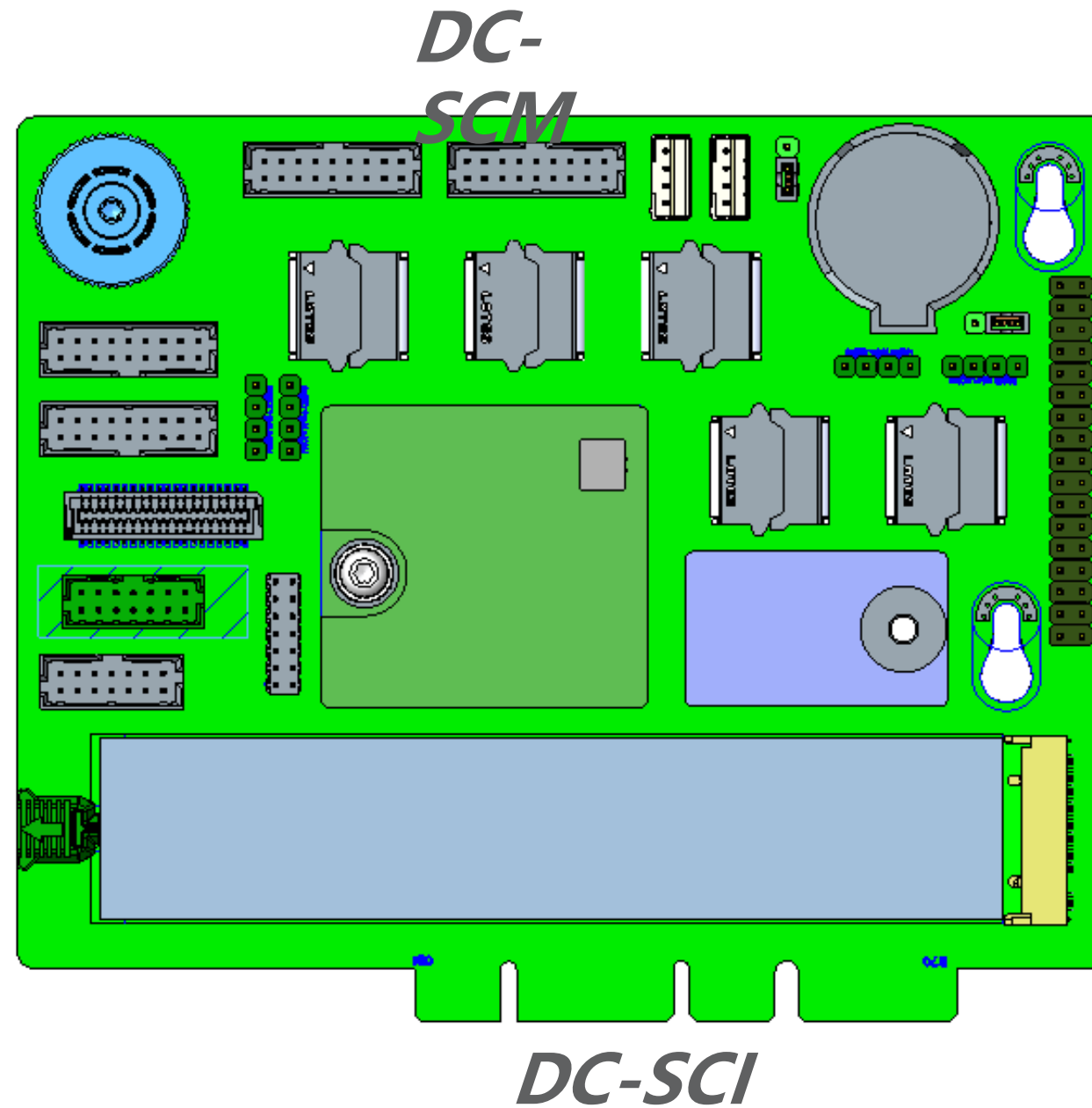
CPU/Mem/IO Module

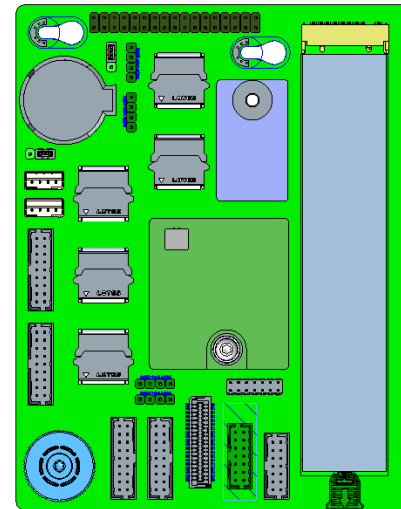
- Just the essential Central Compute Element
- High-speed Memory and
- **IO Connectors Close to the SoC**
- Get ready for *PCIe Gen-5!*



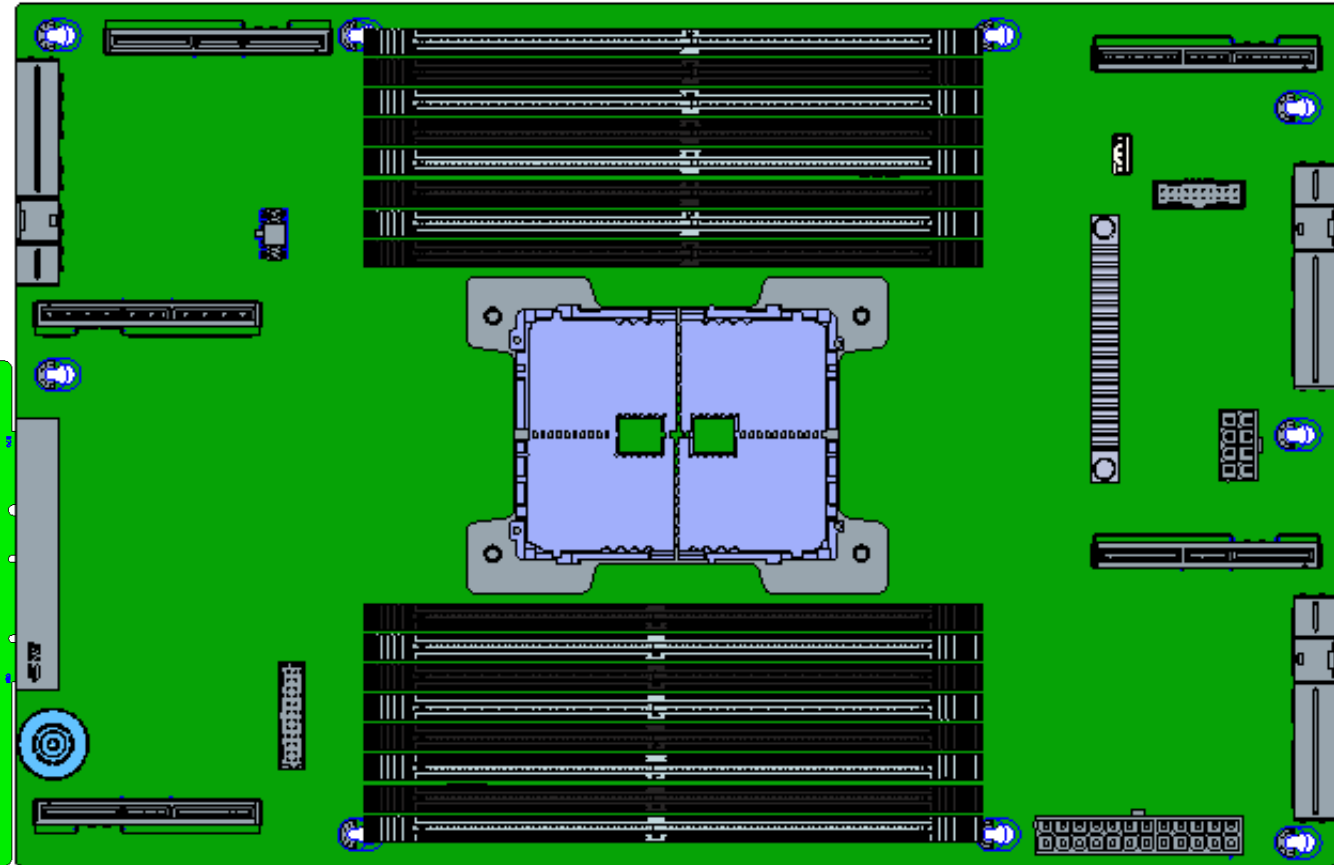
DC-SCM

- Everything Else!
- Security, Control, Management

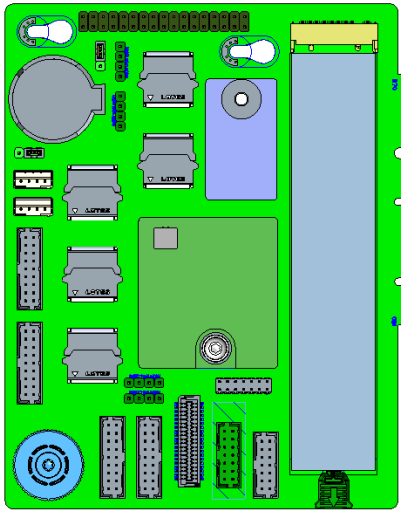




*DC-
SCM*

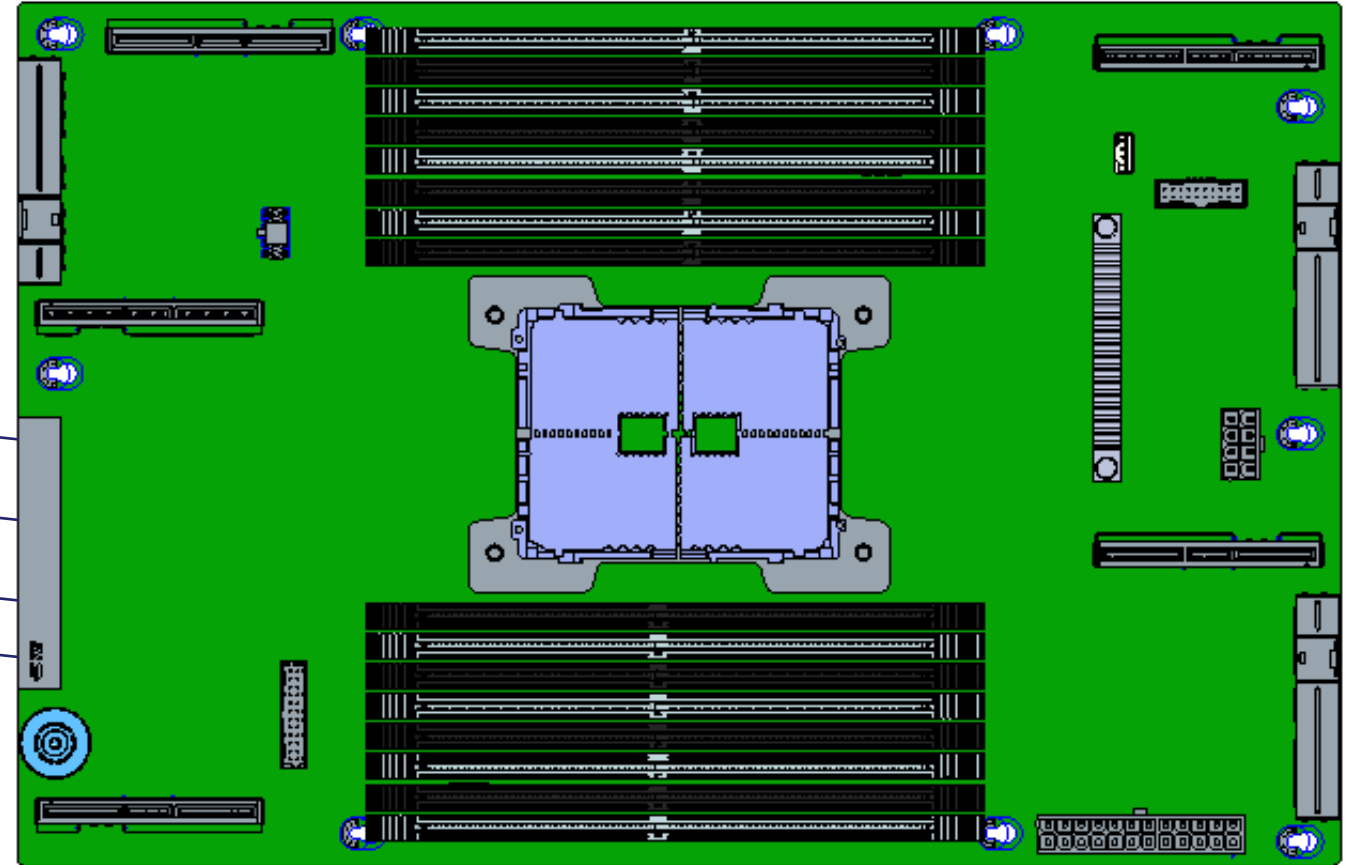


*CPU/Memory/I
O*



*DC-
SCM*

Interconnect

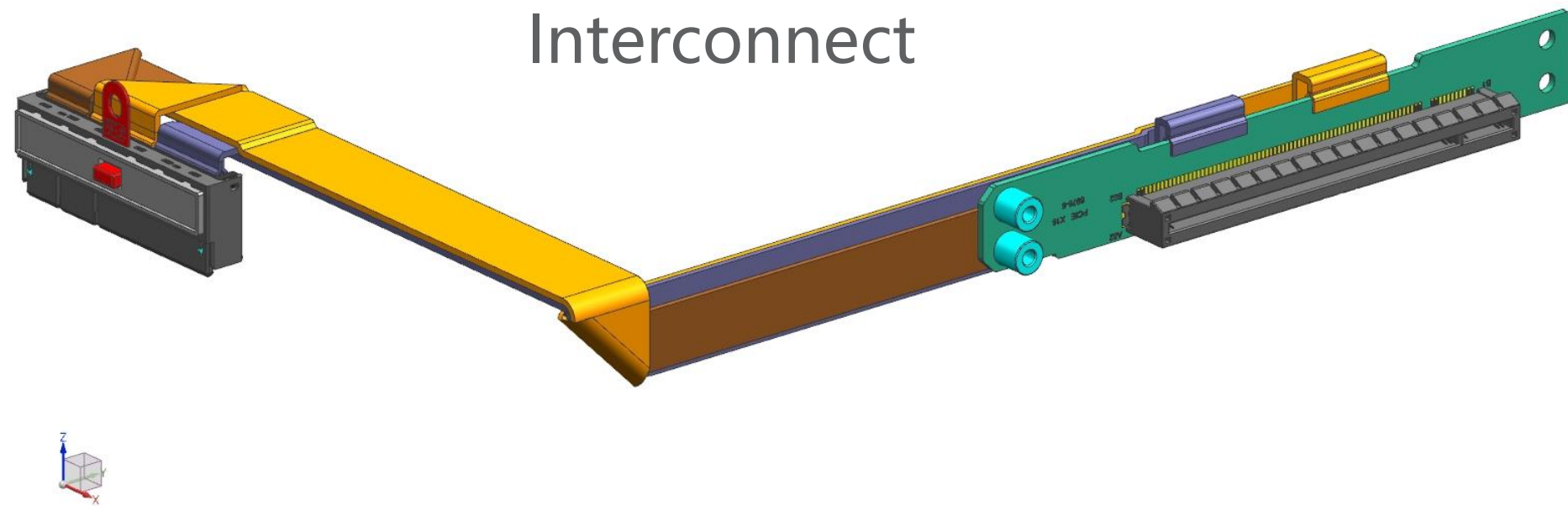


*CPU/Memory/I
O*

Add-in Card (AIC) Attachment

IO Slot to CPU Board Cable Harness

Ready for High-speed!



AIC Attachment...Cont'd

IO Slot to CPU Board Cable Harness



**Gen-Z 4C Connector for
attachment to
CPU/Memory Module**

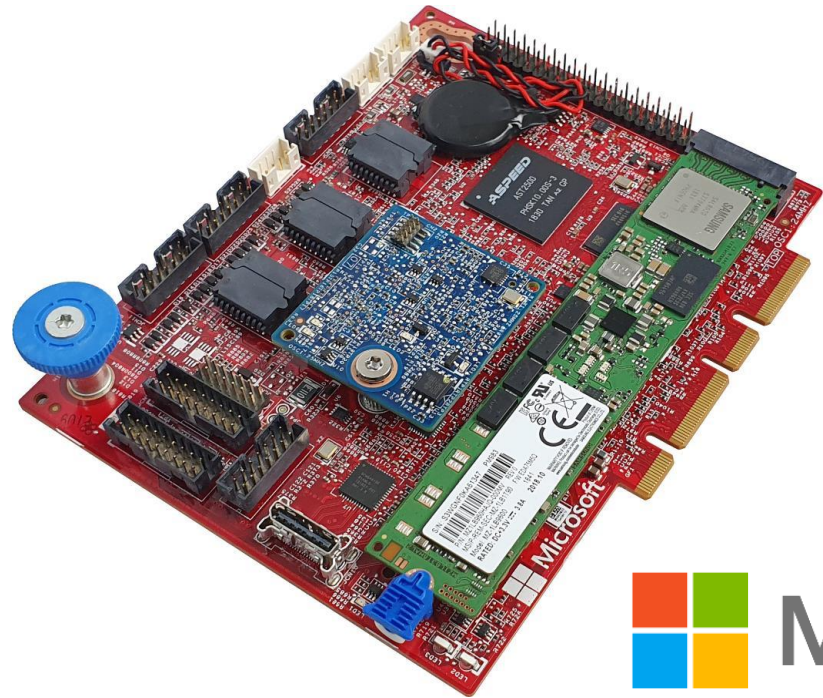
**SFF-TA-1002 4C Scalable
Connector**

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

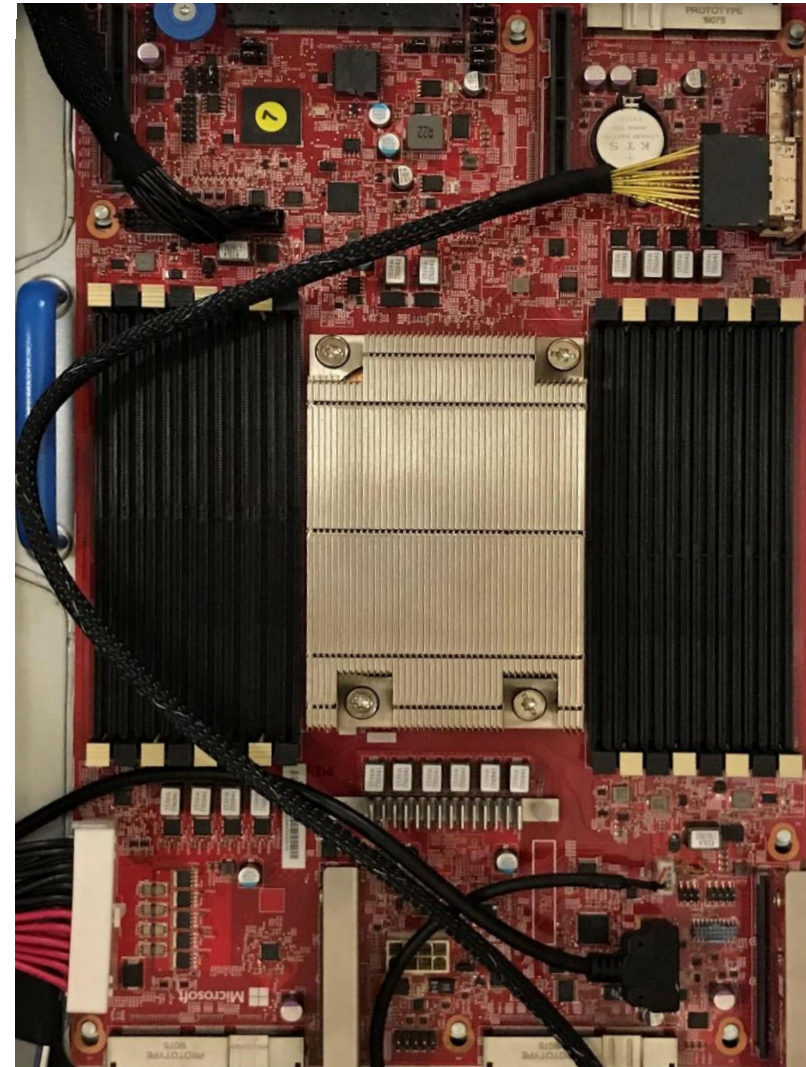
Consume. Collaborate. Contribute.

CPU/Mem/IO + DC-SCM

Realized the concept



Single socket processor and SCM as separate elements



PCIe Slot to Gen-Z Pin Map

(out for review – not final yet!)

ASSEMBLY PINOUT TABLE			ASSEMBLY PINOUT TABLE		
PCIe Side-A		Gen-Z Side-A	PCIe Side-B		Gen-Z Side-B
P1	Description	P2	P1	Description	P2
A1	PRSN1_1	A1	B1, B2, B3	P12V	B1/B2/B3/B4/B5/B6
A2, A3	P12V	B1/B2/B3/B4/B5/B6	B4	GND	GND
A4	GND	A5	B5	SMCLK	A7
A5	JTAG2	A42	B6	SMDAT	A8
A6	JTAG3	A2	B7	GND	B13
A7	JTAG4	A3	B8	P3.3V	A69/B68/B69
A8	JTAG5	A4	B9	JTAG1	A68
A9, A10	P3.3V	A69/B68/B69	B10	P3.3V_AUX	B11
A11	PWRGD	B10	B11	WAKE	A70
A12	GND	A6	B12	CLKREQ	A11
A13	REFCLK_P	B15	B13	GND	B16
A14	REFCLK_N	B14	B14	HS0N_0(TX)	B17
A15	GND	A13	B15	HSOP_0(TX)	B18
A16	HSIN_0(RX)	A17	B16	GND	GND
A17	HSIP_0(RX)	A18	B17	NC_PRSNT_2_B17	NC
A18	GND	A16	B18	GND	B19
A19	NC_RSVD_1	NC	B19	HS0N_1(TX)	B20
A20	GND	A19	B20	HSOP_1(TX)	B21
A21	HSIN_1(RX)	A20	B21, B22	GND	B22
A22	HSIP_1(RX)	A21	B23	HS0N_2(TX)	B23
A23, A24	GND	A22	B24	HSOP_2(TX)	B24
A25	HSIN_2(RX)	A23	B25, B26	GND	B25
A26	HSIP_2(RX)	A24	B27	HS0N_3(TX)	B26
A27, A28	GND	A25	B28	HSOP_3(TX)	B27
A29	HSIN_3(RX)	A26	B29	GND	B28
A30	HSIP_3(RX)	A27	B30	PWRBRK	B8
A31	GND	A28	B31	PRSN1_2_B31	A12
A32	NC_RSVD_2	NC	B32	GND	B29
A33	NC_RSVD_3	NC	B33	HS0N_4(TX)	B30
A34	GND	A29	B34	HSOP_4(TX)	B31
A35	HSIN_4(RX)	A30	B35, B36	GND	B32
A36	HSIP_4(RX)	A31	B37	HS0N_5(TX)	B33
A37, A38	GND	A32	B38	HSOP_5(TX)	B34
A39	HSIN_5(RX)	A33	B39, B40	GND	B35
A40	HSIP_5(RX)	A34	B41	HS0N_6(TX)	B36
A41, A42	GND	A35	B42	HSOP_6(TX)	B37
A43	HSIN_6(RX)	A36	B43, B44	GND	B38
A44	HSIP_6(RX)	A37	B45	HS0N_7(TX)	B39
A45, A46	GND	A38	B46	HSOP_7(TX)	B40
A47	HSIN_7(RX)	A39	B47	GND	B41
A48	HSIP_7(RX)	A40	B48	PRSN1_2_B48	B42
A49	GND	A41	B49	GND	B43
A50	NC_RSVD_5	NC	B50	HS0N_8(TX)	B44
A51	GND	A43	B51	HSOP_8(TX)	B45
A52	HSIN_8(RX)	A44	B52, B53	GND	B46
A53	HSIP_8(RX)	A45	B54	HS0N_9(TX)	B47
A54, A55	GND	A46	B55	HSOP_9(TX)	B48
A56	HSIN_9(RX)	A47	B56, B57	GND	B49
A57	HSIP_9(RX)	A48	B58	HS0N_10(TX)	B50
A58, A59	GND	A49	B59	HSOP_10(TX)	B51
A60	HSIN_10(RX)	A50	B60, B61	GND	B52
A61	HSIP_10(RX)	A51	B62	HS0N_11(TX)	B53
A62, A63	GND	A52	B63	HSOP_11(TX)	B54
A64	HSIN_11(RX)	A53	B64, B65	GND	B55
A65	HSIP_11(RX)	A54	B66	HS0N_12(TX)	B56
A66, A67	GND	A55	B67	HSOP_12(TX)	B57
A68	HSIN_12(RX)	A56	B68, B69	GND	B58
A69	HSIP_12(RX)	A57	B70	HS0N_13(TX)	B59
A70, A71	GND	A58	B71	HSOP_13(TX)	B60
A72	HSIN_13(RX)	A59	B72, B73	GND	B61
A73	HSIP_13(RX)	A60	B74	HS0N_14(TX)	B62
A74, A75	GND	A61	B75	HSOP_14(TX)	B63
A76	HSIN_14(RX)	A62	B76, B77	GND	B64
A77	HSIP_14(RX)	A63	B78	HS0N_15(TX)	B65
A78, A79	GND	A64	B79	HSOP_15(TX)	B66
A80	HSIN_15(RX)	A65	B80	GND	B67
A81	HSIP_15(RX)	A66	B81	PRSN1_2_B81	B70
A82	GND	A67	B82	GND	GND
NC	NC_MGMT_RST	A9	NC	NC_MFG	B7
NC	NC_LED/ACTIVITY	A10	NC	NC_DUALPORTEN	B9
NC	NC_REFCLK1_P	A14	NC	NC_PWRDIS	B12
NC	NC_REFCLK1_N	A15			

Ground pin
Power pin
High speed pin
Detect
Other aux
Reserved

Zero volt reference, all tied together

Supplies power to the card

High speed signals

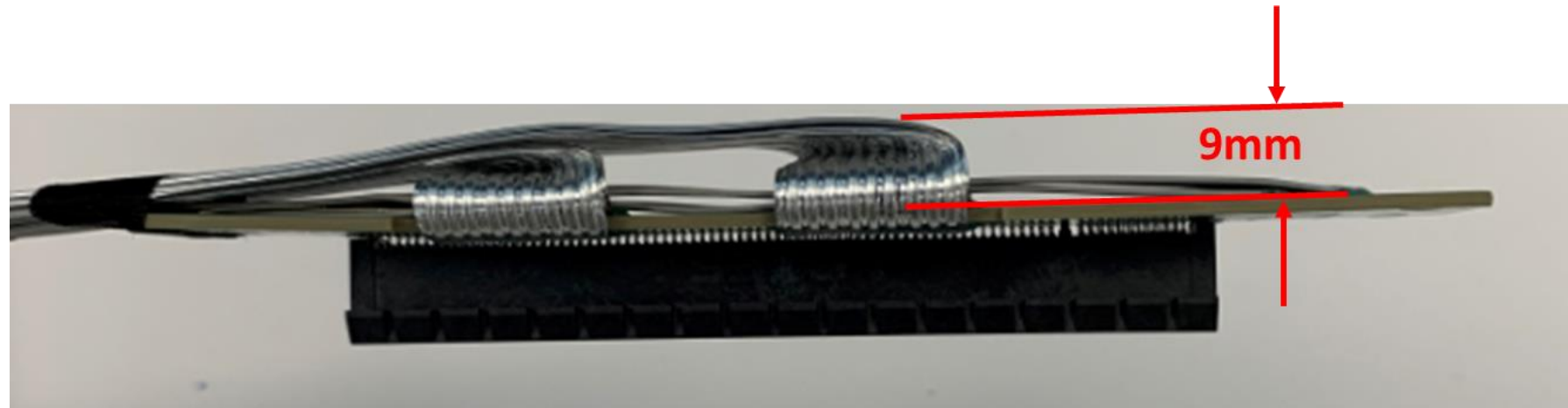
Sense Pin

May be pulled low or sensed by multiple cards

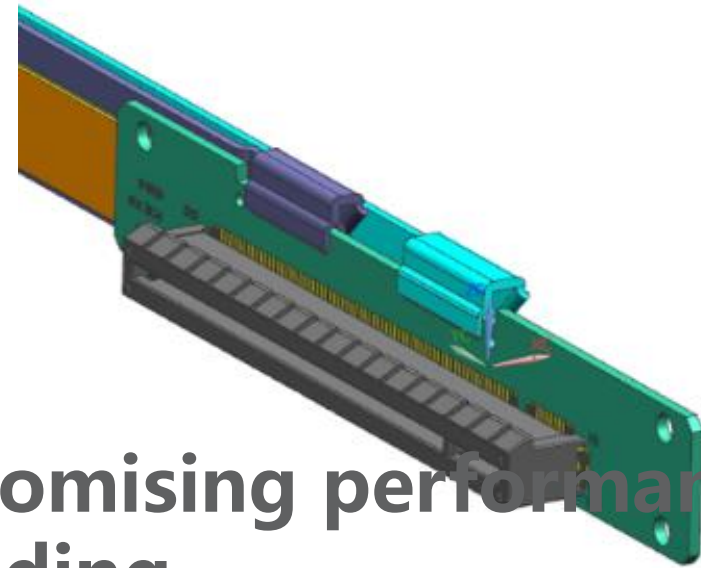
Reserved for future use and no connect



3M™ Twin Ax Assembly



3M



Cable compresses down to 7 mm without compromising performance
Insertion loss is not significantly impacted by folding

Adding Capacitors to the PCB

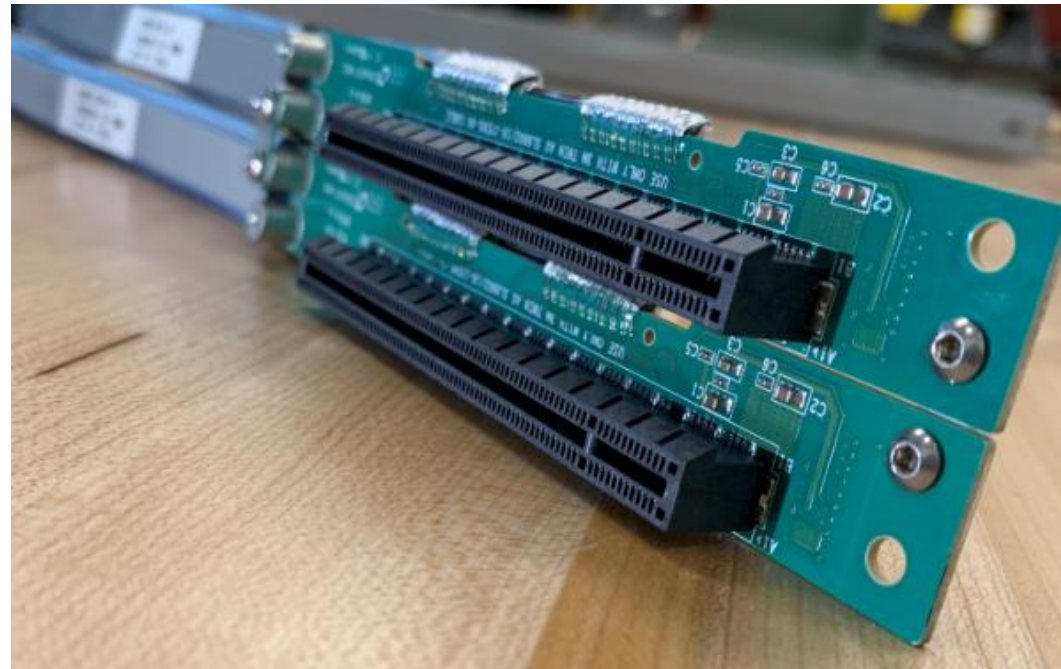
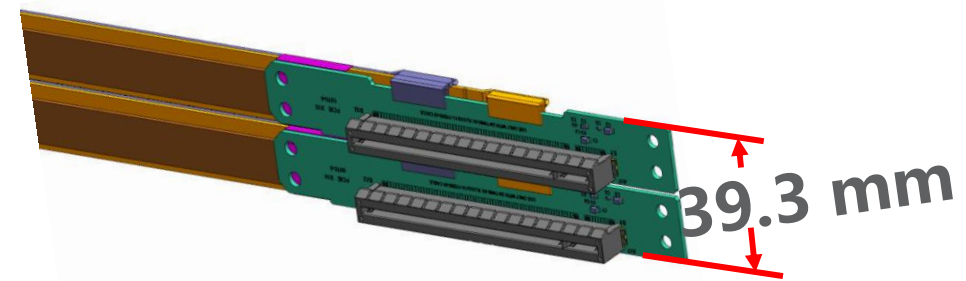


- Added high-frequency by-pass capacitors to power rails
 - to improve Power Integrity and
 - to reduce Simultaneous Switching Noise (SSN) effects



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



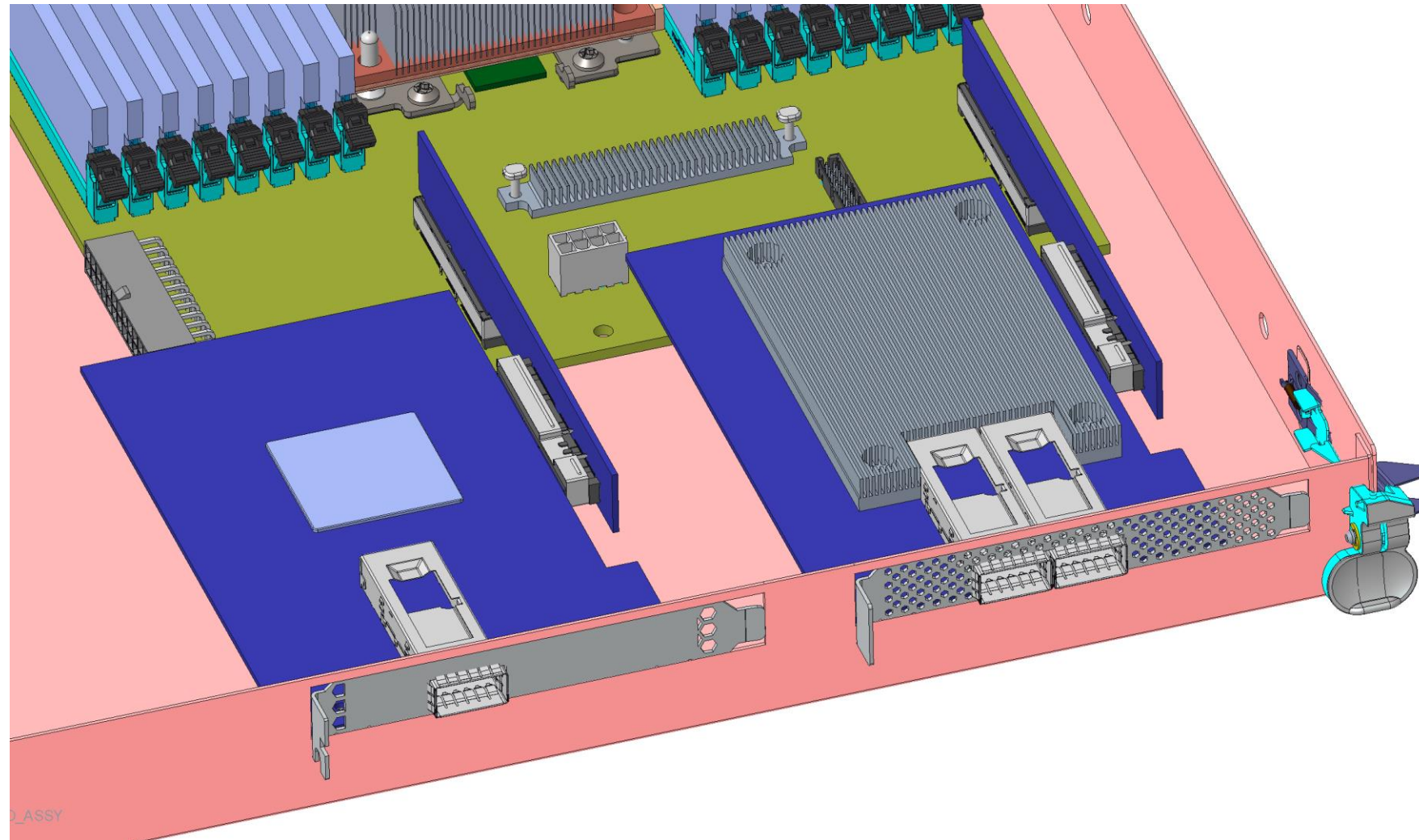
Consume. Collaborate. Contribute.



OPEN
Compute Project

Enabled Options

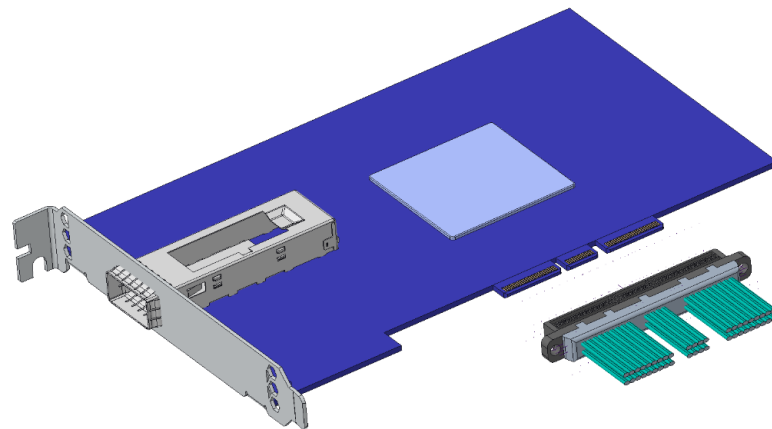
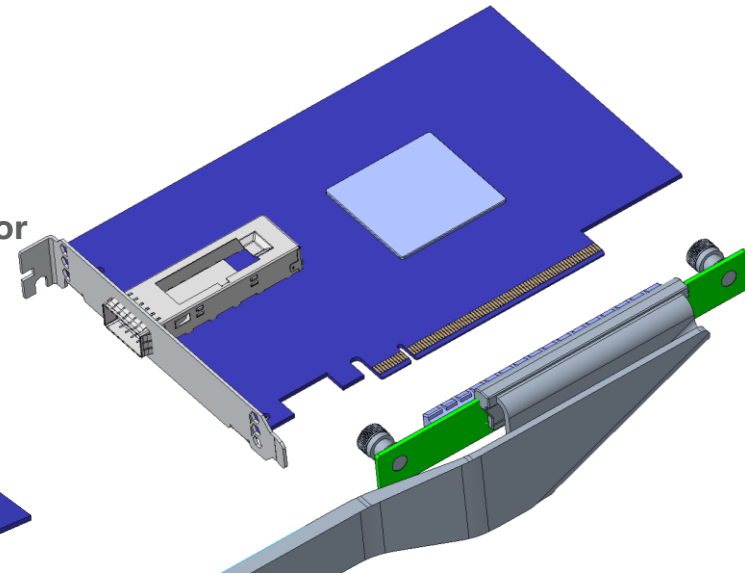
Add-in Card (**Riser** Attached)



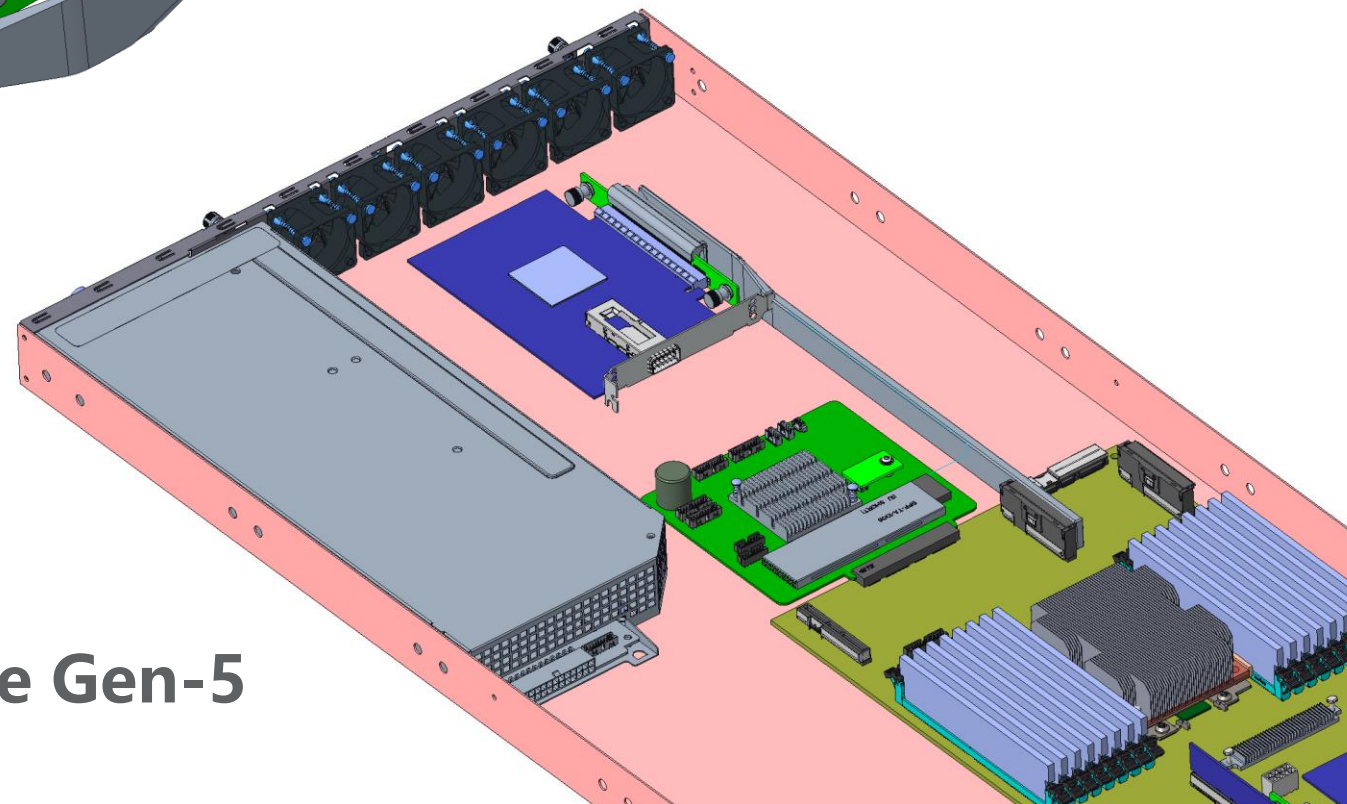
Enabled Options

Add-in Card (**Cable** Attached) *Ready for High-speed!*

PCIe CEM
Cabled Connector



Native
SFF-TA-1002 4C
Cabled Receptacle



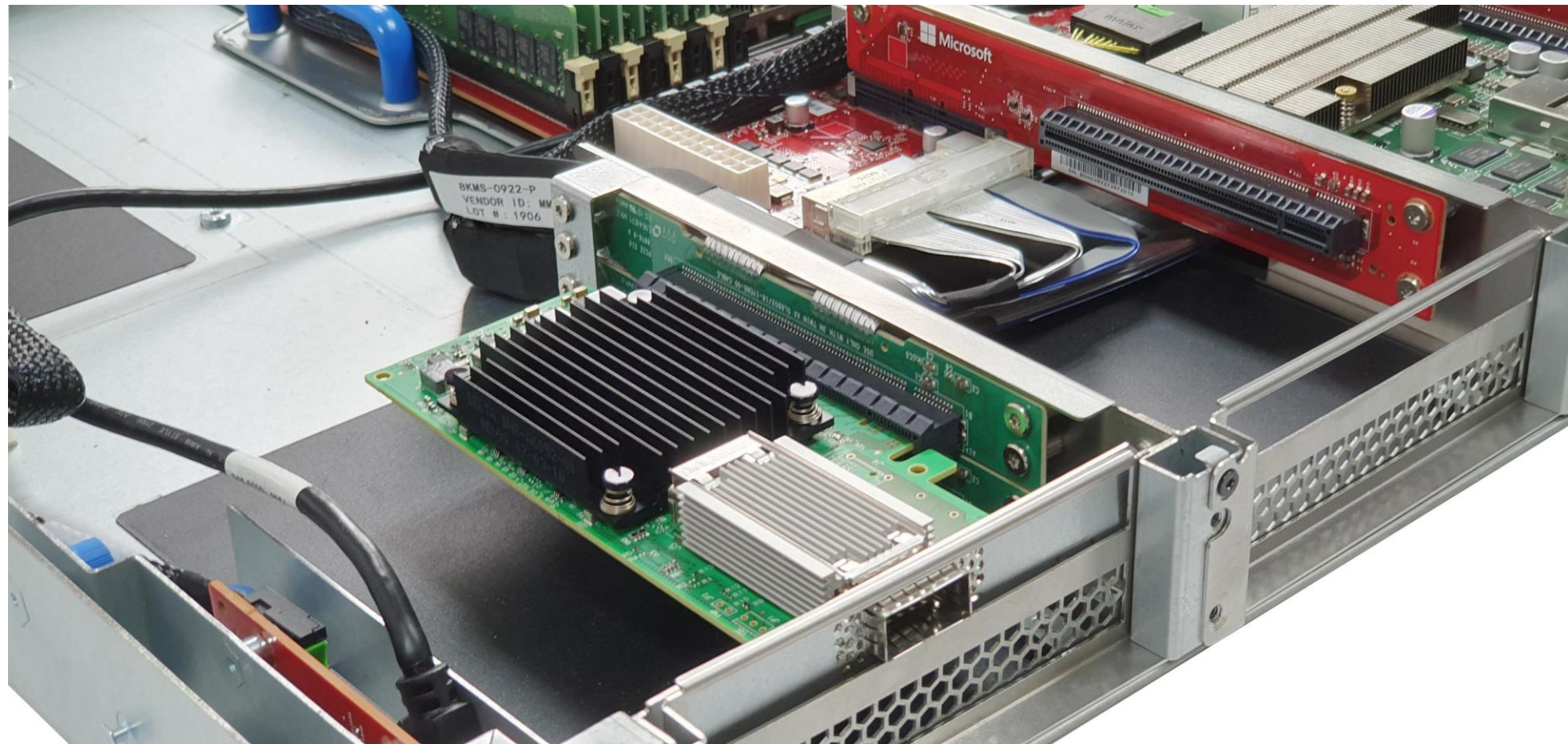
Closer approximation of PCIe Gen-5

Consume. Collaborate. Contribute.



Realized Both Options (for Add-in Card attachment)

3M





- 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

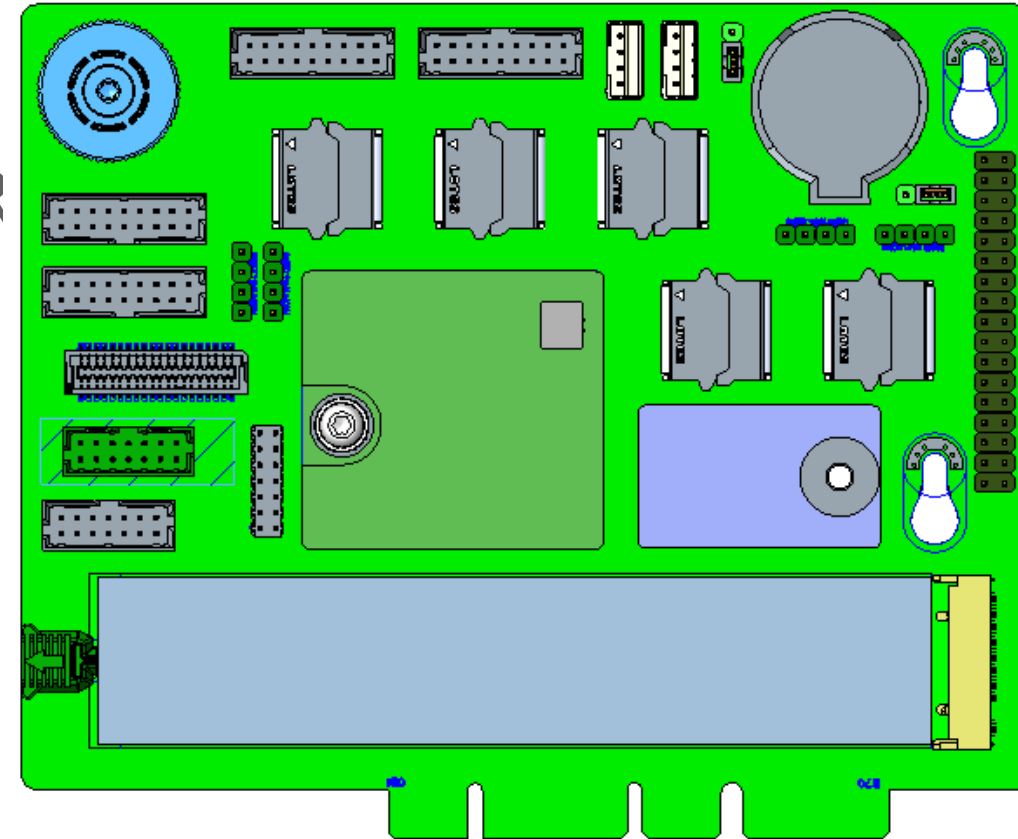
DC-SCM

- Receives Power
- Remote Control at Cloud Scale
 - CPU/Memory/IO Module (Xeon, EPYC, ARM)
 - 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

DC-SCM

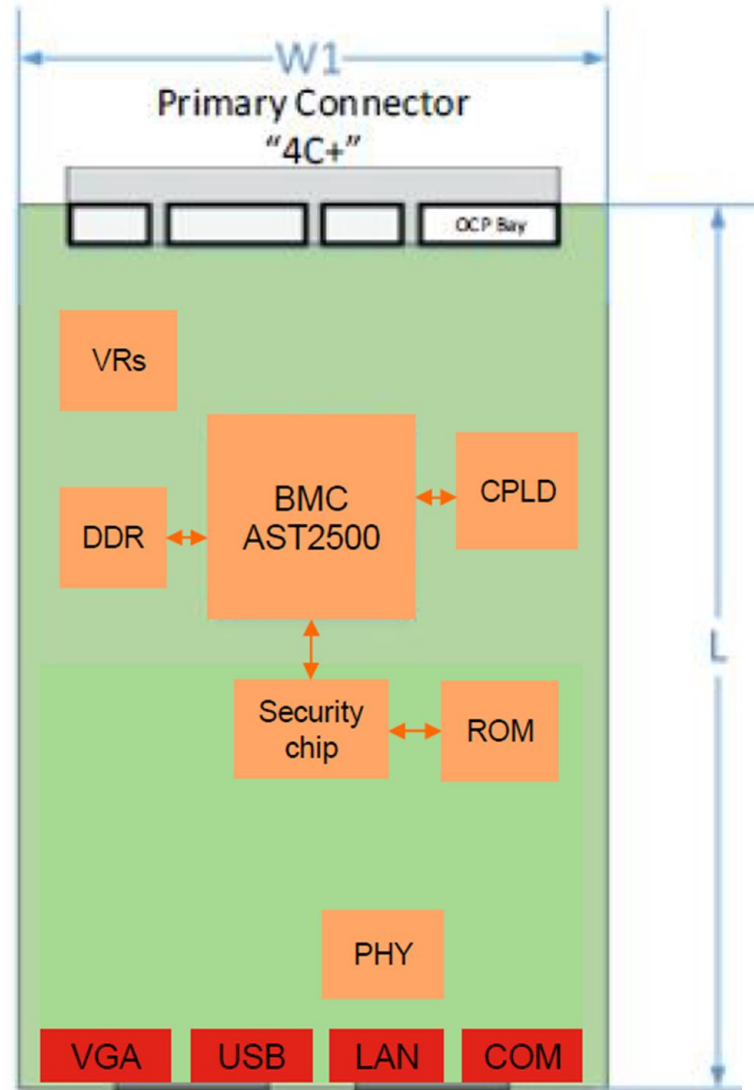


DC-SCI

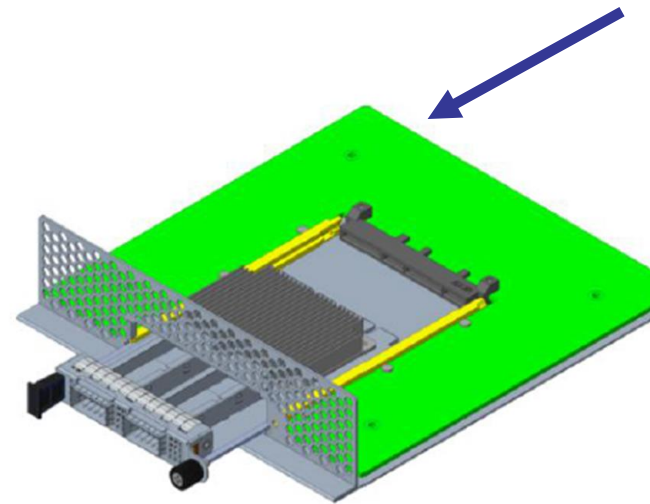
SFF-TA-1002 4C+
168-pin, Scalable
Connector

Another Example of DC-SCM

(OCP NIC3 Form Factor)



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



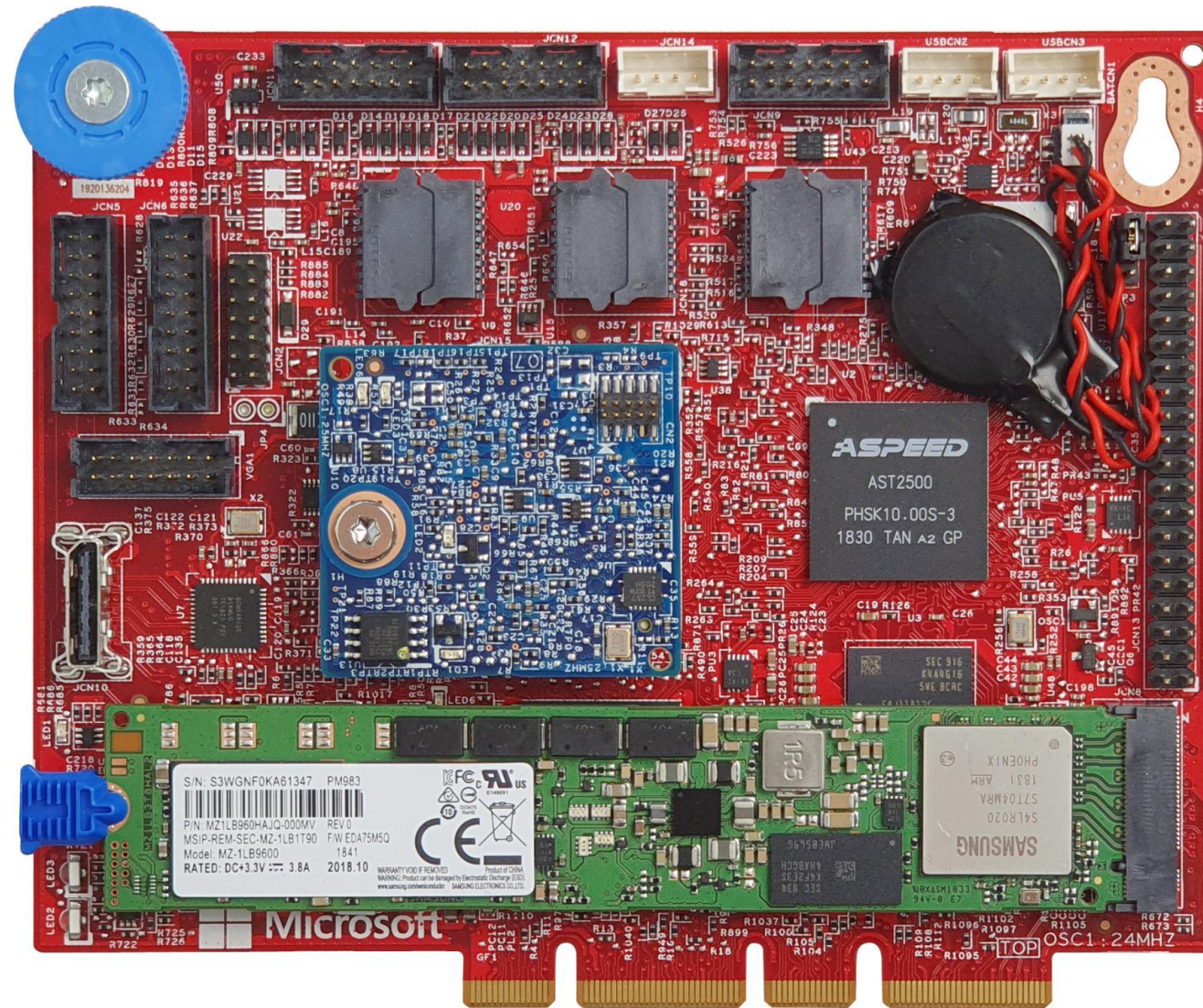
**With a mechanical key to
avoid plugging in the NIC 3.0
connector & vice-versa**

Form Factor	Width	Depth	Primary Connector	Secondary Connector
SFF	W1 = 76 mm	L = 115 mm	"4C+" 168 pins	N/A

DC-SCM

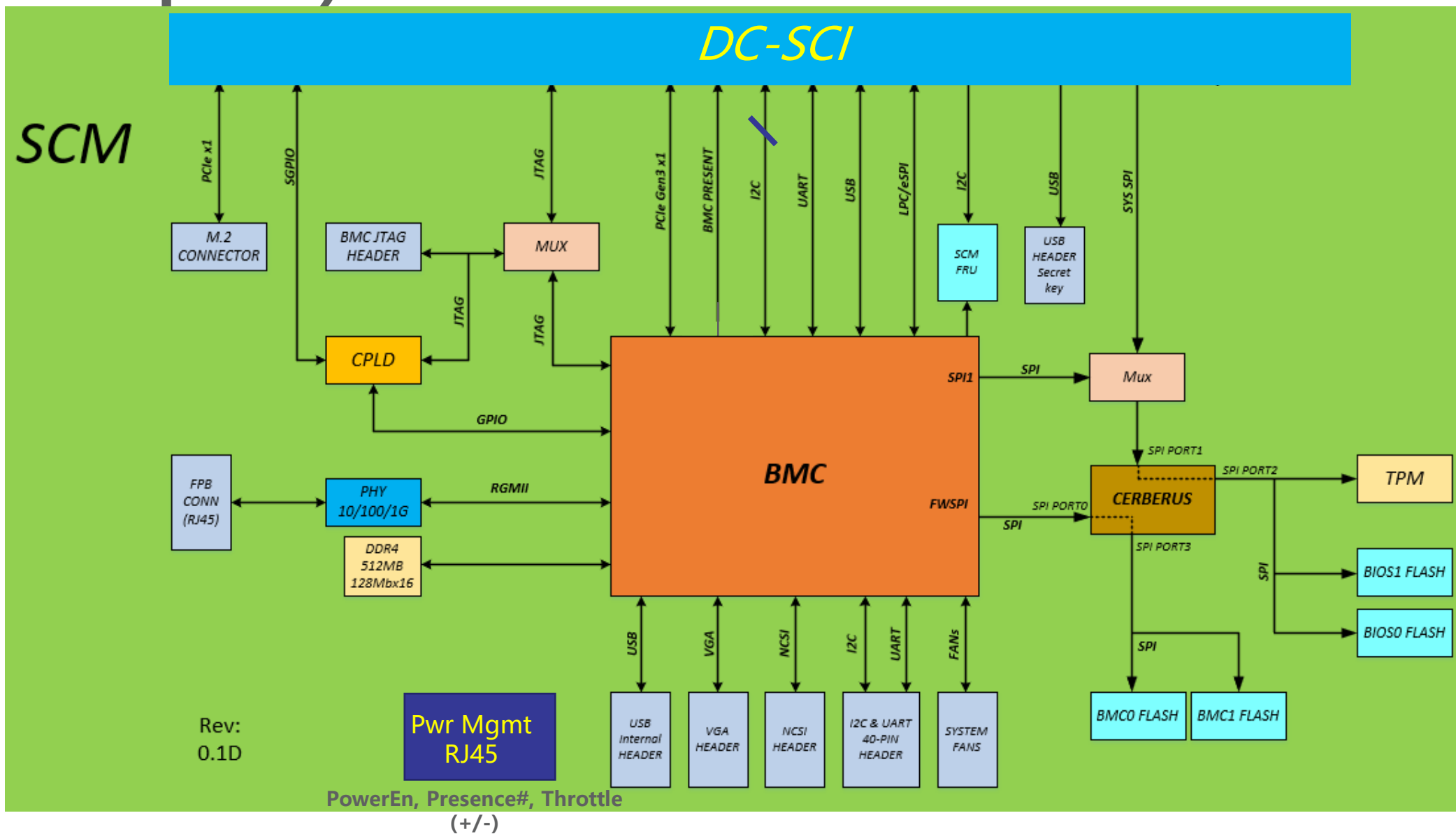
DC-SCM

Realized, Wiwynn has made



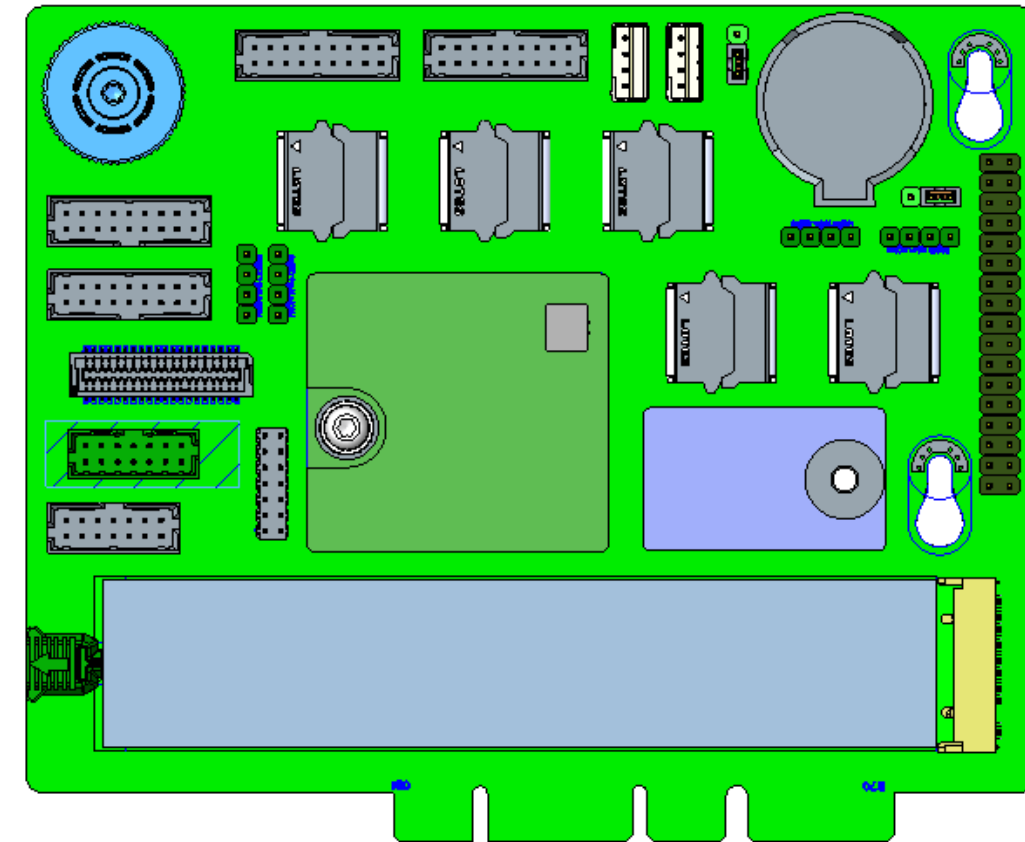
DC-SCI

DC-SCM (Block Diagram Example)



An Example of SCM Expander Connectors

Function	Qty
M2 socket	1
TPM connector	1
SPI socket	5
RoT connector	1
VGA cable connector	1
NCSI cable connector	1
Front panel cable connector	1
FAN cable connector	2
BMC debug UART Pin header	1
HOST UART Pin header	1
Auxiliary UART Pin header	1
Reserved UART Pin header	1
Pin header	2
PSU cable connector	2
JATG cable connector	1
Reserved USB cable header	1
ID LED header	1
Battery	1
I2C Header	1
Golden Finger	1



**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

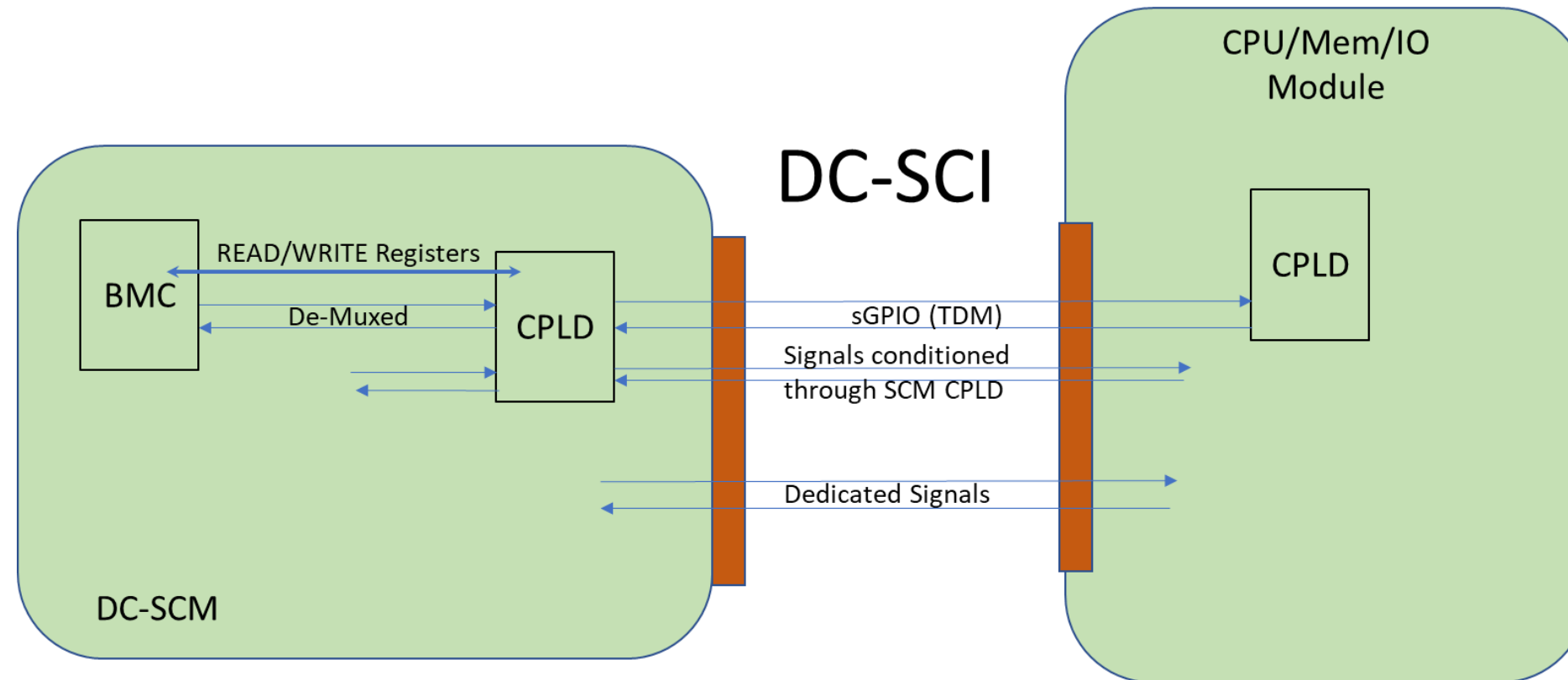
Signals conditioned through the SCM CPLD

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



SFF-TA-1002 4C+
168-pin, Scalable Connector
Gold-finger Plugin or Cabled

SFF-TA-1002 4C+
168-pin, Scalable
Connector
Right-angled or Vertical

Consume. Collaborate. Contribute.

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-

Thanks!

Consume. Collaborate. Contribute.



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

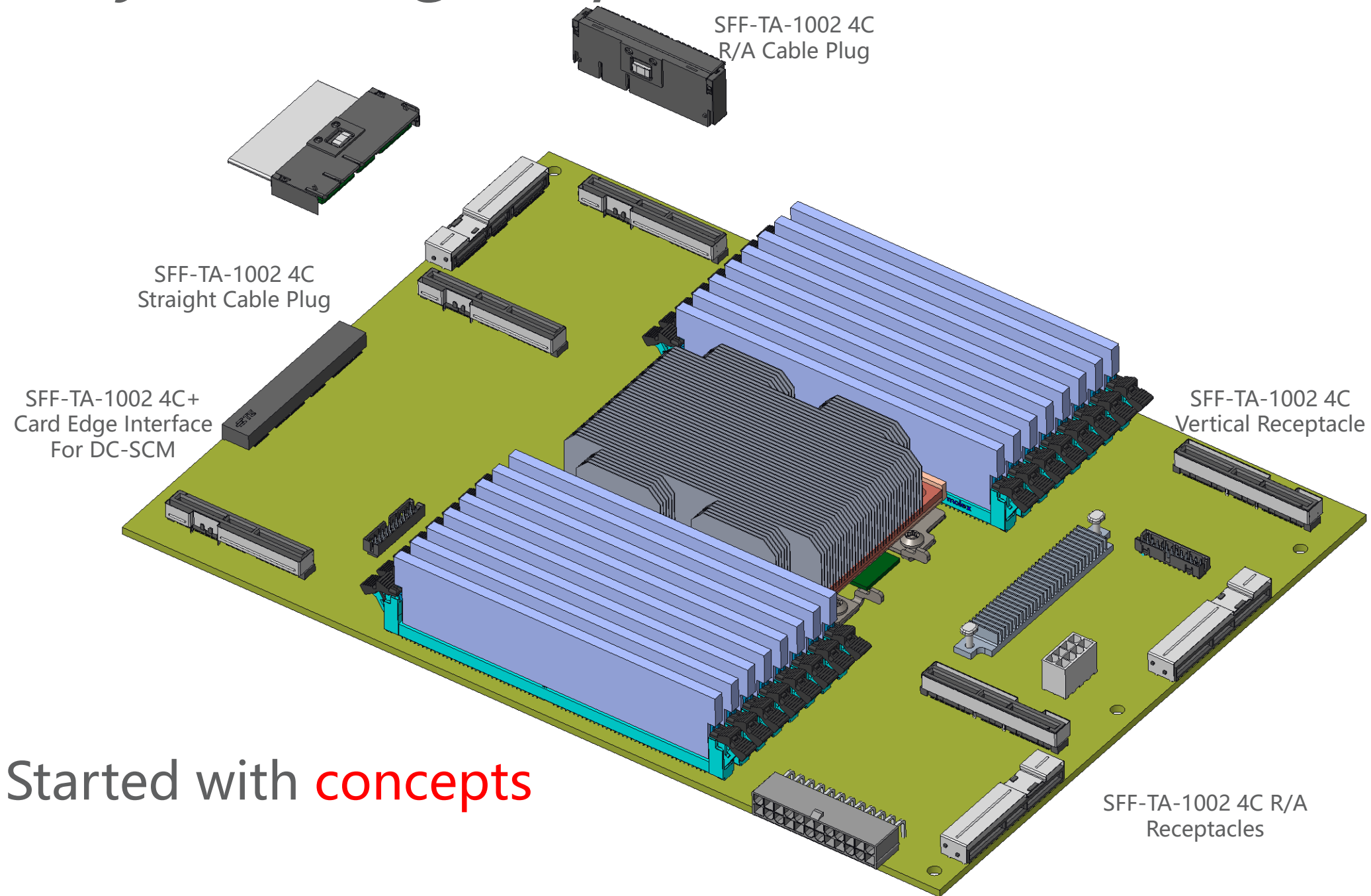
Function	168	Comment
BIOS SPI	6	Serboard SPI
BMC_PCIE	7	PCIE to BMC
CPLD_SGPIO	9	CPLC communication interface between MB/SCM
Critical GPIO	30	The critical event dedicate pin
GND	23	GND
I2C ALERT	13	I2C ALERT signal
JTAG	5	JTAG to MB
LPC & ESPI	12	LPC or ESPI
M2_CLK	2	PCIE clock to M.2 connector
M2_PCIE	4	PCIE to M.2 connector
PECI	2	PECI for Intel platform
POWER	5	3*P12V_STBY, 1* for LPC/ESPI power difference
Sequence	3	For sequence control reference
I2C	26	I2C BUS
SPI	0	SOC/PCH SPI to SCM board
UART	2	SOC/PCH UART to BMC
USB	4	SOC/PCH USB to SCM
Remote Debug	4	For Intel/AMD remote debug
TPM	2	TPM IRQ & chip select
FPGA SPI	4	Reserve for the FPGA FW
RSVD	5	Reserved
PSU Connector	0	PSU Connector signals

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!

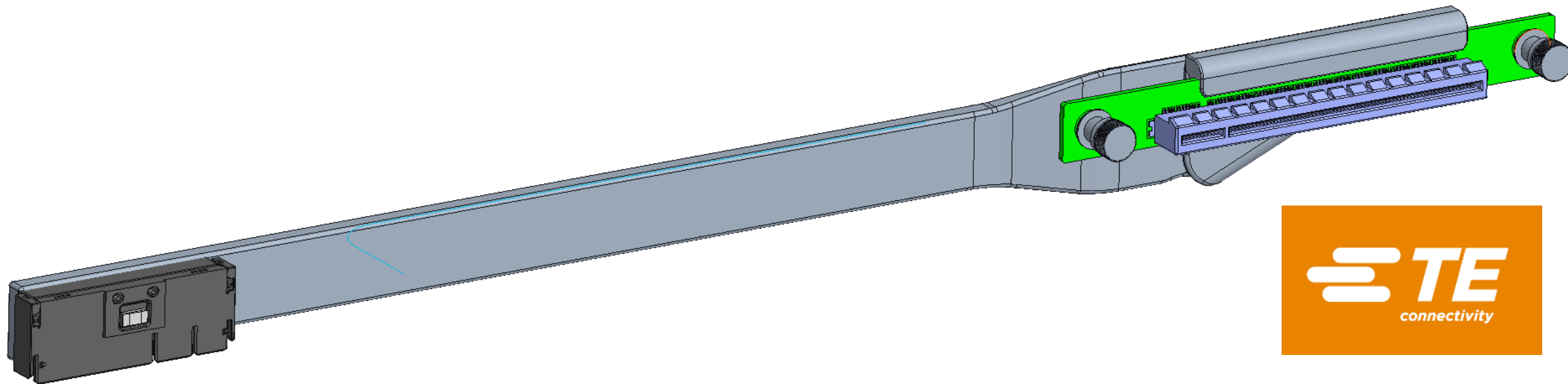


Started with **concepts**

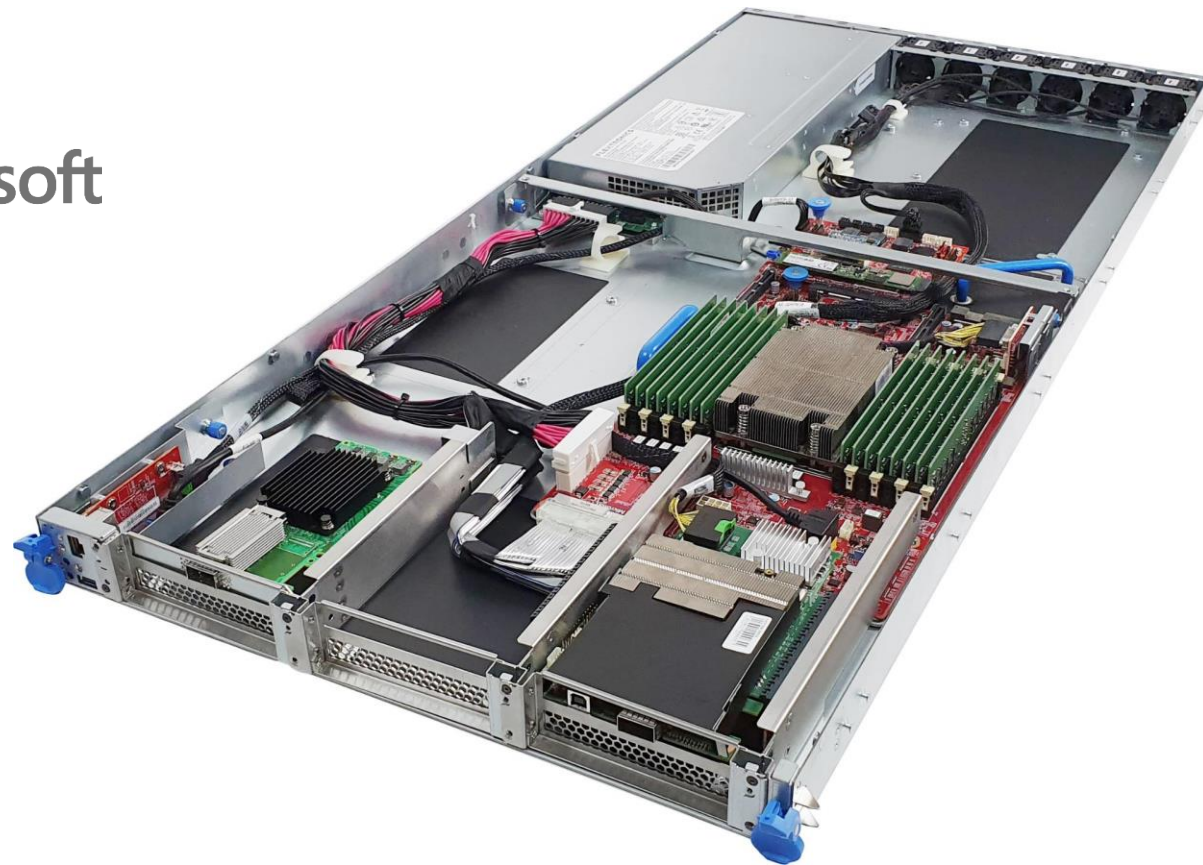
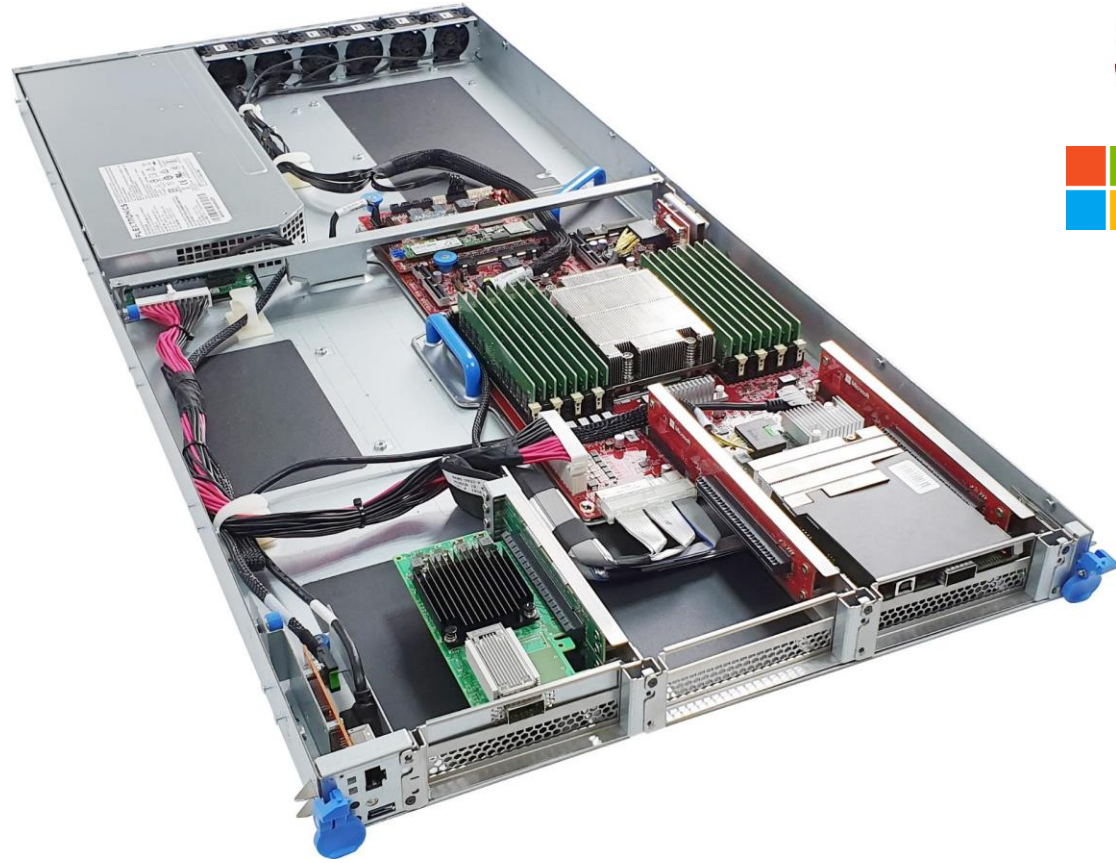
Consume. Collaborate. Contribute.

Add-in Card (AIC) Attachment

IO Slot to CPU Board Cable Harness



Examples of Modular Designs

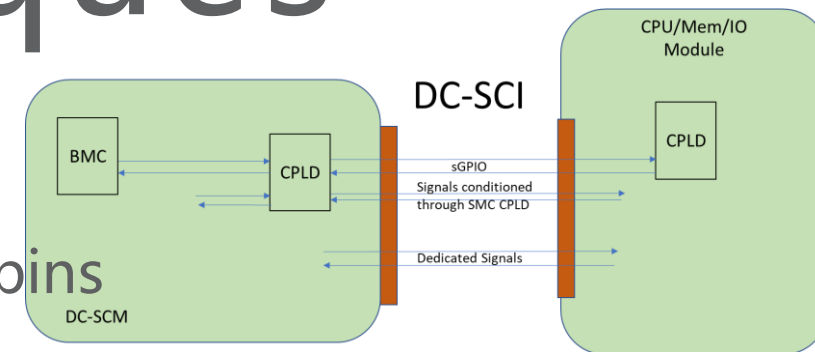


Consume. Collaborate. Contribute.



OPEN
Compute Project

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)