4U 20" Deep GPU Server Proposal

Hardware Spec Summary

- 20" short-depth rugged design
- Dual AMD Genoa MB with 24x DDR5 DIMM slots + 160x G5 PCIe lanes
- UP to 8x double-wide 10.5" size GPU slots
- 2 or 3x 2200W CRPS AC/DC or DC/DC power modules for up to 4400W
- Two 8bay front-accessible Gen5 NVMe canisters with one boot-disk included with HW encryption key
- Optional liquid-cooler for 400W TDP CPUs
- Optional 2x G5 X16 slots for two 200Gb NICs in the front
- Symmetrical architecture with 8x SSDs + 4x GPUs per one CPU (+ one optional 200Gb NIC)
- Support PCIe P2P data transfer if SSD with CMB is used
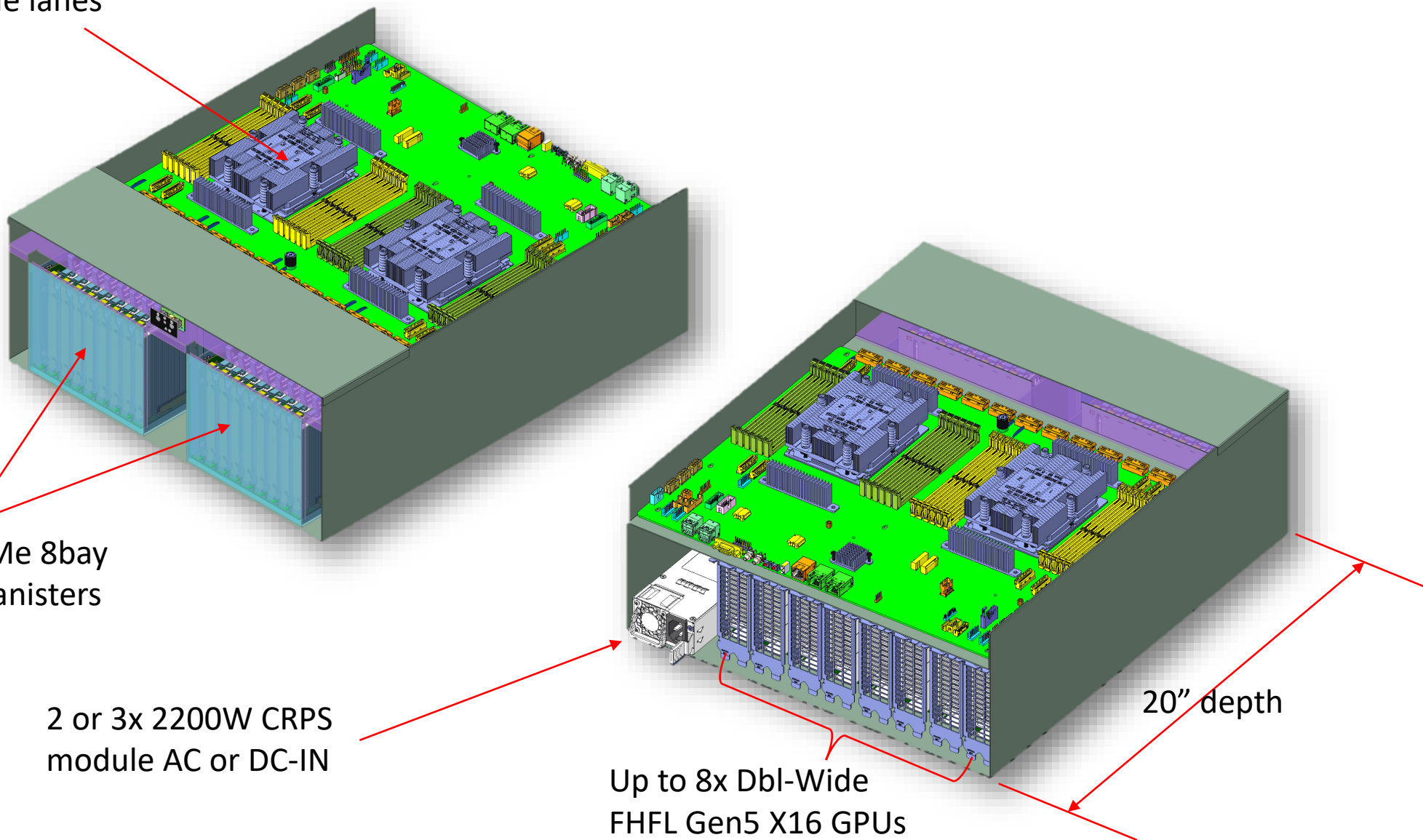- UP to 8x Nvidia NVLink H100 GPUs (4x pairs) can be supported

Dual AMD Genoa MB with
160x Gen5 PCIe lanes

2x Gen5 NVMe 8bay
removable canisters

2 or 3x 2200W CRPS
module AC or DC-IN

Up to 8x Dbl-Wide
FHFL Gen5 X16 GPUs

20" depth

Optional Liquid-Cooler for 400W TDP CPUs

Optional 2x Gen5 X16 front-facing NIC slots

Gen5 X16 NIC slot would be off a Gen5 X64 switch shared with the 8x SSD bays designed for a PCIe P2P architecture (requires SSD with Controller Memory Buffer – CMB)

CMB Application and Use Case

- CMB enables high performance RDMA transfers by reducing Host overhead

- CMB:
  - Data buffer in SSD for storing Read/Write data
  - Implemented in CM6 (Target : 256MB)

- Major use case: NVMe-oF
  - Reduction of CPU overhead → Enables implementation of NVMe-oF storage with less expensive CPUs
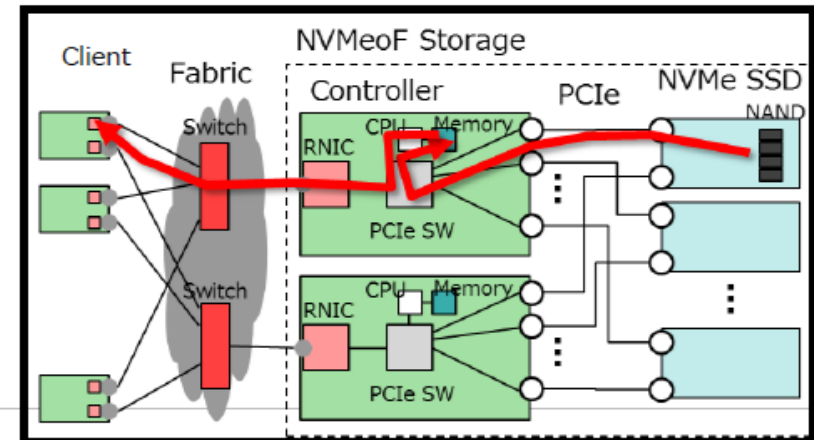
Example: Read data flow

With CMB
Client Read/Write operations directly access CMB
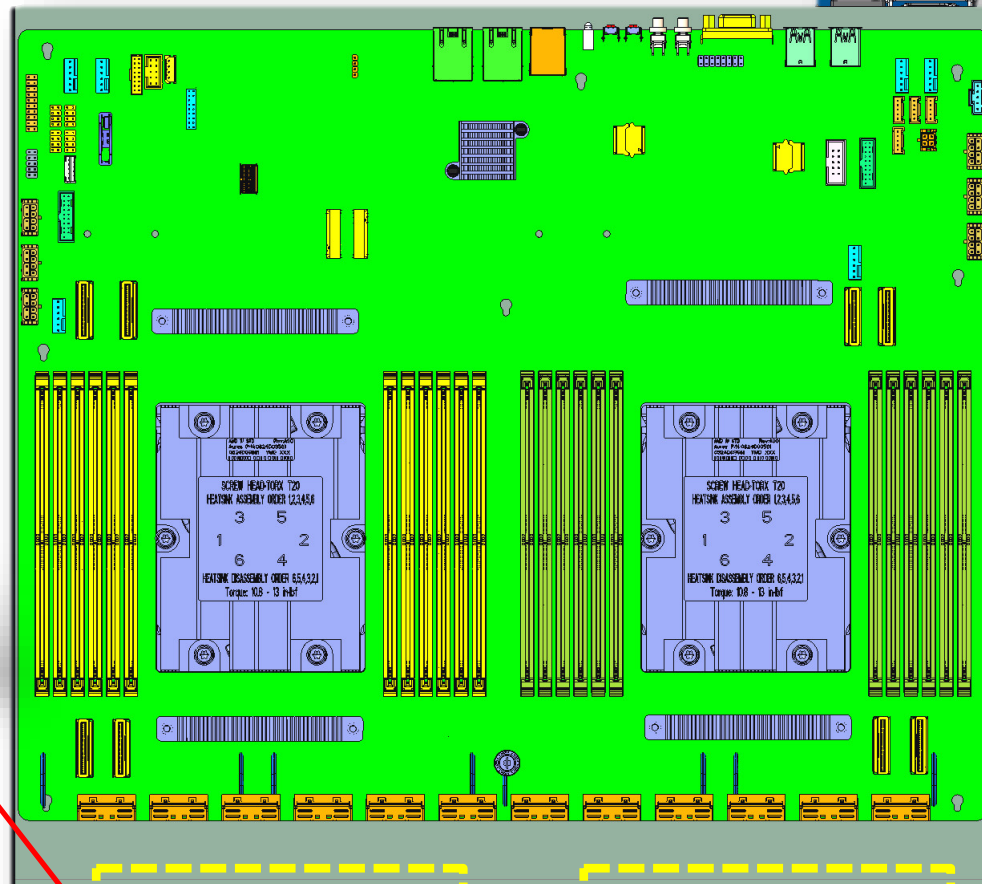without Controller CPU/Memory resource.
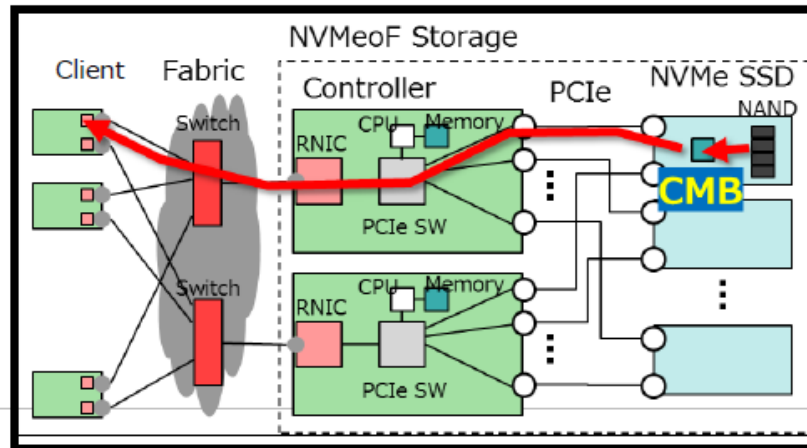
Without CMB
Read data stored in Controller Memory, then sent to Client.

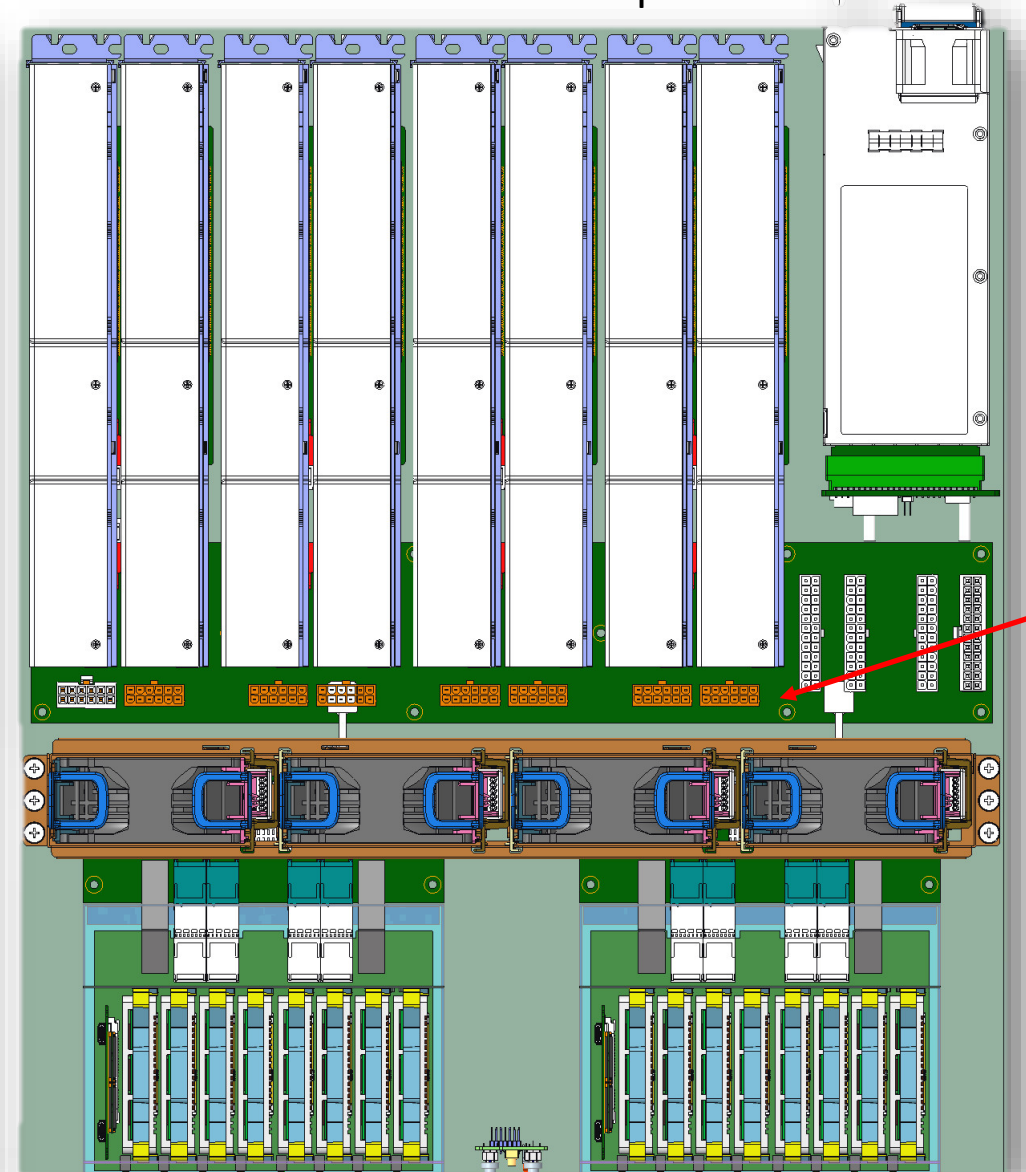MB mounted in upper desk 1U space

8x GPUs in the lower 3U space

PDB

Optional Liquid-Cooler for 400W TDP CPU

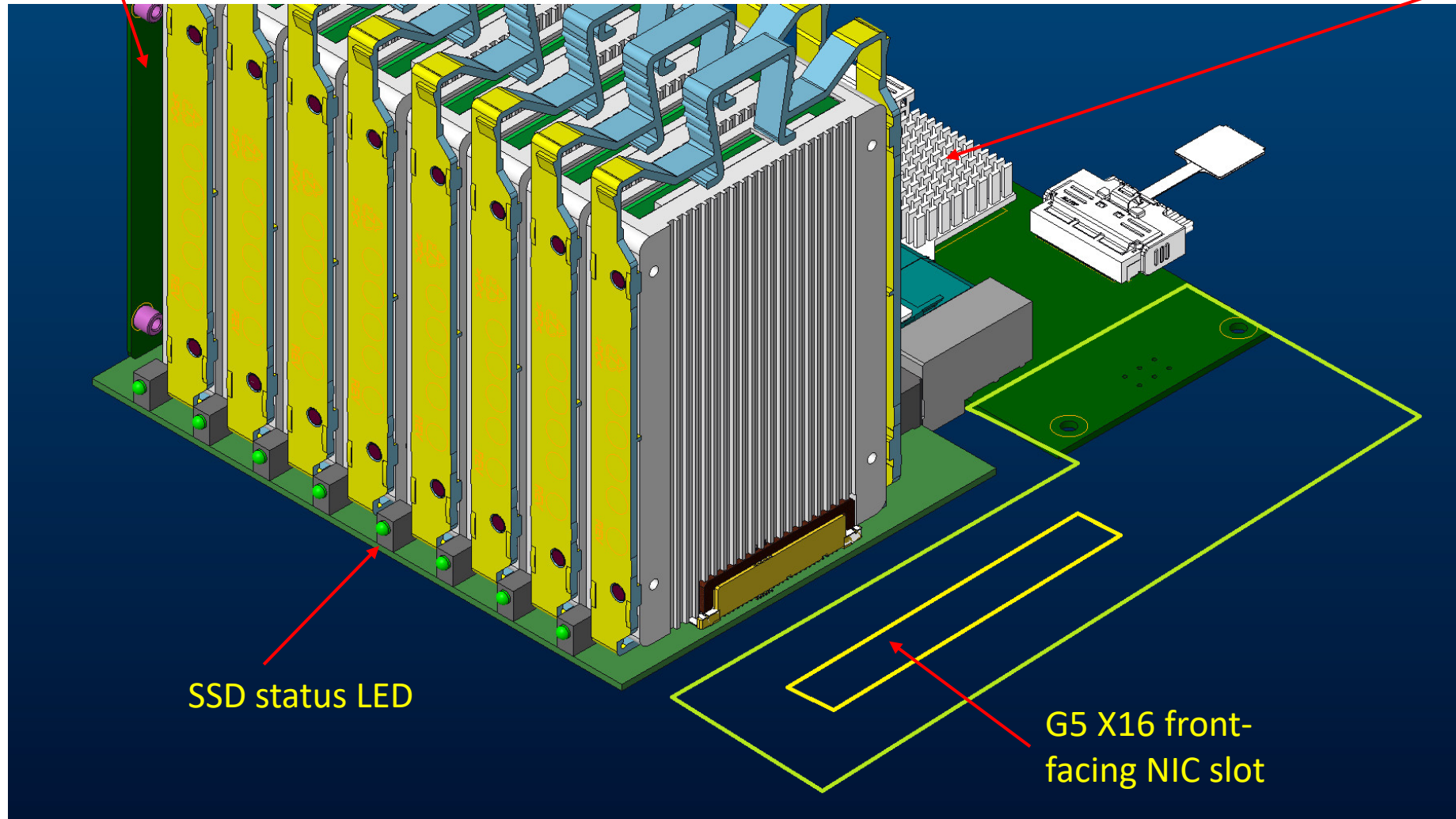Optional Liquid-Cooler for 400W TDP CPU

Canisters removable from the front

2.5" SSDs are top-loaded into canister with retainer bar to secure them

Boot drive with HW encryption key

Tool-less plastic SSD holder

PCIe G5 X64 switch
- G5 X16 uplink to CPU
- 8* G5 X4 to SSDs
- G5 X16 slot for NIC

Note: if front-facing NIC is not desired, NIC card can be installed in the GPU slot in the rear

SSD status LED

G5 X16 front-facing NIC slot