



Case Study: NVIDIA GB200/300

John Norton, Mechanical Engineer | Hot Chips, August 24, 2025



Agenda

GB200/300 and the MGX Architecture

- MGX Architecture

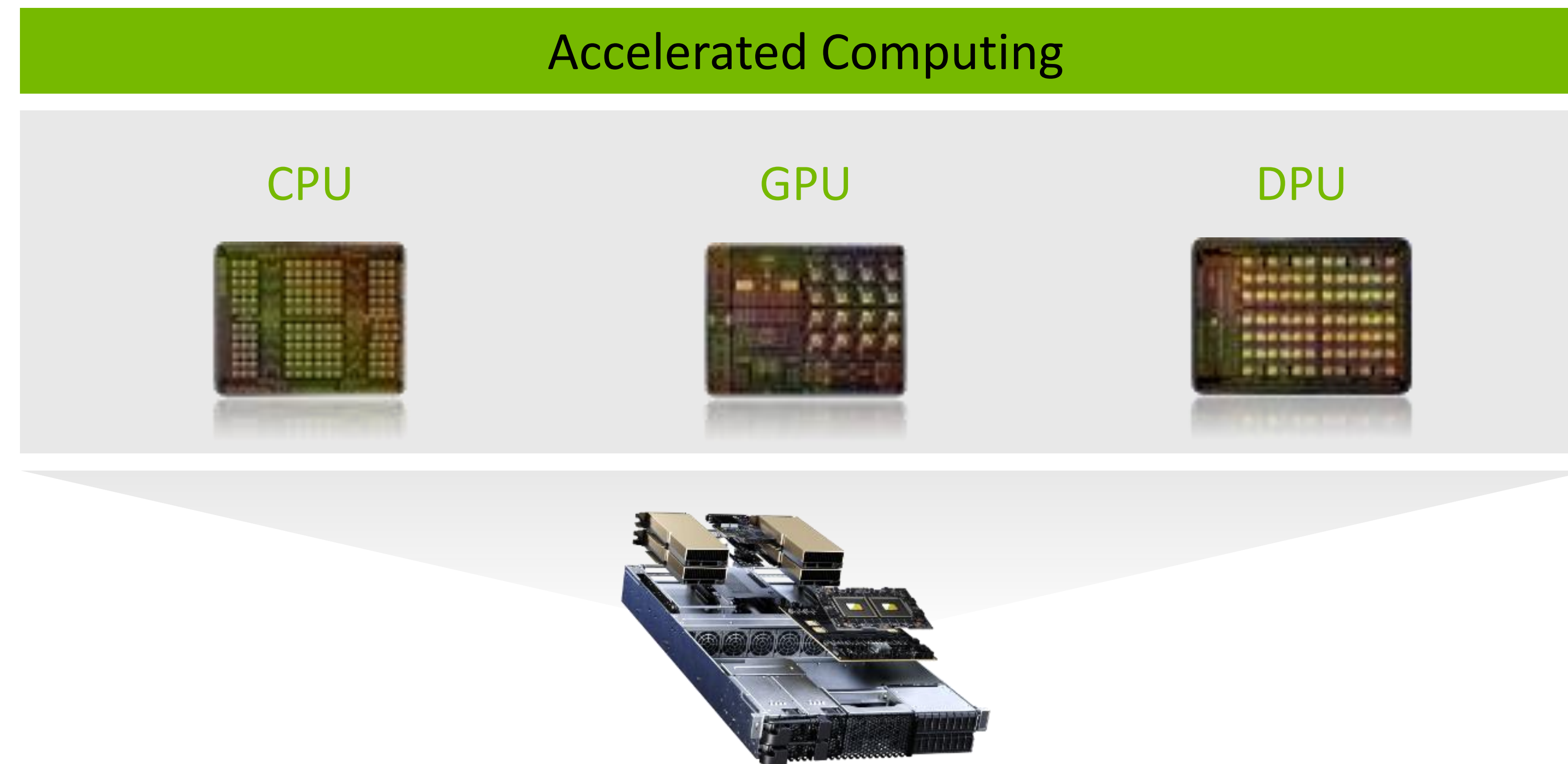
- OCP Contribution

- GB200/300 Challenges

- Product Implementation

MGX Introduction

A solution to tackle the challenge of scaled compute: GPU-centric modular reference architecture

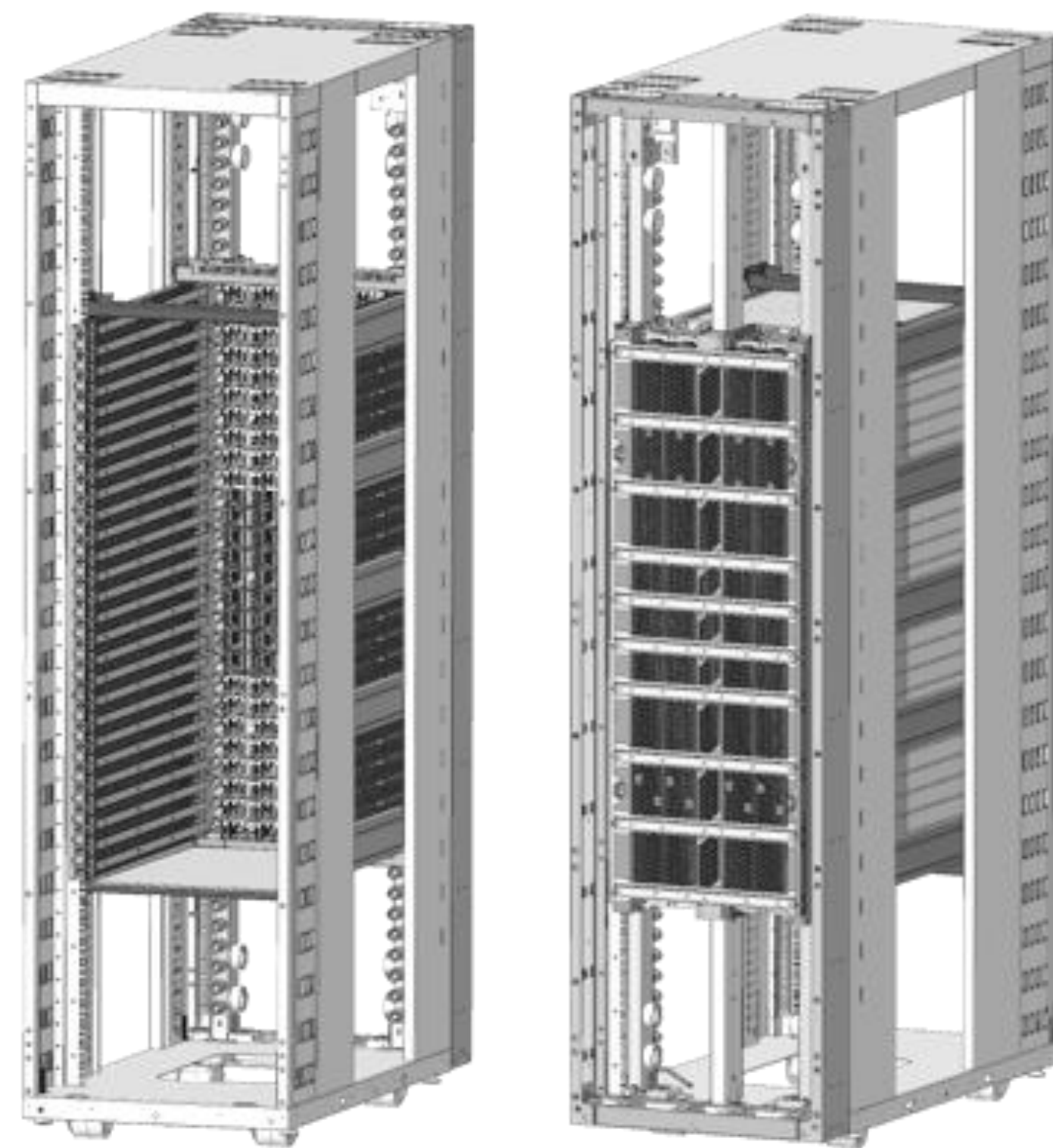


- Time to Market
- Multi-gen Compatibility
- Open & Flexible



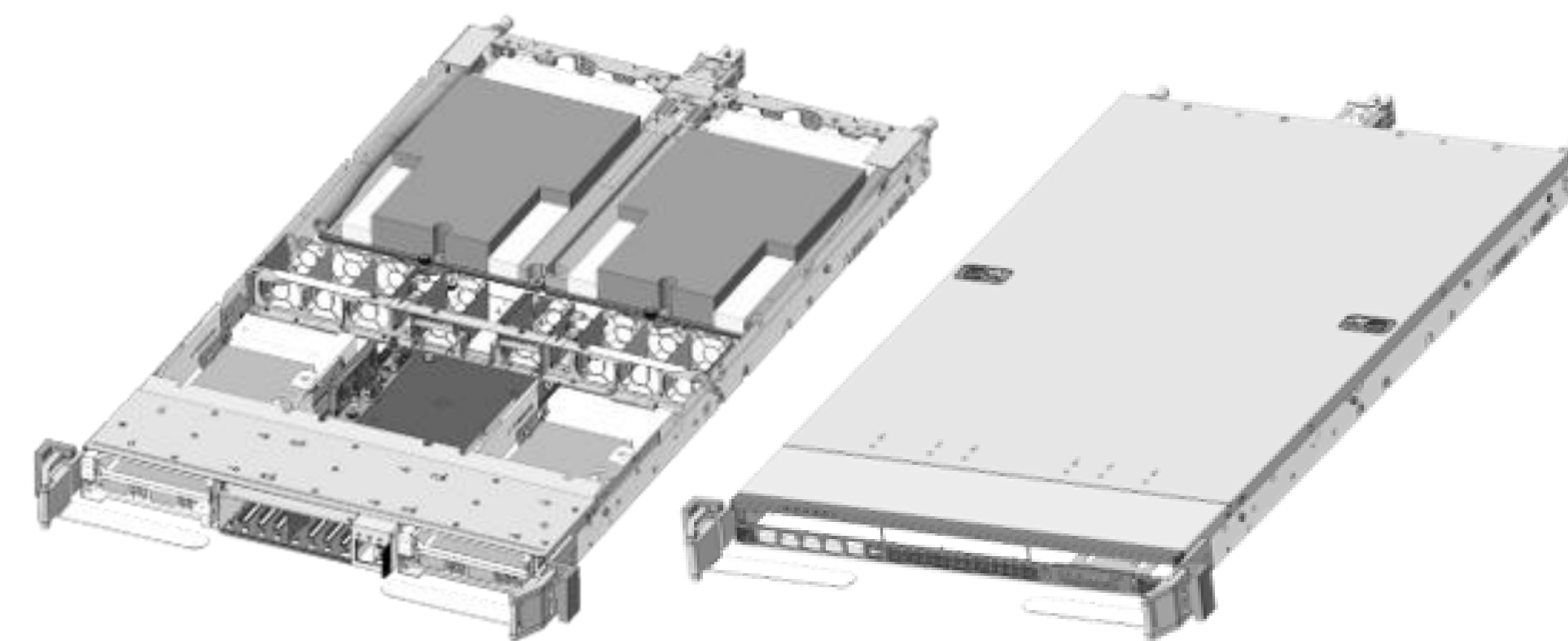
MGX Accelerated Computing Rack & Tray Specification

An OCP contribution derived from MGX architecture/built on top of OCP specifications



MGX Rack Infrastructure

19" EIA support and expanded front cable volume
1400A bus bar
Blind mate rear cables and liquid cooling



Compute and Switch Tray

Modular front IO bay design
New compute board form factors
Common chassis design



GB200/300 System View

Rack Layout



Front View

Top Of Rack Switches

Power Supplies

10 Compute Trays

9 Switch Trays

8 Compute Trays

Power Supplies

NVLINK Spine

Bus Bar

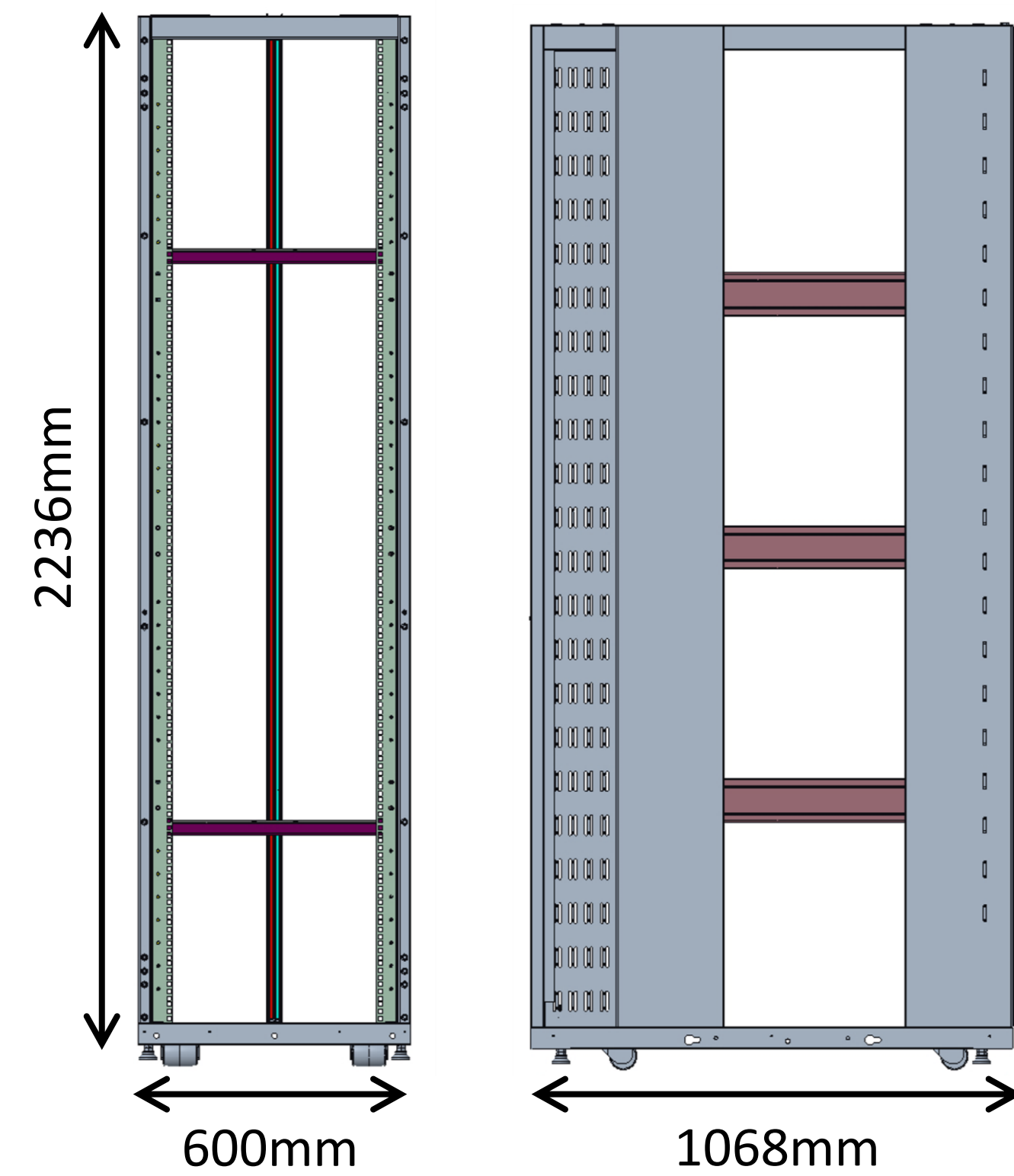


Rear View

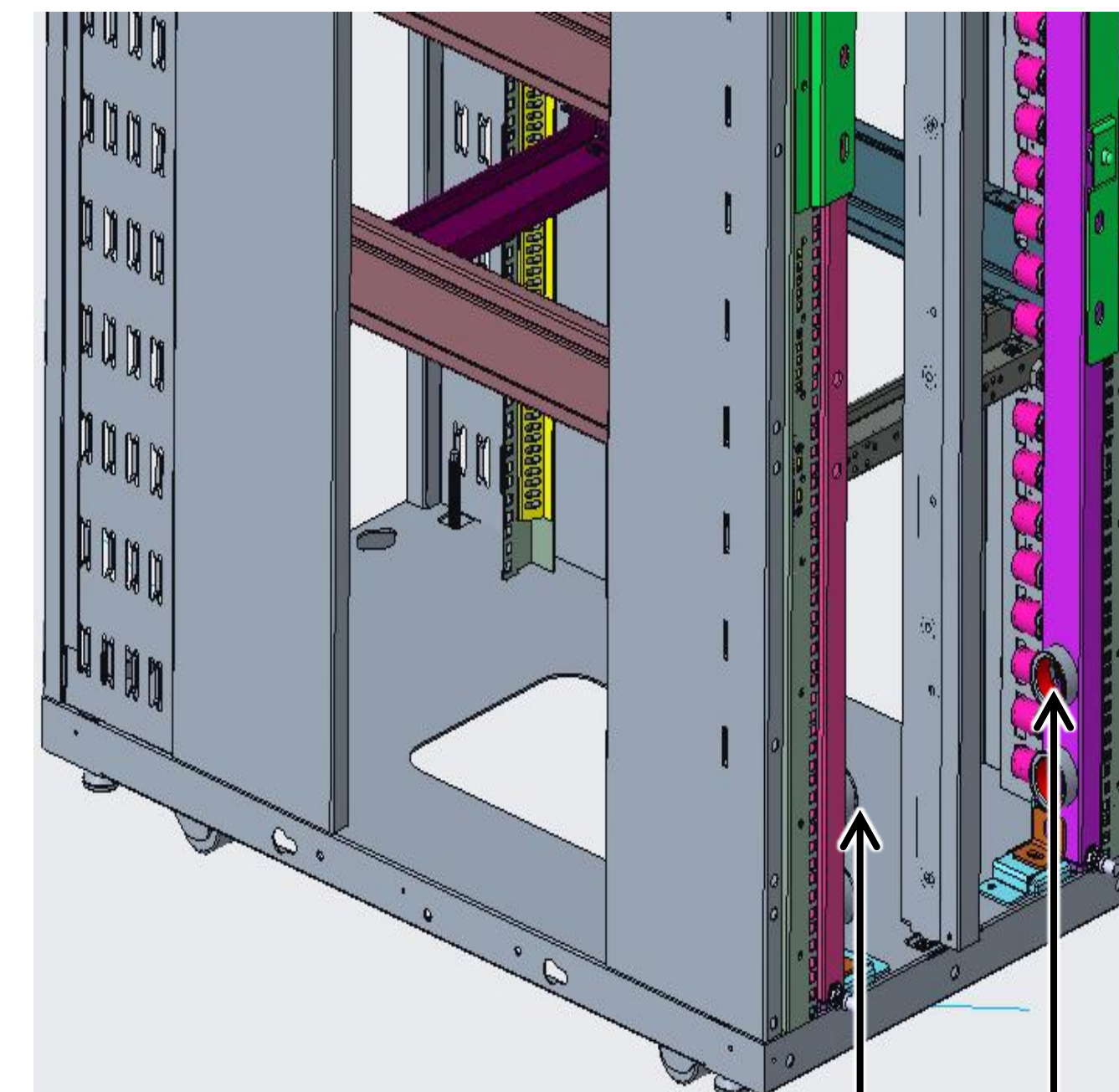
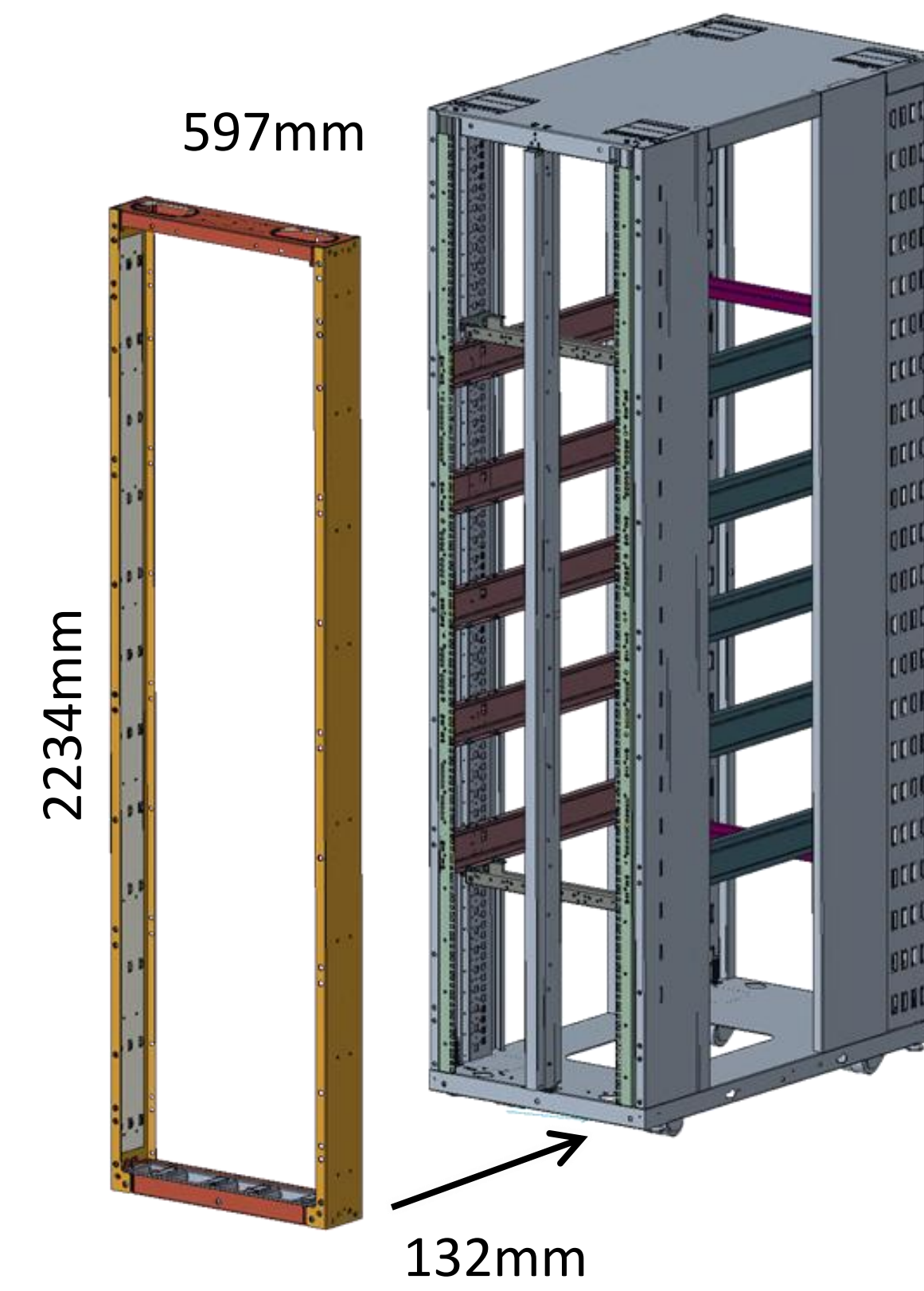
GB200/300 Rack Overview

Built on top of OCP and EIA industry standards

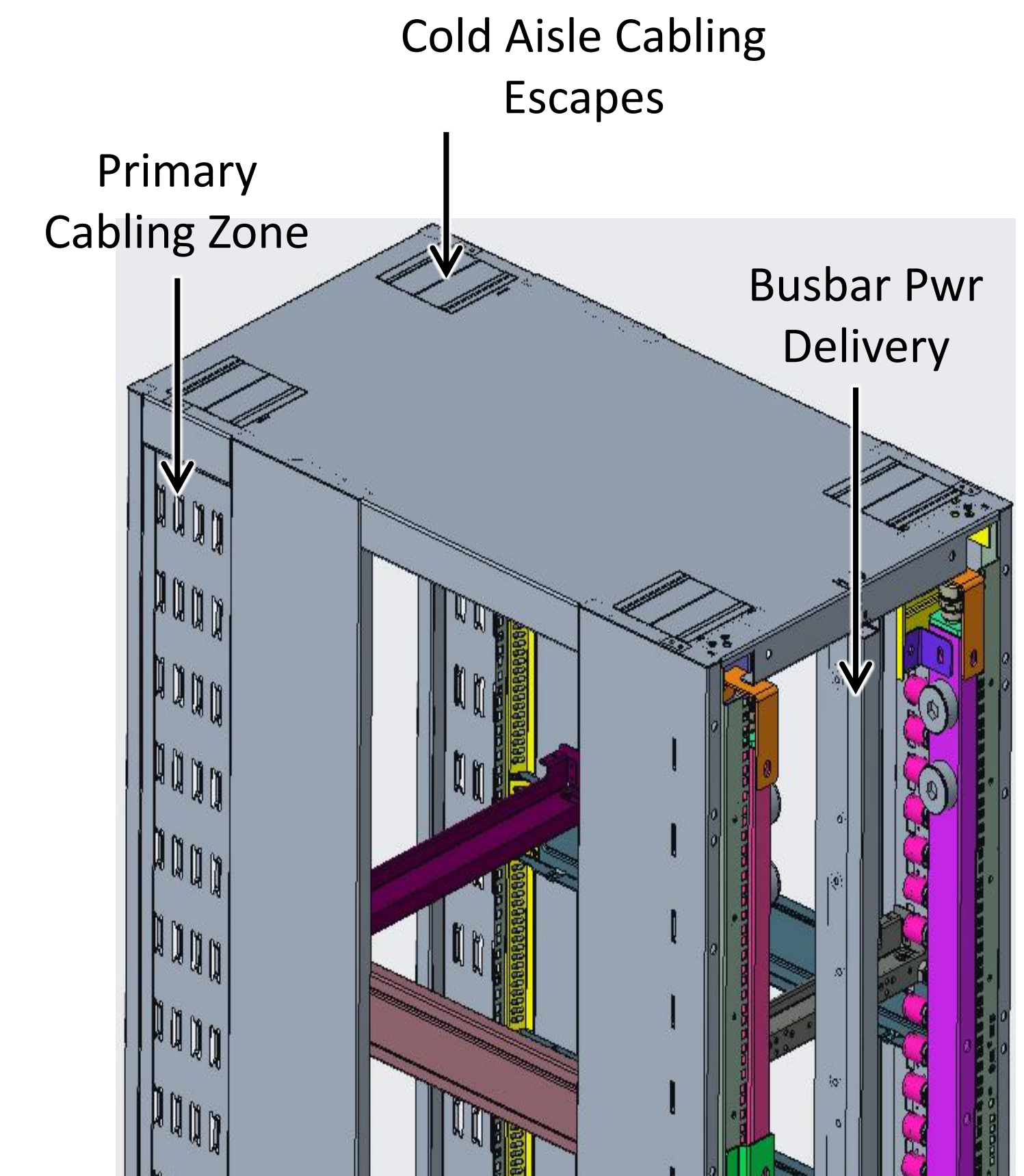
MGX Rack (19" RU Pitch)



Extension Frame



Blind Mate Manifolds

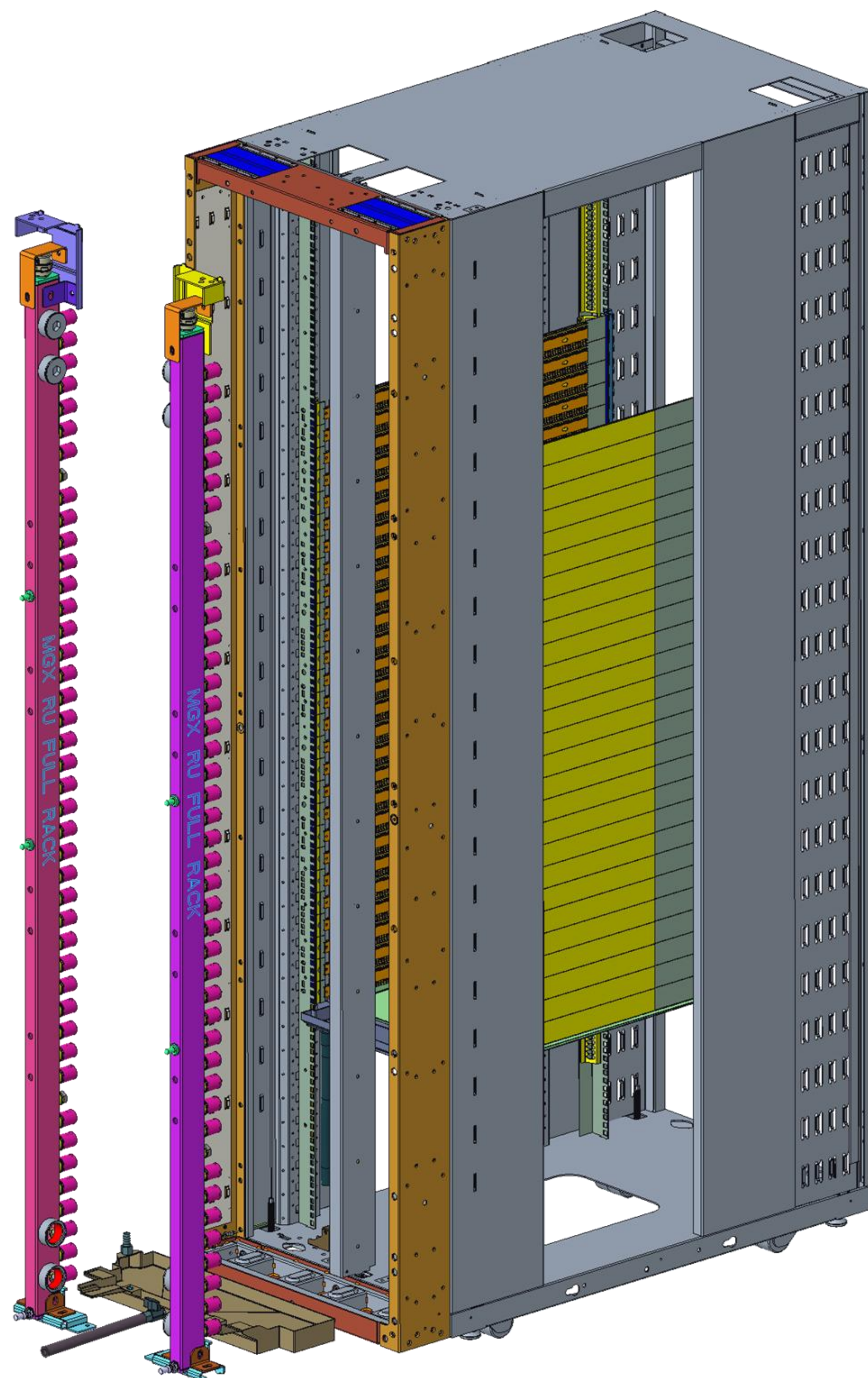


- NVIDIA open contribution to OCP
- 1RU trays for density and front cable volume
- Blind mate liquid cooling manifold

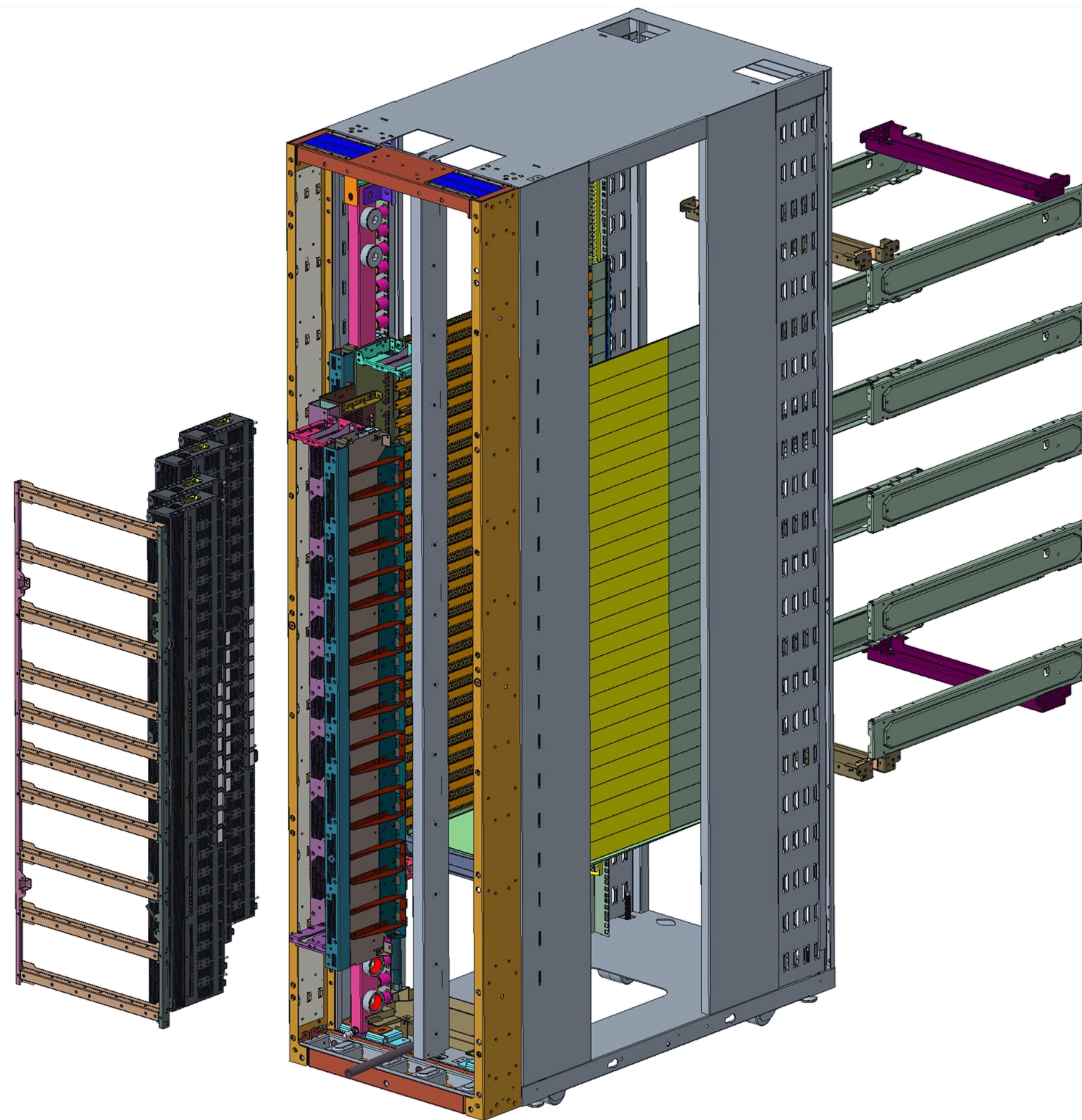
- 1400A bus bar
- NVLINK spine support
- Deployable in enterprise EIA 19" ecosystems

Liquid Cooling and NVLINK Spine

Enabling density and copper interconnects



Blind Mate Rack Manifolds



NVLINK Spine Integration



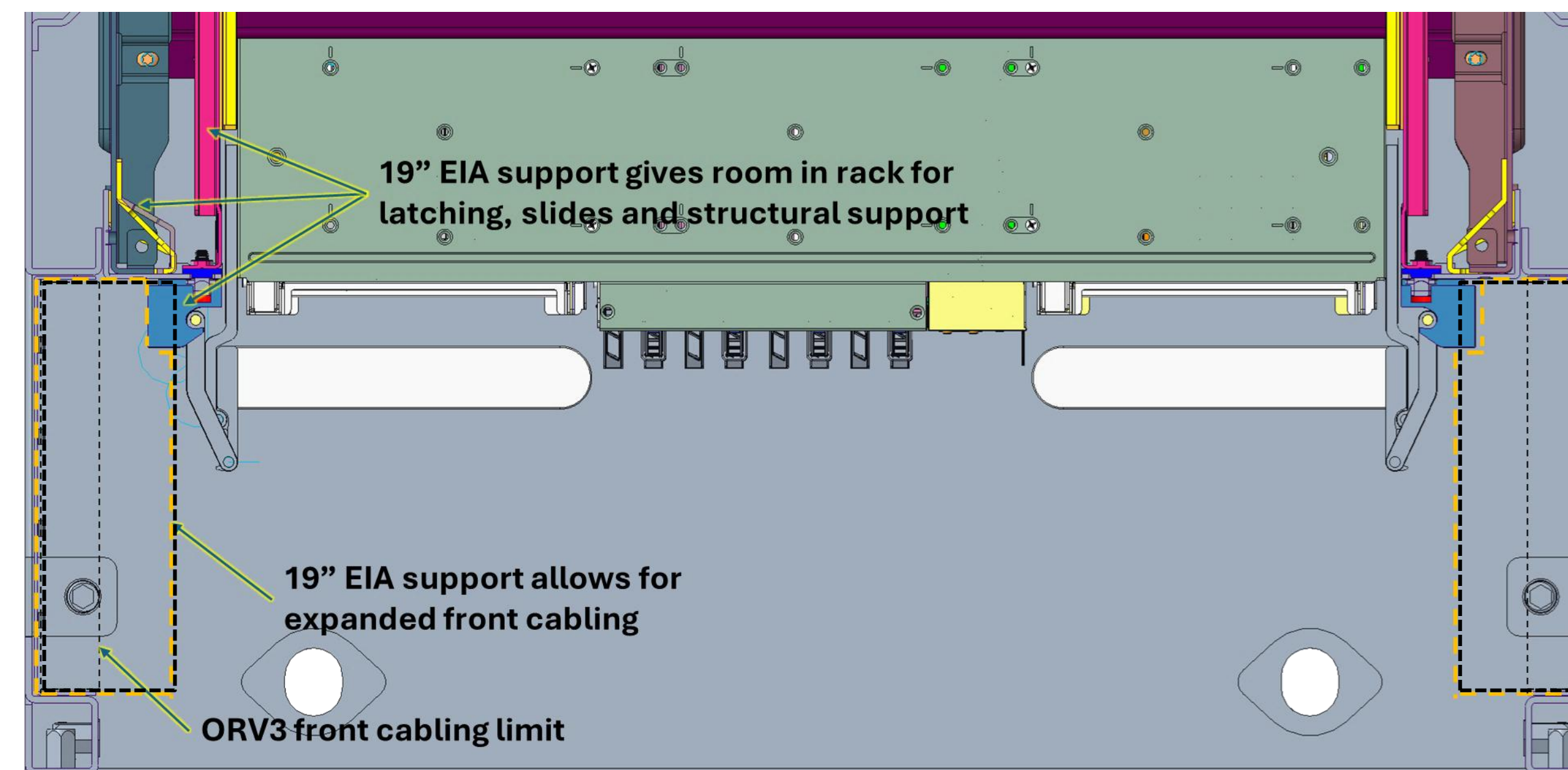
Complete GB200 Rack

Compute and Cable Density

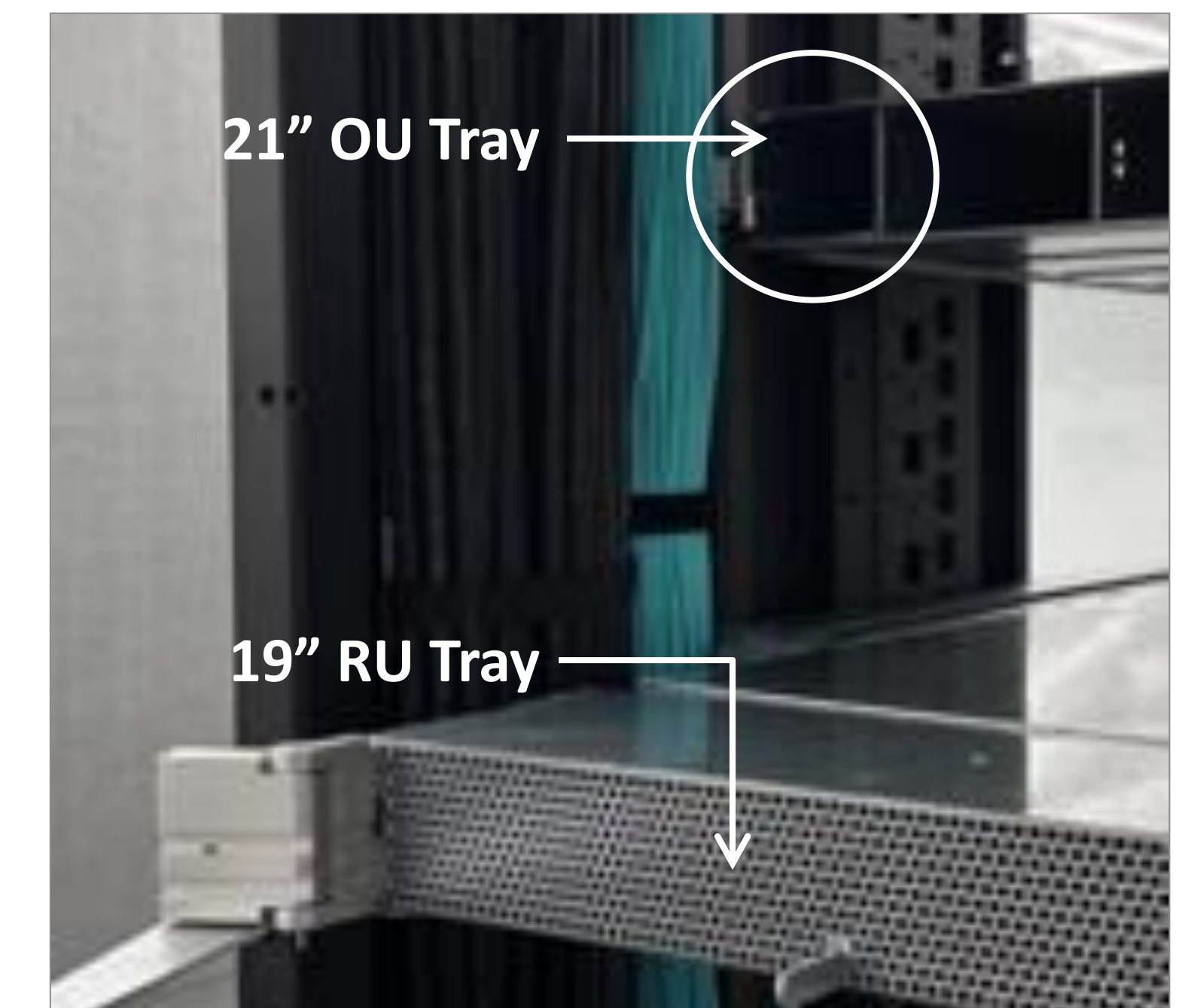
Advantages of 19" RU for accelerated computing



- 48RU vs 440U units yields more node density
- Tighter spacing of ~120mm across 27RU GB 200/300 L1 domain improves NVLINK signal integrity



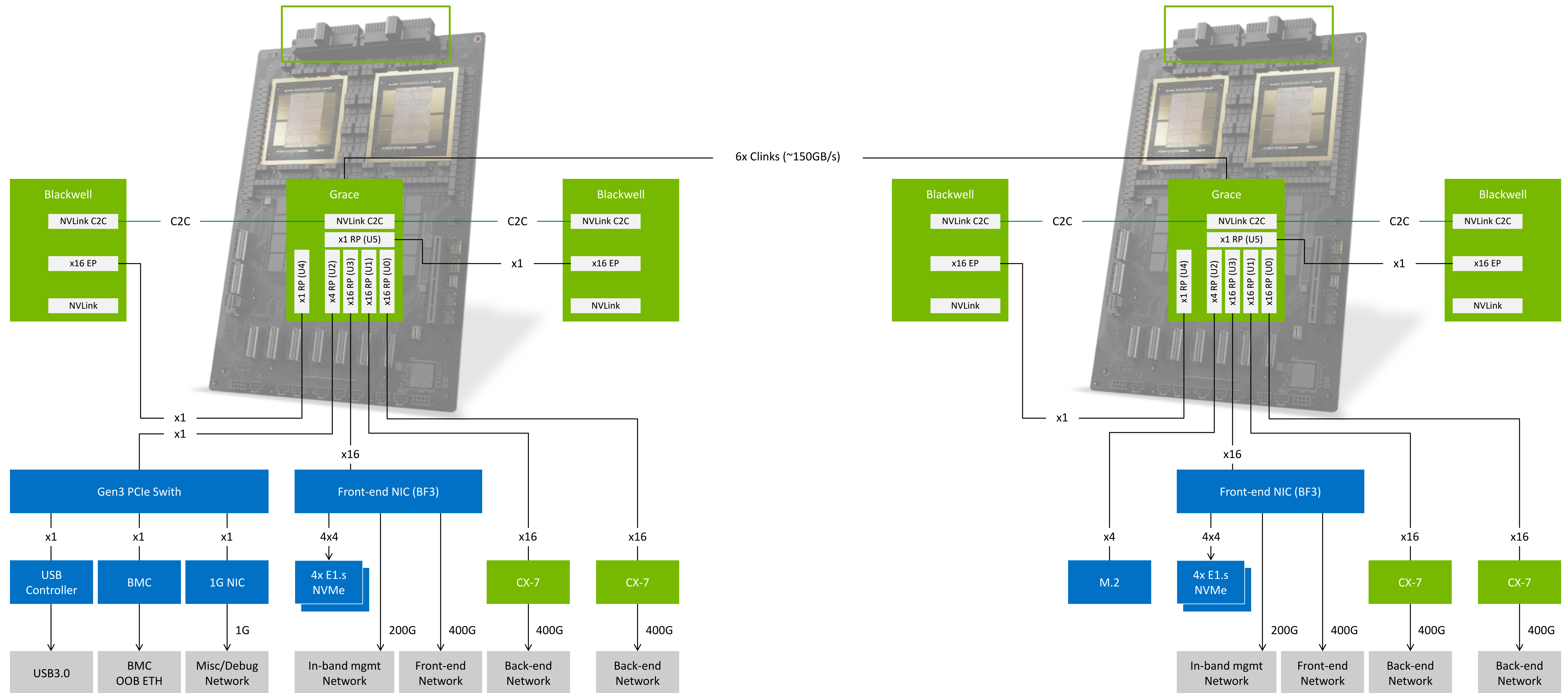
- Volume for latching and structural support
- 19" Systems can fit into EIA and OCP infrastructure



- Expanded front cabling volume

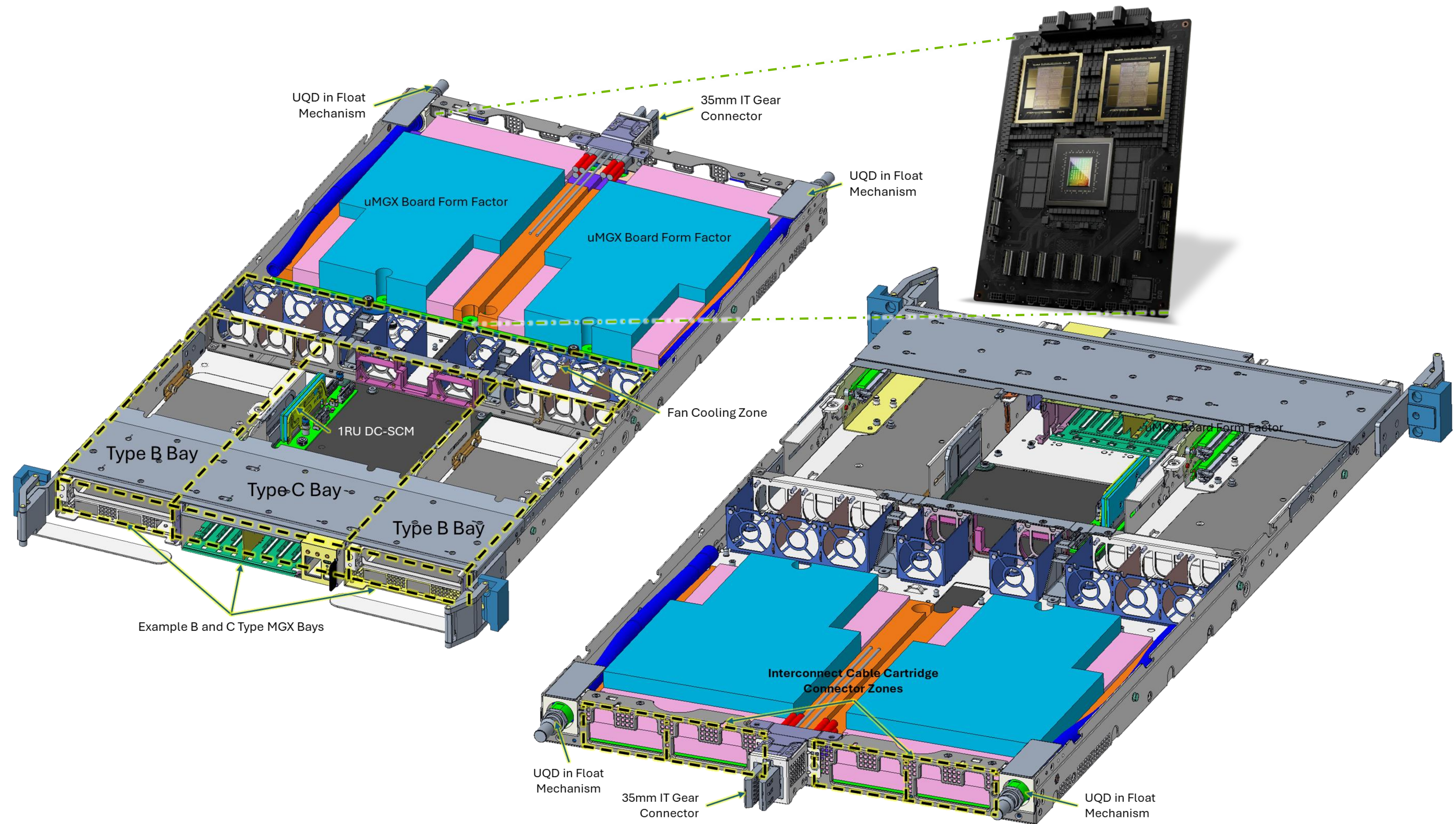
GB200/300 NVL Compute Tray

2P:4GPU compute tray PCIe topology



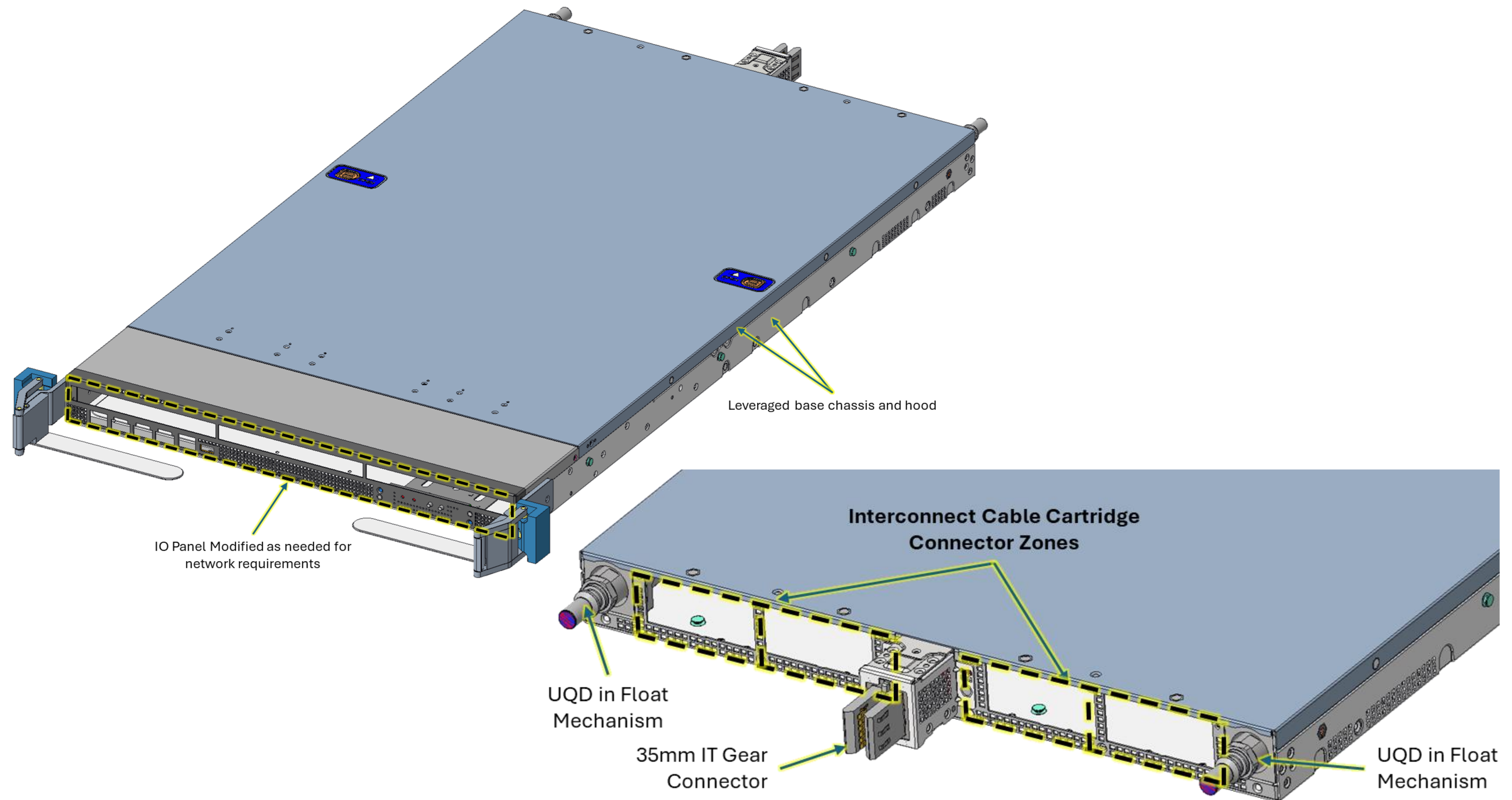
MGX Accelerated Computing Trays

Compute tray



MGX Accelerated Computing Trays

Switch tray



What's Next?

Evolving the architecture

- Open and implemented in GB200/300
 - Future Products
- Pushing Boundaries
 - Density
 - Power
 - Cooling
- NVLINK Fusion
 - Open interconnect to enable other accelerators in this open rack architecture



