
OPEN
Compute Project
OCP U.S. SUMMIT 2017
Santa Clara, CA



Zaius & Barreleye G2

48-volt Google / Rackspace OpenPOWER Platform

Aaron Sullivan | Distinguished Engineer | Rackspace
Adi Gangidi | Systems Design Engineer | Rackspace
Rob Lippert | Staff Engineer | Google

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.



OPEN Compute Project

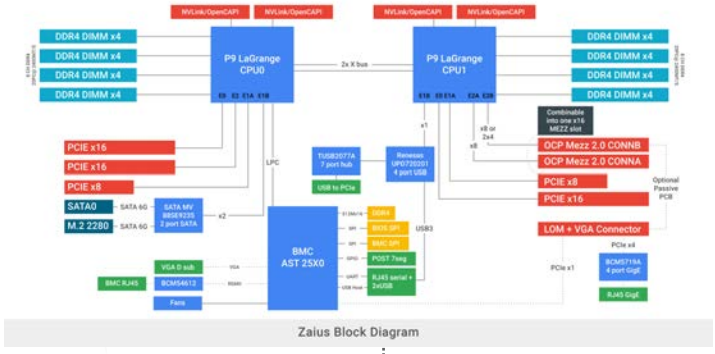

WE’RE TAKING A TOUR...

Introducing Zaius

Google is excited about POWER9 and our collaborative OCP platform codenamed Zaius!

- This platform is a result of a groundbreaking collaboration between the OpenPOWER Foundation, Open Compute Project, OpenCAPI, IBM, Google and Rackspace.
- POWER9 provides a significant upgrade in compute capability and memory capacity.
- Massive I/O connectivity with 80x gen4 PCIe lanes.
- Innovative OpenCAPI expansion ports for new class of accelerators.
- Openness in hardware, firmware, software helps us move faster.

2016 OCP Announcement Block Diagram



Zaius Block Diagram

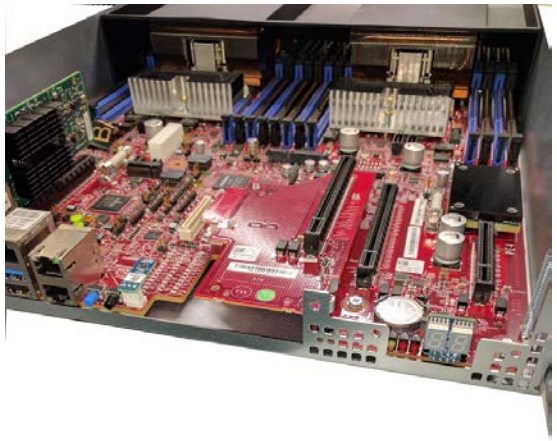
2017 OCP Working Prototype

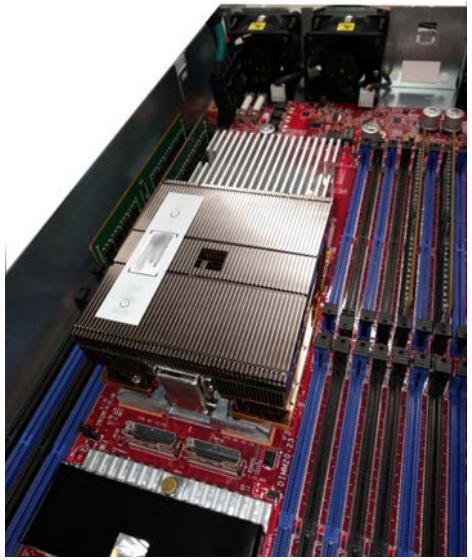


Google Zaius Hardware Quick View

"Zaius" form factor sled – the Google version





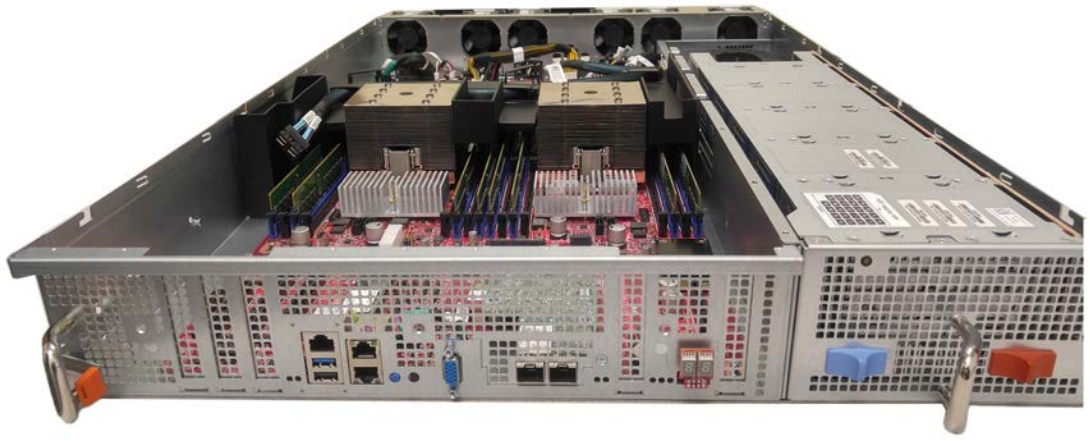


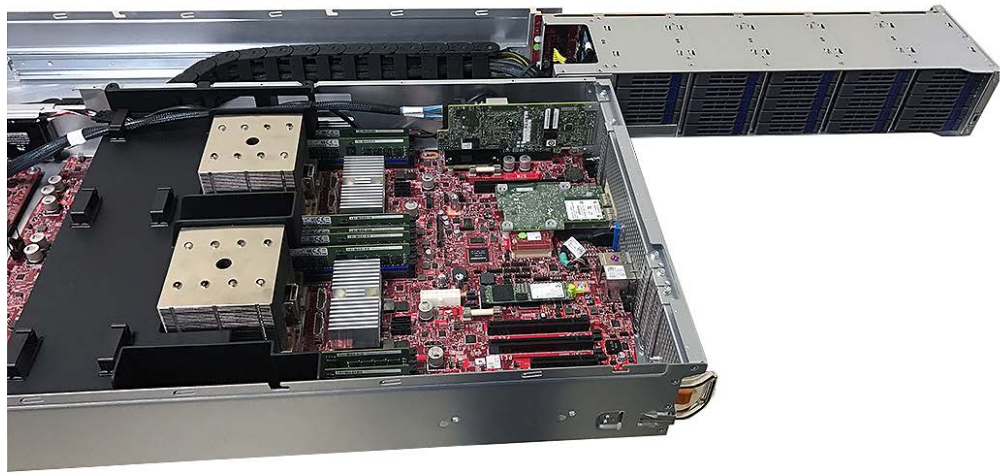
- Designed for Google's datacenter environment
- Compatible with 48V open rack v2
- All I/O and service access in front
- Compact, simple, low cost sheet metal
- Short 1.5OU height requires slightly shorter than standard HH PCIe cards
- Shelf with horizontal bus bar allows flexibility in deployments, multiple expansion options
- Many small Google-specific touches to enhance connectivity, debug-ability

Rackspace Barreleye G2 Hardware Quick View

Zaius motherboard in a new and improved Barreleye chassis

- Built for full-depth 48V open rack v2
- Front IO & service access, hot swap fans
- Supports full height full length cards
- Low profile chassis option (1.25OU)
- High capacity chassis option (20U)
- User-friendly physical service
- High density storage bay





LOTS OF GROUND TO COVER



Topics We're Covering

Zaius / Barreleye G2 Family Tree

PCIe Gen 4, OCP Mez on Gen4, Host & BMC IO Integration

Coherent Interconnects: 25G OpenCAPI and NVLink

Barreleye G2 Storage Design

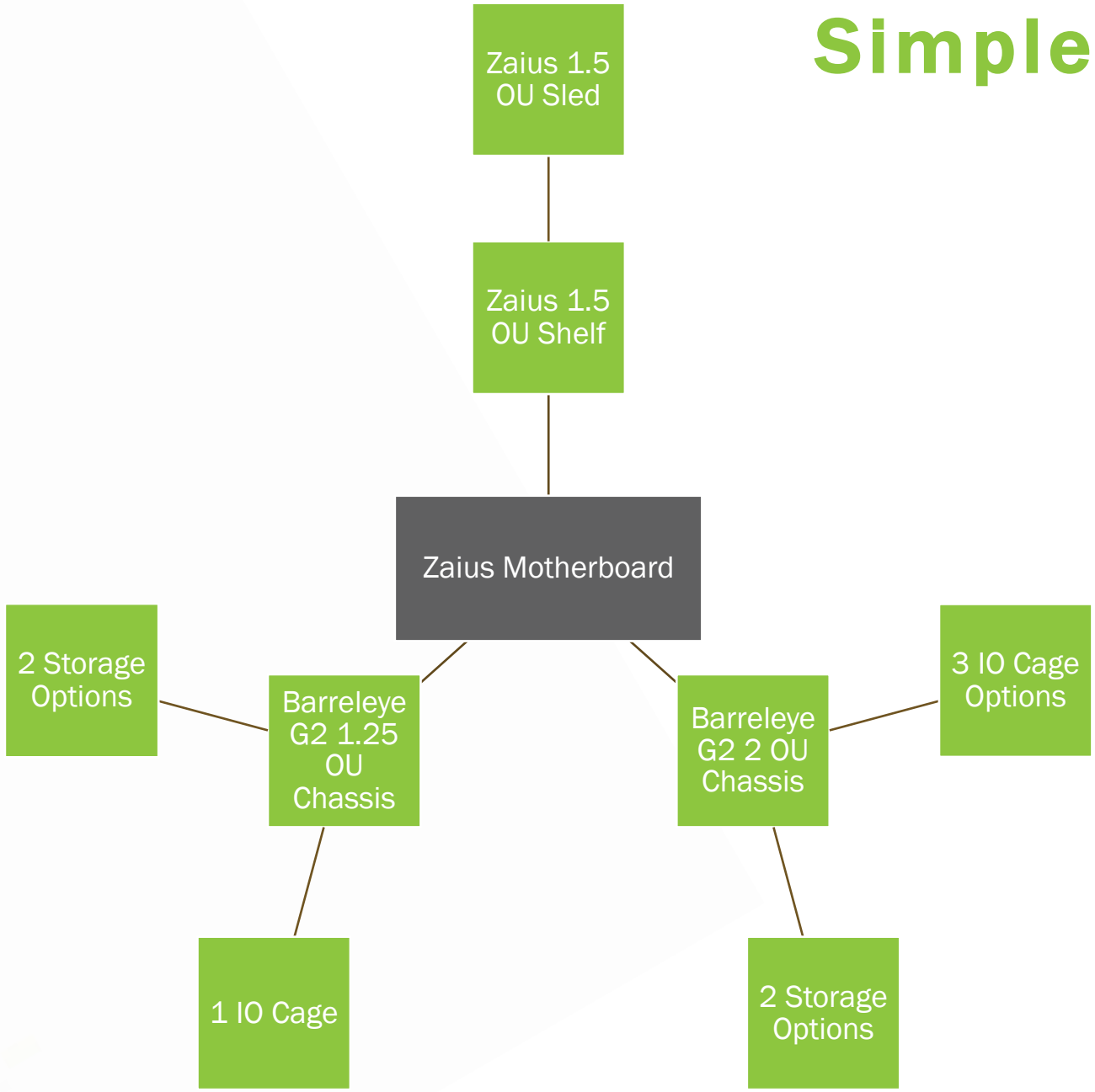
48-volt Point of Load in Zaius, Barreleye G2

48-volt OpenRack v2 & Benchtop Solutions

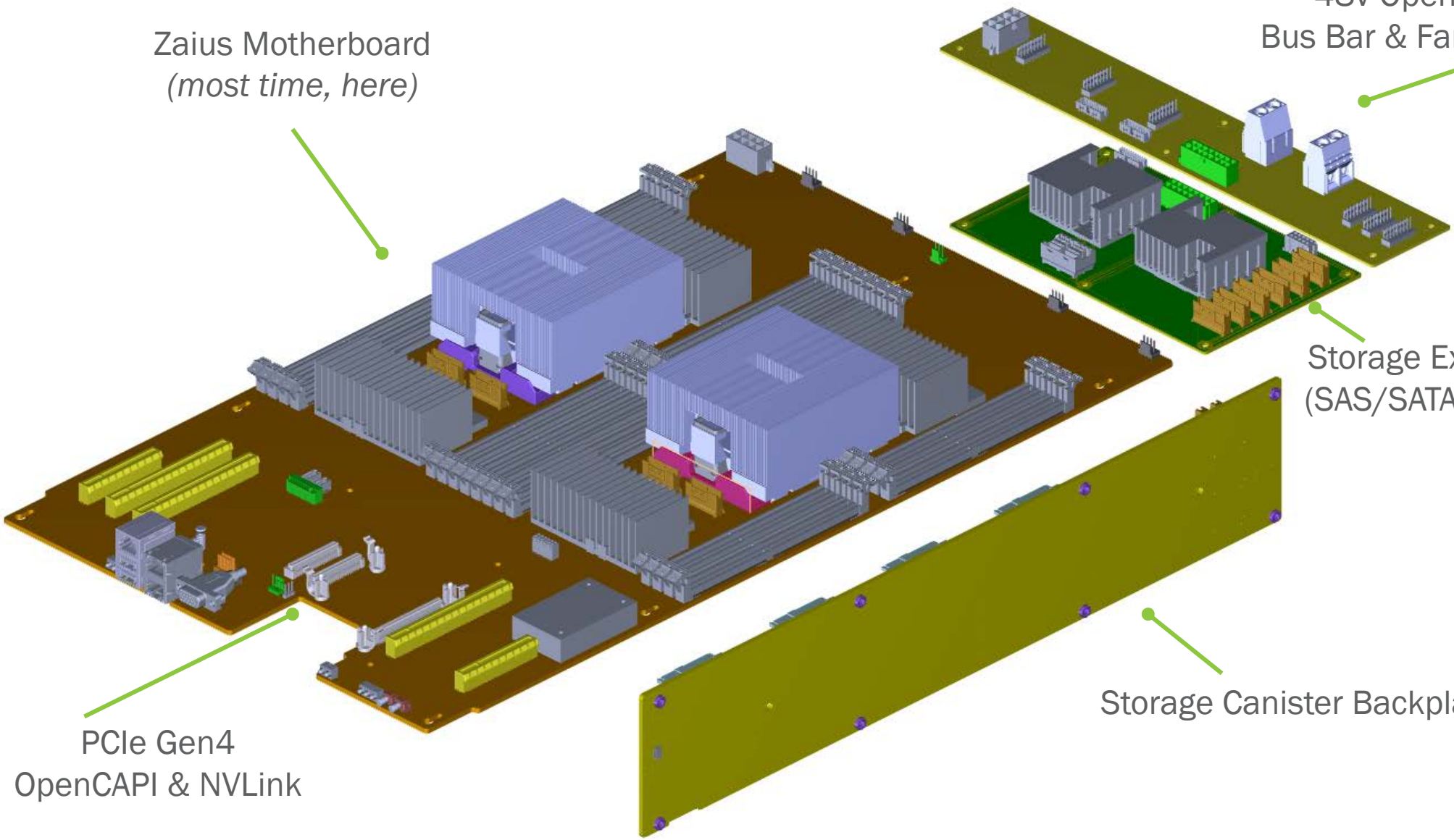
Firmware (OpenPOWER / OpenBMC)

Family Tree

Simple Family Tree



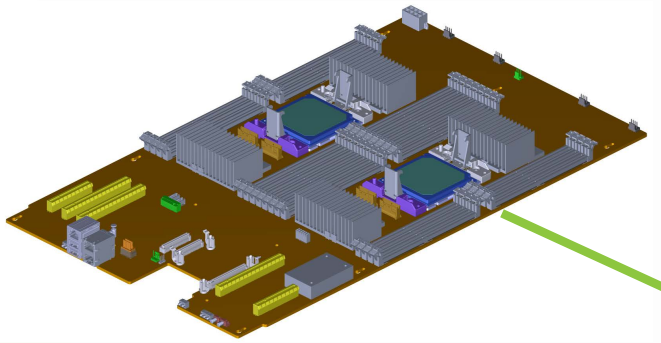
Tour Map



OPEN
Compute Project

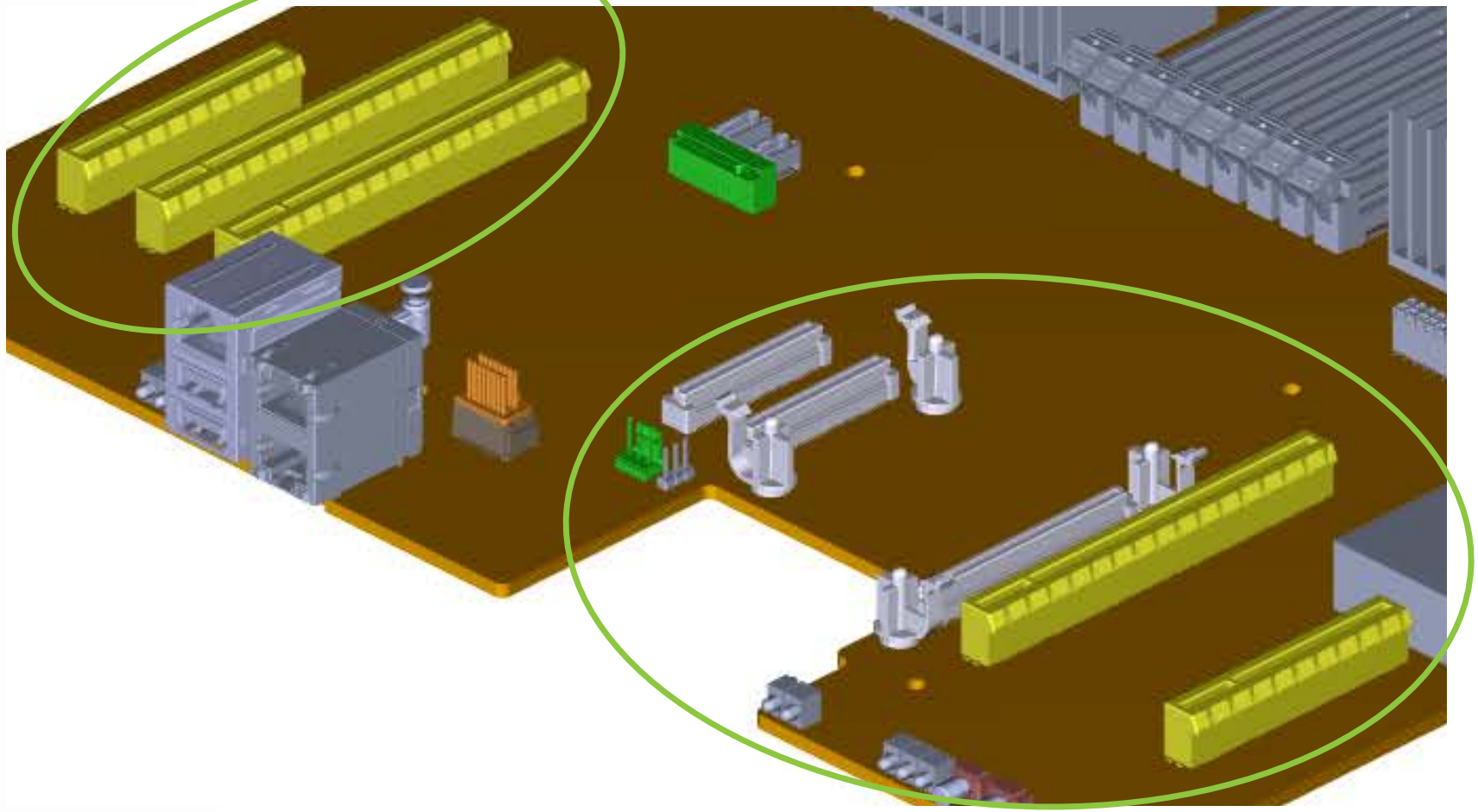
PCIe G4 and OCP Mez 2.0

IO Slots: Standard G4, G4 Mez 2.0, LOM+VGA

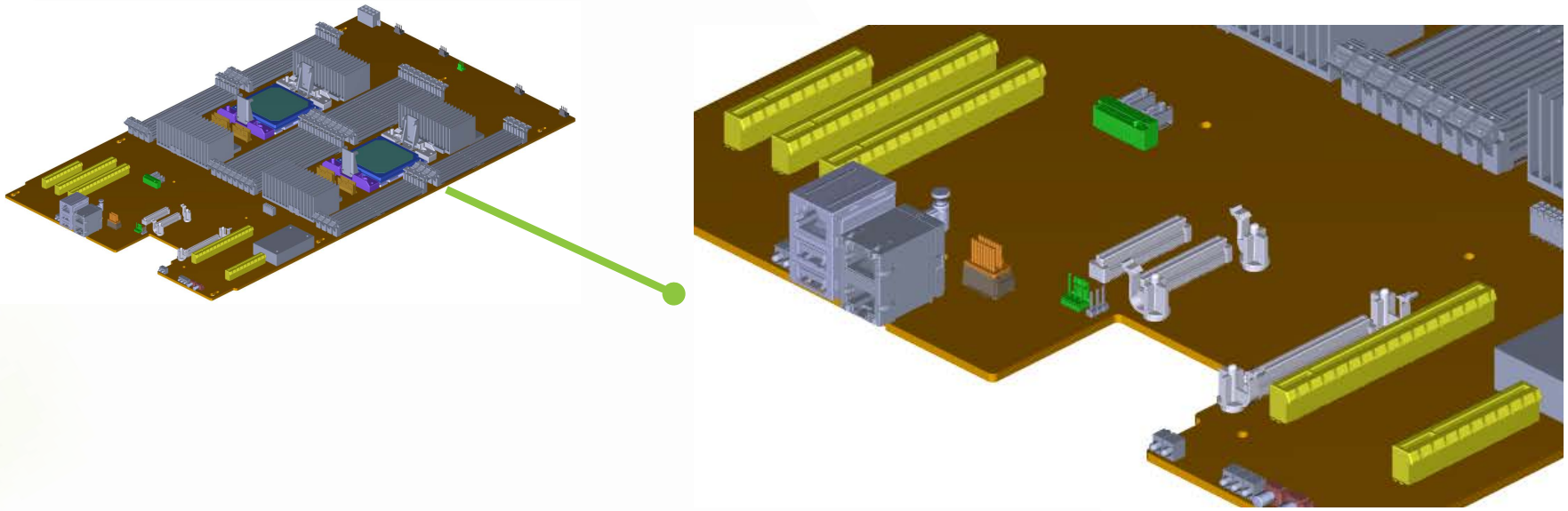


5 – 6 PCIe G4 slots

2 x8 slots
3 x16 slots
1 x8 / x16 OCP Mez 2.0 slot
(x8 if using LOM+VGA mux card)



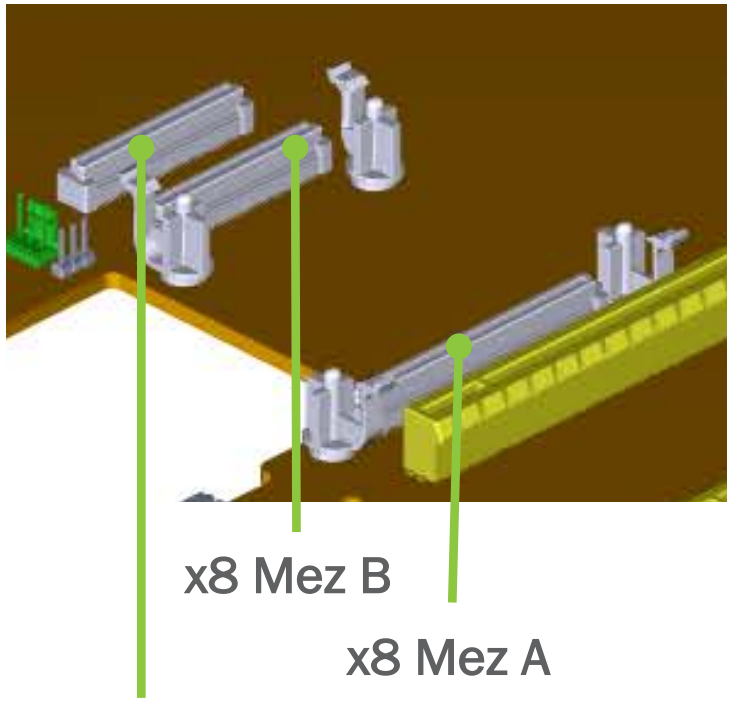
PCIe G4 Slots & Bandwidth

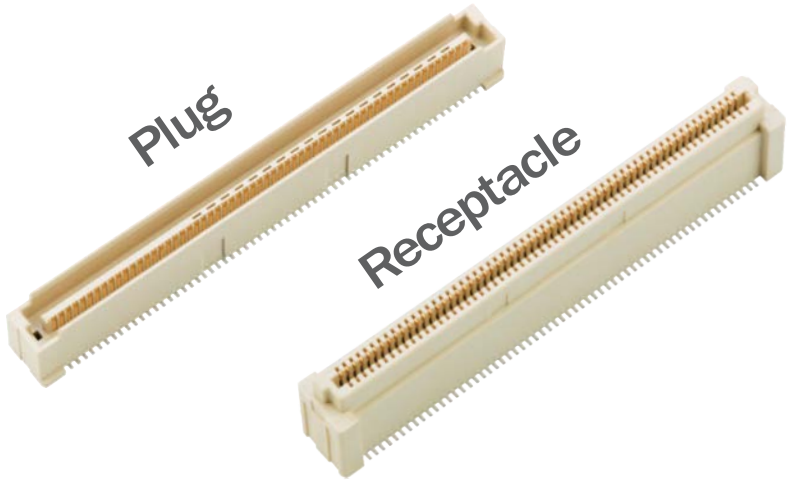


Slot Type	QTY	TOTAL THROUGHPUT (UNI-DIR)
x16 Gen4	3	31.5 Gbytes / sec (252 Gbps)
x8 Gen4	2	15.75 Gbytes / sec (126 Gbps)
x16 Mez A/B Gen4	1	31.5 Gbytes / sec (252 Gbps)
or		
x8 Mez A Gen4*		15.75 Gbytes / sec (126 Gbps)

* Required if using LOM+VGA pass-through, as Mez B x8 connector is used by interposer card.

PCIe G4 Mez 2.0 & LOM+VGA Pass-through





BergStak+™
16Gb/s Gen4 support
Gen3 OCP Mez 2.0 Compatible
5 / 8 / 12 / 16 mm stack heights

HBA & SHARED LOM+VGA	
x8 HBA	Mez A
LOM+VGA Connector	Mez B, Pass-through
VGA & LOM	Host + BMC Access

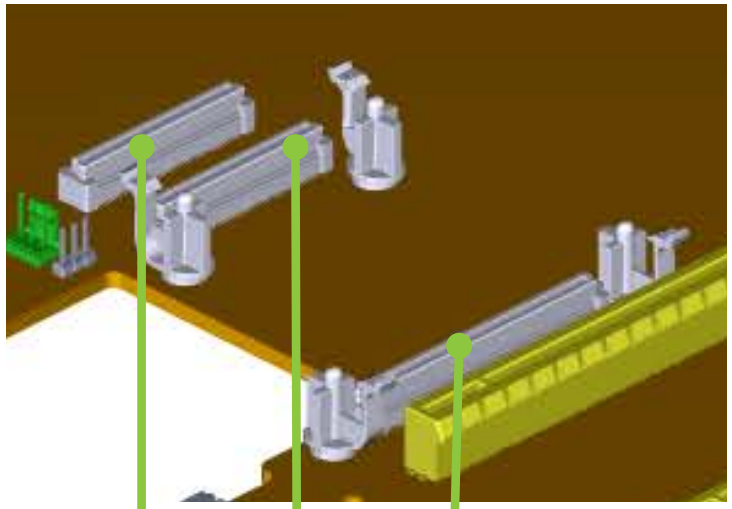
x16 HBA (INDEPENDENT LOM, NO VGA)	
x16 HBA	Mez A, Mez B
Host LOM	Host Access Only
BMC LOM	BMC Access Only

x8 HBA (INDEPENDENT LOM, NO VGA)	
x8 HBA	Mez A (Mez B Empty)
Host LOM	Host Access Only
BMC LOM	BMC Access Only

LOM + VGA Pin Assignments

Connect to LOM BCM 5719 via PCIex4

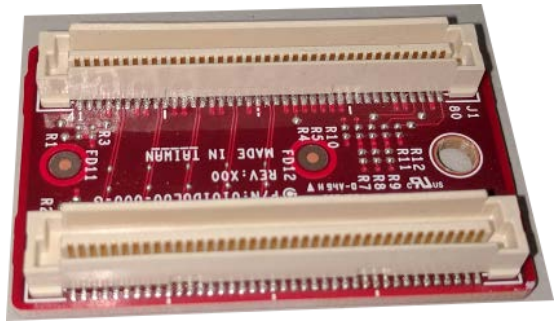
Connect to BMC AST2500 via PCIex1



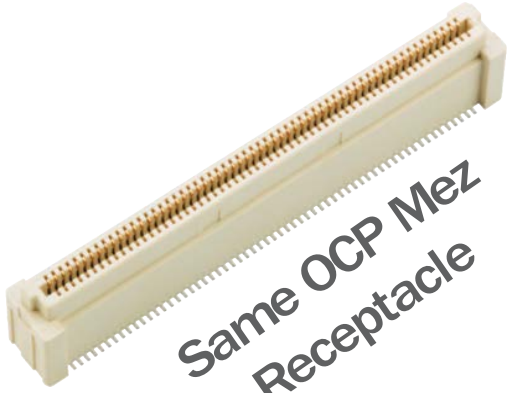
x8 Mez B

x8 Mez A

LOM+VGA Pass-through



LOM + VGA
Interposer Card



Same OCP Mez
Receptacle

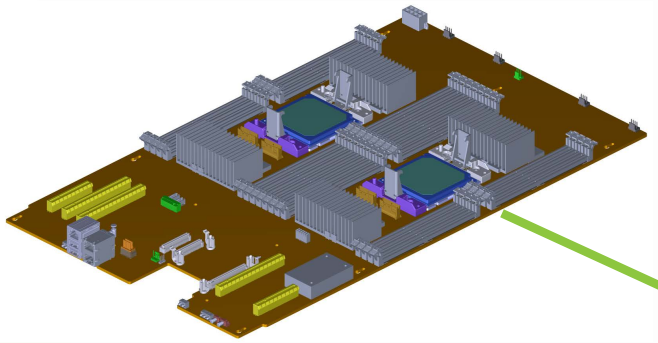
Why This Matters

- Gen4 adapters can be fit into OCP mez 2.0 form factor
- This is possible, but not necessarily optimal
  - May not optimal for low profile chassis solutions
  - Little space for NICs with extra co-processor & memory footprint
- Now is a good time to revisit OCP mezzanine spec



25G OpenCAPI & NVLink

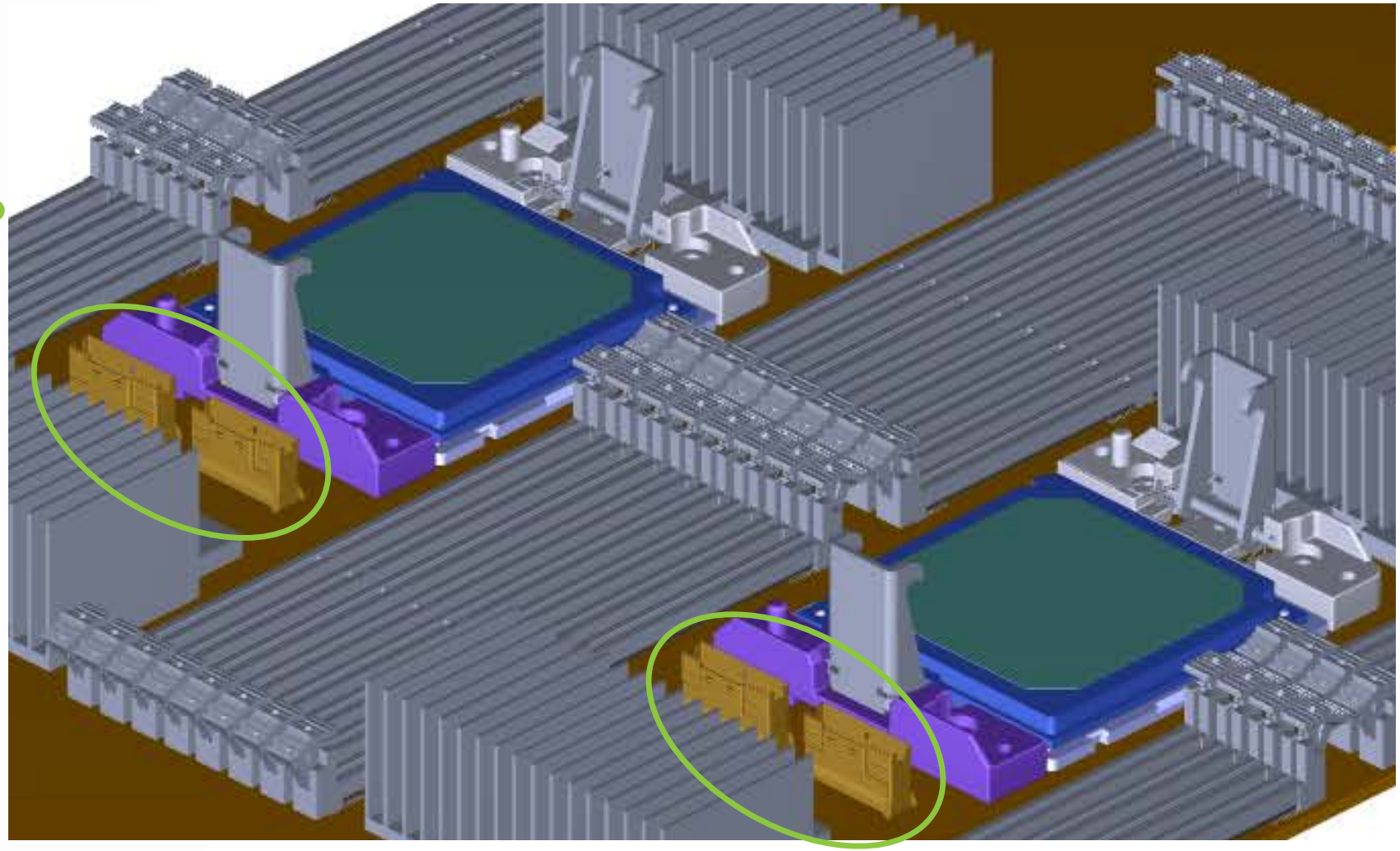
OpenCAPI / NVLink Connector Location & Type



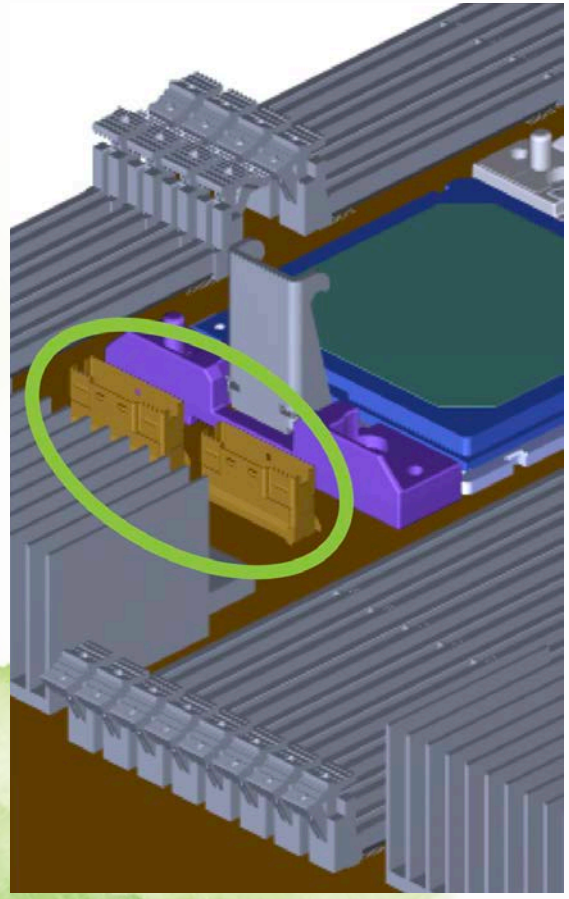
2 “Bricks” per CPU Socket

24G Slimline SAS Connectors
(25GByte per Brick)

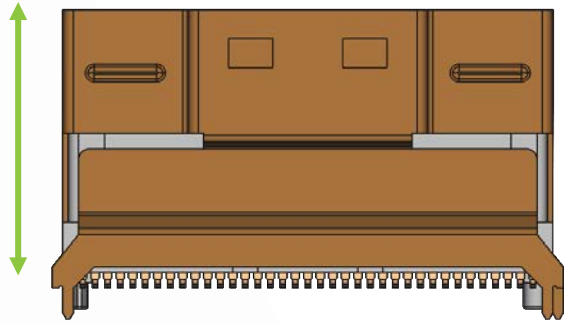
Follows SFF-8654 Standard



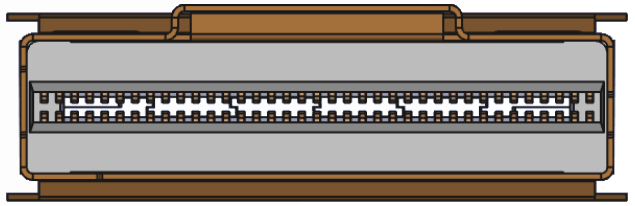
Zaius / BG2 OpenCAPI & NVLink Interface



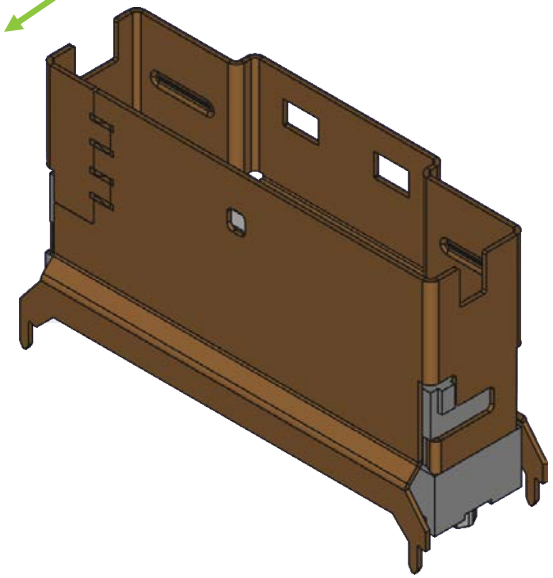
13mm
(23.5 mated)



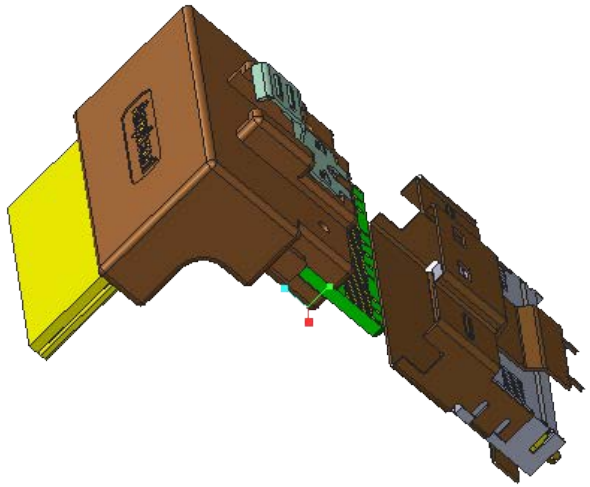
23.5mm
(25.95 mated)



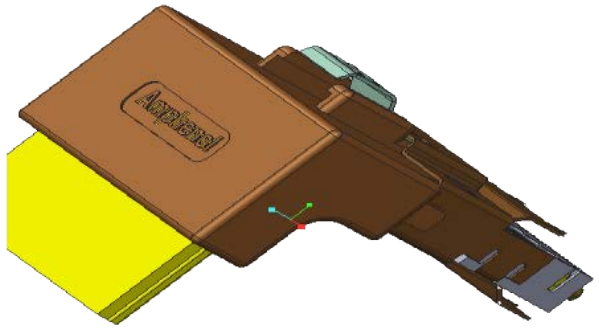
7.1mm
(9.24 mated)



Zaius / BG2 OpenCAPI & NVLink Cable



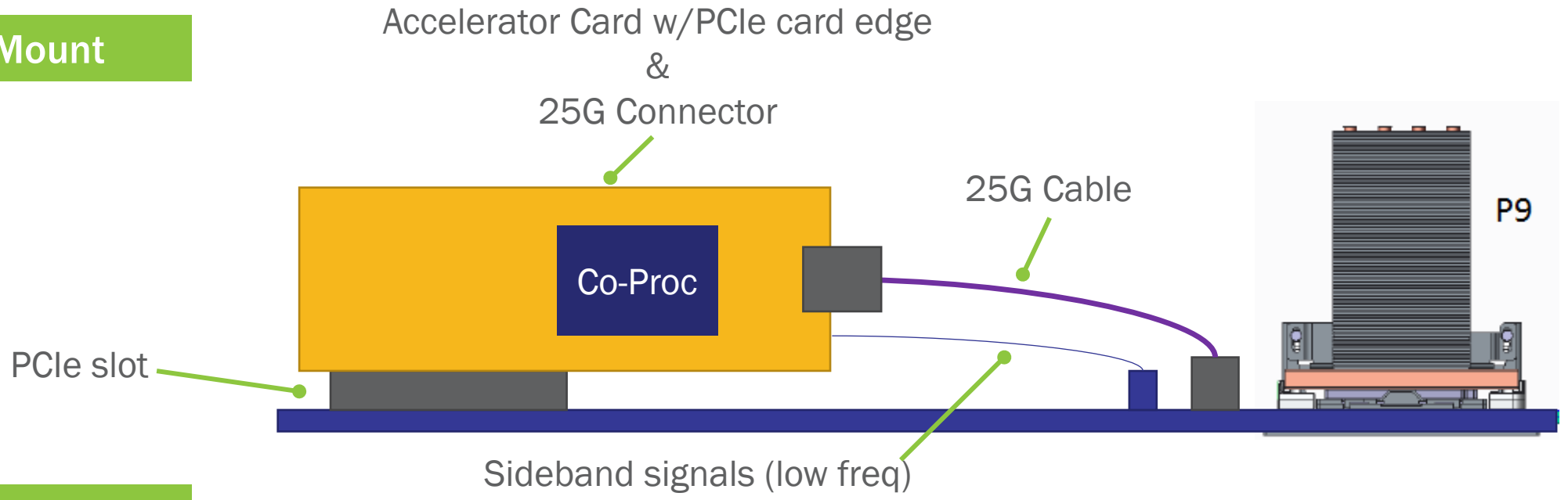
SPECS	MEASUREMENT
RECEPTACLE WIDTH:	23.5mm
PLUG WIDTH:	25.95mm
MATED HEIGHT:	22.4mm
MATED LENGTH:	9.24mm
30AWG DIAMETER:	0.255mm
PCB THICKNESS:	1.0mm



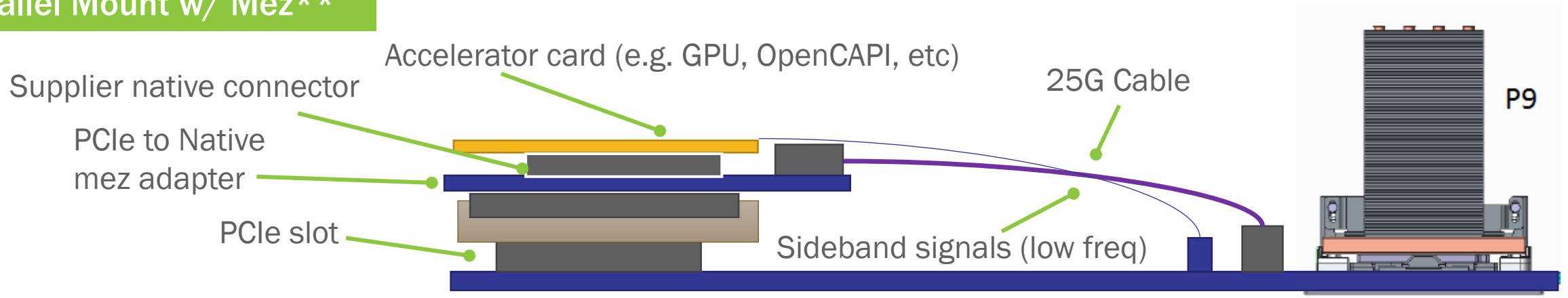
DESCRIPTION	QTY	TOTAL THROUGHPUT (UNI-DIR)
NVLink 2.0 / OpenCAPI 3.0 Brick	-	25 Gbytes / sec (200 Gbps)
Bricks per Socket	2	50 Gbytes / sec (400 Gbps)
Bricks per Server	4	100 Gbytes / sec (800 Gbps)

Example Card and Cable Mount Options*

Perpendicular Mount



Parallel Mount w/ Mez**



* Possible approach to co-processor card, mount, and cable designs.

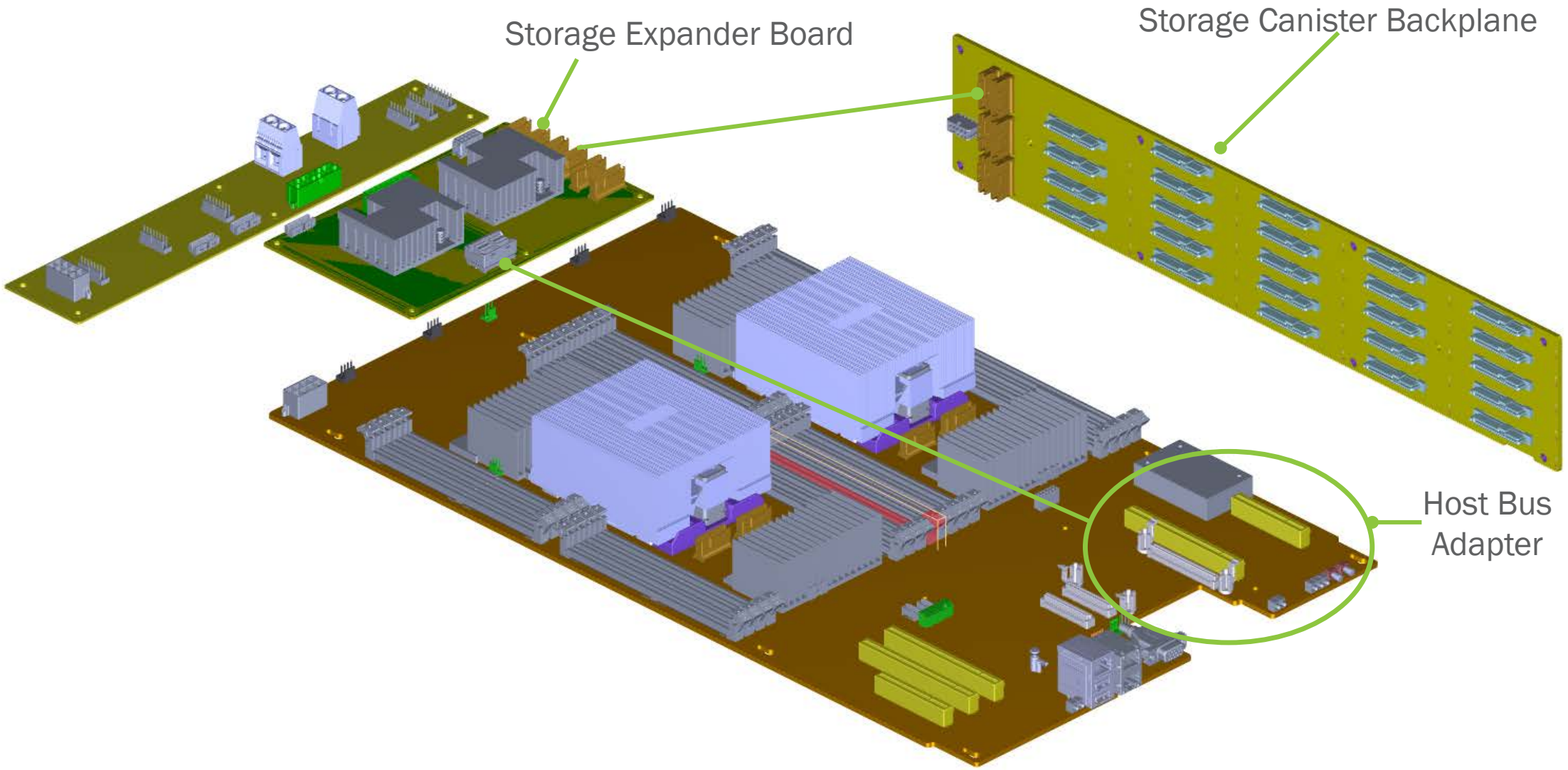
** Does not refer to OCP Mez standard.

Why This Matters

- Coherent attach is powerful – soon to be common
- Too many wire & connector standards
- Gratuitous differentiation at work here
- Could fragment motherboard & chassis specs
- Rackspace considers SFF-8654 (“Slimline SAS”) to be a good choice for OCP
- Let’s work together with processor & co-processor companies to unify on a common standard
  - Include PCI-SIG, CCIX, OpenCAPI

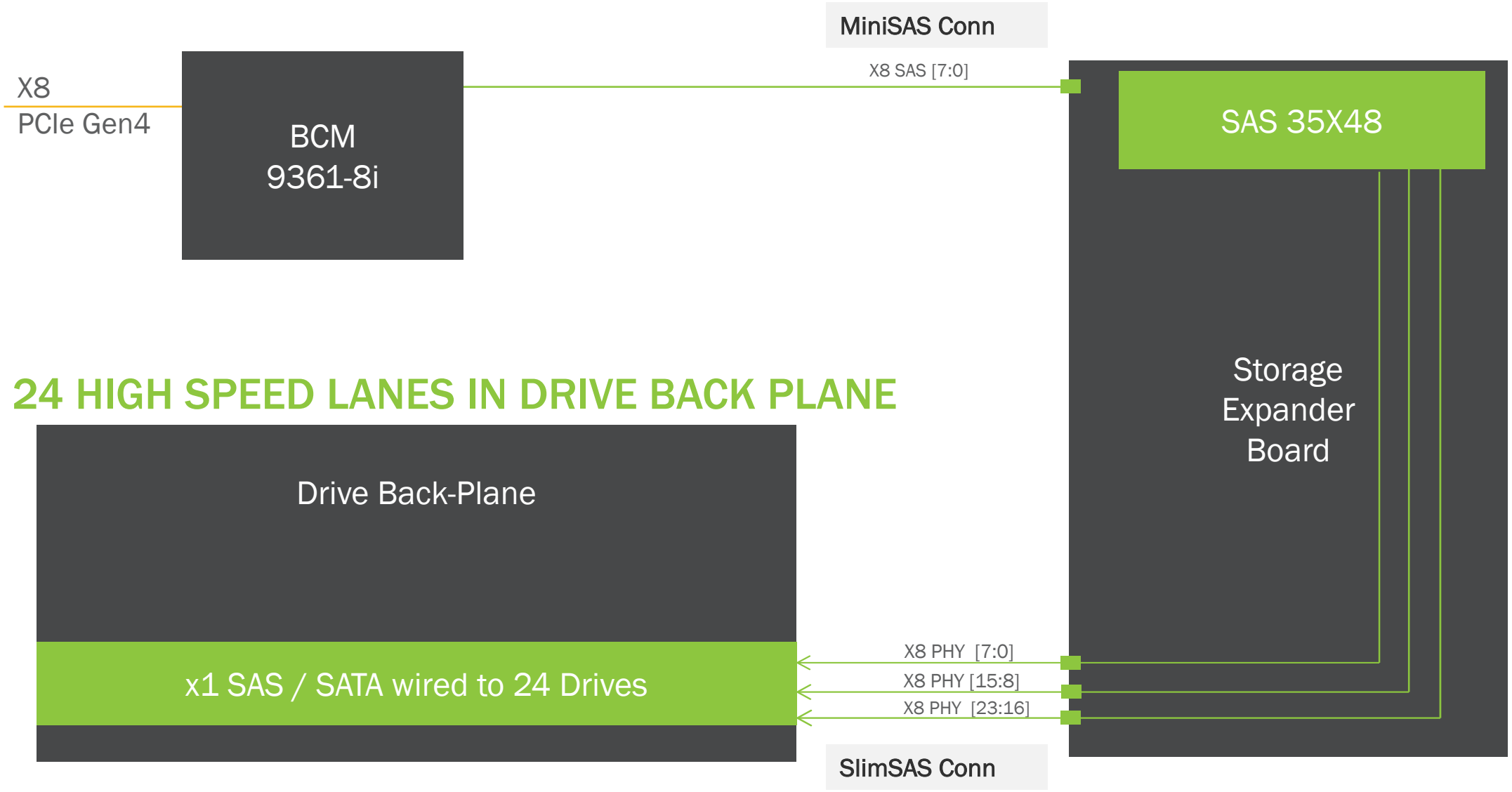
Storage Design

Barreleye G2 Storage Overview

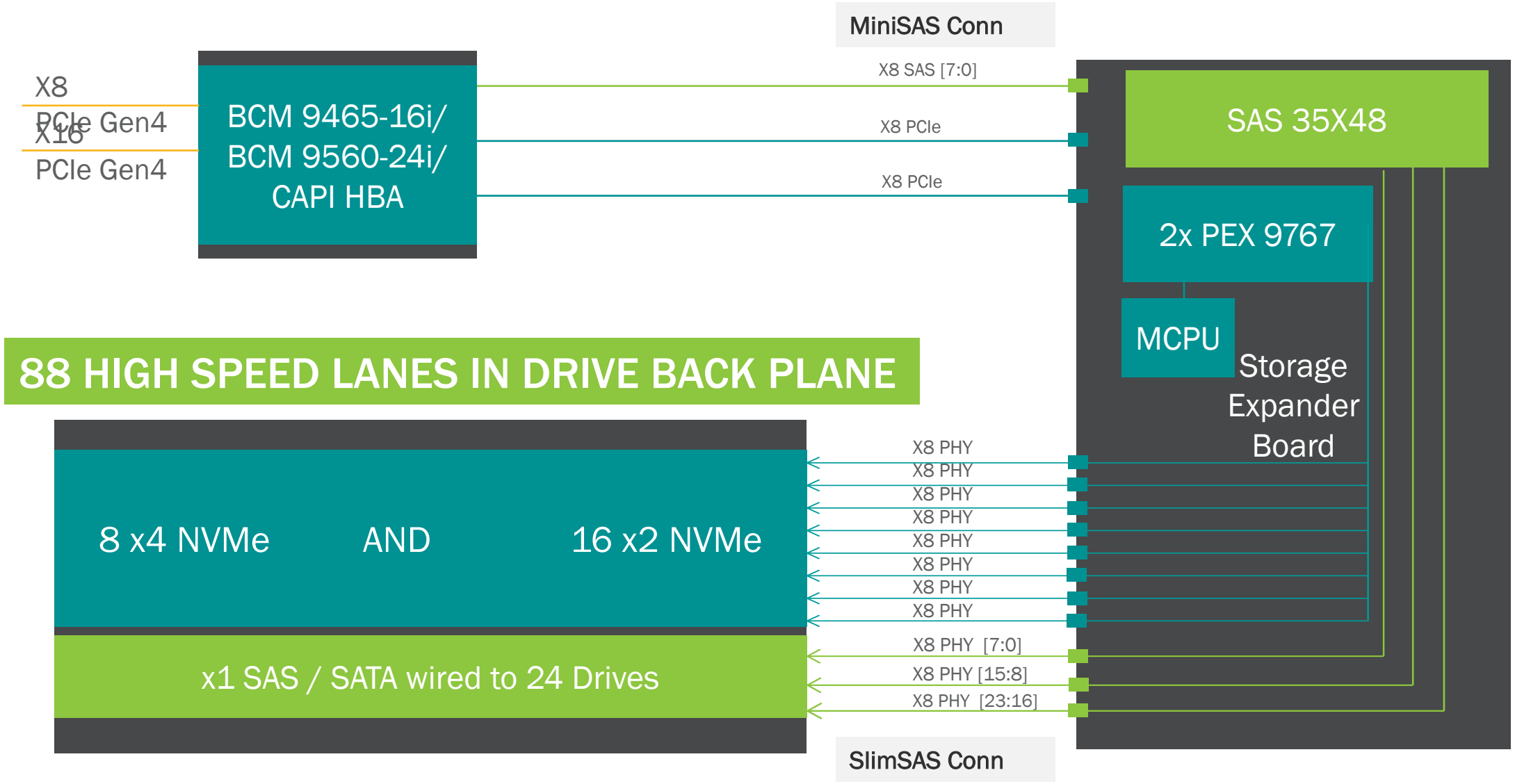


OPEN Compute Project

BG2 EVT Storage Architecture (Today)



BG2 DVT Storage Architecture (NextGen)

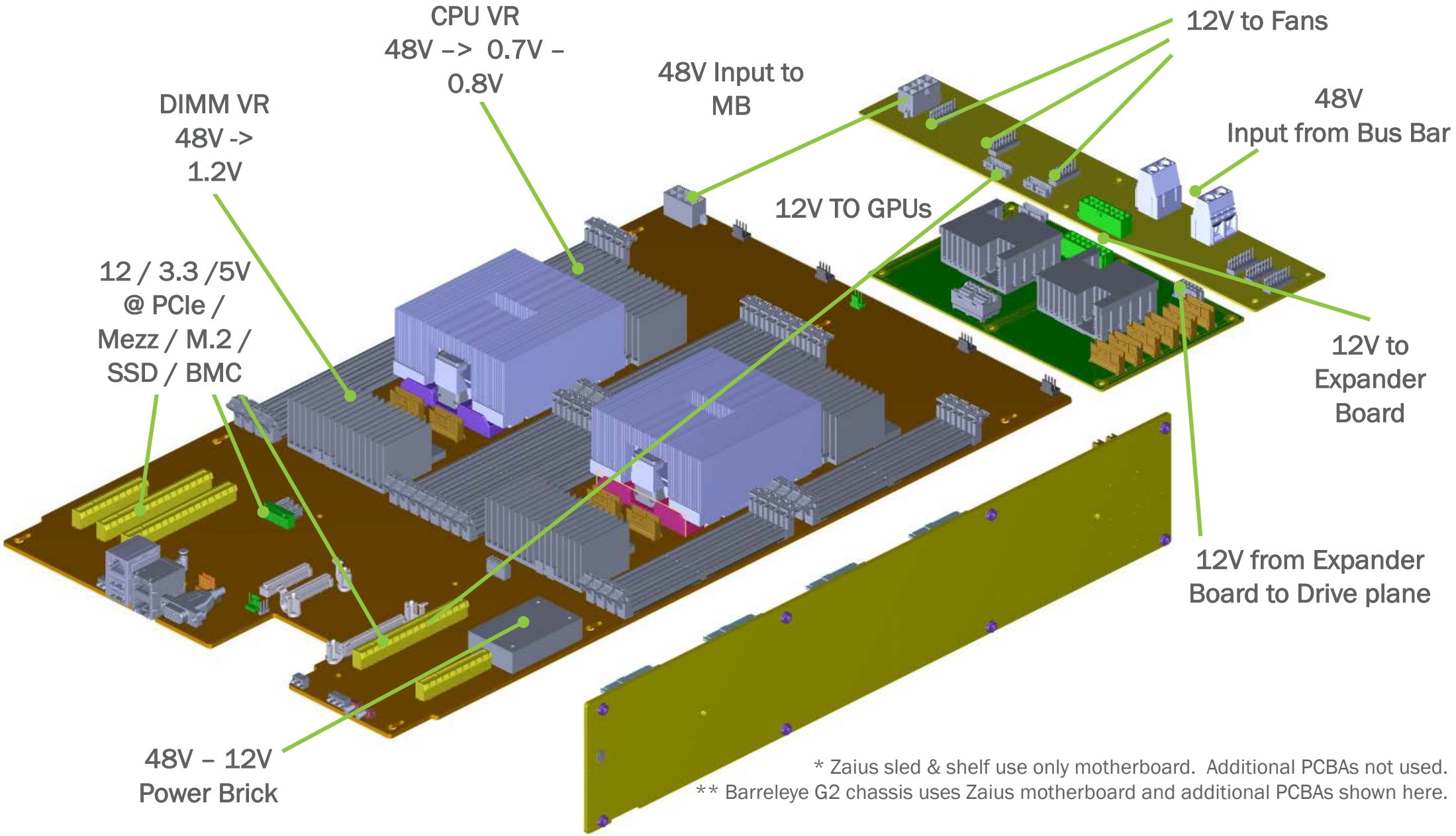


Takeaways

- Barreleye G2 2 O.U will support SAS / SATA / NVMe in each slot
- Drive backplane design is getting complicated
  - Need for next generation connector and protocol standards to ease transition
- Tri-mode Transition
  - U.3 connector proposed to follow U.2
  - Tri-mode Expander
  - U.3 Drives

48 Volt & Point of Load VR

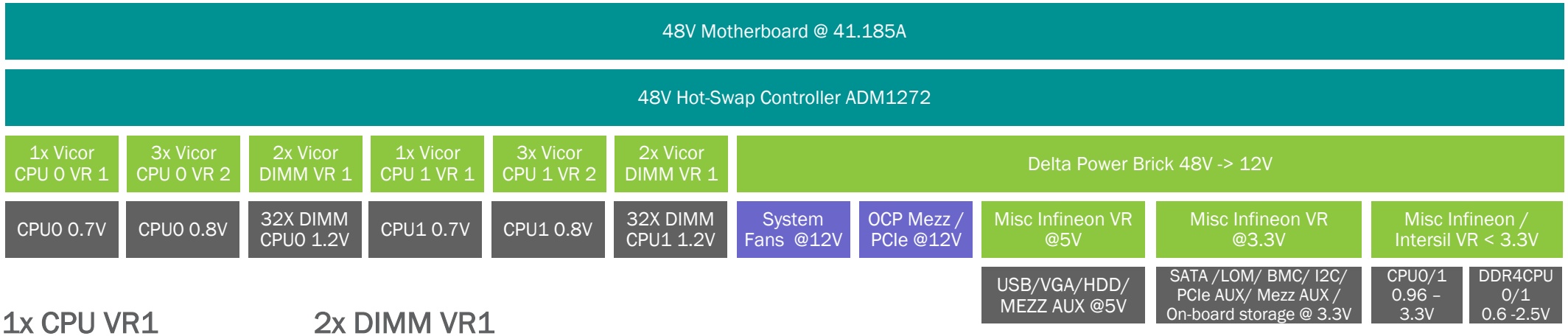
48V Barreleye G2 Voltage Distribution



* Zaius sled & shelf use only motherboard. Additional PCBAs not used.
** Barreleye G2 chassis uses Zaius motherboard and additional PCBAs shown here.

OPEN Compute Project

Zaius Chassis / Motherboard 48V Step Down



1x CPU VR1

- 0.7V @ 74.4 A
- 90% Efficiency

2x DIMM VR1

- 1.2V @ 95.54A each
- 90% Efficiency

3xCPU VR2

- 0.8V @ 297 A
- 90% Efficiency

AVSBus

- Finer-grained CPU power

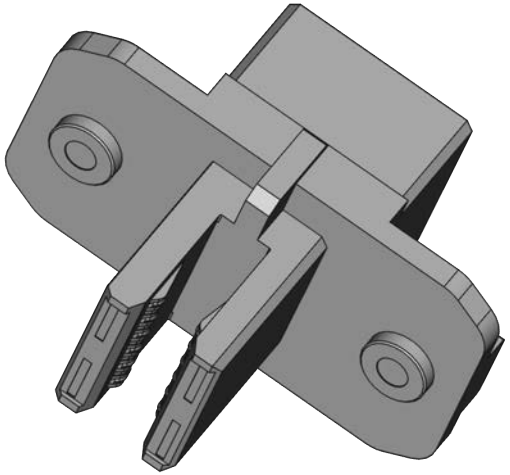
48V VR COMPONENTS 12V <5V

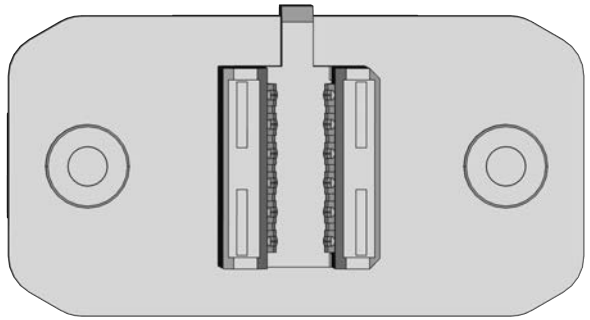
Takeaways

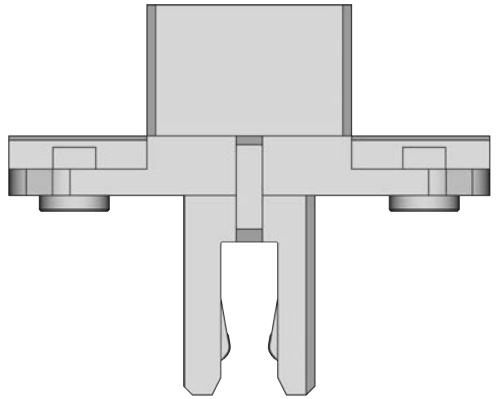
- 48V POL Regulators & AVSBus
  - Compact solution with high efficiency
- Barreleye G2 Add-on power board
  - Provides choice of high-wattage sub-systems (GPU / Storage) without penalty of routing power through Motherboard

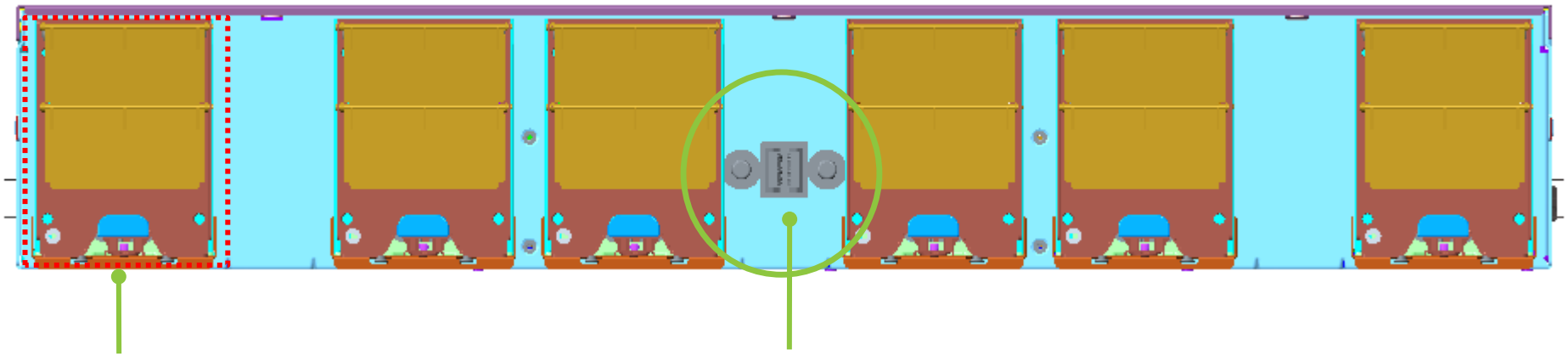
48 Volt Rack & Power Shelf

48-Volt Server-to-Bus-Bar Clip



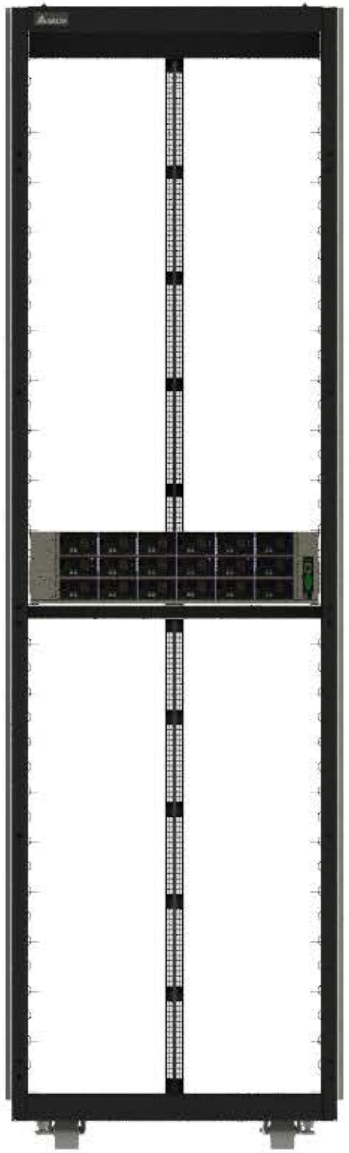






48-Volt Open Rack v2 for Barreleye G2





OCP 48V Rack

+48VDC bus bar
(one power zone)

30U 30kW power shelf
(36kW without redundancy)

41 - 42OU total
38 - 39OU with power shelf

48-Volt Rack Power System





12 PCS 3kW PSU

RMC

6 PCS 6kW ATS

30 kW Power Shelf

3 OU height, 2x – 4x 60A AC input

4x – 6x C14 socket 10A PDU (rear)

12x 3kW PSU Modules

6x ATS Modules (1 per 2 PSUs)

10+2 Redundancy (if ATS fails)

Rack Management Controller

48v Lunchbox (~1700w)

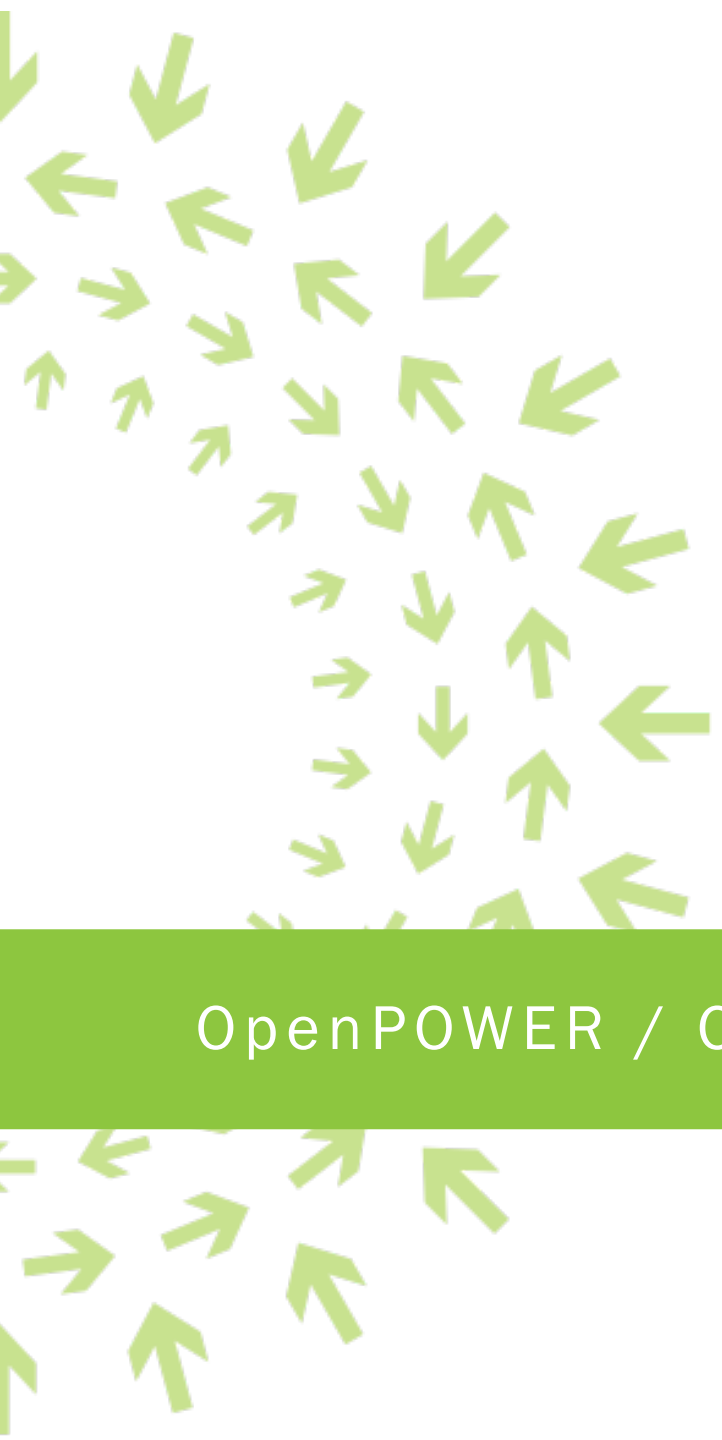


48v Output and Cable



110 – 240v Input and Cables

OPEN
Compute Project



OpenPOWER / OpenBMC

OpenPOWER/OpenBMC



Zaius is part of the OpenPOWER community and is planned to be a fully compliant/tested OpenPOWER system.

- Committed to providing a fully open-source firmware and management stack to go along with server – no black boxes.
- External contributors welcome and encouraged to contribute!

OpenPOWER boot firmware (aka BIOS) fully supported and booting now:

- *git clone --recursive https://github.com/open-power/op-build*
- *op-build zaius_defconfig && op-build*
- Zaius work in progress:
    - Stability, error logging/reporting, slot bifurcation, …

OpenBMC on Zaius AST2500/2520 working to initialize hardware and start booting the machine:

- *git clone https://github.com/openbmc/openbmc (Advances Facebook BMC with new kernel and drivers)*
- *export TEMPLATECONF=meta-openbmc-machines/meta-openpower/meta-ingrasys/meta-zaius/conf; . openbmc-env*
- *bitbake obmc-phosphor-image*
- Zaius work in progress:
    - Sensors, fan control, NC-SI muxing, LEDs, POST card, boot time, …

It Takes a Village: Partner Appreciation

BUS BAR SOLUTIONS	TE connectivity	Amphenol	
NETWORKING	Mellanox TECHNOLOGIES	BROADCOM	
STORAGE	BROADCOM	SAMSUNG	Micron
POWER9	IBM	OpenPOWER™	
OPENCAPI / NVLINK 25G CONNECTIVITY	Amphenol	OpenCAPI™	OpenPOWER™
ACCELLERATORS	NVIDIA	XILINX ALL PROGRAMMABLE	OpenCAPI™
MEMORY	SAMSUNG	Micron	
MOTHERBOARD, PCBA, CHASSIS, INTEGRATION	ingrasys®		
RACK & POWER SHELF	ingrasys®	DELTA	
VOLTAGE REGULATORS	VICOR The Power Behind Performance	DELTA	

JOIN US



Want to Participate?

Inquire About Development Samples

OpenCAPI & NVLink Solution Developers Also Welcome to Join

OPEN
Compute Project

Want to Know More?

OCP Community Preview
http://files.opencompute.org/oc/public.php?service=files&t=4302505c9d59160ec03aaf4ea67b146c&download

OCP Server Mailing List
http://lists.opencompute.org/mailman/listinfo/opencompute-server

OCP Server Mail Archives
http://lists.opencompute.org/pipermail/opencompute-server/

OPEN
Compute Project

Design Package

Available on OCP Server Wiki and GitHub

Enhancements Coming Throughout 2017

GitHub
https://github.com/opencomputeproject/zaius-barreleye-g2 (don't forget to install Git LFS)

OCP
http://www.opencompute.org/wiki/Server/Working#Open_Rack

OPEN
Compute Project

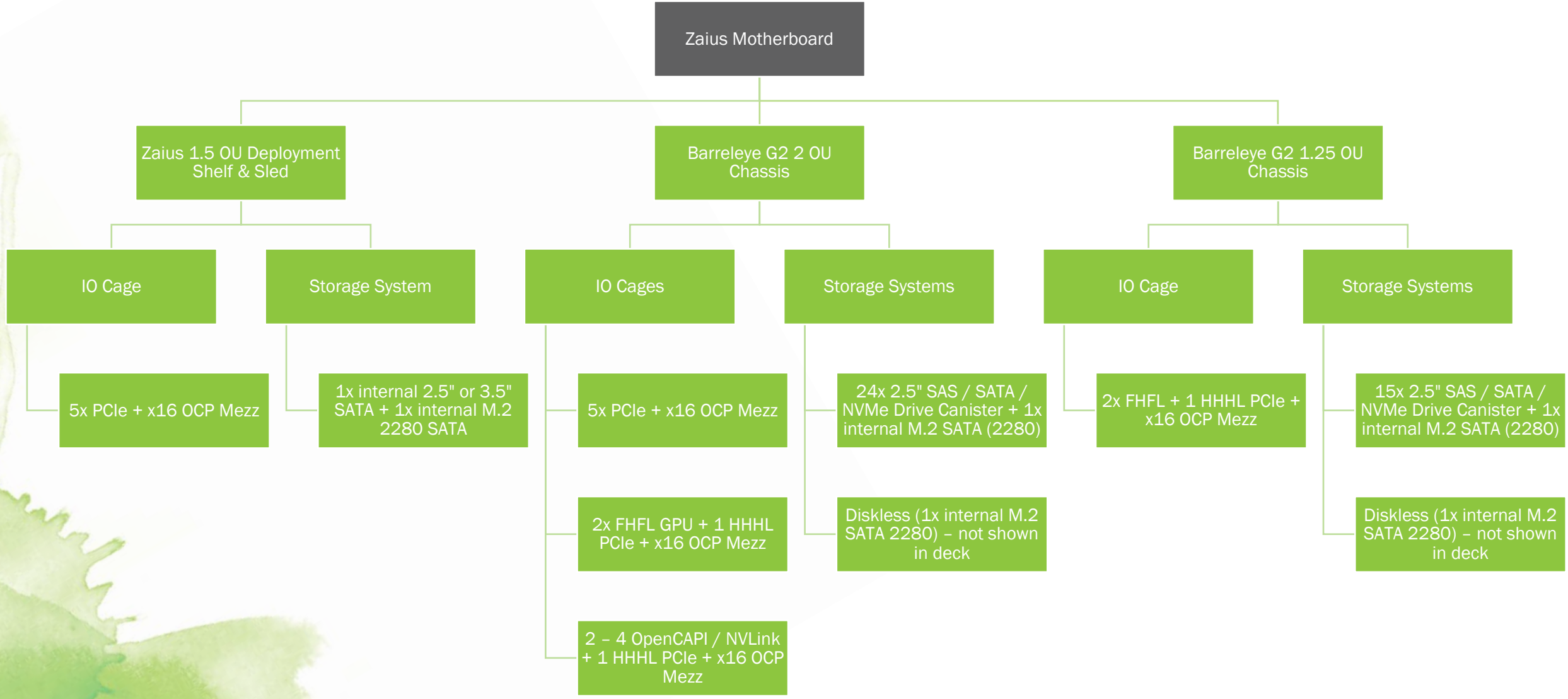
THANK YOU



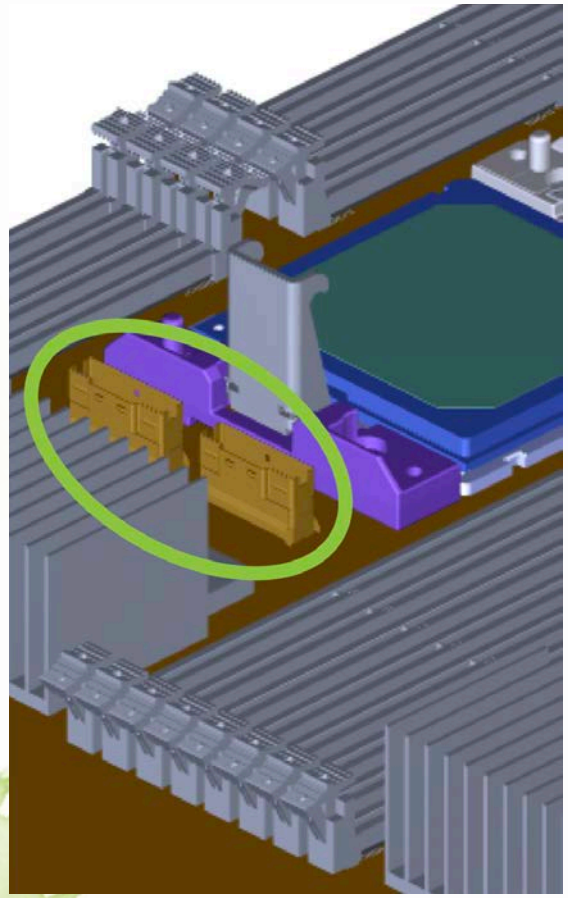
Archive Material

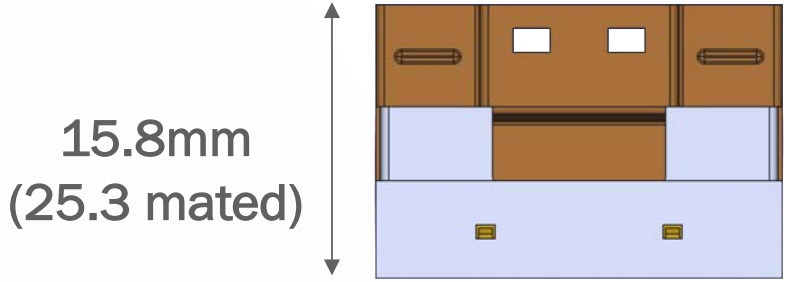
OPEN
Compute Project

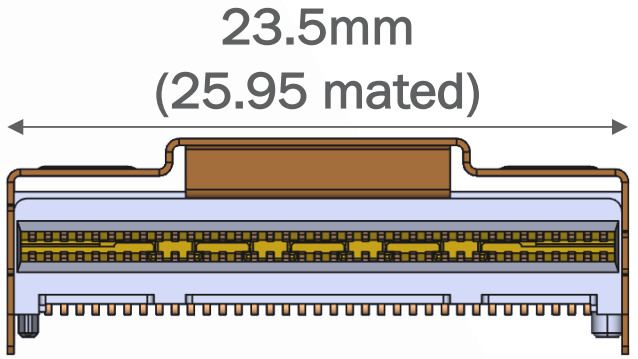
Detailed Family Tree

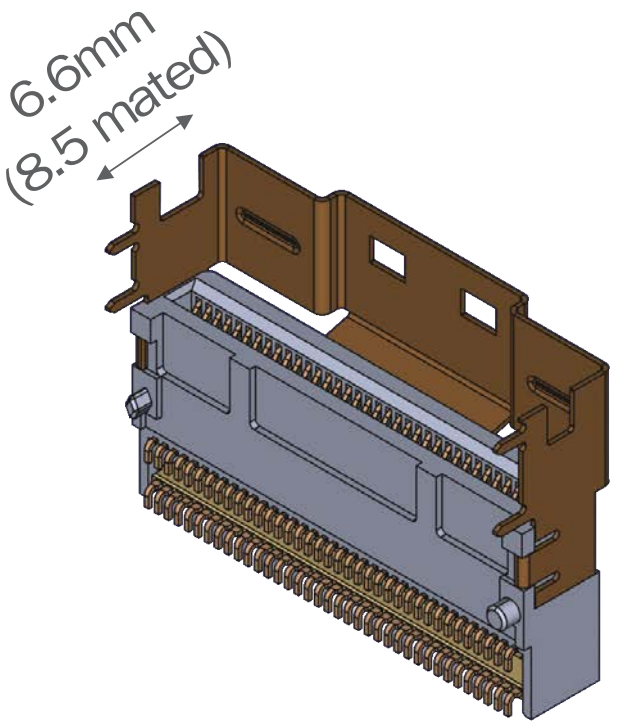


Other NVLink & OpenCAPI Interface Option*



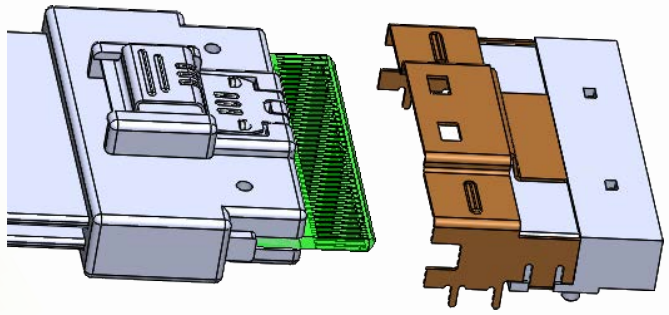


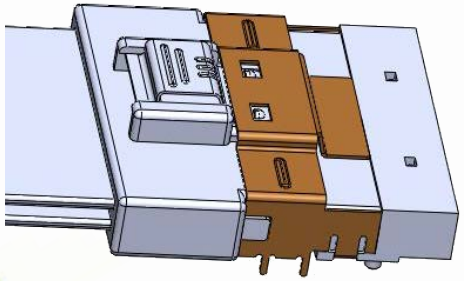




* For illustrative purposes only. Could be used in cards, but is not mounted on motherboard

Other NVLink & OpenCAPI Interface Option*



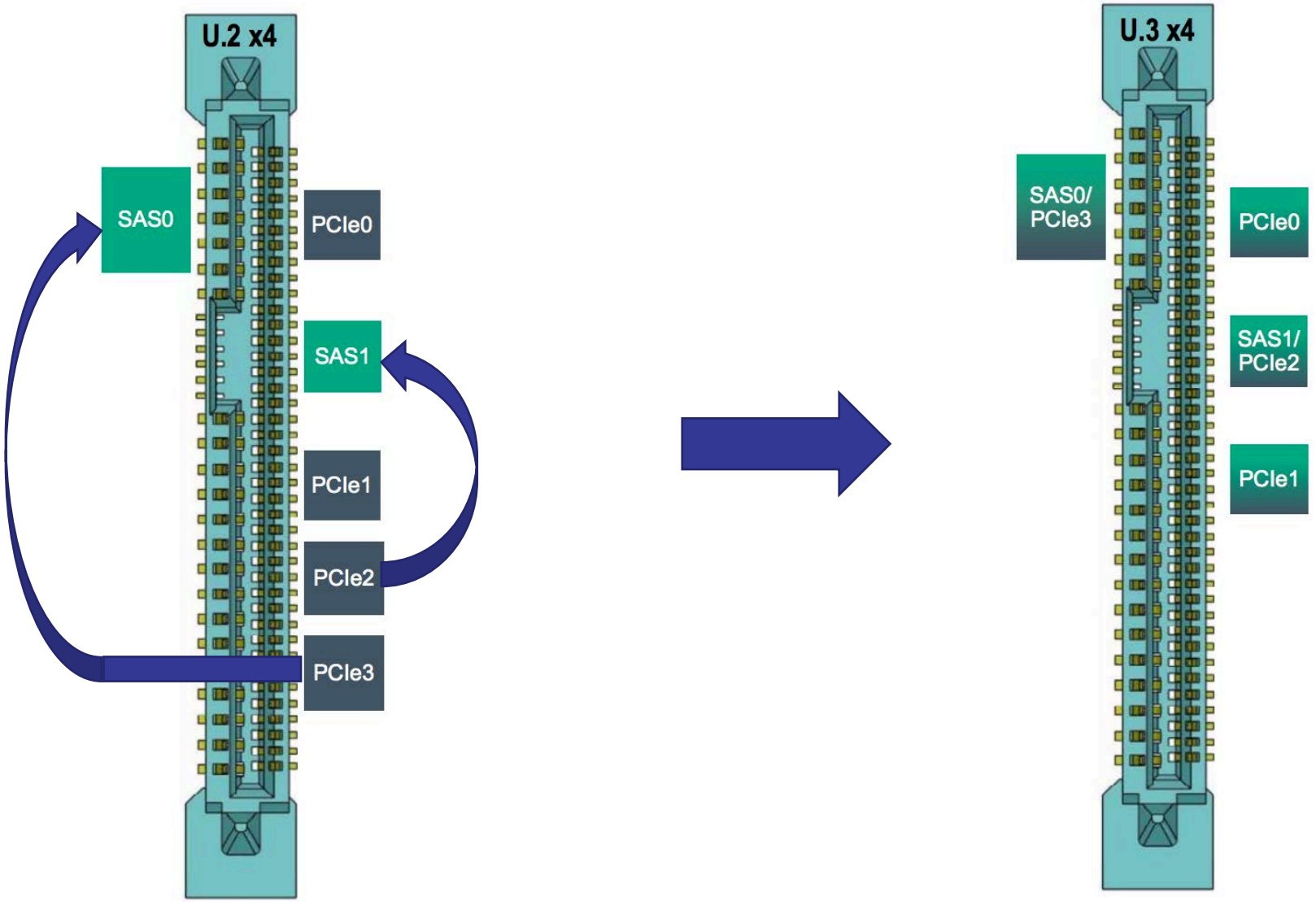


SPECS	MEASUREMENT
RECEPTACLE WIDTH:	23.5mm
PLUG WIDTH:	25.95mm
MATED HEIGHT:	8.55mm
MATED LENGTH:	25.3mm
30AWG DIAMETER:	0.255mm
PCB THICKNESS:	1.0mm

DESCRIPTION	QTY	TOTAL THROUGHPUT (UNI-DIR)
NVLink 2.0 / OpenCAPI 3.0 Brick	-	25 Gbytes / sec (200 Gbps)
Bricks per Socket	2	50 Gbytes / sec (400 Gbps)
Bricks per Server	4	100 Gbytes / sec (800 Gbps)

* For illustrative purposes only. Could be used in cards, but is not mounted on motherboard

BG2 Storage Limitations: U.2 vs U.3



OPEN
Compute Project

Barreleye G2 48V Step Down

