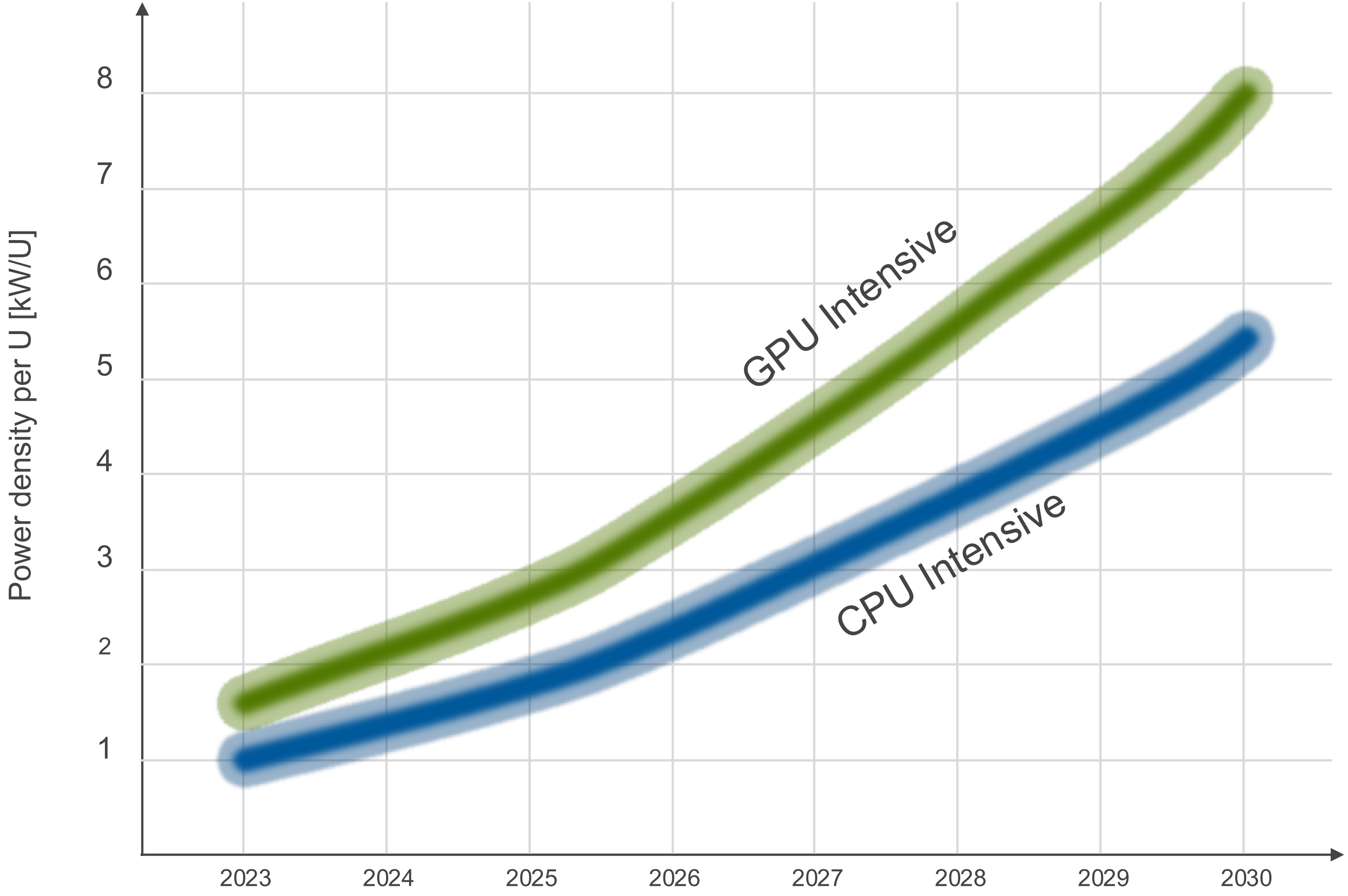
Open Rack v3
A new form-factor for high density computing

Paolo Bianco
EMEA SOUTH HPC&AI Business Developement
Paolo.bianco@dell.com

DELL Technologies

Rack Power Density Trends

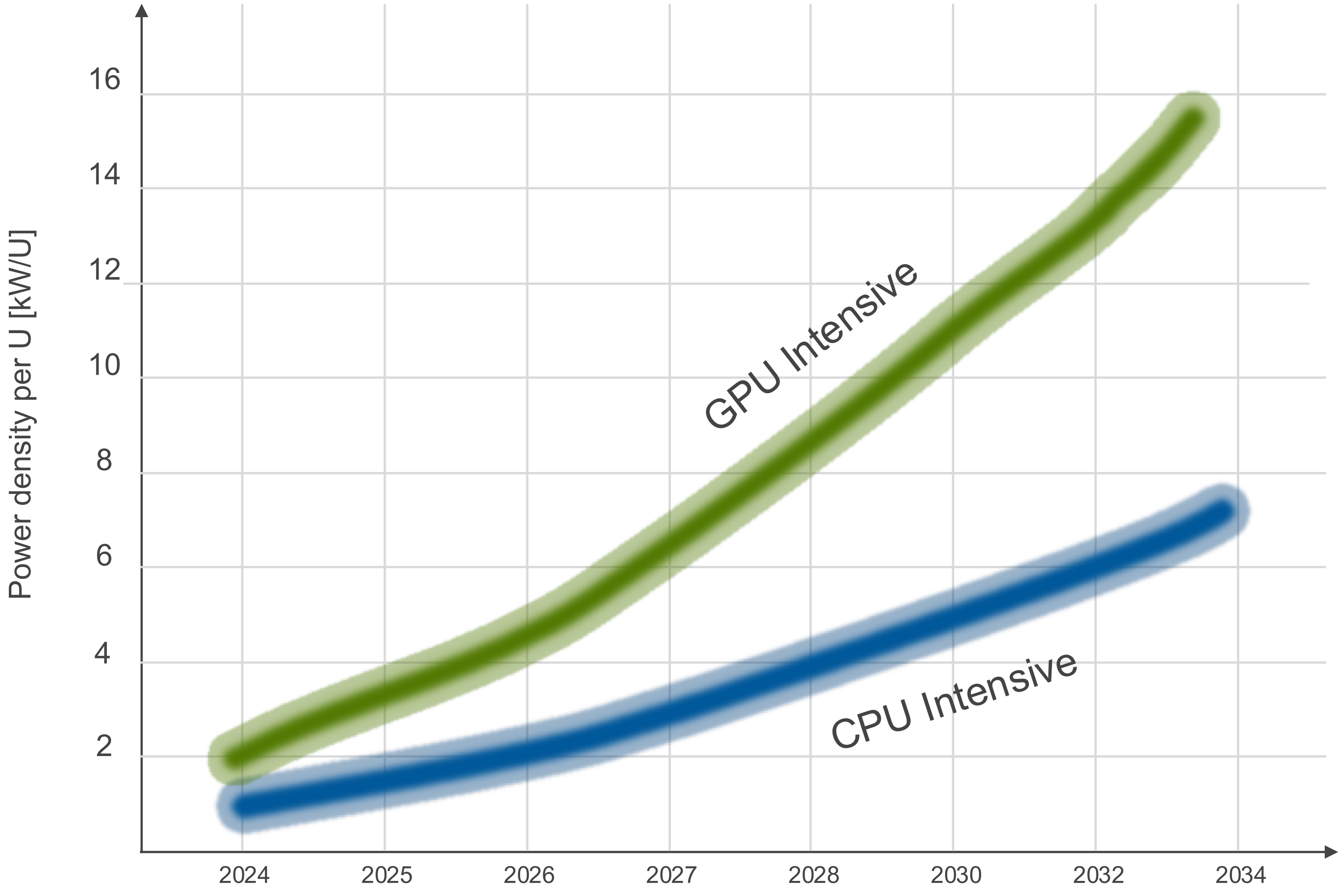
Guidance on rack power density based on chipsets power trends.
(Uncertainty increases with time)



DELLTechnologies

Rack Power Density Trends *revised!*

Guidance on rack power density based on chipsets power trends.
(Uncertainty increases with time)

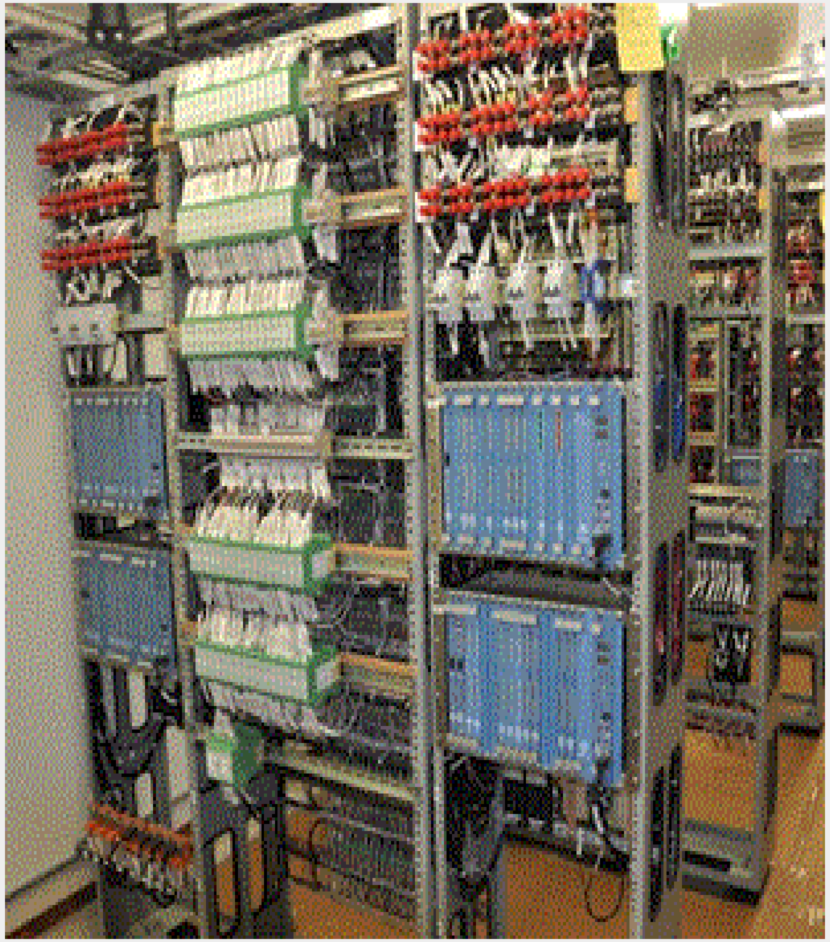


DELL Technologies


Physics
Contemporary System Design

Traditional designs not suitable anymore

- Traditional 19” rack form factor design (EIA 310-D) traces its origins back to railroad signaling relays.
- It wasn’t designed for data centers and scale computing in mind (1950!)
- EIA 310-D standardized only the width between the inner rails in a rack (19”), then in 1992 the Rack Unit height was standardized (1RU = 44,45mm - 1,75”)
- Common 19” racks weaknesses (at today):
  - Insufficient width for high density computing (ie AMD 4N2U)
  - Insufficient space for proper cooling
  - Power distribution complex to implement for high power draws.
  - Cabling coexistence with DCLC can be complex

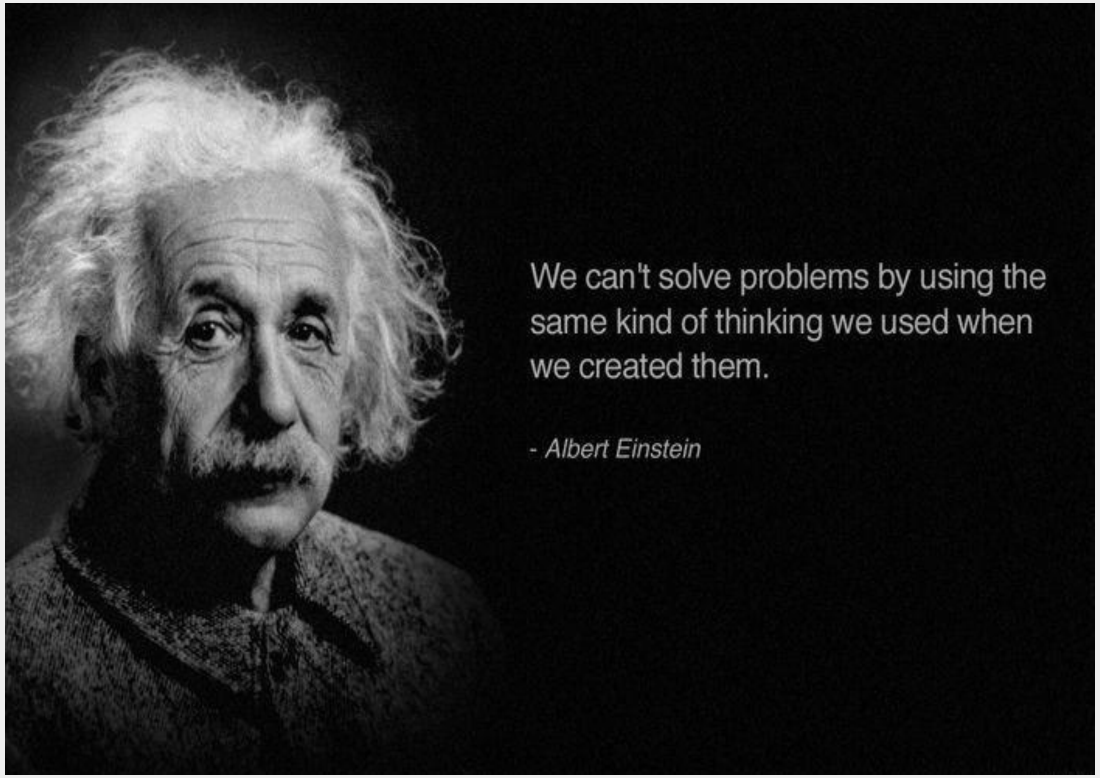


HPC/AI Computing Landscape is changing

- Upcoming AI and HPC chipsets are driving increasing power requirements.
- New AI workloads need compute at scale
- Given the widening adoption of Liquid Cooling, node density demand is increasing again.

NEED FOR A CHANGE IN THE WAY WE DESIGN COMPUTE SYSTEMS!!

- When dealing with scale compute ecosystems, you have to design for the data center. The motherboard is just a module in a server, the server is just a node in a rack.
- When you engineer for density, you can do the wrong thing. Simplicity is hard.



DELL Technologies

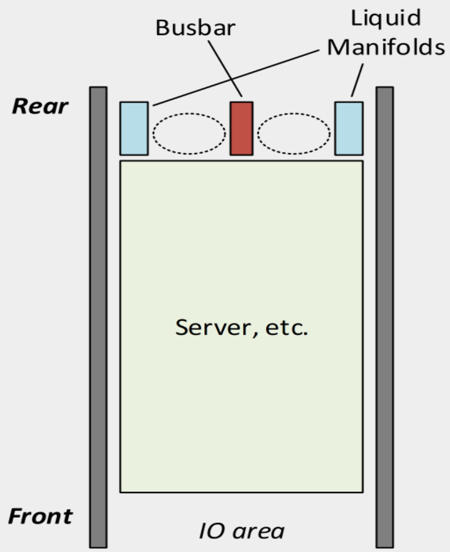
A viable alternative: OpenRack v3

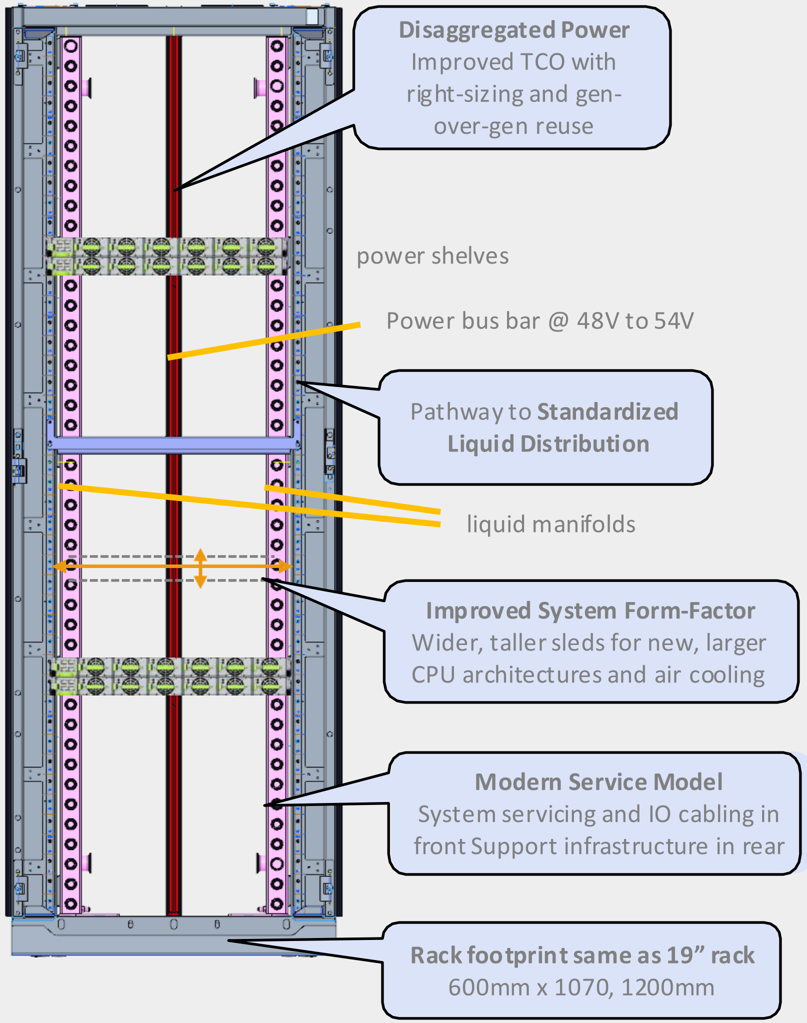
What is Open Rack?

- Open Rack is an Open Compute Project standard for a **new rack and power delivery architecture** and an **efficient, scalable** alternative to the EIA-310D rack.
- The Open Rack features a simple design, built with the scale compute space at its core.
- When you engineer for density, you can do the wrong thing. The Open Rack aims for practical density.
- ORv3 spec defines IT envelope and interfaces, liquid manifold position, and busbar location. All other “soft” dimensions are considered configurable

What are the Key Features vs EIA-310?

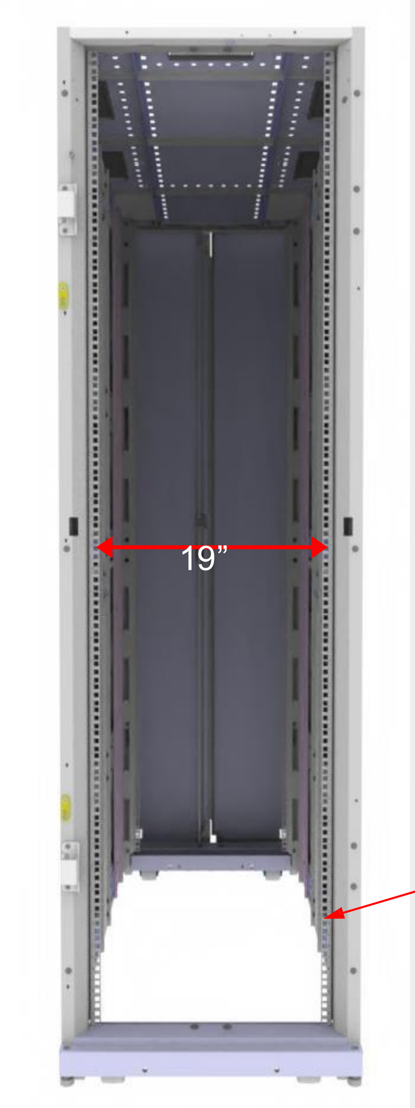
- **Disaggregated power** solution with DC distribution
- 480VAC in, 100s of kW (1.2MW?)
- Increased **Usable Equipment width** in same external cabinet width ***(~21.22” vs. ~17.75”)***
- Increased **usable height** per U **(1OU=48mm, 1 RU = 44.5mm)**
- Adoption of Front IO and Blind Plug Direct Liquid Cooling (DLC).
- Fixed depth to dock with power bus bar & DLC manifolds.
- Cold Aisle Maintenance: all cables and IT gear hot pluggable and serviceable **from the front**





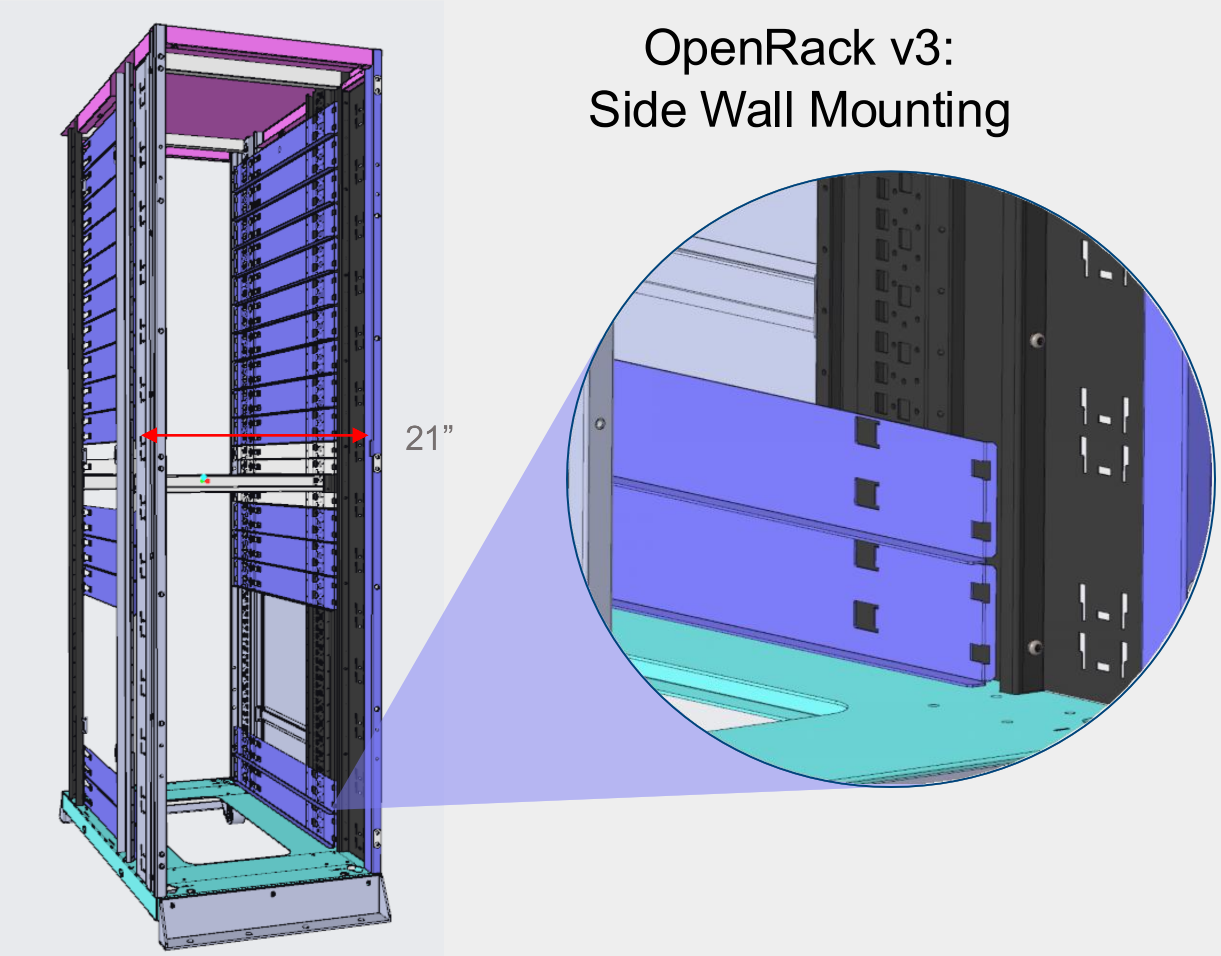
DELL Technologies

How do we gain width? EIA vs OCPv3



Traditional Rack:
Front Justified
Mounting

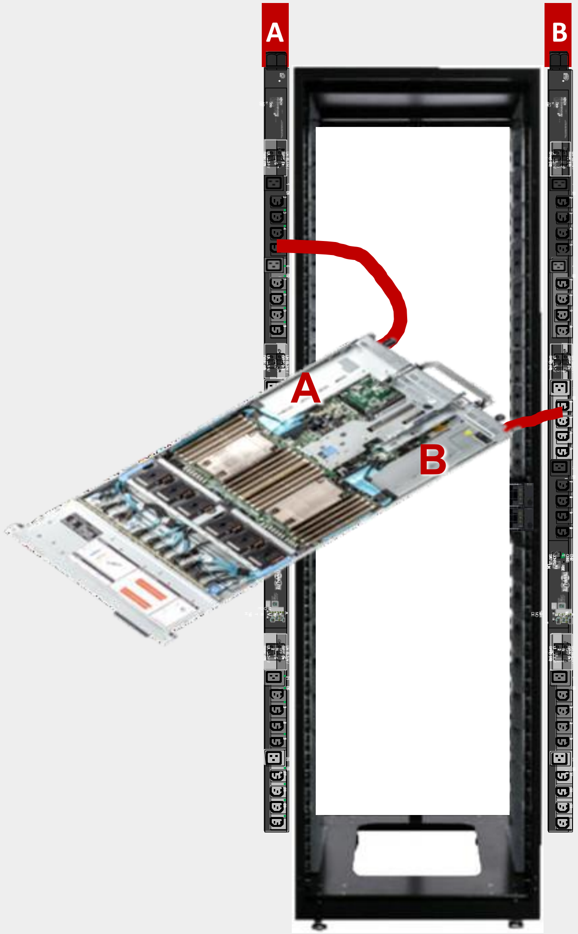
Front and back flange
for mounting



OpenRack v3:
Side Wall Mounting

ORv3 - Distributed Power

Traditional Rack Power Distribution



Primary Feeds
- 1+1 input redundancy
- 1-phase AC or DC inputs

Power Distribution
- 2-4 PDUs/rack

Per Server
- 1+1 input redundancy
- N+1/N+N PSU redundancy
- 2..6 x PSUs
- 2..6 x power cords

PSU
- 3.2kW max PSU
- 12Vdc or 54Vdc output options
- Titanium efficiency

Disaggregated Rack Power Distribution

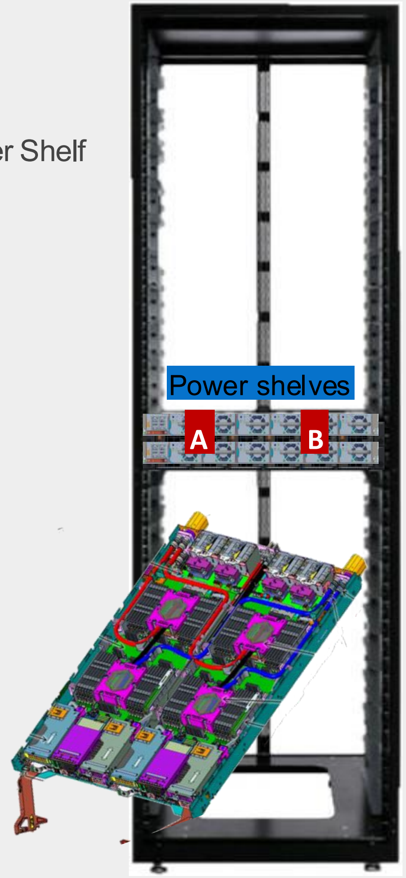
Primary Feeds
- 1+1 input redundancy
- **3-Phase AC** or DC inputs

Power Distribution
- Primary feeds connect to Power Shelf
- **48Vdc busbar**

Power Shelf
- 6 x PSUs / OU
- **Up to 33kW / OU**
- 1+1 input redundancy
- **N+1 PSU redundancy**

PSU
- **PSU w/ integrated ATS**
- 1+1 input redundancy
- **97%+ efficiency**

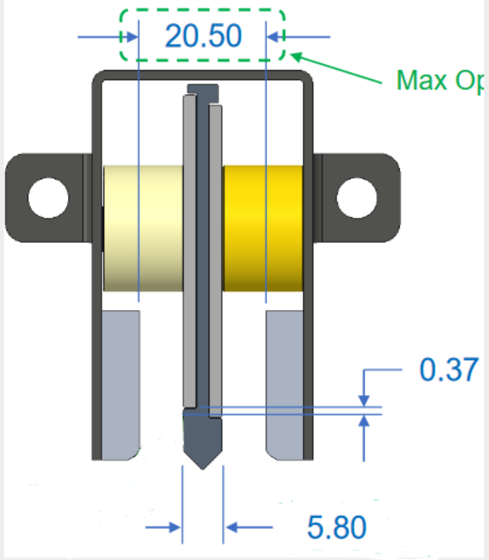
Storage Platforms
- **Busbar Adaptor**

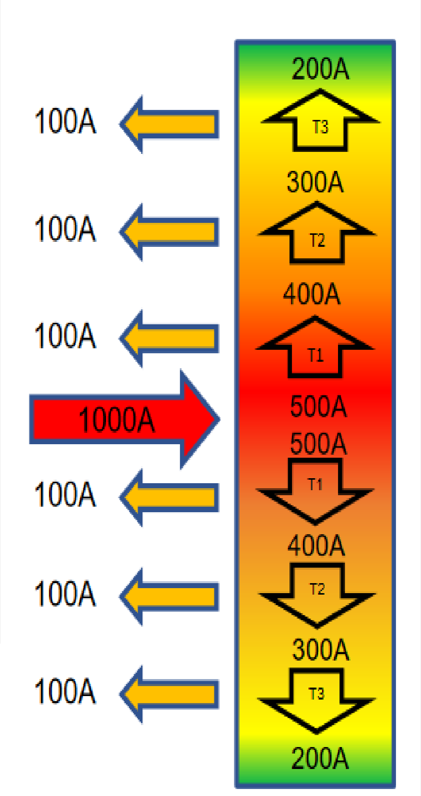


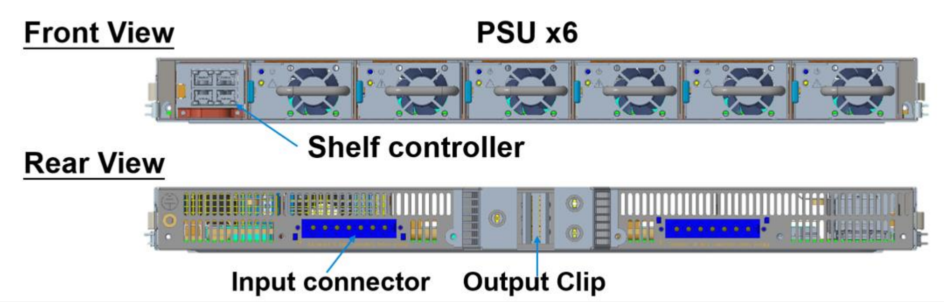
DELL Technologies

Power System

- 1600A busbar (3200A in roadmap)
- Hot-plug “clip” connectors
- Multiple PSUs power shelves
  - 48V DC, 33kW each at today
  - Plan to transition to 400V DC (1MW+ per rack!)
  - Integrated ATS
- Multiple power Shelves can be distributed on the busbar based on system needings



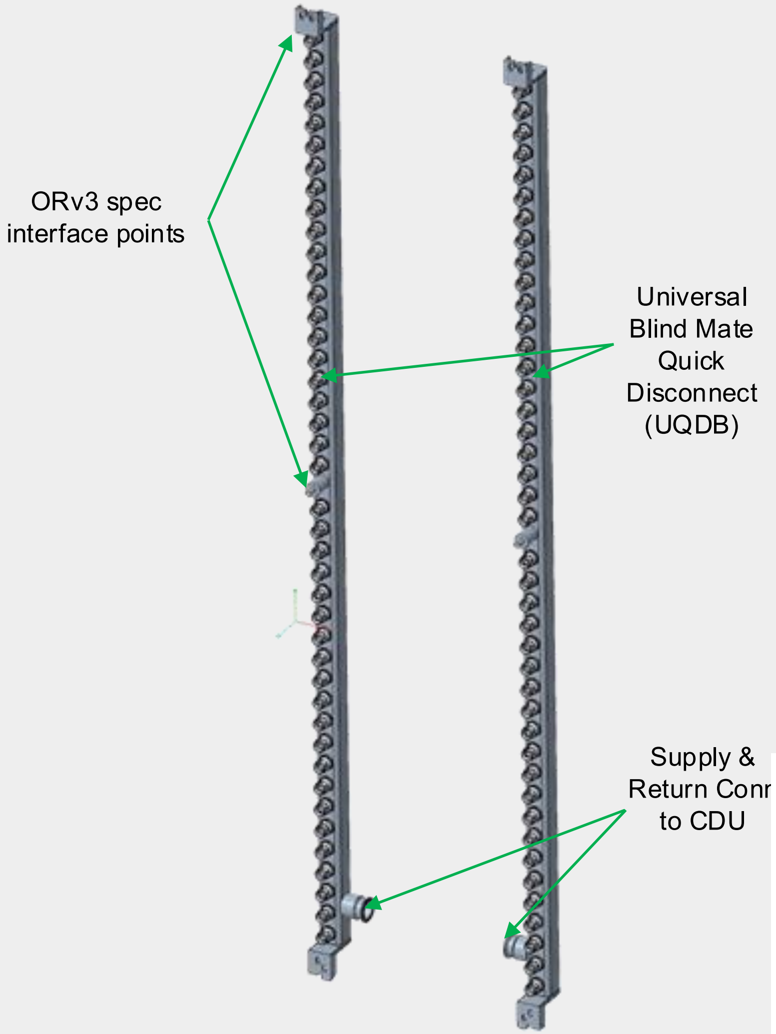


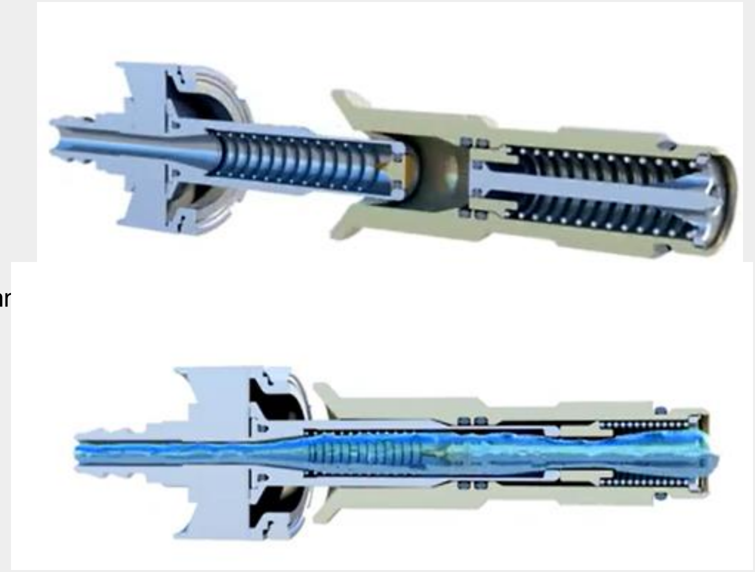


DELLTechnologies

Cooling

- Hybrid-Cooled (DLC + Air)
- Blind-Mate Quick Disconnect with alignment cup
- Floating blind-mate node connectors
- In-Rack or In-Row CDUs

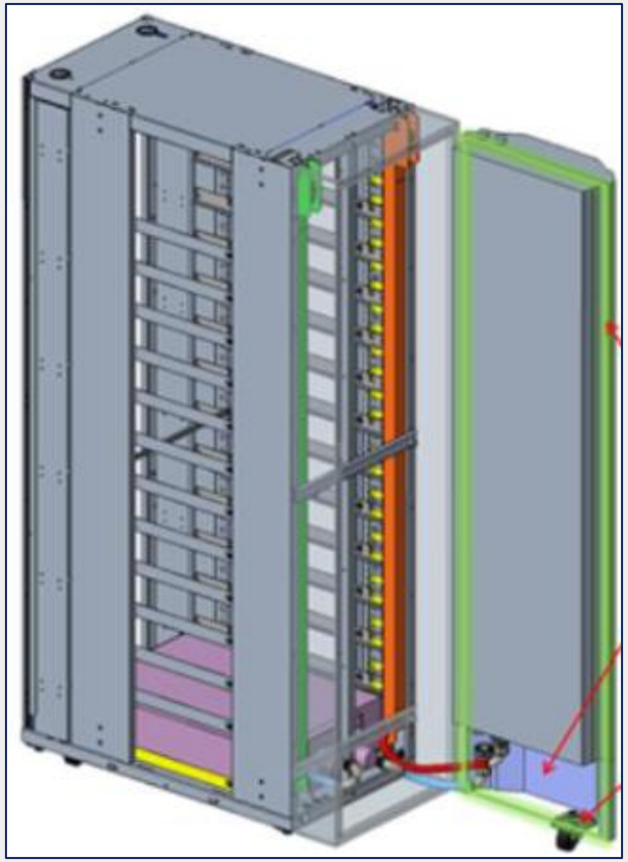




DELLTechnologies

Room-neutral implementations

- Heat capture at rack level either through:
  - Rear Door Heat eXchangers (RDHx)
  - Closed-Loop Heat Capture Cabinets (HCC)
- OpenRack Specs provide RDHx designs, however rack size is standard so it can mount any market-available RDHx
- Dell HCC can collect 100% of the heat in 32°C water (ASHRAE W3 compliant)

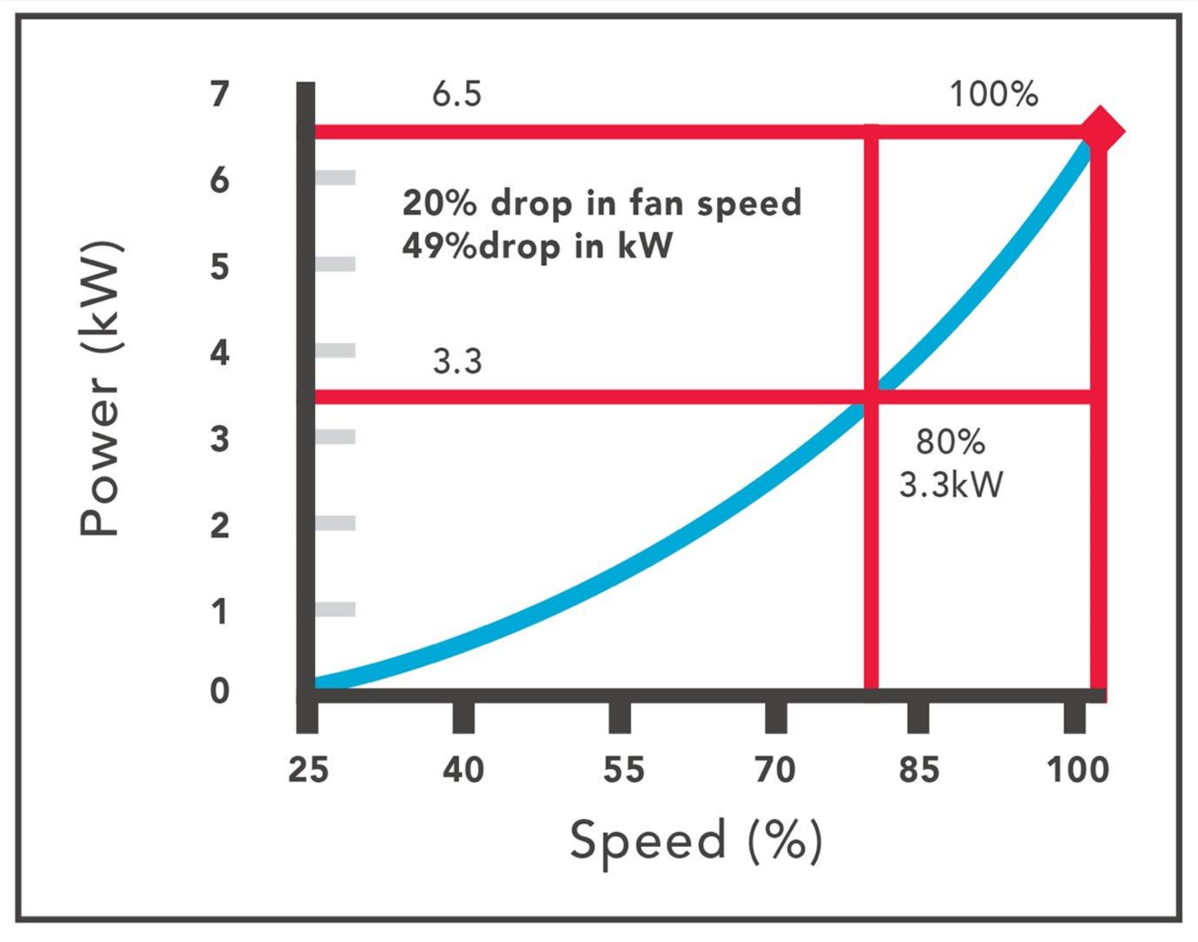


OpenRack V3 RDHx
(from OpenRack docs)

DELL Heat Capture Cabinet
(concept)

Efficiency: Fans & Liquid Cooling

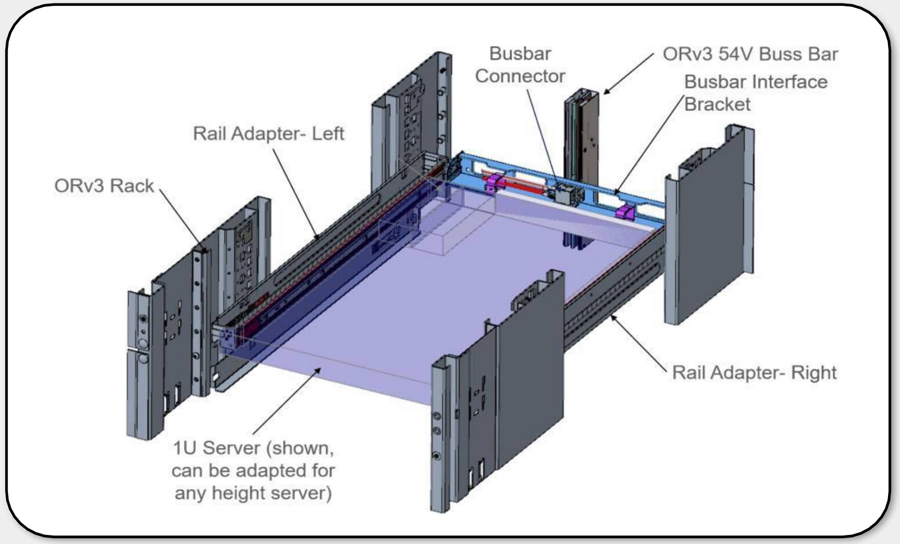
1. On high power compute systems, the biggest fraction of non-production related power of a node is draw by fans
2. Power consumption of a fan is a function of the cube of the fan speed.
3. Fan speed (air flow) is directly proportional to the quantity of heat to be removed (CPU/GPU heat production)
4. If we remove CPU+GPU heat through liquid, the fan speed drops dramatically, thus leading to significative power savings.
5. Removing the remaining node heat through liquid does not reduce significantly the overall fan power draw



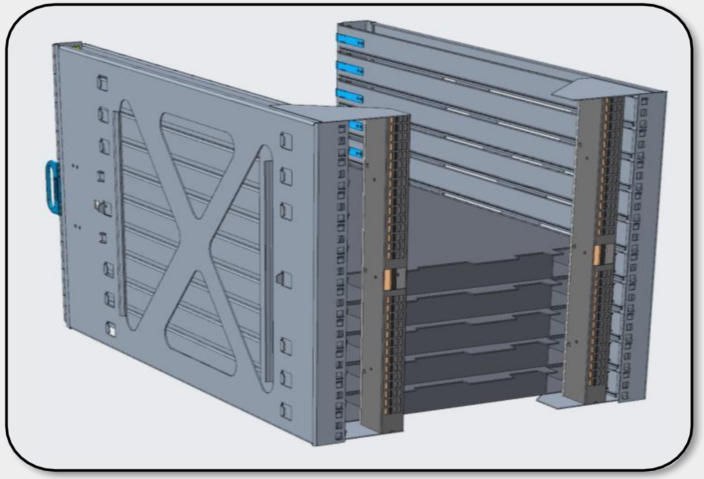
DELL Technologies

19” Rack Gear in ORv3

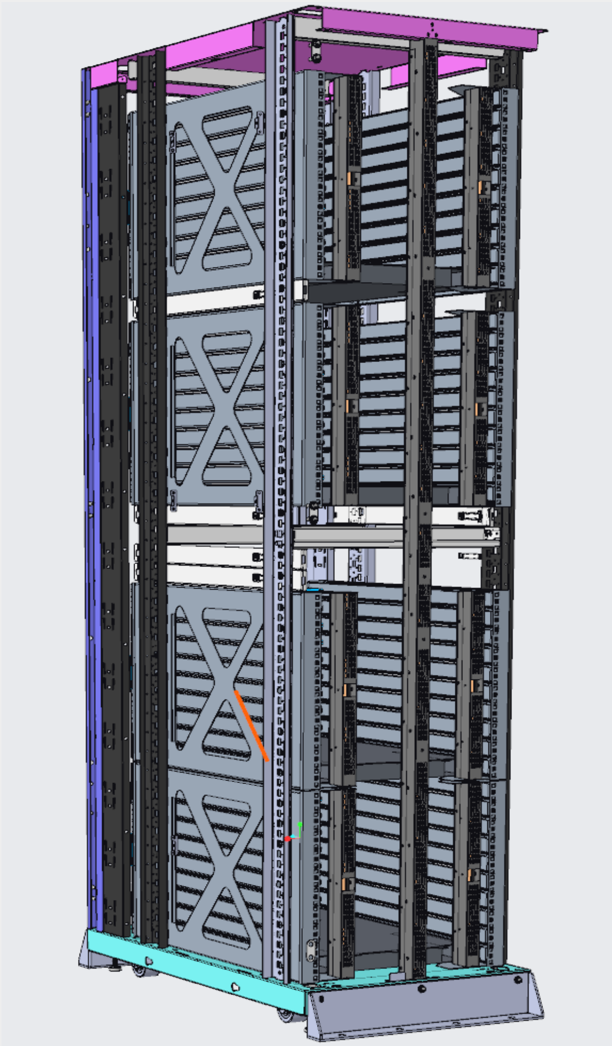
- Most ToR Switches and Storage solutions are still 19” gear.
  - Many of these solution do not have 48v options.
- Mixing PDUs and rear cabling with DLC manifolds and bus bars is challenging.



- 19” gear requires an adapter to go into an ORv3 Rack.
  - Commonly a drawer with bus bar clip.
  - Power pigtail into 48v PSU



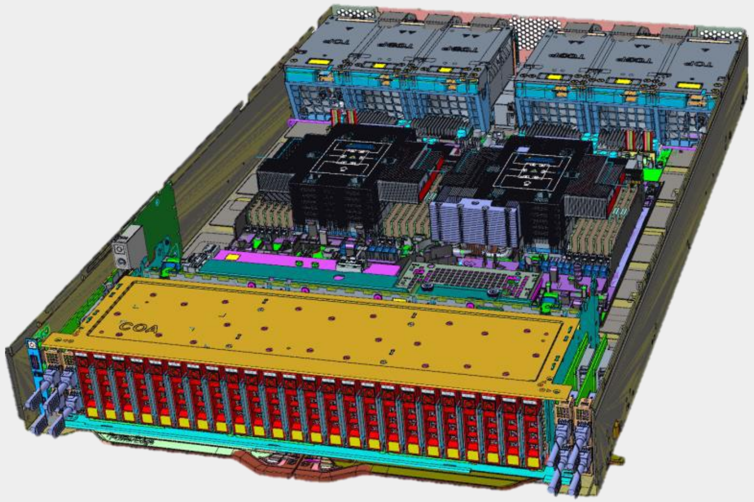
- Concept
  - 10U “Chunk”
  - 48v Server PSUs dock to adjustable depth side mount bus bars.

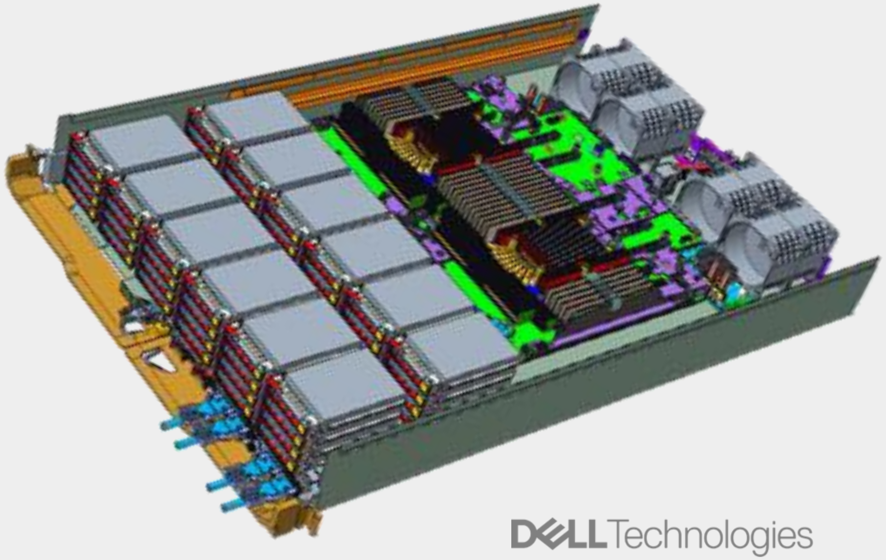


DELL Technologies

Storage in ORv3

- Storage challenges with ORv3
  - Storage typically fills the front of a chassis with media to maximize density.
    - Front I/O erodes storage density.
  - High-Availability Storage Arrays have redundant controllers that are serviced from the rear.
    - The center bus bar for distributed power make it difficult to remove
      controllers from the rear.
    - Service for ORv3 is intended to be cold aisle service.
  - Storage media does not have a viable method for liquid cooling yet.

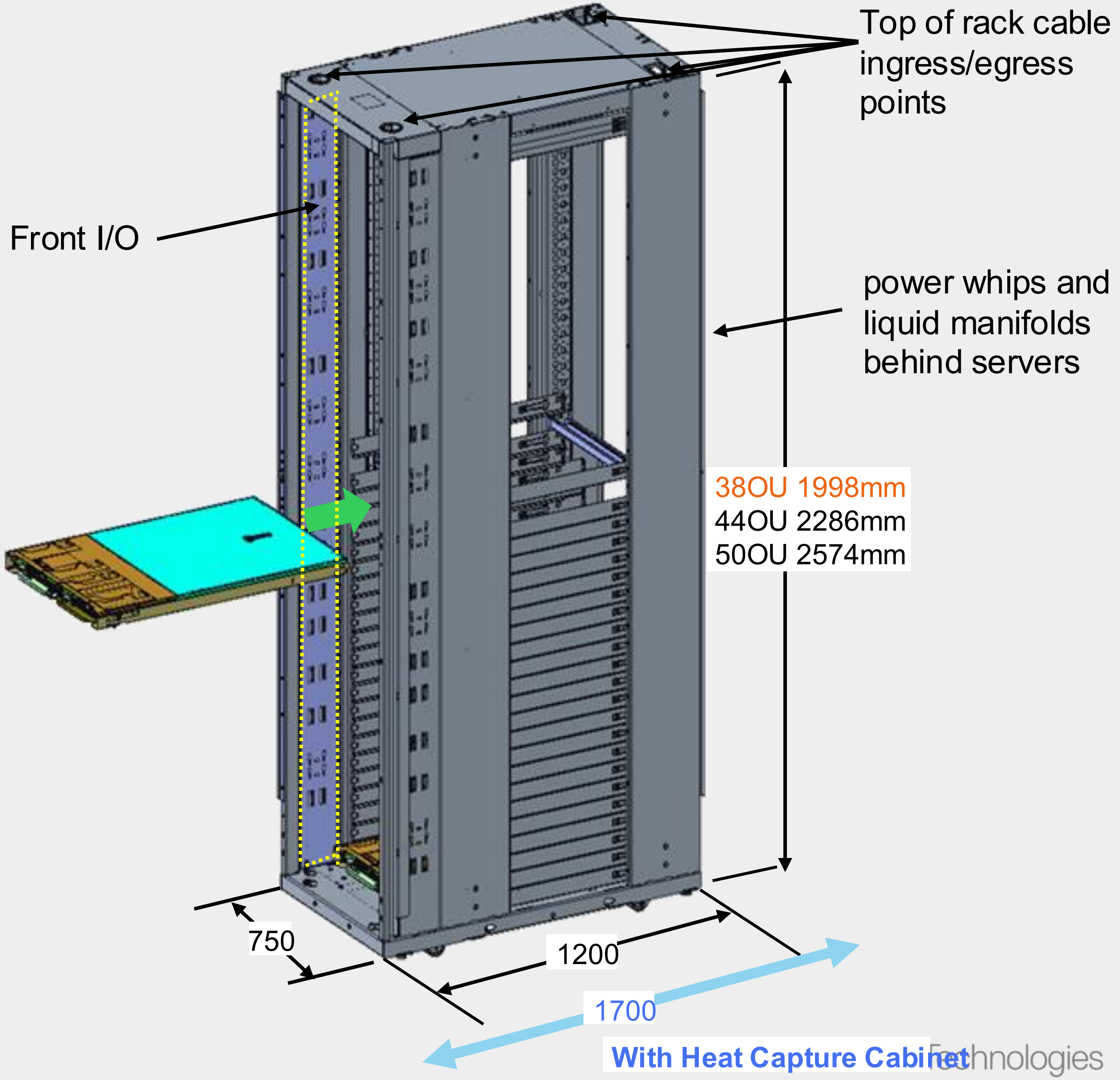




Dell Integrated Rack Family: IR7000

Our ORv3 Interpretation!

- 750x1200 mm footprint: same of standard “wide+deep” racks.
- 100% compliant with ORv3 specs
- Designed to scale up to 480kW+ total power
- Rackmount CDUs compatible
- Three configs available at today:
  - Up to 72x2S nodes per 50OU rack (27k+ AMD cores, around 1.8PF FP64!) - M7725
  - GB200 NVL72 (entire rack) - XE9712
  - GB200 NVL4 (node) - XE8712



Technologies

DataCenter Design Recommendations for ORv3

- Plan for 750x1500mm racks (1700mm depth if considering HCC)
- Plan for minimum 1800kg/sqm floor load
- Plan for multiple 3 ph Rack Power Lines
  - At least (5+5) x 64A (105kW+105kW) per rack
  - Better if (10+10) x 64A (210kW+210kW) per rack
- Plan for water to the rack
  - Power and Network from the ceiling, water from the floor recommended (but we can support any direction)!
  - One cold -20°C- or two water circuits (tempered for DCLC and HCC – max 32°C -, cold – 20°C- for RDHx)
- Plan for good primary water quality
  - A Must for DLC to work properly
- Plan for additional space between racks for CDUs
  - In-Rack CDUs may have limited heat removal capabilities, ORv3 footprint-compatible In-Row CDUs available

DELL Technologies

Thank You

Dell Technologies