

Network Reference Design

K&K Data GPU Hall — NVIDIA GB300 NVL72 network deployment

Edition:	Public Edition P2
Date:	2026-10-05
Publisher:	K&K Data Service Inc.
Canonical location:	https://www.kkdatasvc.com/lab/downloads/
Enquiries:	inquiry@kkdatasvc.com

Contents

- Part I — Network design brief (A4)
- Part II — Network drawing book, 30 A3 sheets N-000-EN to N-951-EN
- Part III — Implementation supplement (A4)

Public reference design for study and evaluation. Not a construction issue, procurement release or equipment selection; no engineer-of-record seal. Engineering inputs pending confirmation and field-verification items: §11.

Foreword

This edition has completed design-document review.

Engineering inputs pending confirmation and field-verification items are listed in §11.

This document is a public reference design for study and evaluation. It is not a construction issue, a procurement release or an equipment selection, and it carries no engineer-of-record seal. Quantities, parameters and conditions are reference design values. Channel lengths are calculated routed lengths, not installed measurements, and the document reports no executed test results.

Edition information

Item	Value
Title	Network Reference Design — K&K Data GPU Hall
Edition	Public Edition P2 (PE-P2)
Date	2026-10-05
Previous edition	Public Edition P1 (PE-P1), 2026-09-28
Publisher	K&K Data Service Inc.
Enquiries	inquiry@kkdatasvc.com
Canonical location	https://www.kkdatasvc.com/lab/downloads/

Engineering pending conditions

Engineering inputs pending confirmation (U-01 to U-15) and field-verification items (FV-01 to FV-12) are consolidated in the register of §11. Where a value depends on one of these items it is marked HOLD or CONFIRM at the point of use. For such values the document states the formula, threshold or condition and makes no compliance judgement. The reliability figure of 99.999 % monthly availability is a design target, not a measured achievement.

Document structure

Part	Content
Part I	Network design brief (A4), Sections 1 to 11, including the register of pending engineering inputs and field-verification items (§11)
Part II	Network drawing book: 30 A3 sheets, N-000-EN to N-951-EN
Part III	Implementation supplement (A4), Sections IS1 to IS6: storage access and capacity model, optical channel components and verification, operating configuration and network acceptance, clarification register

The status marks FACT, DESIGN, CONFIRM and HOLD are defined in Section 1 of the brief. The supplement adds STUDY for example and conditional values (Section IS1.2).

Part I — Network Design Brief

K&K Data GPU Hall — Network Reference Design Brief

Public Edition P2 · 2026-10-05

Prepared by K&K Data Service Inc. · inquiry@kkdatasvc.com · <https://www.kkdatasvc.com/lab/downloads/>

1 Overview and design basis

This brief presents the reference network deployment for the K&K Data GPU Hall: the network architecture, its physical placement in the row, maintenance access, cabling and optics, and the requirements that the network equipment places on the existing facility. The deploying organisation defines the final network solution. Public Edition P2 is a reference design for study and evaluation; it is not a construction, procurement or equipment-selection release and carries no engineer-of-record seal. It applies the clarifications CR-01 to CR-10 registered in the implementation supplement (Part III, Section IS5); the reference topology, quantities, routes, optical module counts and U positions are unchanged from Public Edition P1. The configuration items in Section 9 remain open. Engineering inputs pending confirmation and field-verification items are consolidated in Section 11.

The reference design serves 90 GPU cabinets (NVIDIA GB300 NVL72, OEM rack envelope 600 × 1200 mm) with 40 network cabinets: 30 unit cabinets placed in the GPU row and 10 central cabinets. Existing power supply, room diffuse cooling, building structure, CDUs and fan walls are inputs. Where they appear in the drawings they are spatial context and are labelled as such; this design does not verify or redesign those systems.

Values carry one of four status marks, used consistently in the brief and drawings:

Status	Meaning
FACT	Sourced fact: vendor document or fixed project input
DESIGN	Reference design value or derivation from the reference layout and port table
CONFIRM	Condition to be confirmed by the manufacturer, the engineer of record or the deploying organisation
HOLD	Input not yet available or item not selected; no compliance judgement is made

This document binds the brief (Part I), the drawing book (Part II: 30 A3 sheets, N-000-EN to N-951-EN) and the implementation supplement (Part III, Sections IS1 to IS6) in one PDF. The supplement adds the mark STUDY for example and conditional values. The reference port-level link table (39,066 records) is not published in this edition; drawings and supplement use its identifiers unchanged. Sheet numbers carry an "-EN" suffix.

2 Reference network architecture

2.1 Compute fabric

The compute fabric is a dual-plane, rail-optimised two-tier leaf–spine network. Each compute tray has four ConnectX-8 NICs. Each 800G OSFP NIC port is split into 2 × 400G: half-port P1/1 connects to plane A and half-port P1/2 to plane B, each to the rail leaf whose rail number equals the NIC number. Each leaf group serves three GPU cabinets (54 trays) with four rail leaves per plane.

Item (per plane)	Value	Status
Compute trays / CX-8 NICs	1,620 / 6,480 (each NIC 1 × 400G into the plane)	FACT
Rail leaves (SN5610)	120 (30 leaf groups × 4 rails); 54 down / 54 up × 400G; 20 unused (not reserved) ports per leaf	DESIGN
Spines (SN5610)	54 (3 cabinets × 18); 120 ports used, 8 unused (not reserved) per spine	DESIGN
Leaf oversubscription	1.0 : 1 (non-blocking)	DESIGN
Leaf → spine uplink capacity	2,592 Tb/s per plane (120 × 54 × 400G)	DESIGN

Plane A

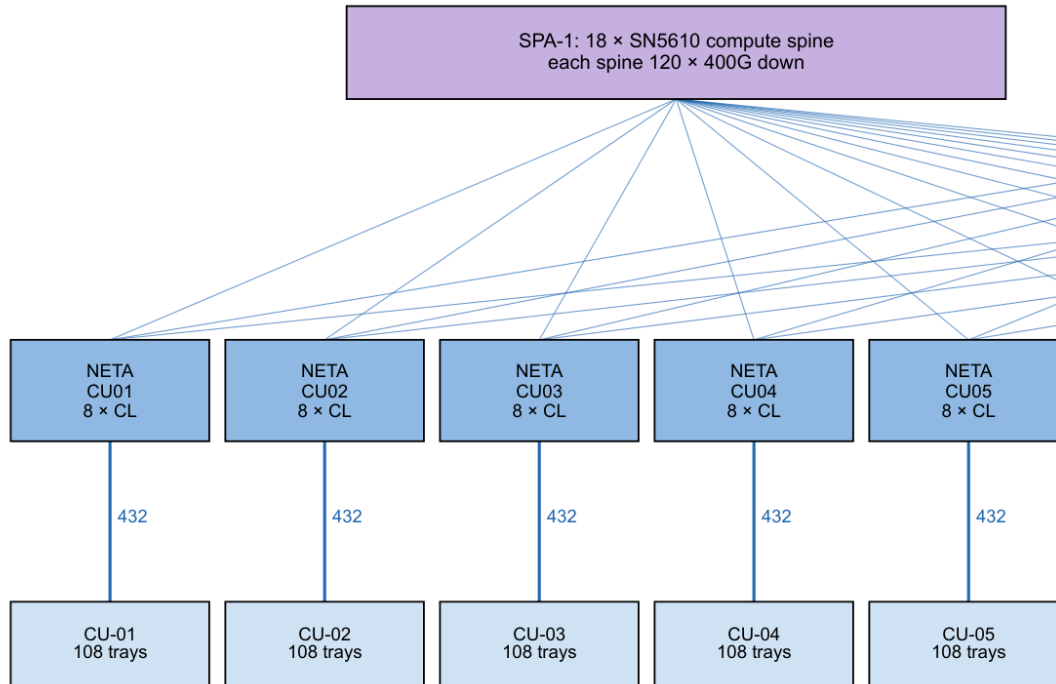


Figure 1. Compute fabric, plane A extract: unit cabinets NETA-CU01 to NETA-CU05 (of 15) and spine cabinet SPA-1 (of 3). Each unit cabinet has $144 \times 400\text{G}$ links to each spine cabinet; the figures on the vertical links are the tray NIC → leaf links per unit cabinet (from N-101-EN; plane B is identical with NETB and SPB)

Every leaf has one 400G link to each of the 54 spines of its plane. Plane-A leaves and spines are in NETA and SPA cabinets; plane-B leaves and spines are in NETB and SPB cabinets. Topology and port mapping: N-101-EN and N-102-EN.

2.2 North–south (DPU) network

Each compute tray has one BlueField-3 (B3240) DPU. Its two 400G ports connect to the A-side and B-side north–south (NS) leaves. Each side has 30 NS leaves (SN5610, U2–U5 of the unit cabinets) with 54 downlinks and 28 uplinks (1.929 : 1), and 7 north–south spines (CVS, SN5610). Each CVS uses 120 ports and has 8 unused (not reserved) 400G ports, 56 per side and 112 in total. These ports reach the external interfaces EXT-IF-NS-A and EXT-IF-NS-B through the 4U border MPO panel in each CVS cabinet. Each NS leaf also has 46 unused ports. Unused ports are not dedicated expansion, storage or border capacity. Border use and the CVS-connected storage options SA-1 and SA-2 would draw on the same 56 unused CVS ports per side; the distributed storage option SA-3 would use unused NS leaf ports (46 per NS leaf) and does not consume the CVS ports. The storage options are not selected, and each assignment is a configuration item of the deploying organisation, checked with the storage and border capacity model of supplement Section IS2. The port positions of the 4U border panel are passive connection points; they do not add switch-port capacity. Upstream routers, carriers, interface rates and optics beyond this boundary are defined by the deploying organisation (N-111-EN).

2.3 Out-of-band management

Inside each GPU cabinet, 18 compute trays × 3, 9 NVLink switch trays × 2 and 8 power shelves × 1 give $80 \times 1\text{G}$ ports on two OEM SN2201 switches. These switches are GPU cabinet equipment, and their power is within the GPU cabinet TDP. Each OEM SN2201 has $2 \times 100\text{G}$ uplinks (MMA1B00-C100D), one to the A-side and one to the B-side OOB aggregation. Each network cabinet has one SN2201 for the management ports of its switches and rack PDUs, with $2 \times 100\text{G}$ uplinks to OA-A-4 and OA-B-4. OOB aggregation uses four SN4600C per side; unused 100G ports reach the site management interface EXT-IF-OOB (N-121-EN). The OOB 100G channels use MPO-12 UPC connectors (Section 5.2).

2.4 Network equipment quantities

Equipment	Quantity	Location	Counted in
SN5610 (rail leaf, spine, NS leaf, CVS)	422	30 unit cabinets, 6 spine cabinets, 2 CVS cabinets	Network load
SN4600C (OOB aggregation)	8	OOB-A, OOB-B	Network load
SN2201 (network cabinet management)	40	One per network cabinet	Network load
SN2201 (GPU cabinet OEM)	180	Two per GPU cabinet	GPU cabinet TDP
MPO patch panels 7U / border 4U	30 / 2	Unit cabinets / CVS cabinets	—
Rack PDUs, 3-phase 4-wire wye 400Y/230 V (400 V line-to-line, 230 V line-to-neutral)	156	40 network cabinets	PDU losses in network load

The reference link table holds 39,066 records: link IDs L000001–L038990 and 76 added PDU management links (L038991–L039066). It is not published in this edition; the link groups on N-901-EN and N-902-EN and the representative port tables are taken from the reference link summary and representative port tables.

3 Physical deployment

GPU and network cabinets stand in a single row between Y2150 and Y3350 (cabinet depth 1,200 mm, fronts facing +Y). Each unit CU-01 to CU-15 contains six GPU cabinets with a plane-A cabinet NETA-CUxx and a plane-B cabinet NETB-CUxx. The central zone between X45,495 and X55,145 holds OOB-A, CVS-A, SPA-1..3, SPB-1..3, CVS-B and OOB-B (N-002-EN, N-201-EN to N-203-EN).

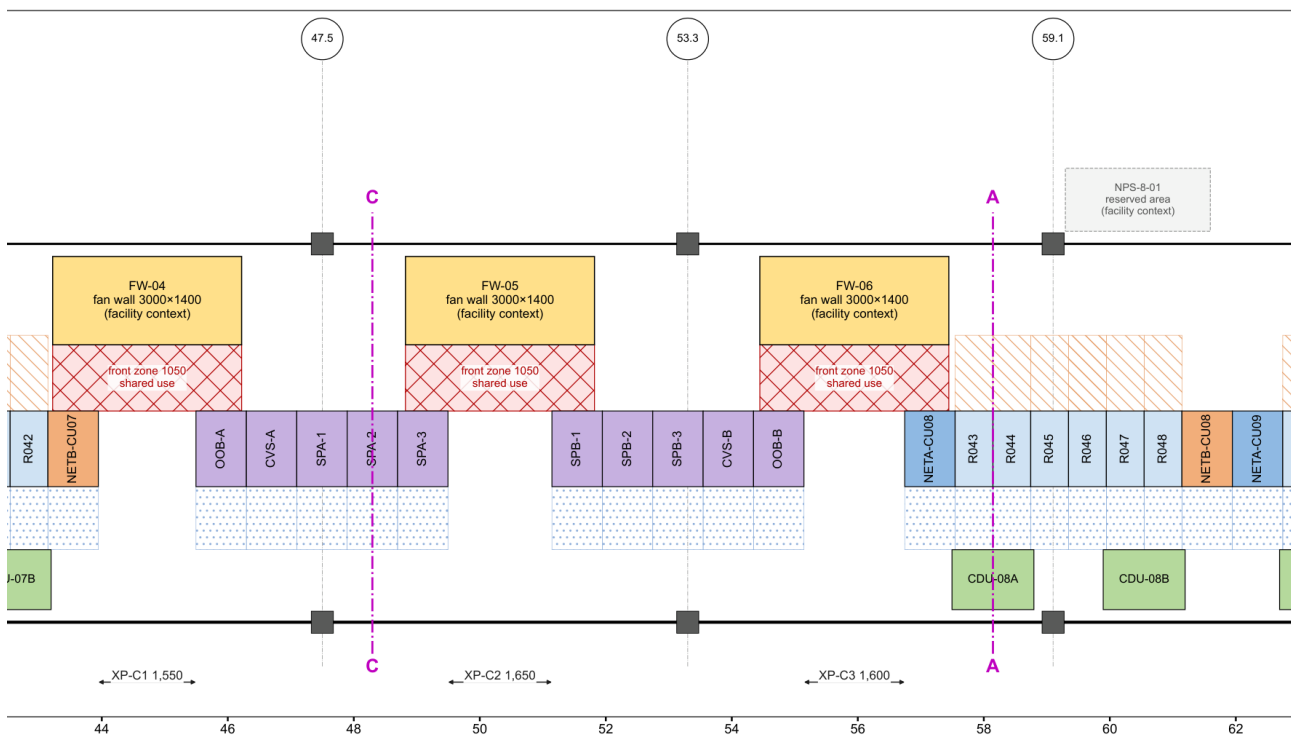


Figure 2. Deployment plan extract, central zone and unit CU-08 (from N-202-EN)

The reference row layout differs from the previous layout: cross aisles are widened by 2,240 mm in total and the row starts at X1155, so that every IT support rail clears the main-beam zone (axis ± 160 mm). Cabinet widths, count, order and Y positions are unchanged. This is a difference between two reference row layouts; the building geometry is unchanged (N-204-EN).

Y dimension	From	To	Clear (mm)	Basis
Cabinet front (cold aisle) to fan-wall rear / column-flange line	Y3350	Y5800	2,450	2,650 to the partition axis Y6000
Network cabinet front to fan-wall face, where a fan wall faces	Y3350	Y4400	1,050	Fan wall Y4400–5800, 1,400 deep
Cabinet rear to CDU front (CDU envelope O4)	Y1150	Y2150	1,000	CDU ≤ 950 deep from Y200

Y dimension	From	To	Clear (mm)	Basis
Network cabinet rear to wall zone, where no CDU is behind	Y200	Y2150	1,950	Y200 wall-zone line

Network cabinets are 800 × 1,200 × 2,294 mm, 48U with 46U usable. The cabinet stands on IT support rails whose top is at Z1128, 98 mm above the floor-board top (Z1030); the cabinet height is 2,294 mm and its top is at absolute elevation Z3422. Switches use the NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A, which fits post spacing of 600–800 mm (FACT). The design uses 750 mm between posts with the front post 100 mm behind the front door. The SN5610 is 796 mm deep including fan handles and extends 46 mm behind the rear post, leaving 304 mm at the rear for power cords and fibre bends. The 1,200 mm cabinet depth is not the rail dimension (N-401-EN, N-402-EN).

4 Maintenance access

The network cabinets that face a fan wall have a 1,050 mm front zone. Switch front withdrawal needs 1,000 mm and fan-wall service needs 850 mm. Both activities use the same zone and are scheduled at different times. This applies to 18 network cabinets in front of fan walls FW-01 to FW-09 (N-312-EN, N-951-EN).

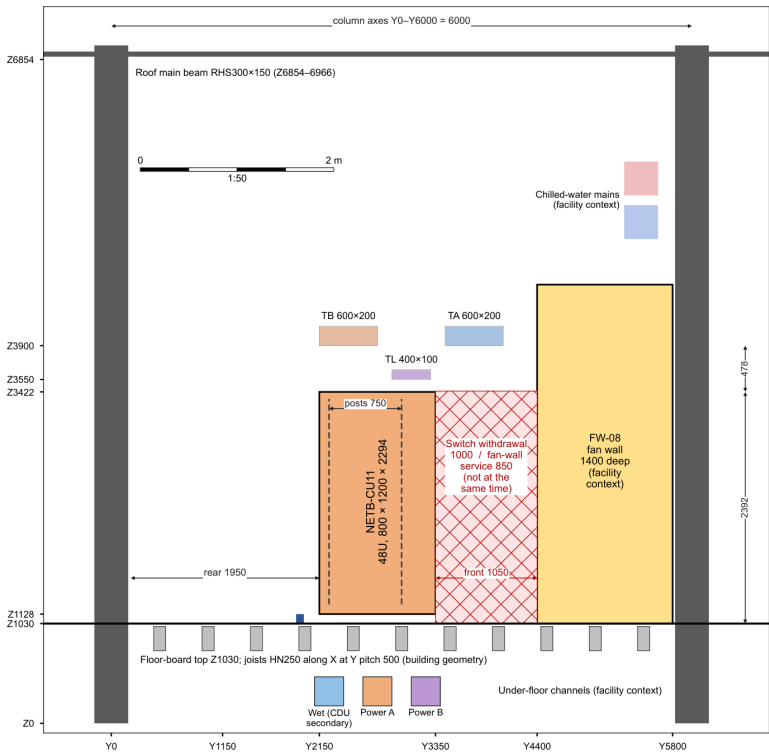


Figure 3. Section B-B through NETB-CU11 and fan wall FW-08 (from N-312-EN)

Access requirements differ by location:

- Network cabinet front: switch withdrawal 1,000 mm. At fan walls, the 1,050 mm zone is shared in time with fan-wall service; elsewhere the cold aisle is 2,450 mm.
- Network cabinet rear: PSUs and fans are replaced from the rear, within the 1,000 mm zone to the CDU front or the 1,950 mm zone to the wall.
- GPU cabinet front: tray withdrawal needs 1,200 mm (Y3350–4550) in the cold aisle. No fan wall faces a GPU cabinet.
- GPU cabinet rear: access within the 1,000 mm zone between cabinet rear (Y2150) and CDU front (Y1150).

Door positions in the external wall and partition are not supplied and are not shown. An earlier door drawing placed doors where they would conflict with four fan-wall or electrical-zone positions; it does not describe the current condition. Current door positions are to be confirmed against the fan-wall and front-zone locations before the layout is confirmed (FV-01).

5 Cabling, trays and optics

5.1 Selected media

The initial project premise was short copper cables from the GPU cabinets to the switches. The reference design uses optical NIC and DPU uplinks. Each ConnectX-8 port is an 800G flat-top OSFP host port that must be split into two 400G links to switch ports in two different cabinets (NETA and NETB). No validated copper part is listed for this host-800G → 2 × switch-400G direction; the listed Ethernet breakout copper parts run in the opposite direction. Direct routes would be 4.27–7.05 m (calculated) and cross up to five GPU cabinets. This means that no adoptable validated part is available; it does not mean that copper is physically impossible.

For the BlueField-3 ports, a listed part exists for the switch-800G → 2 × host-400G direction, but 690 of 1,620 trays would exceed 5 m and 660 routes would cross the GPU tray-withdrawal zone. The DPU uplinks therefore also use fibre, giving one medium, one maintenance method and one tray. Calculated channel lengths are 5.20–10.07 m for NIC → leaf and 6.09–10.34 m for DPU → leaf, through the TL tray. Channel lengths in this brief are calculated routed lengths from the reference layout model (DESIGN), including the design allowances of the reference model; they are not installed measurements. Installed lengths and OLTS losses are recorded per channel at installation and replace the calculated values in the as-built record (FV-11).

Copper remains for cabinet-local 1G management (Cat6A). If the OEM certifies a copper or active-copper part for the host-800G → 2 × 400G direction, units within its certified length and outside the tray-withdrawal zone can be revised together with the module, port and tray-occupancy tables.

5.2 Trays and cable units

Tray	Size (mm)	Y range	Bottom	Carries	Max. occupancy	Cables	Fill	Cable kg/m
TL	400 × 100	Y2900–3300	Z3550	NIC and DPU jumpers, both planes	X2755	540	9.5 %	6.48
TA	600 × 200	Y3450–4050	Z3900	Plane-A trunks and OOB uplinks	X49095	341	25.5 %	39.48
TB	600 × 200	Y2150–2750	Z3900	Plane-B trunks and OOB uplinks	X54745	336	25.4 %	39.42

Plane-A trunks use TA only, plane-B trunks use TB only, and NIC/DPU jumpers use TL. The three trays are separated in Y and do not cross in plan; TA and TB bottoms (Z3900) are 250 mm above the TL top (Z3650). Power channels A (Y2750) and B (Y3250) and the wet channel (Y2250) are below the floor boards (N-501-EN, N-502-EN).

Cable counting units: jumper MPO-12 APC, OD 3.0 mm, for the NIC, DPU, leaf–spine and NS channels (OC-1 to OC-4); trunk 18 × MPO-12 legs / 216 fibres, OD 13.0 mm envelope (product CONFIRM). The OOB 100G channels (OC-5) connect MMA1B00-C100D modules, whose ports are MPO UPC (NVIDIA product page). Every module-facing cord end is an MPO-12 UPC female connector; in the current direct assembly, with no intermediate mated pair, both ends of each cord are therefore UPC. Any intermediate mated pair added later needs complementary pin/hole gender and compatible end-face polish and key on its two faces; UPC throughout is the design choice for these channels. The counting unit for OC-5 is MPO-12 UPC, OM4 (product CONFIRM); the per-strand pin and key map remains HOLD until the cable assemblies are selected. Bend radius 10 × OD static and 20 × OD during installation; fill limit 40 % of the cross-section area (engineer of record / deploying organisation CONFIRM). The design uses 840 trunks and 14,640 MPO-12 legs (N-503-EN).

5.3 Optical modules and reach

Module	Manufacturer statement (FACT)	Design use	Longest calculated channel
MMA4Z00-NS / -FLT	2 × 400G SR4 OSFP, 50 m; reach stated with two patch panels	45 m design changeover length to single-mode cages (design rule)	44.35 m
MMA1Z00-NS400 family (exact OPN and shell variant HOLD)	400G SR4 QSFP112, 50 m (page retrieved for MMA1Z00-NS400-T)	BF3 side; 45 m design changeover length	10.34 m
MMS4X00-NS	2 × 400G DR4 OSFP single-mode, 100 m; two patch panels	Cages whose longer half-port exceeds 45 m	61.40 m (range 43.95–61.40)
MMA1B00-C100D	100G SR4 QSFP28, 100 m on OM4; MPO UPC	OOB 100G uplinks, checked against 100 m	63.95 m

All 31,280 optical links are within the reach stated for the selected module; channels have at most one patch panel. This is a distance check only. Insertion loss is counted by mated pairs: a pass-through adapter panel contributes one mated pair, a cassette with connectors on both faces contributes two. The manufacturer's reach statement with

two patch panels does not define a channel loss allowance. 45 m is the design changeover length between multimode and single-mode cages (manufacturer reach 50 m on OM4, stated with two patch panels); it is a design rule, not a manufacturer limit or a measured value. The longest 400G/800G multimode channel is 44.35 m. The OS2 single-mode module (100 m) serves channels up to 61.40 m. OM4 100 m applies only to the selected 100G SR4 module on the OOB uplinks (longest 63.95 m); it is not a general OM4 reach for other modules. Both half-ports of a cage use the same module type, so a cage is single-mode when its longer half-port exceeds 45 m; 68 single-mode links are themselves shorter than 45 m. The BF3-side module is a 400G SR4 QSFP112 multimode module of the MMA1Z00-NS400 family; the exact OPN and shell variant for the B3240 cage are HOLD pending OEM/NVIDIA confirmation (U-10).

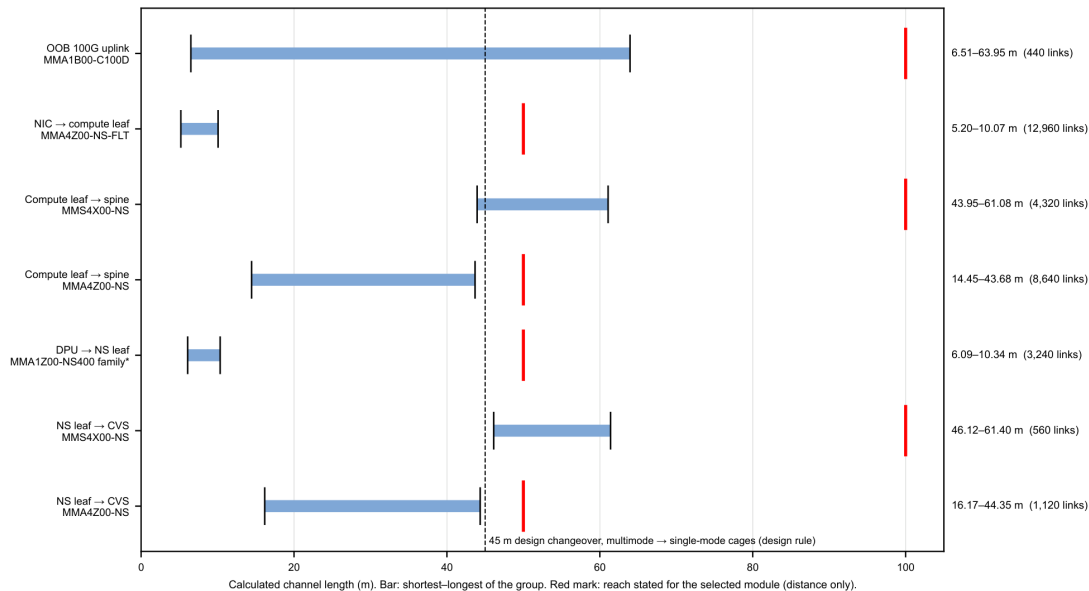


Figure 4. Calculated channel length range of each link group against the reach stated for the selected module; bars show shortest to longest, not a distribution (from N-902-EN)

The manufacturer pages state no channel insertion-loss budget. The insertion-loss estimates in the reference link table (OM4 3.0 dB/km, OS2 0.5 dB/km, MPO 0.35 dB per mated pair) are prior reference estimates only. They are not maximum specifications of selected products and do not define a channel allowance; the approved allowance, component maxima and test limits come from the selected module and cabling vendors (CONFIRM). The channel loss allowance, component maxima and instrument uncertainty are HOLD (U-07, U-08); the margin method and test stages are set out in supplement Sections IS3.3 and IS3.6.

6 Network equipment interface demands

The network equipment places the following demands on the existing facility. Facility capacity, protection, cooling capability and structural capacity are inputs to this design; no spare capacity and no expansion capacity are claimed.

Demand	Value	Condition	Status
Network input power, typical / planning	808.0 / 1,010.0 kW	Planning = typical × 1.25 (no published SN5610 maximum)	DESIGN
Heat to room air (planning, incl. PDU losses)	1,012.4 kW	Front-to-rear airflow; all heat to room air	DESIGN
Airflow at 15 K air temperature rise	208,212 m³/h	Sum of the 40 cabinet values	DESIGN
Switch inlet air	≤35 °C	SN5610 manufacturer inlet limit; front-to-rear airflow	FACT
Rack PDUs	156 (32 A: 132; 63 A: 24)	3-phase 4-wire wye 400Y/230 V (400 V line-to-line, 230 V line-to-neutral), metered, C13/C19 outlets; outlet voltage per the selected PDU product	DESIGN; product CONFIRM
PSU inlets	1,784	C14–C15 cords; 200–240 V device range	DESIGN

Demand	Value	Condition	Status
Unit cabinet NETA/NETB	18.58 / 23.23 kW; 572 kg	11 switches, 4 PDUs, 42 PSU inlets	DESIGN
Spine cabinet SPA/SPB	36.22 / 45.27 kW; 794 kg	19 switches, 4 PDUs, 74 PSU inlets	DESIGN
Heaviest cabinet (SPA-1)	794 kg cabinet mass; ≈ 1.95 kN per support point (static gravity load, equal sharing)	With the weight shared equally by four support points, $794 \text{ kg} \times 9.81 \text{ m/s}^2 \div 4 \approx 1.95 \text{ kN}$. This is not the worst unequal support reaction and not a rail, joist or floor capacity. The 2.92 kN value applies an illustrative factor of 1.5 for unequal sharing; it is not a calculated reaction. Empty cabinet 180 kg	CONFIRM (structure)
Tray loads	Cable up to 39.48 kg/m plus tray 15 kg/m	Hangers from roof purlins, 1.5 m spacing	CONFIRM (structure)

Each SN5610 has four PSUs with at least two operating (FACT): PSU1/PSU2 connect to PDU-A1/A2 and PSU3/PSU4 to PDU-B1/B2. SN4600C and SN2201 connect PSU1 to A1 and PSU2 to B1. The A and B PDUs of each network cabinet are fed from two different existing UPS paths; supply capacity, protection and spare circuits of the existing system are inputs (FV-10). The network load counts the switches in network cabinets, with power per installed optics, and PDU losses only. GPU-side optical modules (MMA4Z00-NS-FLT 6,480 and MMA1Z00-NS400 family 3,240; exact BF3-side OPN and shell variant HOLD) are not counted in network load; whether the GPU cabinet TDP includes them is to be confirmed by the OEM (FV-09). Support reactions and the capacity of rails, joists and floor are confirmed by the structural engineer (FV-02); the per-support-point value above is a static load statement, not an additional structural condition of this design.

7 Consolidated spatial conditions

The conditions below are the spatial and facility coordination reference for this design. Items marked CONFIRM rest with the manufacturer or the engineer of record; they are not network configuration items (Section 9). The CONFIRM items are carried in the field-verification register of Section 11.

Ref.	Condition	Drawing	Status
SC-01	1,050 mm front zone at 18 network cabinets facing fan walls: switch withdrawal (1,000) and fan-wall service (850) at different times	N-312-EN, N-951-EN	DESIGN
SC-02	GPU tray withdrawal 1,200 mm in front of every GPU cabinet; no fan wall faces a GPU cabinet	N-311-EN	DESIGN
SC-03	Rear access 1,000 mm to CDU front, or 1,950 mm to the wall zone	N-201-EN to N-203-EN	DESIGN
SC-04	Row adjustment of 2,240 mm in cross aisles relative to the previous layout; building geometry unchanged	N-204-EN	DESIGN
SC-05	Current door positions not supplied; to be checked against fan walls and front zones	N-001-EN	CONFIRM
SC-06	Cabinet 800 × 1,200 × 2,294 mm on rails at Z1128; cabinet top Z3422; rail posts 750 mm	N-401-EN	DESIGN
SC-07	Tray levels TL Z3550, TA/TB Z3900; clear height above cabinet top to roof beam underside more than 3 m	N-502-EN	DESIGN
SC-08	IT support rails, joists, tray hangers and roof purlins: capacity by the structural engineer; the 30 mm floor board is not an equipment support	N-311-EN	CONFIRM
SC-09	Partition penetrations for power branches into the under-floor A/B channels	N-501-EN	CONFIRM
SC-10	Door perforation and blanking panels for front-to-rear airflow	N-401-EN	CONFIRM

8 Reliability target and common causes

The monthly availability figure of 99.999 % is a design target. It corresponds to 24.192, 25.056, 25.920 and 26.784 s for months of 28, 29, 30 and 31 days. It is not a measured achievement, and dual network planes are not a demonstrated service-level guarantee. Monthly availability is measured per the supplement measurement rule (Section IS4.5): union of overlapping impaired intervals, stated measurement resolution and clock uncertainty, and no exclusion of common-cause events by definition. It is not derived from equipment MTBF or from soak-test duration. Topological reachability is not job continuity; convergence time, host failover time and job restart time are measured at commissioning (IST; FV-12).

Independent elements: plane-A and plane-B leaves, spines, NS leaves, CVS, OOB aggregation, cabinets and the TA/TB trunk trays. Each network cabinet's PDUs A1/A2 and B1/B2 are fed from two UPS paths. In the reference connectivity mapping of network cabinets to UPS paths, the loss of any one of the 13 UPS paths leaves every NIC

with both planes. This is a mapping result, not an operating test or a facility guarantee.

Common causes that remain in the reference design:

- One CX-8 NIC carries both half-ports in one OSFP cage. A NIC or cage failure removes that NIC from both planes (physical endpoint loss).
- The TL jumper tray carries the NIC and DPU jumpers of both planes. Loss of the whole TL section of unit CU-08 leaves 432 NICs unreachable on both planes and 108 DPU trays with no path while the hosts remain powered.
- In 27 of 40 network cabinets both UPS paths belong to the same generator segment. Loss of a segment after battery depletion is a risk case that removes both paths.

Event	NICs: one plane unreachable	NICs: both planes unreachable	BF3 trays: no path
One NIC → leaf jumper broken	1	0	0
One CX-8 NIC failed (both half-ports in one cage)	0	1	0
One compute leaf failed	54	0	0
One plane-A network cabinet lost (NETA-CU08)	432	0	0
Entire plane A lost	6,480	0	0
TA cut at X30000	2,160	0	0
TL section of unit CU-08 lost	0	432	108
Any one of 13 UPS paths lost (reference connectivity mapping)	0	0	0

"One plane unreachable" means that the other plane still carries a path; it does not mean zero interruption. The full event table, including spine, NS leaf and generator-segment cases and the commissioning measurements, is on N-801-EN. The TA cut at X30000 is represented in the supplement by fault case FT-16, a logical isolation of the affected links at the switch side without unplugging cords or touching the tray; its isolation set and acceptance basis are HOLD, and no physical test has been executed.

9 Network configuration items

The configuration items below remain open for the deploying organisation; they are unchanged from Public Edition P1, and where earlier wording differs, the wording of this edition applies. The items are:

- Network architecture: rail-optimised in-row leaf with central spine, dual planes, 1:1 compute leaf oversubscription, NS leaf 1.929:1.
- Switch platform and quantities: SN5610, SN4600C and SN2201 as listed; free U and unused ports are reported, not reserved for expansion, storage or border use.
- External north–south interface: number of unused CVS ports assigned (border use and the CVS-connected storage options SA-1/SA-2 share the 56 unused CVS ports per side; the distributed option SA-3 would use NS leaf ports instead), rate, optics and upstream equipment.
- Site management interface: upstream network, addressing and security zoning.
- Media: optical NIC/DPU uplinks; copper for cabinet-local 1G management; copper alternative only with OEM-certified parts.
- Optics and cabling products: module selection, cable and termination products, loss budget and test acceptance.
- Equipment dimensions and interfaces: switch dimensions and mounting, power inputs (PSU inlets, cords, rack PDU requirement), heat, airflow direction and inlet requirements.
- Service and withdrawal envelope of the selected network equipment (front withdrawal, rear PSU and fan access). The shared 1,050 mm front-zone restriction of Section 4 (SC-01) applies to the network cabinets facing fan walls.
- Network service measurement needs: measurement scope and service accounting for the monthly target, convergence, host failover and job restart.

Facility conditions such as door positions, structure and partition penetrations remain in the spatial and facility coordination reference of Section 7 and the register of Section 11; they are not network configuration items.

10 References and drawing register

10.1 Sources

Reference	Use	Status
Fixed project inputs and scope decisions of 28 September 2026	Scope and fixed quantities	FACT
Existing electrical design (current issue)	Identity of existing UPS paths feeding network cabinets	FACT (mapping)
Building and structure geometry (reference issue)	Axes, columns, joists, IT support rails, roof beams	FACT (geometry)
Port-level connection table and routing (reference design; not published)	Links, device identities, module selection	DESIGN
NVIDIA SN5000 / SN4000 / SN2201 hardware manuals	Ports, PSUs, power, mass, airflow	FACT
NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A (web page, accessed 28 Sep 2026)	Rail for 600–800 mm post spacing	FACT
NVIDIA MMA4Z00-NS and MMS4X00-NS (web pages, accessed 28 Sep 2026)	50 m multimode and 100 m single-mode reach, stated with two patch panels	FACT
NVIDIA MMA1Z00-NS400-T and MMA1B00-C100D product pages (supplement SRC-03, SRC-04)	BF3-side module family; OOB module MPO UPC end-face	FACT (exact BF3-side OPN HOLD)
Implementation supplement (Part III of this document)	Clarification register CR-01 to CR-10; storage capacity model; optical, configuration and acceptance conditions	DESIGN / HOLD
OEM GB300 NVL72 rack documentation	GPU cabinet envelope 600 × 1200 mm, trays, in-rack SN2201	FACT

10.2 Drawing register

Sheet	Title	Scale
N-000-EN	Cover, drawing register and revision	NTS
N-001-EN	Design basis, status key, symbols and abbreviations	NTS
N-002-EN	Hall key plan and unit index	1:200
N-101-EN	Compute fabric: dual-plane rail topology, quantities and oversubscription	NTS
N-102-EN	Compute fabric: rail port mapping, leaf group LG15	NTS
N-103-EN	Compute fabric: independent and shared failure domains	NTS
N-111-EN	North–south (DPU) topology to external interface	NTS
N-121-EN	Out-of-band management hierarchy	NTS
N-201-EN	Deployment plan X 0–36 m	1:100
N-202-EN	Deployment plan X 35–72 m	1:100
N-203-EN	Deployment plan X 71–107 m	1:100
N-204-EN	Reference row adjustment relative to previous layout	1:300
N-301-EN	Typical unit CU-08 detail: plan and X–Z elevation	1:50
N-302-EN	Central spine zone detail: plan and X–Z elevation	1:75
N-303-EN	Far-end unit CU-15 detail: plan and X–Z elevation	1:50
N-311-EN	Section A-A: GPU cabinets and CDU	1:50
N-312-EN	Section B-B: network cabinet and fan wall	1:50
N-313-EN	Section C-C: spine cabinet and trunk entry	1:50
N-401-EN	Network cabinet elevations: unit cabinets NETA/NETB-CU08	1:15
N-402-EN	Network cabinet elevations: SPA-1, CVS-A, OOB-A, SPB-3	1:20
N-403-EN	GPU cabinet network interfaces (OEM reference)	1:15
N-411-EN	Network cabinet schedule (1 of 2)	NTS
N-412-EN	Network cabinet schedule (2 of 2)	NTS
N-501-EN	Pathway plan: fibre trays and under-floor channels	1:150
N-502-EN	Tray sections: maximum and typical occupancy, crossing and drop	1:10 / 1:50
N-503-EN	Cable units, trunk quantities and tray fill	NTS
N-801-EN	Failure domains, service boundaries and commissioning measurements	NTS
N-901-EN	Cable, media and optical module schedule; labelling	NTS

Sheet	Title	Scale
N-902-EN	Channel length against optical module reach	NTS
N-951-EN	Network equipment interface requirements	NTS

10.3 Document contents

Part	Content
Part I	This design brief (A4), Sections 1 to 11
Part II	Drawing book: 30 A3 sheets N-000-EN to N-951-EN
Part III	Implementation supplement (A4, English), Sections IS1 to IS6: storage access and capacity model, optical channel components and verification, operating configuration and network acceptance, clarification register

The reference port-level table and the reference data files are not part of this document.

10.4 Changes in Public Edition P2

Item	Change	Locations
CR-01	Support-point load stated as equal-sharing static load; ×1.5 value identified as illustrative	§6; N-411-EN, N-412-EN, N-951-EN
CR-02	400Y/230 V written out as line-to-line and line-to-neutral voltages	§2.4, §6; N-951-EN
CR-03	Channel lengths identified as calculated routed lengths, not installed measurements	§5.1, §5.3; N-301-EN, N-303-EN, N-901-EN, N-902-EN
CR-04	Spare ports renamed unused (not reserved); 56 per side, 112 in total; CVS panel named the border panel	§2.1, §2.2, §2.3, §9; N-101-EN, N-111-EN, N-121-EN, N-402-EN, N-411-EN, N-412-EN, N-951-EN
CR-05	OOB 100G channels: MPO-12 UPC module-facing ends; OC-1 to OC-4 remain APC	§2.3, §5.2; N-121-EN, N-403-EN, N-503-EN, N-901-EN
CR-06	Distance check separated from loss counting by mated pairs	§5.3; N-902-EN
CR-07	Reference loss values identified as prior estimates, not an allowance	§5.3; N-902-EN
CR-08	45 m identified as a design changeover rule	§5.3; N-902-EN
CR-09	Availability measurement rule referenced	§8; N-000-EN, N-801-EN
CR-10	BF3-side module stated as MMA1Z00-NS400 family; exact OPN and shell variant HOLD	§5.3, §6; N-403-EN, N-901-EN, N-902-EN
Data basis	Link-group ranges and representative ports from the reference link summary and port tables; N-902-EN shows ranges instead of a distribution	N-101-EN, N-102-EN, N-111-EN, N-121-EN, N-301-EN, N-303-EN, N-403-EN, N-901-EN, N-902-EN
Supplement	Implementation supplement added as Part III; the separate technical appendix of Public Edition P1 is not part of this edition	Part III
Register	Pending engineering inputs U-01 to U-15 and field-verification items FV-01 to FV-12 consolidated in one register	§11

11 Register of pending engineering inputs and field-verification items

This section consolidates every engineering input pending confirmation (U-01 to U-15, referenced from supplement Section IS6) and every field-verification item of this edition (FV-01 to FV-12). No item is confirmed or measured in this document. Until an item is closed, the dependent checks state formulas, thresholds and conditions only and make no compliance judgement; dependent values keep their HOLD or CONFIRM mark at the point of use.

11.1 Engineering inputs pending confirmation

ID	Input	Sections	Source of information	Confirmation method
U-01	Dataset size, distribution method (F), local cache capacity and hit fraction; warm-up deadline	IS2.4–IS2.6, IS4.3	Deploying organisation	Workload statement of the deploying organisation; recalculation with the capacity calculator; PT-06 records data volume, completion time and effective throughput and compares them with the warm-up deadline
U-02	Concurrent jobs and tray counts, checkpoint size / interval / deadline, recovery-read method and deadline	IS2.5–IS2.6, IS4.3	Deploying organisation	As U-01; PT-06 and PT-08 record checkpoint and recovery data volume, completion time and effective throughput and compare them with the deadlines

ID	Input	Sections	Source of information	Confirmation method
U-03	Storage system model and read/write throughput (C5); attachment option SA-1, SA-2 or SA-3	IS2.2, IS2.5	Deploying organisation / storage vendor	Vendor performance documents; PT-06 measurement
U-04	Storage and border port assignment k, k_border	IS2.4	Deploying organisation	Port assignment table
U-05	Storage-client multipath and transfer factor ρ	IS2.4, IS4.4	Storage vendor	Client documentation; FT-14
U-06	Payload efficiency η and reserve r	IS2.4	NIC / storage vendors, deploying organisation	Protocol overhead statement; PT-06
U-07	Channel allowance A_ch, assignment of end-interface loss, component maximum losses	IS3.2–IS3.3	Module and cable vendors	Written specifications
U-08	OLTS method and uncertainty U	IS3.3, IS3.6	Test provider	Instrument calibration and method records
U-09	Per-strand pin and key order of cable assemblies; patch-panel type	IS3.2, IS3.4	Cable vendor	Polarity drawings and ordering drawings
U-10	Exact OPN and shell variant of the B3240-side module (MMA1Z00-NS400 family)	IS3.1, IS5	OEM / NVIDIA	Written confirmation (CR-10)
U-11	All CFG items on HOLD (versions, FEC, MTU, routing, QoS, NCCL, host presentation)	IS4.1–IS4.2	Deploying organisation, OEM, vendors	Confirmation-method column of the configuration matrix
U-12	Performance and fault acceptance basis; throughput recovery fraction x (including the PT-06/PT-08 deadline comparison and FT-16)	IS4.3–IS4.4	Deploying organisation	Acceptance rules
U-13	Availability measurement object, treatment of planned maintenance, Δ and ε	IS4.5	Deploying organisation	Measurement rules; measured clock synchronisation
U-14	Vendor thresholds for Q2/Q3 (FEC, DDM)	IS3.6	Module / equipment vendors	Documents for the installed versions
U-15	Tray range and overlap of client groups; SA-3 storage placement and replica distribution; storage port distribution and session redirection when a CVS is lost	IS2.4–IS2.6	Deploying organisation, storage vendor	Job and tray allocation table; storage-client documentation; FT-09 / FT-14

11.2 Field-verification and confirmation items

ID	Item	Reference	Responsible confirmation	Status
FV-01	Door positions in the external wall and partition: not supplied; check against fan walls and the 1,050 mm front zones before the layout is confirmed	§4, SC-05; N-001-EN, N-204-EN	Facility design	CONFIRM
FV-02	Support reactions and capacity of IT support rails, joists, tray hangers, seismic bracing and roof purlins; the 30 mm floor board is not an equipment support; heaviest cabinet 794 kg (equal-sharing static load ≈ 1.95 kN per support point)	§6, SC-08; N-311-EN, N-411-EN, N-412-EN, N-502-EN, N-951-EN	Structural engineer	CONFIRM
FV-03	Tray loads: cable up to 39.48 kg/m plus tray self-weight 15 kg/m; hangers from roof purlins at 1.5 m spacing	§6; N-502-EN	Structural engineer	CONFIRM
FV-04	Partition penetrations for power branches into the under-floor A/B channels	SC-09; N-501-EN	Facility design	CONFIRM
FV-05	Door perforation and blanking panels for front-to-rear airflow	SC-10; N-401-EN	Product	CONFIRM
FV-06	Tray fill limit 40 % of the cross-section area	§5.2; N-502-EN, N-503-EN	Engineer of record / deploying organisation	CONFIRM
FV-07	Cable products: trunk 18 × MPO-12 legs / 216 fibres, OD 13.0 mm envelope; OC-5 counting unit MPO-12 UPC, OM4; per-strand pin and key map HOLD until the cable assemblies are selected (U-09)	§5.2; N-503-EN, N-901-EN	Cable vendor	CONFIRM / HOLD
FV-08	Rack PDU product: 3-phase 4-wire wye 400Y/230 V, metered, C13/C19 outlets; outlet voltage per the selected product	§6; N-951-EN	Product	CONFIRM

ID	Item	Reference	Responsible confirmation	Status
FV-09	Inclusion of GPU-side optical modules (MMA4Z00-NS-FLT 6,480; MMA1Z00-NS400 family 3,240) in the GPU cabinet TDP	§6; N-901-EN	OEM	CONFIRM
FV-10	Existing facility inputs: supply capacity, protection and spare circuits of the existing UPS paths; room diffuse cooling; CDU and fan-wall objects. These are inputs and are not verified by this design	§1, §6; N-951-EN	Facility engineer of record	CONFIRM
FV-11	Installed channel lengths and OLTS losses, recorded per channel at installation; they replace the calculated lengths in the as-built record	§5.1, §5.3; IS3.6; N-301-EN, N-303-EN, N-901-EN, N-902-EN	Installation and test provider	Field measurement
FV-12	Commissioning measurements: route convergence time, host failover time and job restart time (IST); performance tests PT-01 to PT-08 and fault tests FT-01 to FT-16 are not executed in this document	§8; N-801-EN; IS4.3–IS4.4	Commissioning	Field measurement

K&K Data GPU Hall

Network Reference Design — Drawing Book

Edition: Public Edition P2 · 2026-10-05

Prepared by K&K Data Service Inc. · inquiry@kkdatasvc.com · <https://www.kkdatasvc.com/lab/downloads/>

Sheet	Title	Scale	Content change in P2
N-000-EN	Cover, drawing register and revision	NTS	CR-09; register; edition
N-001-EN	Design basis, status key, symbols and abbreviations	NTS	HOLD key; data note
N-002-EN	Hall key plan and unit index	1:200	—
N-101-EN	Compute fabric: dual-plane rail topology, quantities and oversubscription	NTS	CR-04
N-102-EN	Compute fabric: rail port mapping, leaf group LG15	NTS	data note
N-103-EN	Compute fabric: independent and shared failure domains	NTS	—
N-111-EN	North–south (DPU) topology to external interface	NTS	CR-04
N-121-EN	Out-of-band management hierarchy	NTS	CR-04; CR-05
N-201-EN	Deployment plan X 0–36 m	1:100	—
N-202-EN	Deployment plan X 35–72 m	1:100	—
N-203-EN	Deployment plan X 71–107 m	1:100	—
N-204-EN	Reference row adjustment relative to previous layout	1:300	data note
N-301-EN	Typical unit CU-08 detail: plan and X–Z elevation	1:50	CR-03
N-302-EN	Central spine zone detail: plan and X–Z elevation	1:75	—
N-303-EN	Far-end unit CU-15 detail: plan and X–Z elevation	1:50	CR-03
N-311-EN	Section A-A: GPU cabinets and CDU	1:50	—
N-312-EN	Section B-B: network cabinet and fan wall	1:50	—
N-313-EN	Section C-C: spine cabinet and trunk entry	1:50	—
N-401-EN	Network cabinet elevations: unit cabinets NETA/NETB-CU08	1:15	—
N-402-EN	Network cabinet elevations: SPA-1, CVS-A, OOB-A, SPB-3	1:20	CR-04
N-403-EN	GPU cabinet network interfaces (OEM reference)	1:15	CR-05; CR-10
N-411-EN	Network cabinet schedule (1 of 2)	NTS	CR-01; CR-04
N-412-EN	Network cabinet schedule (2 of 2)	NTS	CR-01; CR-04
N-501-EN	Pathway plan: fibre trays and under-floor channels	1:150	—
N-502-EN	Tray sections: maximum and typical occupancy, crossing and drop	1:10 / 1:50	—
N-503-EN	Cable units, trunk quantities and tray fill	NTS	CR-05
N-801-EN	Failure domains, service boundaries and commissioning measurements	NTS	CR-09; FT-16 ref.
N-901-EN	Cable, media and optical module schedule; labelling	NTS	CR-03; CR-05; CR-10
N-902-EN	Channel length against optical module reach	NTS	CR-03/06/07/08/10; range chart
N-951-EN	Network equipment interface requirements	NTS	CR-01; CR-02; CR-04

Revision	Date	Description
PE-P1	2026-09-28	Public Edition P1: reference design brief, drawing book and technical appendix.
PE-P2	2026-10-05	Public Edition P2. Applies clarifications CR-01 to CR-10 of the implementation supplement (support-load basis, voltage notation, calculated lengths, unused ports, OOB 100G UPC end-faces, mated-pair counting, reference loss values, 45 m changeover rule, availability measurement, BF3 module family). Topology, quantities, routes, optics counts and U positions are unchanged.

Notes

- This drawing book presents the reference network deployment, spatial provisions and network equipment interface requirements for study and evaluation as a public reference design. The deploying organisation defines the final network solution; this book is not a construction, procurement or equipment-selection release and carries no engineer-of-record seal.
- Reference design: rail-optimised in-row leaf with central spine, 90 GPU cabinets (GB300 NVL72) and 40 network cabinets (30 in-row unit cabinets and 10 central cabinets).
- Status key used on every sheet: FACT = sourced fact (vendor document, fixed input); DESIGN = reference design value or derivation; CONFIRM = condition to be confirmed by the manufacturer, engineer of record or deploying organisation; HOLD = input not yet available or item not selected, no compliance judgement made.
- Existing power supply, room diffuse cooling, building structure and CDU/fan-wall objects are inputs. Where shown, they are spatial context only and are labelled "facility context".
- The monthly 99.999 % figure is a design target (allowance 24.192 / 25.056 / 25.920 / 26.784 s for 28 / 29 / 30 / 31-day months), not a measured achievement. Dual network planes are not a demonstrated service-level guarantee (N-103-EN, N-801-EN). Monthly availability is measured per the supplement measurement rule (Section IS4.5): union of overlapping impaired intervals, stated measurement resolution and clock uncertainty, and no exclusion of common-cause events by definition. It is not derived from equipment MTBF or from soak-test duration.
- Coordinates: local building grid in mm. X runs along the column axes (0–106,600); Y from the external wall line Y=0 to the partition axis Y=6000; Z from column base (floor-board top Z1030).
- Port-level data: the reference port table holds 39,066 link records; it is not published in this edition. Sheets that show link groups or representative ports use the reference link summary and representative port tables.
- This document binds the design brief (Part I), this drawing book (Part II) and the implementation supplement (Part III) in one PDF.

K&K Data Service Inc. kkdatasvc.com/lab/downloads inquiry@kkdatasvc.com	K&K Data GPU Hall — Network Reference Design PUBLIC EDITION P2 Reference design for study and evaluation — not for construction	Cover, drawing register and revision Scale NTS	Public Edition P2 2026-10-05 Units: mm	N-000-EN Sheet 1 of 30
--	--	--	---	--

Source hierarchy, status key, symbols and abbreviations used in this drawing book

Level	Source	Use in this drawing book	Status	Status	Meaning
1	Fixed project inputs and scope decisions (28 Sep 2026)	Scope; fixed quantities (90 GPU cabinets, CDUs, fan-wall positions)	FACT	FACT	Sourced fact or fixed input
2	Existing electrical design (current issue)	Identity of existing UPS paths feeding network cabinets (reference only)	FACT (mapping)	DESIGN	Reference design value or derivation
3	Building and structure geometry (reference issue)	Axes, columns, joists, IT support rails, roof beams (geometry only)	FACT (geometry)	CONFIRM	To be confirmed by manufacturer, engineer of record or deploying organisation
4	Port-level connection table and routing (not published)	Link groups and representative ports via the reference link summary	DESIGN	HOLD	Input not yet available or item not selected; no compliance judgement
4a	Implementation supplement (Part III of this document)	Corrections CR-01..10, optical, configuration and acceptance conditions	DESIGN / HOLD		
5	NVIDIA SN5000 / SN4000 / SN2201 manuals; optical module pages	Ports, PSU, power, mass, rail kit, module reach	FACT		
6	OEM GB300 NVL72 rack (rack envelope 600 × 1200)	GPU cabinet envelope, trays, in-rack SN2201	FACT		

Symbols

	GPU cabinet GB300 NVL72 (600 × 1200)		Column HW350 on axis		Plane-A fibre / source-A power
	Plane-A network cabinet NETA-CUxx (800 × 1200)		Electrical-zone reserved area (facility context)		Plane-B fibre / source-B power
	Plane-B network cabinet NETB-CUxx (800 × 1200)		Rear maintenance zone (1000)		TL jumper tray (NIC/DPU, both planes)
	Central network cabinet SPA/SPB/CVS/OOB		GPU tray withdrawal zone (1200)		Out-of-band management 1G / 100G
	CDU envelope O4, 1300 × 950 (facility context)		Network front zone 1050: switch withdrawal / fan-wall service		External interface (outside this design)
	Fan wall FW, 3000 × 1400 (facility context)		Section cut line		

Abbrev.	Meaning
CL	Compute rail leaf (SN5610)
SP	Compute spine (SN5610)
NS	North–south leaf for DPU (SN5610)
CS / CVS	North–south spine (SN5610) / its cabinet
OA	OOB aggregation switch (SN4600C)
NOOB	Network-cabinet management switch (SN2201)
OSW	GPU-rack OEM management switch (SN2201)
CU	Unit: 6 GPU cabinets with NETA + NETB

Abbrev.	Meaning
LG	Leaf group: 3 GPU cabinets, 4 rails
CT / NVS / PS	Compute tray / NVLink switch tray / power shelf
TL / TA / TB	Jumper tray / plane-A trunk tray / plane-B trunk tray
C2P	Front-to-rear (connector-to-power) airflow
PP / BP	Unit patch panel / border patch panel
HCM	Horizontal cable manager
EOR	Engineer of record

Notes

- Coordinates: local building grid (mm). X along the column-axis chain 0–106,600 (floor boards 80–106,520); Y from the external wall line Y=0; Z from column base, floor-board top Z1030.
- Link IDs L000001–L039066 and device, port, cage and cabinet IDs are the reference design identifiers and are used unchanged.
- Door positions in the external wall and partition are not supplied and are not shown. They are to be confirmed against fan-wall and front-clearance zones before layout confirmation.
- Sheet numbers carry an "-EN" suffix (English edition).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Design basis, status key, symbols and abbreviations

Scale NTS

Public Edition P2

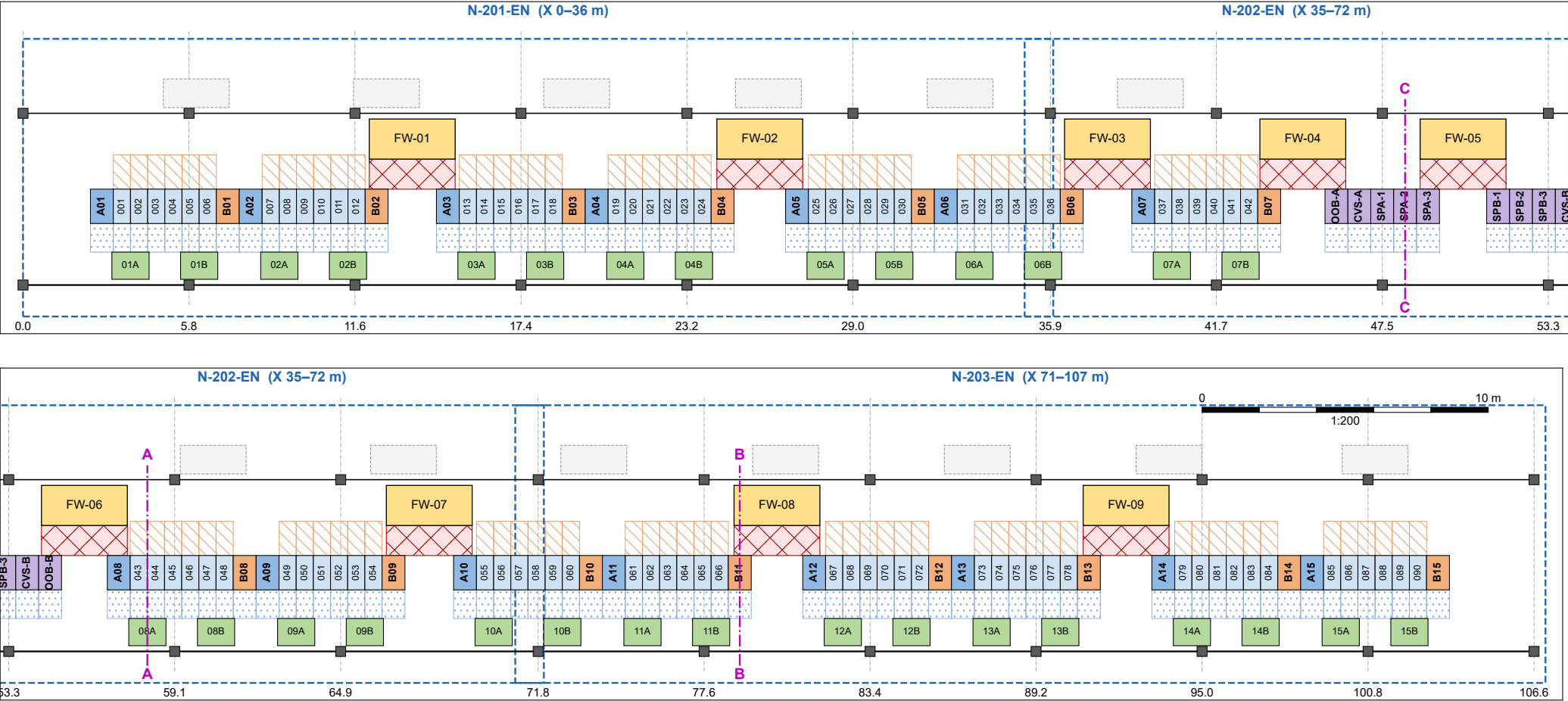
2026-10-05

Units: mm

N-001-EN

Sheet 2 of 30

Whole-hall key plan (1:200, two strips) with the three deployment-plan zones; unit and central-cabinet index. Y=0 external wall line, Y=6000 partition axis; the electrical zone north of Y6000 is facility context.



Axis values in m. Network cabinets: A01 = NETA-CU01, B01 = NETB-CU01; GPU cabinets shown by number (043 = R043); CDUs by unit (08A = CDU-08A).

Unit	GPU cabinets	Leaf groups	X range (mm)	Fan wall at NETA	Fan wall at NETB
CU-01	R001–R006	LG01, LG02	2,355–7,555	—	—
CU-02	R007–R012	LG03, LG04	7,555–12,755	—	FW-01
CU-03	R013–R018	LG05, LG06	14,445–19,645	FW-01	—
CU-04	R019–R024	LG07, LG08	19,645–24,845	—	FW-02
CU-05	R025–R030	LG09, LG10	26,645–31,845	FW-02	—
CU-06	R031–R036	LG11, LG12	31,845–37,045	—	FW-03
CU-07	R037–R042	LG13, LG14	38,745–43,945	FW-03	FW-04
CU-08	R043–R048	LG15, LG16	56,745–61,945	FW-06	—
CU-09	R049–R054	LG17, LG18	61,945–67,145	—	FW-07
CU-10	R055–R060	LG19, LG20	68,845–74,045	FW-07	—
CU-11	R061–R066	LG21, LG22	74,045–79,245	—	FW-08
CU-12	R067–R072	LG23, LG24	81,045–86,245	FW-08	—
CU-13	R073–R078	LG25, LG26	86,245–91,445	—	FW-09
CU-14	R079–R084	LG27, LG28	93,245–98,445	FW-09	—
CU-15	R085–R090	LG29, LG30	98,445–103,645	—	—

Central cabinet	Role	X range (mm)	Fan wall in front
OOB-A	OOB aggregation	45,495–46,295	FW-04
CVS-A	North–south spine	46,295–47,095	—
SPA-1	Compute spine, plane A	47,095–47,895	—
SPA-2	Compute spine, plane A	47,895–48,695	—
SPA-3	Compute spine, plane A	48,695–49,495	FW-05
SPB-1	Compute spine, plane B	51,145–51,945	FW-05
SPB-2	Compute spine, plane B	51,945–52,745	—
SPB-3	Compute spine, plane B	52,745–53,545	—
CVS-B	North–south spine	53,545–54,345	—
OOB-B	OOB aggregation	54,345–55,145	FW-06

Notes

- 1. Row content: 90 GPU cabinets, 40 network cabinets (30 unit + 10 central), 30 CDUs and 9 fan walls (facility context).
- 2. Cross-aisle widening relative to the previous layout totals 2,240 mm (N-204-EN). This is a difference in the reference row layout; the building geometry is unchanged.
- 3. Every fan wall faces a cross aisle and network cabinets only; GPU cabinet fronts keep the 1,200 mm tray-withdrawal zone.
- 4. Door positions are not supplied and are not shown.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Hall key plan and unit index

Scale 1:200

Public Edition P2

2026-10-05

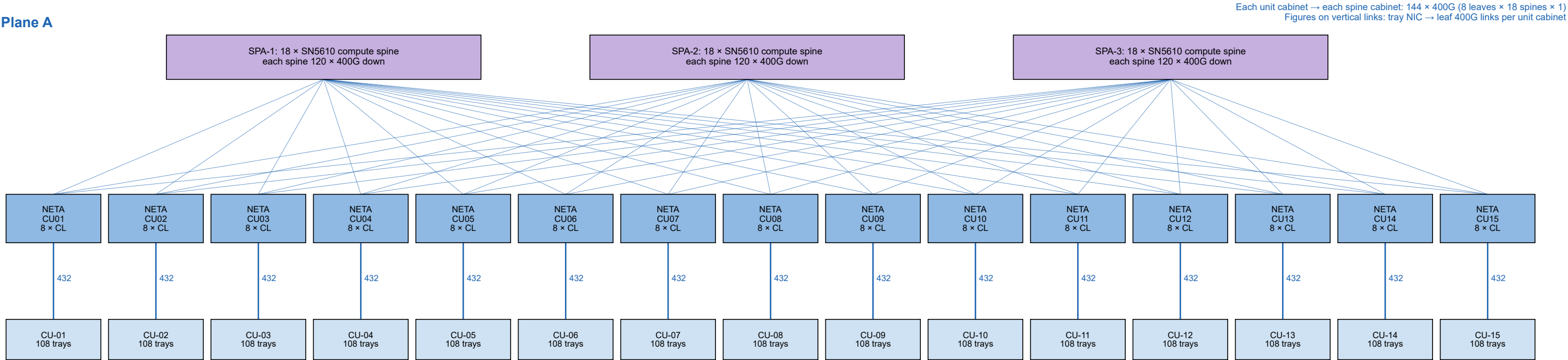
Units: mm

N-002-EN

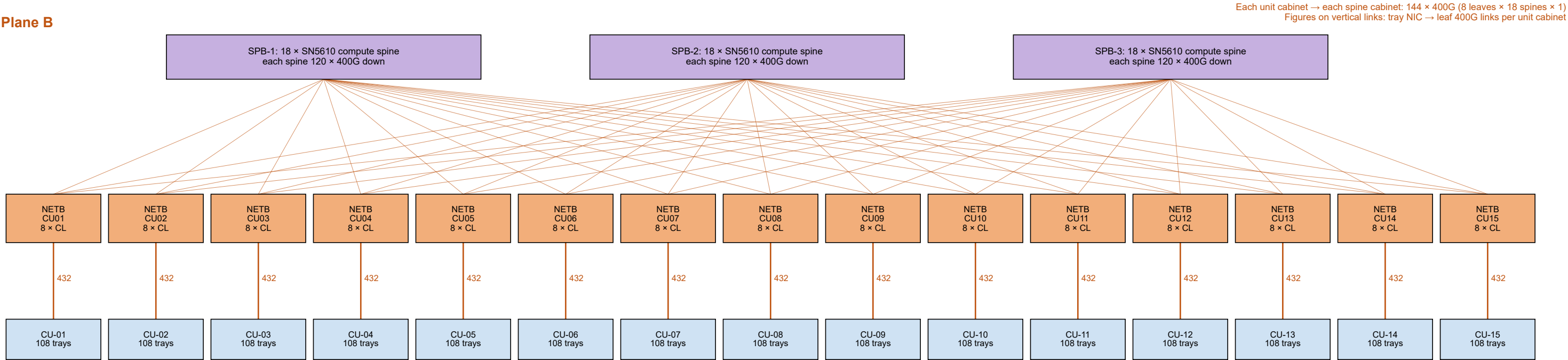
Sheet 3 of 30

Each ConnectX-8 800G port is split into 2 × 400G, one to each plane. Per plane: 120 rail leaves (30 leaf groups × 4 rails) and 54 spines; leaf 54 / 54 down/up (1:1). Values per the reference fabric summary.

Plane A



Plane B



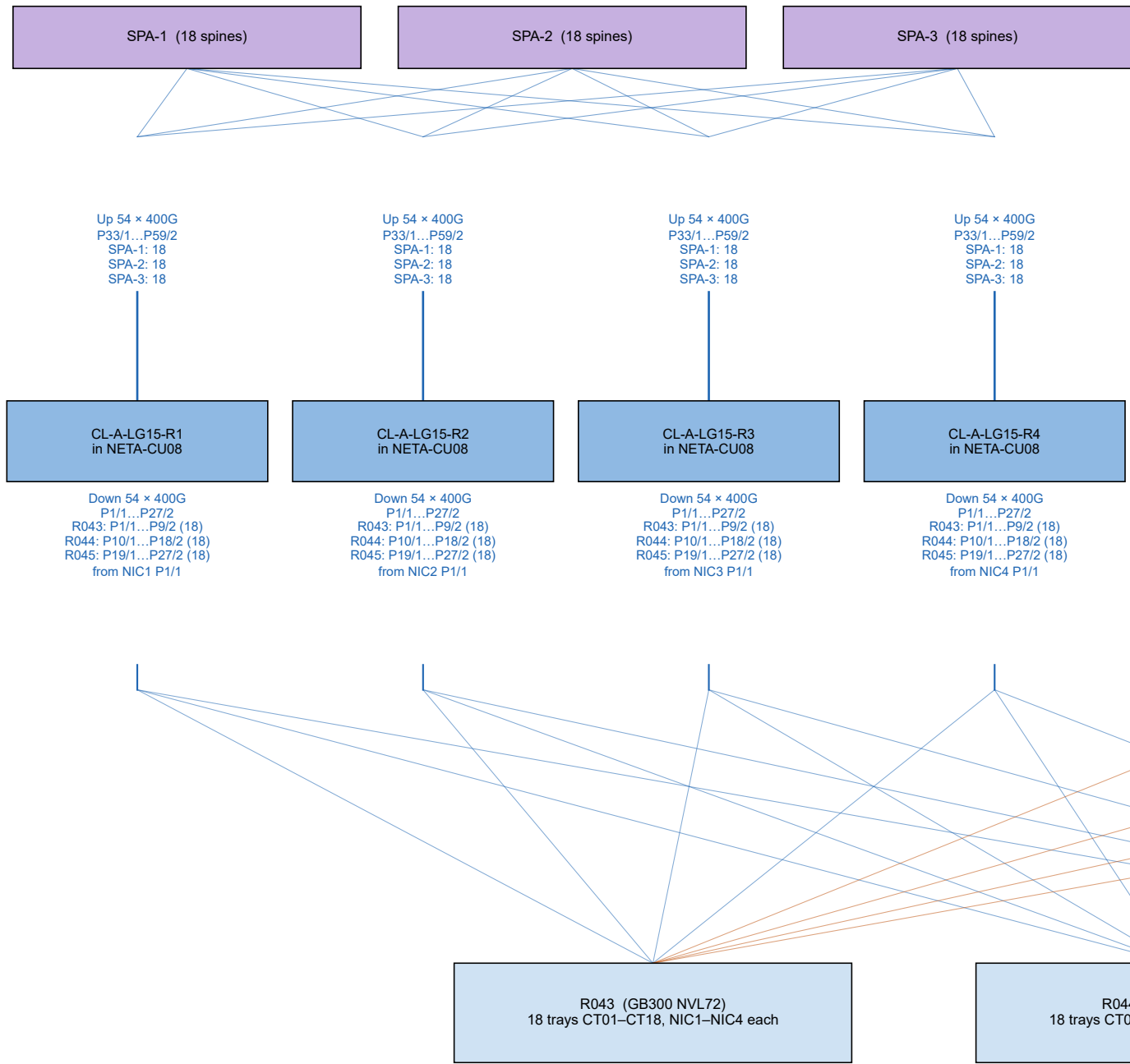
Item	Per plane	Basis	Status
Compute trays / CX-8 NICs	1,620 / 6,480 (each NIC 1 × 400G into this plane)	OEM rack documentation; NVIDIA reference architecture	FACT
Rail leaves (SN5610)	120; down/up 54 / 54; 20 unused (not reserved) ports per leaf	128 × 400G logical ports per switch	DESIGN
Spines (SN5610)	54; 120 ports used, 8 unused (not reserved) per spine	3 cabinets × 18	DESIGN
Leaf oversubscription	1.0:1 (non-blocking)	Downlink / uplink port count	DESIGN
Leaf → spine uplink capacity	2,592 Tb/s per plane	120 × 54 × 400G	DESIGN

- Notes**
- Links: NIC → leaf 12,960 and leaf → spine 12,960 (both planes).
 - Plane A and plane B leaves, spines, cabinets and trunk trays (TA / TB) are separate.
 - The two half-ports of one NIC share one OSFP cage: a NIC or cage failure affects both planes for that NIC (N-103-EN).
 - Unused ports are not dedicated expansion, storage or border capacity; any assignment is a configuration item of the deploying organisation checked with the capacity model of the supplement Section IS2.
 - Link counts per unit and per spine cabinet from the reference port table.

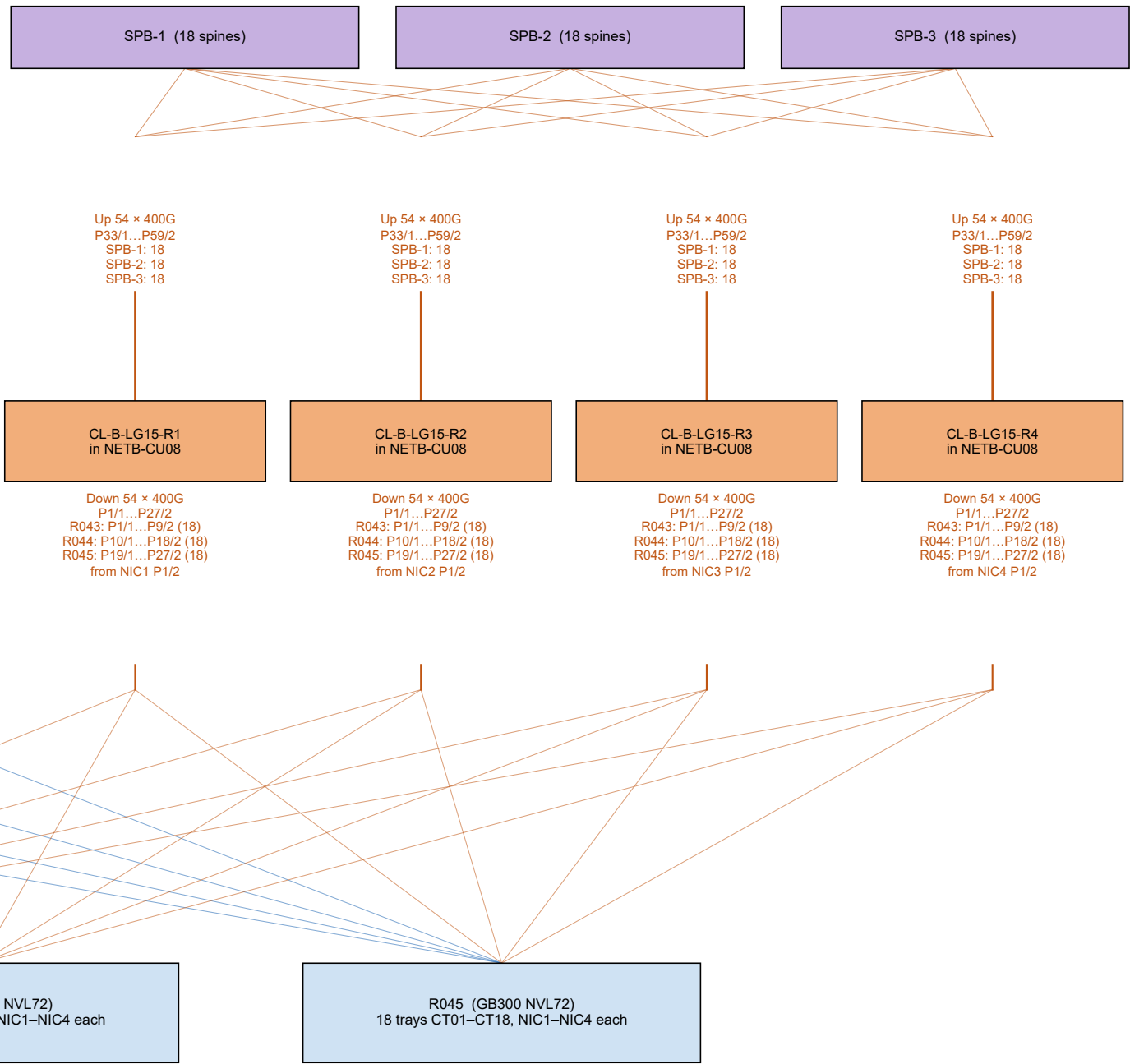
K&K Data Service Inc. kkdatasvc.com/lab/downloads inquiry@kkdatasvc.com	K&K Data GPU Hall — Network Reference Design PUBLIC EDITION P2 Reference design for study and evaluation — not for construction	Compute fabric: dual-plane rail topology, quantities and oversubscription Scale NTS	Public Edition P2 2026-10-05 Units: mm	N-101-EN Sheet 4 of 30
--	---	---	--	--------------------------------------

Leaf group LG15 (unit CU-08, GPU cabinets R043–R045): tray NIC r → leaf CL-x-LG15-Rr; each leaf 54 downlinks and 54 uplinks, one 400G link to each of the 54 spines of its plane. Port ranges from the reference port table.

Plane A



Plane B



Link	End A	End B	Rate	Module A / B	Plane
L007561	R043-CT01:NIC1:P1/1	CL-A-LG15-R1:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	A
L007562	R043-CT01:NIC1:P1/2	CL-B-LG15-R1:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	B
L007563	R043-CT01:NIC2:P1/1	CL-A-LG15-R2:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	A
L007564	R043-CT01:NIC2:P1/2	CL-B-LG15-R2:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	B
L007565	R043-CT01:NIC3:P1/1	CL-A-LG15-R3:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	A
L007566	R043-CT01:NIC3:P1/2	CL-B-LG15-R3:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	B
L007567	R043-CT01:NIC4:P1/1	CL-A-LG15-R4:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	A
L007568	R043-CT01:NIC4:P1/2	CL-B-LG15-R4:P1/1	400G	MMA4Z00-NS-FLT / MMA4Z00-NS	B

Notes

- Each CX-8 NIC OSFP cage (e.g. R043-CT01-NIC1) carries an MMA4Z00-NS-FLT: half-port P1/1 to plane A, P1/2 to plane B.
- On the leaf, the two half-ports of each OSFP cage serve the same plane and rail (same module type per cage, N-901-EN).
- Uplinks: one 400G from each leaf to each spine of its plane; 8 leaves of this group give 432 uplinks.
- Representative ports and links from the reference port table. The complete reference port table (39,066 records) keeps the same identifiers and is not published in this edition.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Compute fabric: rail port mapping, leaf group LG15

Scale NTS

Public Edition P2

2026-10-05

Units: mm

N-102-EN

Sheet 5 of 30

Independent: plane-A and plane-B devices, cabinets and trunk trays. Shared: the two half-ports of one CX-8 NIC (one OSFP cage), the TL jumper tray, and in 27 of 40 network cabinets the generator segment behind both UPS paths.

Level	Independent	Shared (common cause)
NIC / cage	Half-ports P1/1 and P1/2 go to different planes and different leaves	Both half-ports of one CX-8 NIC use one OSFP cage and module: a NIC or cage failure removes the NIC from both planes (physical endpoint loss)
Leaf / spine	Plane A in NETA and SPA cabinets; plane B in NETB and SPB cabinets	One leaf failure: its 54 endpoints lose one plane
Trays	Plane-A trunks only in TA; plane-B trunks only in TB	TL carries NIC and DPU jumpers of both planes. Loss of the TL section of one unit disconnects that unit from both planes (N-801-EN)
Cabinet power	PDU A1/A2 on UPS path A and B1/B2 on UPS path B (existing UPS paths, input)	27 of 40 network cabinets have both UPS paths in the same generator segment: after battery depletion a segment loss removes both
Generator segments	4 segments (existing facility input)	One segment lost (risk case) affects all UPS paths of that segment

Network cabinet source pairing (existing UPS paths and generator segments; input shown for common-cause identification)

Cabinet	UPS A / B	Segment A / B	Same segment
NETA-CU01	UPS7-01 / UPS7-02	SEG-A / SEG-A	yes
NETB-CU01	UPS7-02 / UPS7-03	SEG-A / SEG-A	yes
NETA-CU02	UPS7-01 / UPS7-02	SEG-A / SEG-A	yes
NETB-CU02	UPS7-02 / UPS7-03	SEG-A / SEG-A	yes
NETA-CU03	UPS7-02 / UPS7-03	SEG-A / SEG-A	yes
NETB-CU03	UPS7-03 / UPS7-04	SEG-A / SEG-B	no
NETA-CU04	UPS7-03 / UPS7-04	SEG-A / SEG-B	no
NETB-CU04	UPS7-04 / UPS7-05	SEG-B / SEG-B	yes
NETA-CU05	UPS7-04 / UPS7-05	SEG-B / SEG-B	yes
NETB-CU05	UPS7-05 / UPS7-06	SEG-B / SEG-B	yes
NETA-CU06	UPS7-05 / UPS7-06	SEG-B / SEG-B	yes
NETB-CU06	UPS7-06 / UPS7-01	SEG-B / SEG-A	no
NETA-CU07	UPS7-06 / UPS7-01	SEG-B / SEG-A	no
NETB-CU07	UPS7-05 / UPS7-06	SEG-B / SEG-B	yes

Cabinet	UPS A / B	Segment A / B	Same segment
OOB-A	UPS7-03 / UPS7-04	SEG-A / SEG-B	no
CVS-A	UPS7-01 / UPS7-02	SEG-A / SEG-A	yes
SPA-1	UPS7-01 / UPS7-02	SEG-A / SEG-A	yes
SPA-2	UPS7-03 / UPS7-04	SEG-A / SEG-B	no
SPA-3	UPS7-05 / UPS7-06	SEG-B / SEG-B	yes
SPB-1	UPS8-01 / UPS8-02	SEG-C / SEG-C	yes
SPB-2	UPS8-03 / UPS8-04	SEG-C / SEG-D	no
SPB-3	UPS8-05 / UPS8-06	SEG-D / SEG-D	yes
CVS-B	UPS8-07 / UPS8-01	SEG-D / SEG-C	no
OOB-B	UPS8-03 / UPS8-04	SEG-C / SEG-D	no
NETA-CU08	UPS8-01 / UPS8-02	SEG-C / SEG-C	yes
NETB-CU08	UPS8-02 / UPS8-03	SEG-C / SEG-C	yes
NETA-CU09	UPS8-01 / UPS8-02	SEG-C / SEG-C	yes
NETB-CU09	UPS8-02 / UPS8-03	SEG-C / SEG-C	yes

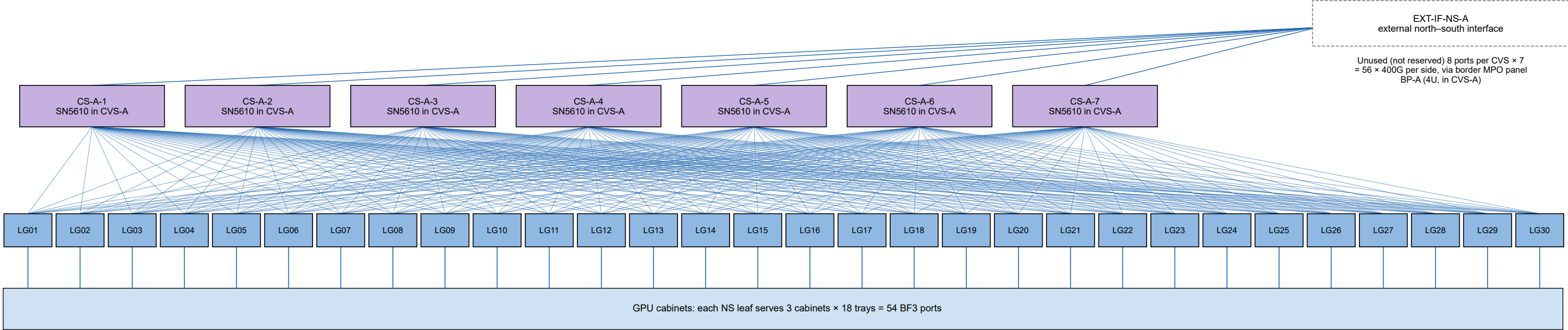
Cabinet	UPS A / B	Segment A / B	Same segment
NETA-CU10	UPS8-02 / UPS8-03	SEG-C / SEG-C	yes
NETB-CU10	UPS8-03 / UPS8-04	SEG-C / SEG-D	no
NETA-CU11	UPS8-03 / UPS8-04	SEG-C / SEG-D	no
NETB-CU11	UPS8-04 / UPS8-05	SEG-D / SEG-D	yes
NETA-CU12	UPS8-04 / UPS8-05	SEG-D / SEG-D	yes
NETB-CU12	UPS8-05 / UPS8-06	SEG-D / SEG-D	yes
NETA-CU13	UPS8-05 / UPS8-06	SEG-D / SEG-D	yes
NETB-CU13	UPS8-06 / UPS8-07	SEG-D / SEG-D	yes
NETA-CU14	UPS8-06 / UPS8-07	SEG-D / SEG-D	yes
NETB-CU14	UPS8-07 / UPS8-01	SEG-D / SEG-C	no
NETA-CU15	UPS8-07 / UPS8-01	SEG-D / SEG-C	no
NETB-CU15	UPS8-01 / UPS8-02	SEG-C / SEG-C	yes

Notes

- 1. Quantified effects of plane, whole-tray and local-cut events are tabulated on N-801-EN.
- 2. Independence of the two PSU inputs is counted at UPS-path level. Two paths in one segment back each other up while utility or generator supply is available; they are not independent at generator-segment level.
- 3. Dual planes provide an alternative network path; they do not by themselves demonstrate a service-level guarantee.

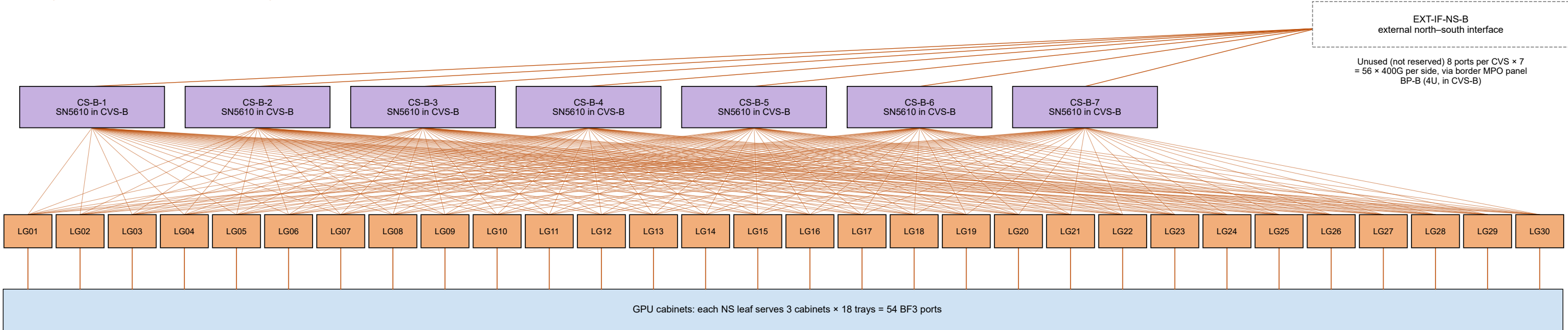
Each tray BlueField-3 (B3240) has two 400G ports, one to the A-side and one to the B-side north–south leaf. NS leaf 54 / 28 down/up (1.929:1) to 7 CVS per side; unused CVS ports reach the external interface via a border patch panel.

A side (NS leaves in NETA-CUxx, U2–U5)



Each NS leaf: 54 BF3 downlinks; uplinks to each of 7 CVS: 4 × 400G, total 28

B side (NS leaves in NETB-CUxx, U2–U5)



Each NS leaf: 54 BF3 downlinks; uplinks to each of 7 CVS: 4 × 400G, total 28

Item	Per side	Note
BF3 ports	1,620 × 400G	One B3240 per tray; two ports to different sides
NS leaves (SN5610)	30; down/up 54 / 28; 1.929:1	46 unused (not reserved) ports per NS leaf
CVS (SN5610)	7; 120 ports used, 8 unused (not reserved)	56 unused per side, 112 in total; border, storage or other use is a configuration item
NS leaf → CVS capacity	336 Tb/s	30 × 28 × 400G

Notes

- Upstream routers, carriers, interface rates and optical modules beyond EXT-IF-NS-A/B are defined by the deploying organisation; no upstream equipment is assumed in this design.
- Configuration item of the deploying organisation: required external bandwidth, number of unused CVS ports to be assigned (border use and the CVS-connected storage options SA-1/SA-2 share the 56 unused CVS ports per side; the distributed option SA-3 would use unused NS leaf ports instead; supplement Section IS2) and module type. Border-panel port positions are passive connection points, not additional switch-port capacity.
- Uplinks per NS leaf to each CVS from the reference port table.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

North–south (DPU) topology to external interface

Scale NTS

Public Edition P2

2026-10-05

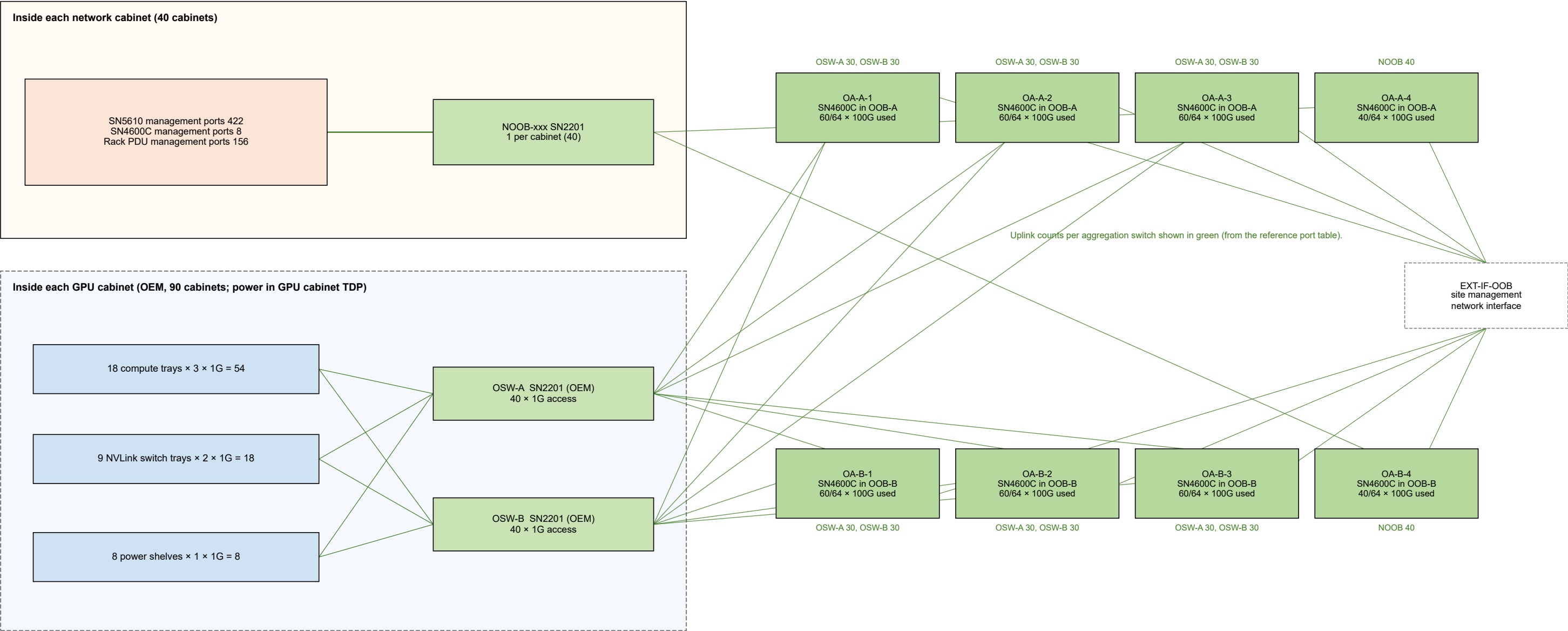
Units: mm

N-111-EN

Sheet 7 of 30

In-rack OEM SN2201 switches in GPU cabinets (part of GPU cabinet TDP) and one SN2201 per network cabinet connect to OOB aggregation SN4600C on sides A and B; aggregation reaches the site management network through EXT-IF-OOB.

Each OSW: 2 × 100G uplinks (MMA1B00-C100D, OM4, MPO-12 UPC), one to side A and one to side B. Each NOOB: 2 × 100G to OA-A-4 and OA-B-4.



Class	From	To	Rate	Links	Scope
MGMT	Compute leaf	Cabinet mgmt switch (NOOB)	1G	240	Network cabinet equipment
MGMT	Compute spine	Cabinet mgmt switch (NOOB)	1G	108	Network cabinet equipment
MGMT	NS leaf	Cabinet mgmt switch (NOOB)	1G	60	Network cabinet equipment
MGMT	CVS spine	Cabinet mgmt switch (NOOB)	1G	14	Network cabinet equipment
MGMT	OOB aggregation	Cabinet mgmt switch (NOOB)	1G	8	Network cabinet equipment
MGMT	Rack PDU	Cabinet mgmt switch (NOOB)	1G	156	Network cabinet equipment
OOB_INRACK	Compute tray	GPU-rack OEM switch (OSW)	1G	4,860	GPU rack OEM equipment (in GPU rack TDP)
OOB_INRACK	NVLink switch tray	GPU-rack OEM switch (OSW)	1G	1,620	GPU rack OEM equipment (in GPU rack TDP)
OOB_INRACK	Power shelf	GPU-rack OEM switch (OSW)	1G	720	GPU rack OEM equipment (in GPU rack TDP)
OOB_UPLINK	GPU-rack OEM switch (OSW)	OOB aggregation	100G	360	GPU rack OEM SN2201 → network OOB aggregation
OOB_UPLINK	Cabinet mgmt switch (NOOB)	OOB aggregation	100G	80	Network cabinet equipment

Notes

- GPU-rack OEM SN2201 switches and their 1G access are OEM equipment counted in GPU cabinet power, not in network load.
- Addressing, upstream management network and security zoning beyond EXT-IF-OOB are defined by the deploying organisation.
- Copper is used for cabinet-local 1G management only (Cat6A in-cabinet patch cords).
- OOB 100G channels (OC-5): MMA1B00-C100D modules have MPO UPC ports. Every module-facing cord end is MPO-12 UPC female; the current direct cords are therefore UPC at both ends. Any intermediate mated pair added later needs complementary pin/hole gender and compatible end-face polish and key; per-strand pin map HOLD (product CONFIRM).
- Unused (not reserved) 100G ports on OA-A/B-1..4 are the EXT-IF-OOB candidates.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Out-of-band management hierarchy

Scale NTS

Public Edition P2

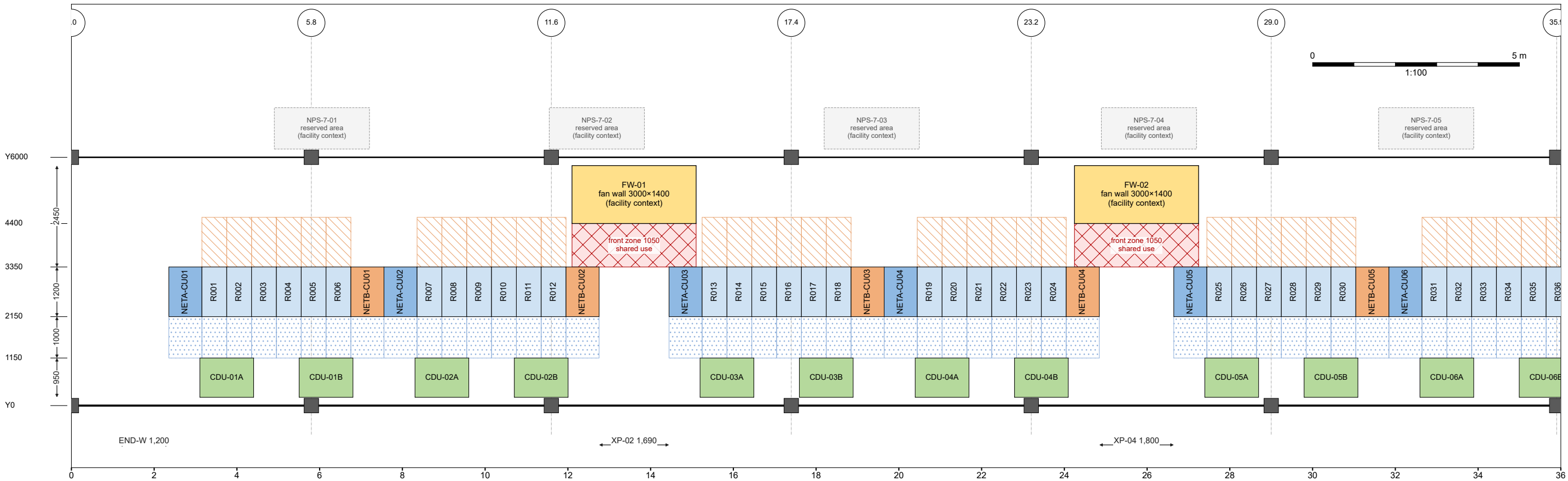
2026-10-05

Units: mm

N-121-EN

Sheet 8 of 30

Deployment plan 1:100, X 0–36 m: all GPU and network cabinet IDs, CDUs and fan walls (facility context), column axes, maintenance zones and section cuts. Door positions are not supplied and are not shown.



Network cabinet	X0–X1	ΔX vs previous	Front clearance
NETA-CU01	2,355–3,155	-3,680	2450 cold aisle
NETB-CU01	6,755–7,555	-3,680	2450 cold aisle
NETA-CU02	7,555–8,355	-3,680	2450 cold aisle
NETB-CU02	11,955–12,755	-3,680	1050 to FW-01
NETA-CU03	14,445–15,245	-3,390	1050 to FW-01
NETB-CU03	18,845–19,645	-3,390	2450 cold aisle

Network cabinet	X0–X1	ΔX vs previous	Front clearance
NETA-CU04	19,645–20,445	-3,390	2450 cold aisle
NETB-CU04	24,045–24,845	-3,390	1050 to FW-02
NETA-CU05	26,645–27,445	-2,990	1050 to FW-02
NETB-CU05	31,045–31,845	-2,990	2450 cold aisle
NETA-CU06	31,845–32,645	-2,990	2450 cold aisle

Network cabinet front zone where a fan wall faces (1,050 mm)

- Switch front withdrawal needs 1,000 mm and fan-wall service needs 850 mm; both use the same 1,050 mm zone and are not performed at the same time (red cross-hatch).
- GPU cabinet fronts keep a 1,200 mm tray-withdrawal zone (orange hatch); fan walls do not face GPU cabinets.

- Notes
- Y chain: wall line Y0 / wall zone Y200 / CDU ≤950 to Y1150 / rear maintenance 1000 to Y2150 / cabinet 1200 to Y3350 / cold aisle 2450 to the fan-wall rear and column-flange line Y5800 (flange Y5825, partition axis Y6000).
 - Rear zone 1000 (blue dots): GPU rear access and network PSU/fan replacement from the rear.
 - Cabinets sit on IT support rails clear of the main-beam zone (axis ±160 mm) at every position.
 - Electrical-zone reserved areas north of Y6000 are facility context; door positions are to be confirmed.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Deployment plan X 0–36 m

Scale 1:100

Public Edition P2

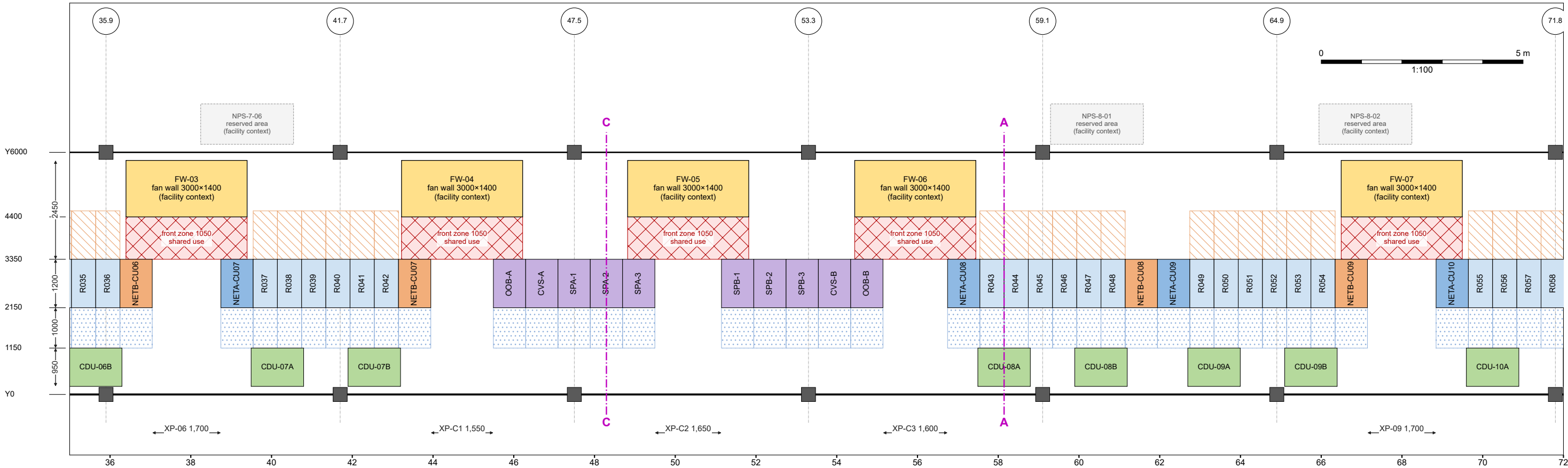
2026-10-05

Units: mm

N-201-EN

Sheet 9 of 30

Deployment plan 1:100, X 35–72 m: all GPU and network cabinet IDs, CDUs and fan walls (facility context), column axes, maintenance zones and section cuts. Door positions are not supplied and are not shown.



Network cabinet	X0–X1	ΔX vs previous	Front clearance
NETB-CU06	36,245–37,045	-2,990	1050 to FW-03
NETA-CU07	38,745–39,545	-2,690	1050 to FW-03
NETB-CU07	43,145–43,945	-2,690	1050 to FW-04
OOB-A	45,495–46,295	-2,690	1050 to FW-04
CVS-A	46,295–47,095	-2,690	2450 cold aisle
SPA-1	47,095–47,895	-2,690	2450 cold aisle
SPA-2	47,895–48,695	-2,690	2450 cold aisle
SPA-3	48,695–49,495	-2,690	1050 to FW-05
SPB-1	51,145–51,945	-2,590	1050 to FW-05

Network cabinet	X0–X1	ΔX vs previous	Front clearance
SPB-2	51,945–52,745	-2,590	2450 cold aisle
SPB-3	52,745–53,545	-2,590	2450 cold aisle
CVS-B	53,545–54,345	-2,590	2450 cold aisle
OOB-B	54,345–55,145	-2,590	1050 to FW-06
NETA-CU08	56,745–57,545	-2,540	1050 to FW-06
NETB-CU08	61,145–61,945	-2,540	2450 cold aisle
NETA-CU09	61,945–62,745	-2,540	2450 cold aisle
NETB-CU09	66,345–67,145	-2,540	1050 to FW-07
NETA-CU10	68,845–69,645	-2,240	1050 to FW-07

Network cabinet front zone where a fan wall faces (1,050 mm)

- Switch front withdrawal needs 1,000 mm and fan-wall service needs 850 mm; both use the same 1,050 mm zone and are not performed at the same time (red cross-hatch).
- GPU cabinet fronts keep a 1,200 mm tray-withdrawal zone (orange hatch); fan walls do not face GPU cabinets.

- Notes
- Y chain: wall line Y0 / wall zone Y200 / CDU ≤950 to Y1150 / rear maintenance 1000 to Y2150 / cabinet 1200 to Y3350 / cold aisle 2450 to the fan-wall rear and column-flange line Y5800 (flange Y5825, partition axis Y6000).
 - Rear zone 1000 (blue dots): GPU rear access and network PSU/fan replacement from the rear.
 - Cabinets sit on IT support rails clear of the main-beam zone (axis ±160 mm) at every position.
 - Electrical-zone reserved areas north of Y6000 are facility context; door positions are to be confirmed.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Deployment plan X 35–72 m

Scale 1:100

Public Edition P2

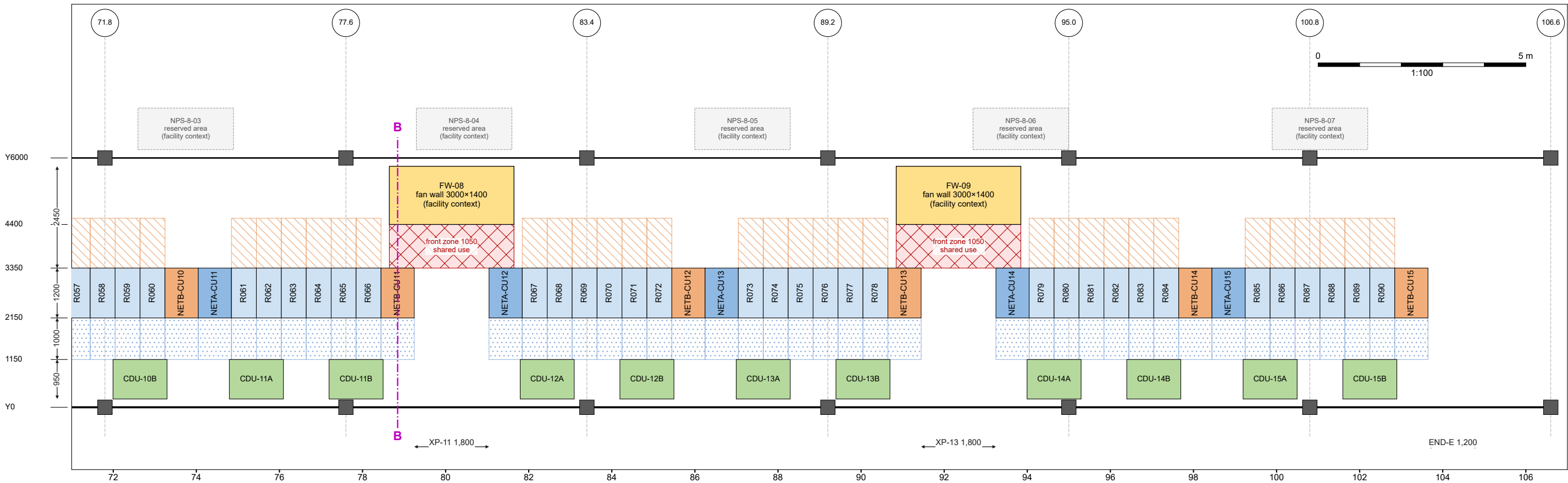
2026-10-05

Units: mm

N-202-EN

Sheet 10 of 30

Deployment plan 1:100, X 71–107 m: all GPU and network cabinet IDs, CDUs and fan walls (facility context), column axes, maintenance zones and section cuts. Door positions are not supplied and are not shown.



Network cabinet	X0–X1	ΔX vs previous	Front clearance
NETB-CU10	73,245–74,045	-2,240	2450 cold aisle
NETA-CU11	74,045–74,845	-2,240	2450 cold aisle
NETB-CU11	78,445–79,245	-2,240	1050 to FW-08
NETA-CU12	81,045–81,845	-1,840	1050 to FW-08
NETB-CU12	85,445–86,245	-1,840	2450 cold aisle
NETA-CU13	86,245–87,045	-1,840	2450 cold aisle

Network cabinet	X0–X1	ΔX vs previous	Front clearance
NETB-CU13	90,645–91,445	-1,840	1050 to FW-09
NETA-CU14	93,245–94,045	-1,440	1050 to FW-09
NETB-CU14	97,645–98,445	-1,440	2450 cold aisle
NETA-CU15	98,445–99,245	-1,440	2450 cold aisle
NETB-CU15	102,845–103,645	-1,440	2450 cold aisle

Network cabinet front zone where a fan wall faces (1,050 mm)

- Switch front withdrawal needs 1,000 mm and fan-wall service needs 850 mm; both use the same 1,050 mm zone and are not performed at the same time (red cross-hatch).
- GPU cabinet fronts keep a 1,200 mm tray-withdrawal zone (orange hatch); fan walls do not face GPU cabinets.

- Notes**
- Y chain: wall line Y0 / wall zone Y200 / CDU ≤950 to Y1150 / rear maintenance 1000 to Y2150 / cabinet 1200 to Y3350 / cold aisle 2450 to the fan-wall rear and column-flange line Y5800 (flange Y5825, partition axis Y6000).
 - Rear zone 1000 (blue dots): GPU rear access and network PSU/fan replacement from the rear.
 - Cabinets sit on IT support rails clear of the main-beam zone (axis ±160 mm) at every position.
 - Electrical-zone reserved areas north of Y6000 are facility context; door positions are to be confirmed.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Deployment plan X 71–107 m

Scale 1:100

Public Edition P2

2026-10-05

Units: mm

N-203-EN

Sheet 11 of 30

Reference row adjustment relative to the previous layout: cross aisles widened and row start moved; cabinet widths, count, order and Y positions unchanged. Building geometry, axes, columns and joists are unchanged.

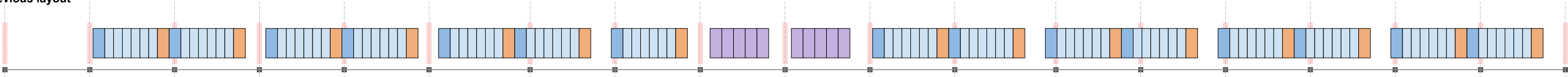
Aisle	Previous X0	Previous width	Current X0	Current width	Widened
END-W	4,835	1,200	1,155	1,200	0
XP-02	16,435	1,400	12,755	1,690	+290
XP-04	28,235	1,400	24,845	1,800	+400
XP-06	40,035	1,400	37,045	1,700	+300
XP-C1	46,635	1,550	43,945	1,550	0
XP-C2	52,185	1,550	49,495	1,650	+100
XP-C3	57,735	1,550	55,145	1,600	+50
XP-09	69,685	1,400	67,145	1,700	+300
XP-11	81,485	1,400	79,245	1,800	+400
XP-13	93,285	1,400	91,445	1,800	+400
END-E	105,085	1,200	103,645	1,200	0
Total					+2,240

Cabinet run (row order)	Cabinets	ΔX (mm)
NETA-CU01 ... NETB-CU02	16	-3,680
NETA-CU03 ... NETB-CU04	16	-3,390
NETA-CU05 ... NETB-CU06	16	-2,990
NETA-CU07 ... SPA-3	13	-2,690
SPB-1 ... OOB-B	5	-2,590
NETA-CU08 ... NETB-CU09	16	-2,540
NETA-CU10 ... NETB-CU11	16	-2,240
NETA-CU12 ... NETB-CU13	16	-1,840
NETA-CU14 ... NETB-CU15	16	-1,440

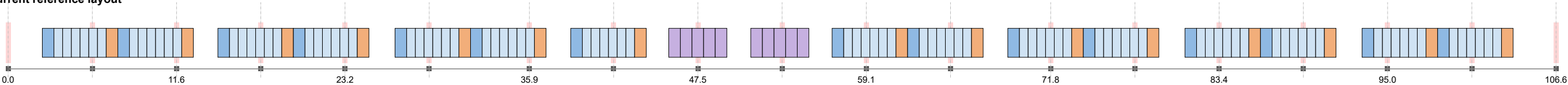
Notes

1. Aisle widening: 2,240 mm in total; row start at X1155.
2. Purpose: every IT support rail under a cabinet clears the main-beam zone (axis ±160 mm). Previous layout: 4 positions in the beam zone; current reference layout: 0.
3. This is a difference between two reference row layouts. It is not a change to the building drawings.
4. Per-cabinet previous and current X values are held in the reference cabinet placement table, which is not published in this edition.
5. Door positions are not supplied. An earlier door drawing is superseded as a current condition; door locations are to be confirmed against fan-wall and 1,050 mm front zones before layout confirmation.

Previous layout



Current reference layout



Pink bands: main-beam zone on each column axis (±160 mm). Cabinet colours as N-001-EN. Strips drawn at 1:300; axis values in m.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Reference row adjustment relative to previous layout

Scale 1:300

Public Edition P2

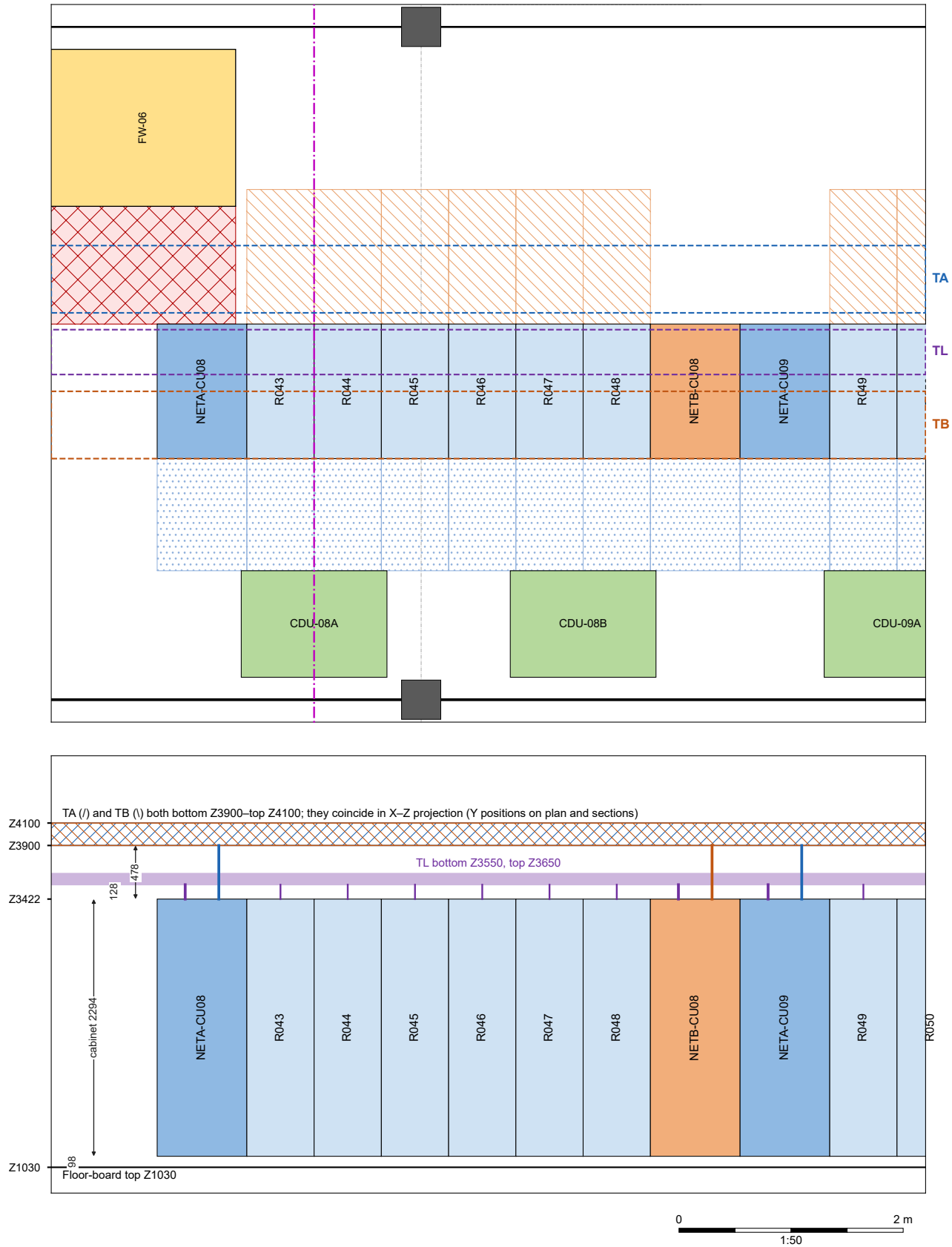
2026-10-05

Units: mm

N-204-EN

Sheet 12 of 30

Typical unit CU-08 (NETA-CU08, R043–R048, NETB-CU08, CDU-08A/B): plan (upper) and X–Z elevation (lower). Fibre drops from the trays onto the cabinet tops. Dimensions in mm.



Jumpers entering the unit cabinets from R043–R048

Enters	Class	Links	Calculated length m	Medium
NETA-CU08	NIC → compute leaf	432	5.65–9.62	OM4 MPO-12 APC jumper
NETA-CU08	DPU → NS leaf	108	6.18–10.25	OM4 MPO-12 APC jumper
NETB-CU08	NIC → compute leaf	432	5.20–10.07	OM4 MPO-12 APC jumper
NETB-CU08	DPU → NS leaf	108	6.09–10.34	OM4 MPO-12 APC jumper

Jumpers per GPU cabinet (NIC + DPU) and calculated length range

GPU cabinet	To NETA-CU08	Calculated m	To NETB-CU08	Calculated m
R043	72 + 18	5.65–7.34	72 + 18	8.65–10.34
R044	72 + 18	6.25–7.94	72 + 18	8.05–9.74
R045	72 + 18	6.85–8.54	72 + 18	7.45–9.14
R046	72 + 18	7.00–9.05	72 + 18	6.40–8.45
R047	72 + 18	7.60–9.65	72 + 18	5.80–7.85
R048	72 + 18	8.20–10.25	72 + 18	5.20–7.25

Notes

- Drops: NIC/DPU jumpers leave TL (bottom Z3550) and enter the cabinet top through a 128 mm drop; leaf/spine trunks leave TA (plane A) or TB (plane B) at bottom Z3900 and enter at the cabinet top rear. Trunk bend radius $\geq 10 \times OD$ (13 mm trunk $\rightarrow \geq 130$ mm).
- Cabinet height 2294 mm stands on the IT support rails (top Z1128); cabinet top is at absolute elevation Z3422.
- Clear height from cabinet top to the roof main-beam underside (about Z6700) exceeds 3 m; tray hangers bear on roof purlins (hanger and purlin capacity: CONFIRM, structural engineer).
- Per GPU cabinet: 72 NIC + 18 DPU jumpers to NETA and the same to NETB (N-102-EN, N-111-EN). Lengths are calculated routed lengths of the reference layout model (DESIGN), not installed measurements; values from the reference port table.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Typical unit CU-08 detail: plan and X–Z elevation

Scale 1:50

Public Edition P2

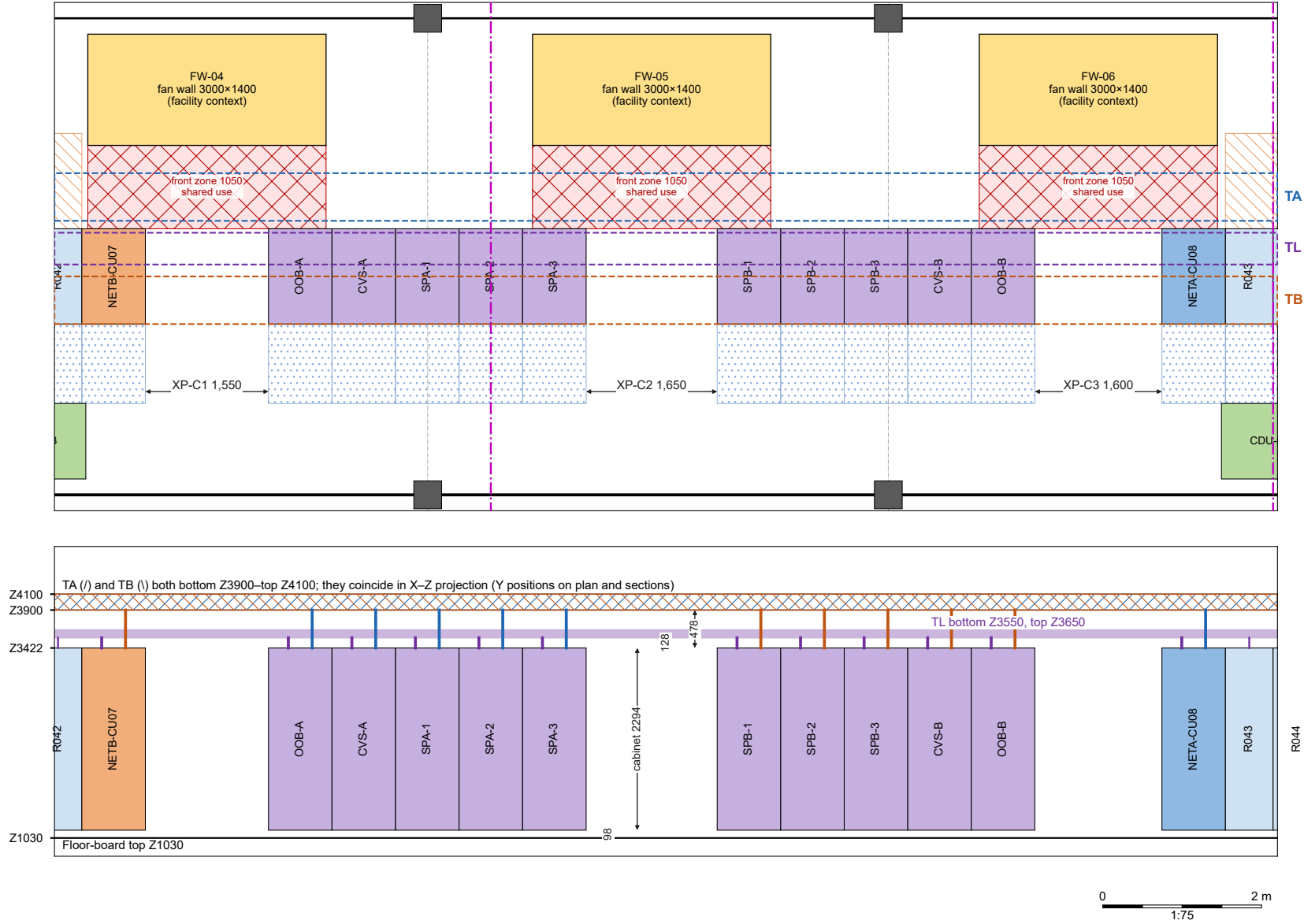
2026-10-05

Units: mm

N-301-EN

Sheet 13 of 30

Central spine zone (OOB-A, CVS-A, SPA-1..3, SPB-1..3, CVS-B, OOB-B): plan (upper) and X–Z elevation (lower). Fibre drops from the trays onto the cabinet tops. Dimensions in mm.



Trunks landing in central cabinets

Cabinet	Tray	Trunks
CVS-A (listed CS-A)	TA	60
CVS-B (listed CS-B)	TB	60
SPA-1	TA	120
SPA-2	TA	120
SPA-3	TA	120
SPB-1	TB	120
SPB-2	TB	120
SPB-3	TB	120

Notes

- Trunk unit: 1 trunk = 18 × MPO-12 legs / 216 fibres (N-503-EN).
- Drops: jumpers from TL (bottom Z3550); trunks from TA/TB (bottom Z3900) at the cabinet top rear; bend radius ≥130 mm for 13 mm trunks.
- Cabinet height 2294 mm on rails at Z1128; cabinet top Z3422.
- Fan walls FW-04, FW-05 and FW-06 face OOB-A, SPA-3/SPB-1 and OOB-B: 1,050 mm front zone shared in time between switch withdrawal (1,000) and fan-wall service (850).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Central spine zone detail: plan and X–Z elevation

Scale 1:75

Public Edition P2

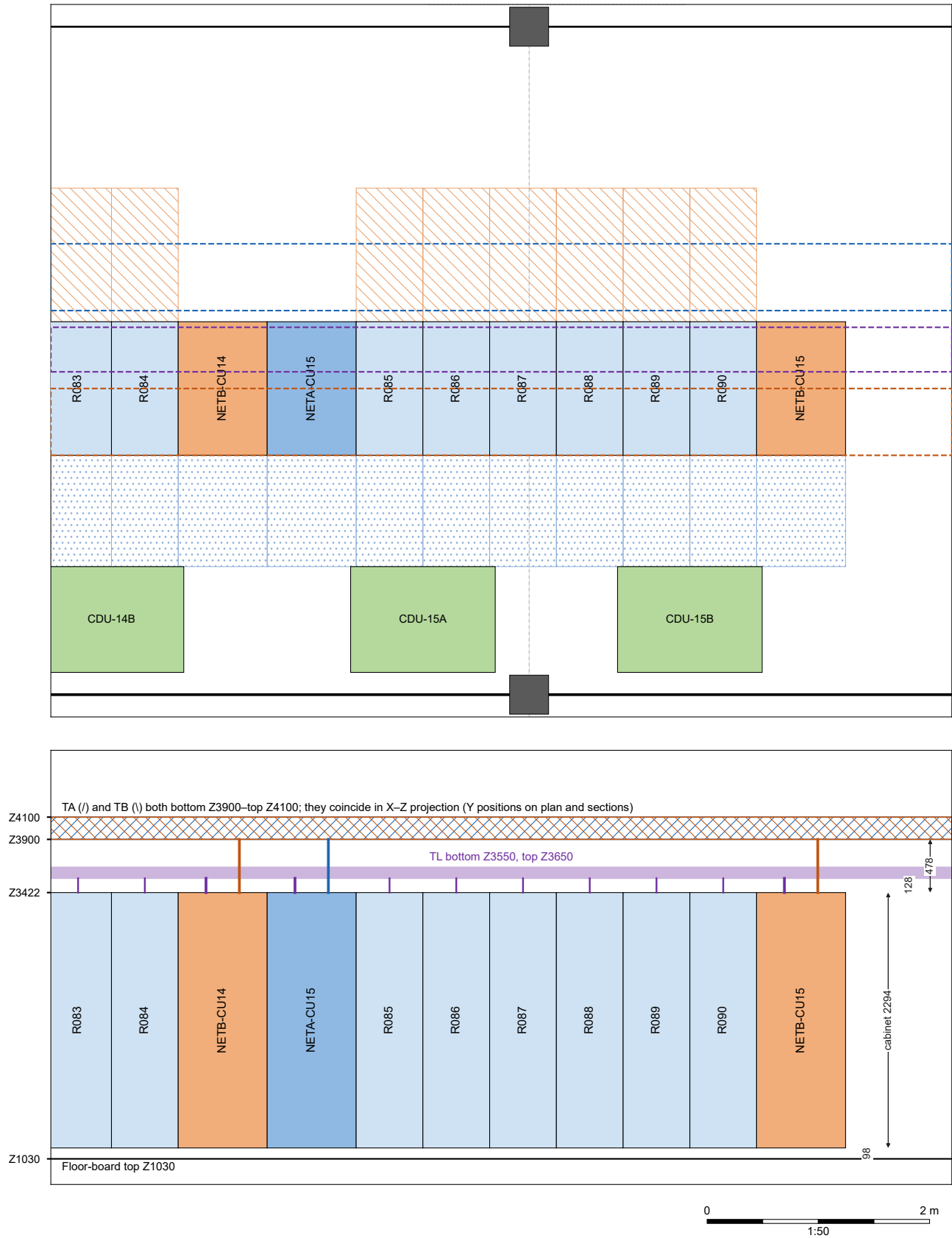
2026-10-05

Units: mm

N-302-EN

Sheet 14 of 30

Far-end unit CU-15 (NETA-CU15, R085–R090, NETB-CU15): plan (upper) and X–Z elevation (lower). Fibre drops from the trays onto the cabinet tops. Dimensions in mm.



Uplinks from the far-end unit cabinets

From	To	Links	Module	Longest calculated m	Panels
NETA-CU15	CVS-A	56	MMS4X00-NS	61.40	1
NETA-CU15	SPA-1	144	MMS4X00-NS	60.33	1
NETA-CU15	SPA-2	144	MMS4X00-NS	59.53	1
NETA-CU15	SPA-3	144	MMS4X00-NS	58.73	1
NETB-CU15	CVS-B	56	MMS4X00-NS	58.95	1
NETB-CU15	SPB-1	144	MMS4X00-NS	61.08	1
NETB-CU15	SPB-2	144	MMS4X00-NS	60.28	1
NETB-CU15	SPB-3	144	MMS4X00-NS	59.48	1

Notes

1. Drops: NIC/DPU jumpers leave TL (bottom Z3550) and enter the cabinet top through a 128 mm drop; leaf/spine trunks leave TA (plane A) or TB (plane B) at bottom Z3900 and enter at the cabinet top rear. Trunk bend radius $\geq 10 \times OD$ (13 mm trunk $\rightarrow \geq 130$ mm).
2. Cabinet height 2294 mm stands on the IT support rails (top Z1128); cabinet top is at absolute elevation Z3422.
3. Clear height from cabinet top to the roof main-beam underside (about Z6700) exceeds 3 m; tray hangers bear on roof purlins (hanger and purlin capacity: CONFIRM, structural engineer).
4. Far-end uplink channels have calculated routed lengths of 58.7–61.4 m (reference layout model, DESIGN; not installed measurements) and use MMS4X00-NS (OS2 single-mode, 100 m stated with two patch panels). Values from the reference port table.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Far-end unit CU-15 detail: plan and X–Z elevation

Scale 1:50

Public Edition P2

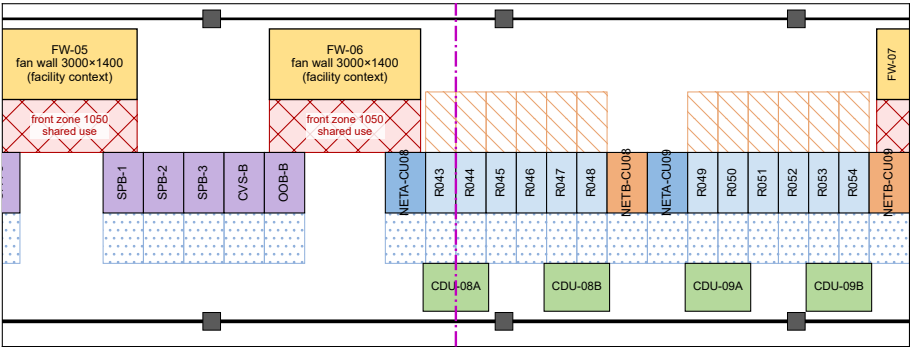
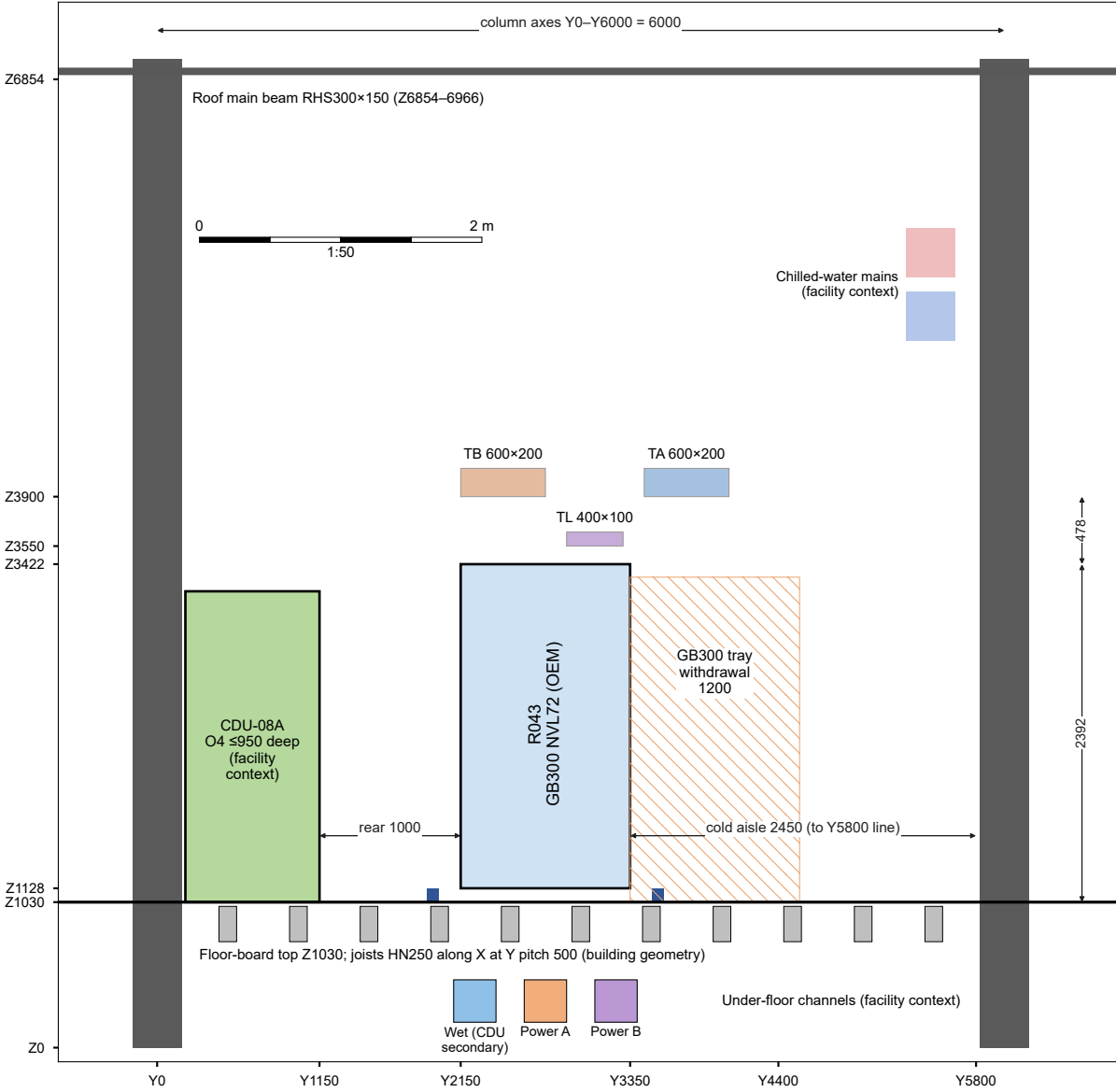
2026-10-05

Units: mm

N-303-EN

Sheet 15 of 30

Section A-A at X58145 through CDU-08A and GPU cabinets R043/R044: GPU front tray withdrawal 1,200 mm in the 2,450 mm cold aisle; rear access 1,000 mm to the CDU front. No fan wall on this cut.



Cut location (1:150): section A-A at X58,145 through CDU-08A and GPU cabinets R043/R044 boundary; no fan wall in the cut.

Notes

1. The section shows only objects that exist on the cut line; equipment at other X positions is not superimposed.
2. Network cabinet: 800 × 1200 × 2294 mm (48U, 46U usable) on IT support rails. Cabinet height is 2294 mm; its top is at absolute elevation Z3422 (rails top Z1128).
3. Rail posts 750 mm apart (NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A, 600–800 mm, FACT); front post 100 mm behind the front door.
4. IT support rails RHS80×40 at Y1950 / Y3550 (dark blue, top Z1128) carry the cabinets; the 30 mm floor board is not an equipment support.
5. Trays TL / TA / TB are separated in Y; TA and TB (bottom Z3900) are 250 mm above the TL top (Z3650). Power and wet channels are below the floor boards.
6. Roof beams, purlins and tray hangers: capacity to be confirmed by the structural engineer (CONFIRM).
7. GPU cabinets: front tray withdrawal 1,200 mm (Y3350–4550); rear access in the 1,000 mm zone between cabinet rear (Y2150) and CDU front (Y1150).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Section A-A: GPU cabinets and CDU

Scale 1:50

Public Edition P2

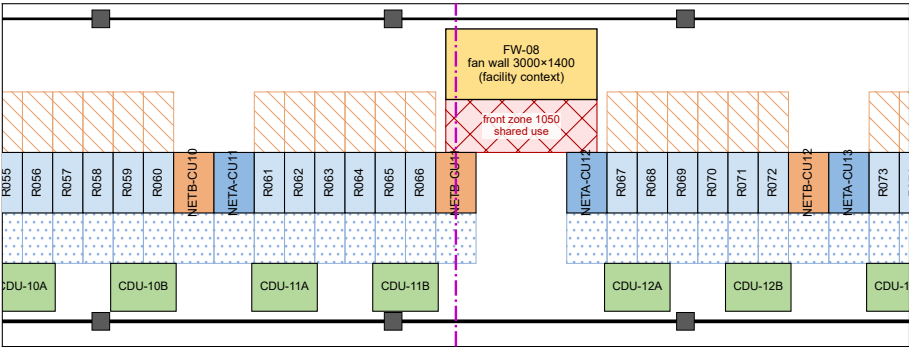
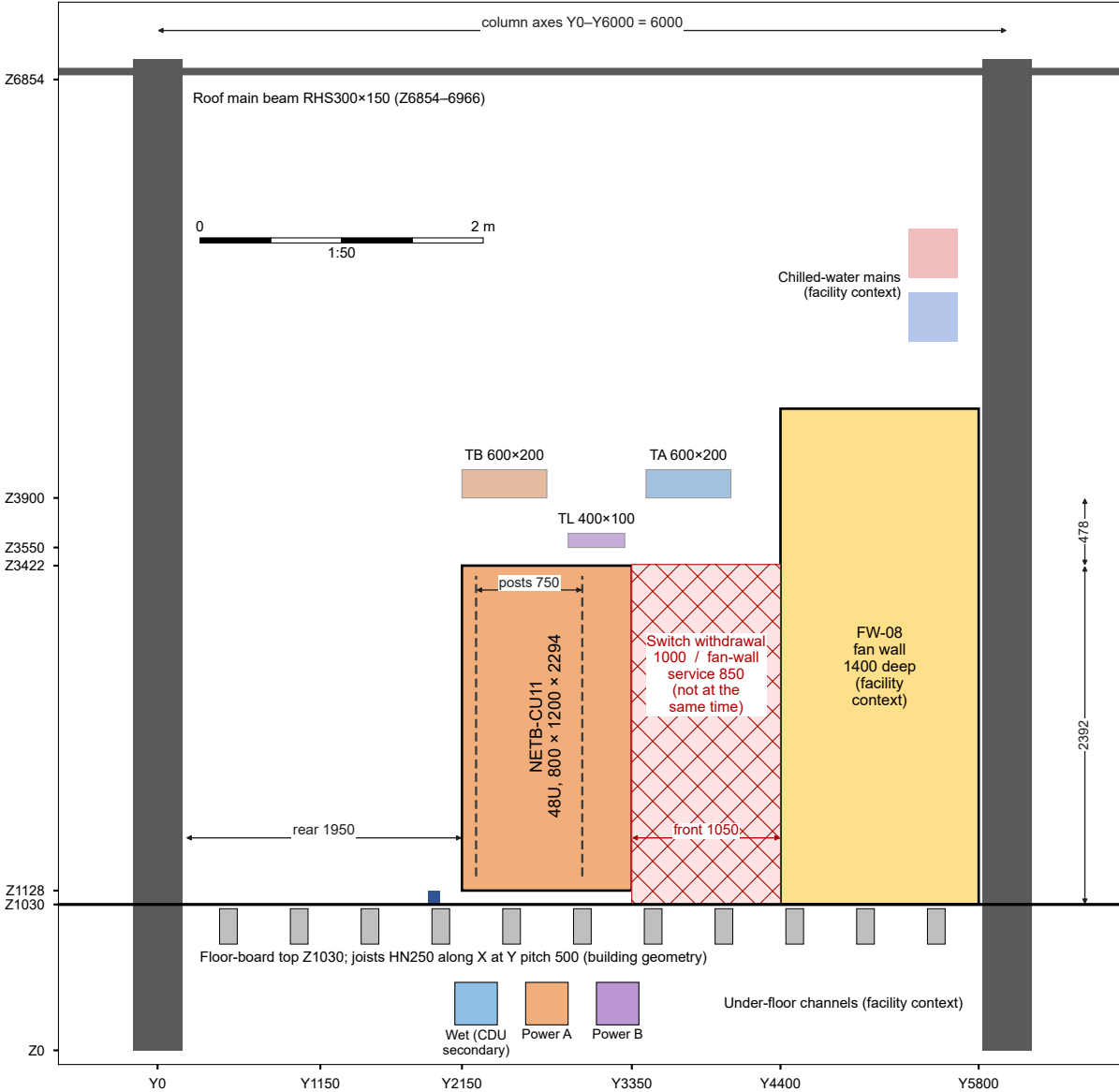
2026-10-05

Units: mm

N-311-EN

Sheet 16 of 30

Section B-B at X78845 through network cabinet NETB-CU11 and fan wall FW-08: the network cabinet front zone is 1,050 mm to the fan-wall face.



Cut location (1:150): section B-B at X78,845 through Network cabinet NETB-CU11 and fan wall FW-08.

Front-zone limitation at NETB-CU11 and every network cabinet facing a fan wall

- 1,050 mm clear between cabinet front (Y3350) and fan-wall face (Y4400).
- Switch front withdrawal (1,000 mm) and fan-wall service (850 mm) use this same zone at different times; they are not scheduled together. Fan-wall service is coordinated with network maintenance.

Notes

1. The section shows only objects that exist on the cut line; equipment at other X positions is not superimposed.
2. Network cabinet: 800 × 1200 × 2294 mm (48U, 46U usable) on IT support rails. Cabinet height is 2294 mm; its top is at absolute elevation Z3422 (rails top Z1128).
3. Rail posts 750 mm apart (NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A, 600–800 mm, FACT); front post 100 mm behind the front door.
4. IT support rails RHS80×40 at Y1950 / Y3550 (dark blue, top Z1128) carry the cabinets; the 30 mm floor board is not an equipment support.
5. Trays TL / TA / TB are separated in Y; TA and TB (bottom Z3900) are 250 mm above the TL top (Z3650). Power and wet channels are below the floor boards.
6. Roof beams, purlins and tray hangers: capacity to be confirmed by the structural engineer (CONFIRM).
7. Rear: switch PSUs and fans are replaced from the rear (1,950 mm to the wall zone where no CDU is behind the cabinet).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Section B-B: network cabinet and fan wall

Scale 1:50

Public Edition P2

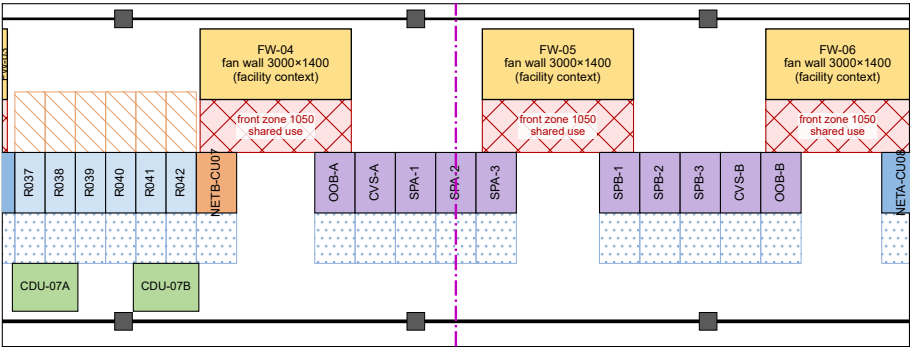
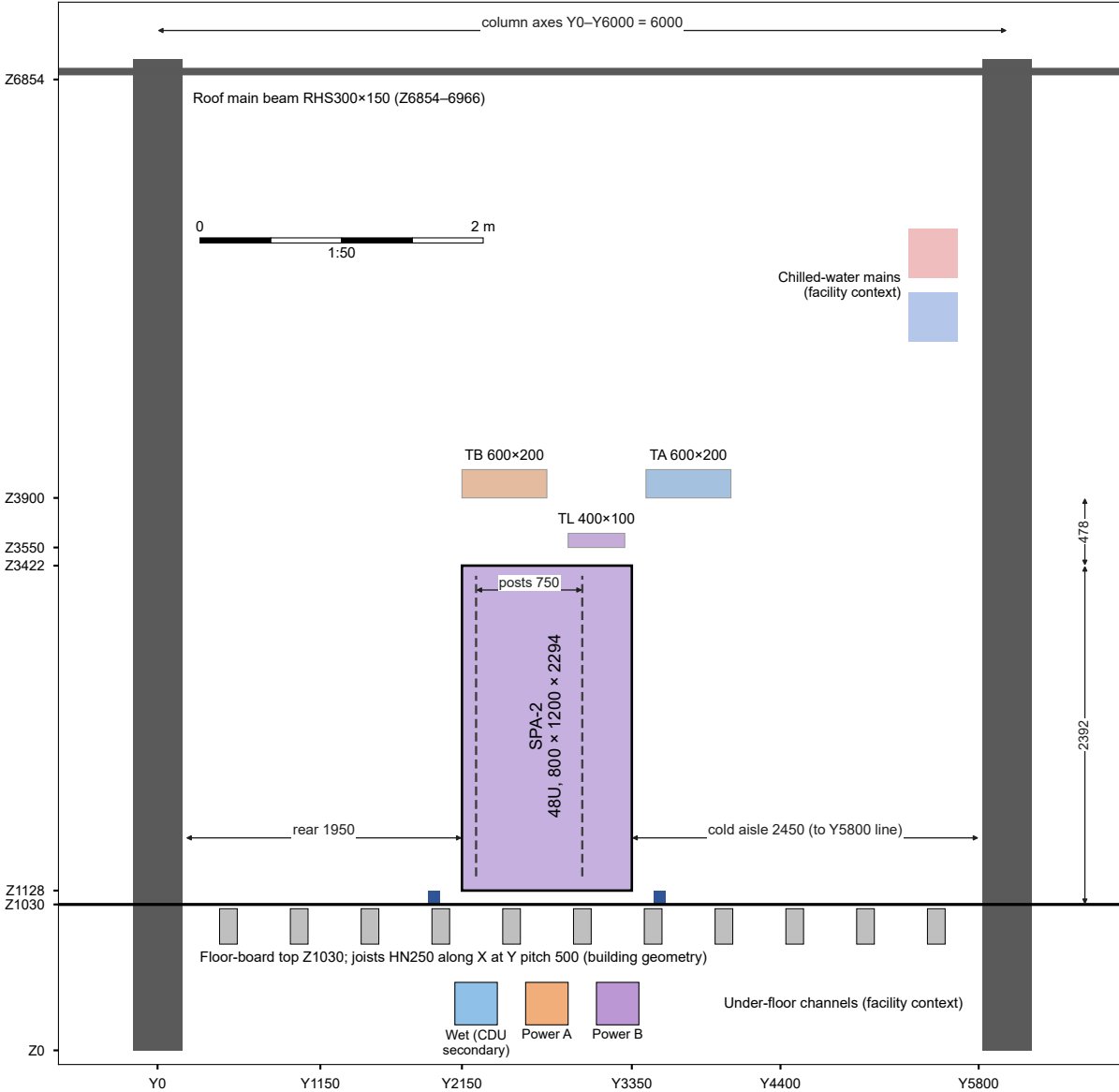
2026-10-05

Units: mm

N-312-EN

Sheet 17 of 30

Section C-C at X48295 through spine cabinet SPA-2: TA and TB trunks enter from the cabinet top rear; the cold aisle in front of SPA-2 is 2,450 mm.



Cut location (1:150): section C-C at X48,295 through Central spine cabinet SPA-2 (TA/TB trunk entry).

Notes

1. The section shows only objects that exist on the cut line; equipment at other X positions is not superimposed.
2. Network cabinet: 800 × 1200 × 2294 mm (48U, 46U usable) on IT support rails. Cabinet height is 2294 mm; its top is at absolute elevation Z3422 (rails top Z1128).
3. Rail posts 750 mm apart (NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A, 600–800 mm, FACT); front post 100 mm behind the front door.
4. IT support rails RHS80×40 at Y1950 / Y3550 (dark blue, top Z1128) carry the cabinets; the 30 mm floor board is not an equipment support.
5. Trays TL / TA / TB are separated in Y; TA and TB (bottom Z3900) are 250 mm above the TL top (Z3650). Power and wet channels are below the floor boards.
6. Roof beams, purlins and tray hangers: capacity to be confirmed by the structural engineer (CONFIRM).
7. Spine cabinets SPA-1..3 / SPB-1..3 receive 120 trunks each from TA (plane A) or TB (plane B); CVS-A/B receive 60 each (N-503-EN).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Section C-C: spine cabinet and trunk entry

Scale 1:50

Public Edition P2

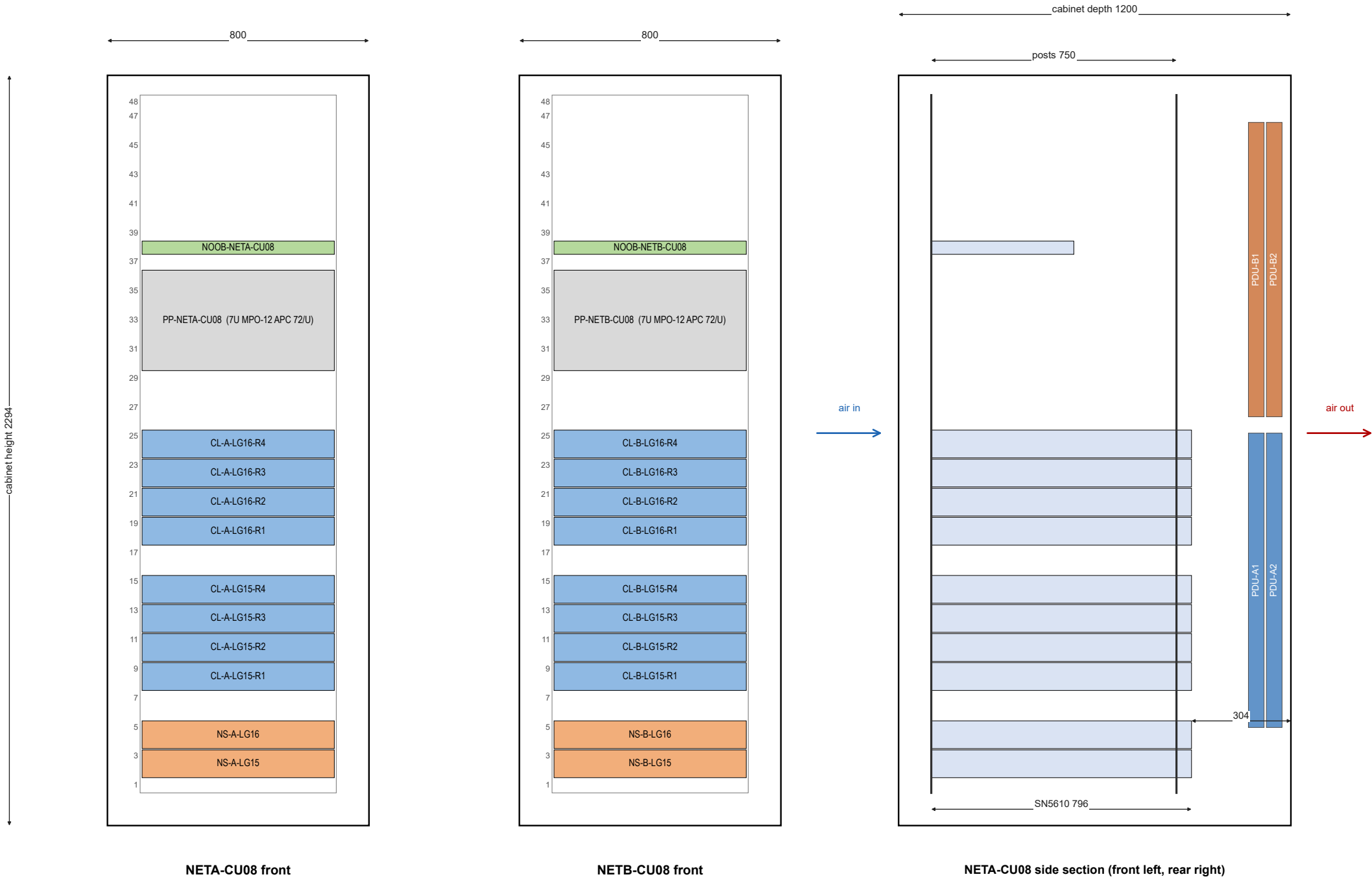
2026-10-05

Units: mm

N-313-EN

Sheet 18 of 30

Network cabinet 800 × 1200 × 2294 mm, 48U (46U usable). SN5610 depth 796 mm incl. fan handles; NVIDIA SN5xxx rail kit 930-9SKIT-00L0-00A fits post spacing 600–800 mm; PSUs and fans are replaced from the rear, switches withdraw to the front.



NETA-CU08 equipment (NETB-CU08 mirrors with plane B)

U	Device	Role	Model	kg
U2–U3	NS-A-LG15	North–south (NS) leaf	SN5610	26.9
U4–U5	NS-A-LG16	North–south (NS) leaf	SN5610	26.9
U8–U9	CL-A-LG15-R1	Compute rail leaf	SN5610	26.9
U10–U11	CL-A-LG15-R2	Compute rail leaf	SN5610	26.9
U12–U13	CL-A-LG15-R3	Compute rail leaf	SN5610	26.9
U14–U15	CL-A-LG15-R4	Compute rail leaf	SN5610	26.9
U18–U19	CL-A-LG16-R1	Compute rail leaf	SN5610	26.9
U20–U21	CL-A-LG16-R2	Compute rail leaf	SN5610	26.9
U22–U23	CL-A-LG16-R3	Compute rail leaf	SN5610	26.9
U24–U25	CL-A-LG16-R4	Compute rail leaf	SN5610	26.9
U30–U36	PP-NETA-CU08	Patch panel	7U MPO-12 APC 72/U	28.0
U38	NOOB-NETA-CU08	Cabinet management switch	SN2201	7.41

Notes

1. Posts: front post 100 mm behind the front door; rear post 750 mm behind the front post (adjustable 600–800 mm). SN5610 extends 46 mm behind the rear post; 304 mm remains at the rear for power cords and fibre bend space. The 1,200 mm cabinet depth is not the rail dimension; the rail fits the post spacing.
2. PDUs: 0U vertical at the rear, both sides. SN5610 PSU1/PSU2 → PDU-A1/A2 and PSU3/PSU4 → PDU-B1/B2 (at least two PSUs operating, FACT); SN4600C / SN2201 PSU1 → A1, PSU2 → B1.
3. Used 28U / free 18U in each unit cabinet (U positions from the reference device table). Door perforation and blanking panels follow the front-to-rear (C2P) airflow requirement (product: CONFIRM).
4. Front withdrawal of switches needs 1,000 mm; at cabinets facing a fan wall the 1,050 mm front zone is shared in time with fan-wall service.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Network cabinet elevations: unit cabinets

NETA/NETB-CU08

Scale 1:15

Public Edition P2

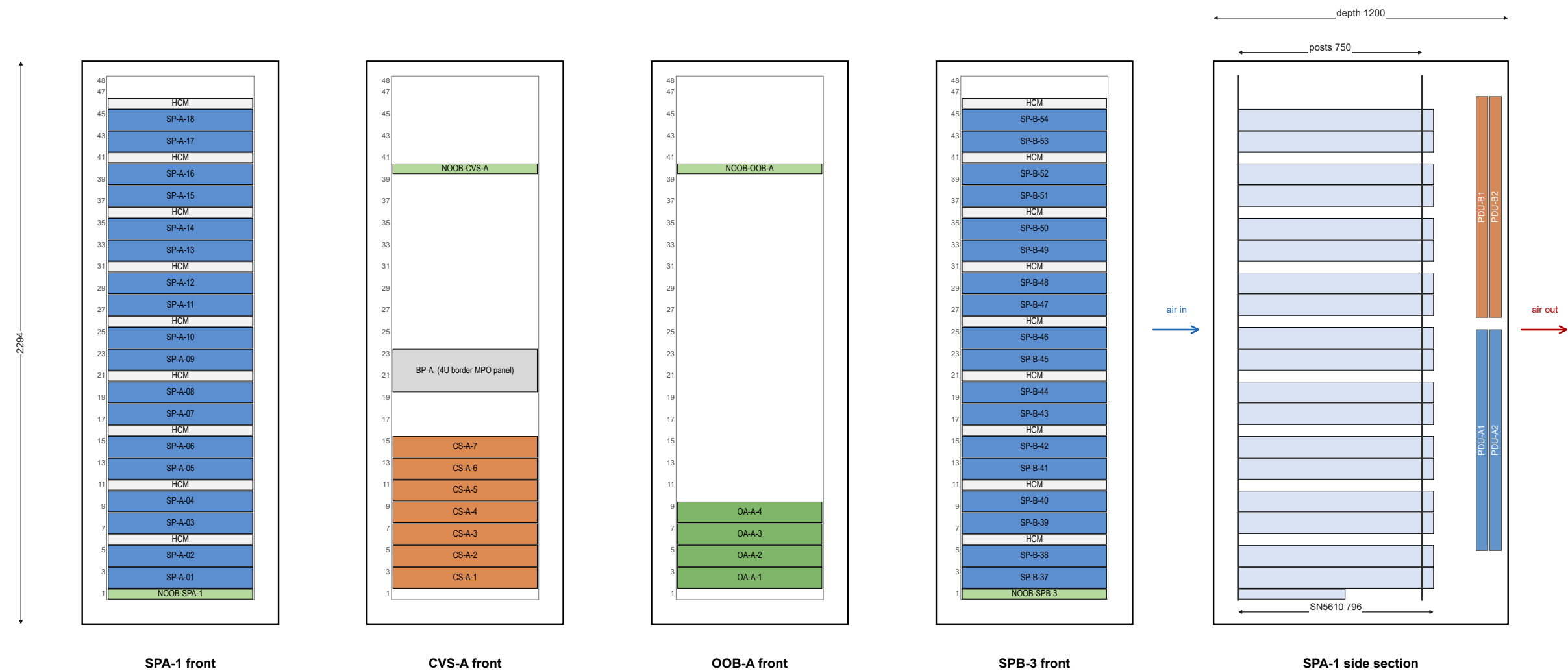
2026-10-05

Units: mm

N-401-EN

Sheet 19 of 30

Central network cabinets: compute spine SPA-1 (plane A) and SPB-3 (plane B), north–south spine CVS-A, OOB aggregation OOB-A. Same cabinet (800 × 1200 × 2294, 48U) and rail geometry as N-401-EN.



Central cabinet composition

Cabinet	Equipment	U used / free	PDUs
OOB-A	SN2201 × 1; SN4600C × 4	9 / 37	2
CVS-A	4U border MPO panel × 1; SN2201 × 1; SN5610 × 7	19 / 27	4
SPA-1	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
SPA-2	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
SPA-3	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
SPB-1	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
SPB-2	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
SPB-3	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18	46 / 0	4
CVS-B	4U border MPO panel × 1; SN2201 × 1; SN5610 × 7	19 / 27	4
OOB-B	SN2201 × 1; SN4600C × 4	9 / 37	2

Notes

- Spine cabinets hold 18 SN5610 each with a 1U horizontal cable manager after every two switches; the cabinet management switch (NOOB, SN2201) is at U1.
- CVS-A/B hold 7 SN5610 (U2–U15) and a 4U border MPO panel (BP, U20–U23) for the external north–south interface EXT-IF-NS-A/B. The 56 CVS ports per side behind it are unused (not reserved); storage use of these ports is a candidate only.
- OOB-A/B hold 4 SN4600C aggregation switches (U2–U9); unused 100G ports reach the site management interface EXT-IF-OOB.
- Free U is an arithmetic total from the device table; it is not guaranteed contiguous and is not reserved space.
- PDU arrangement: cabinets with SN5610 have four PDUs (A1/A2/B1/B2); OOB cabinets have two (A1/B1).
- Front zone: OOB-A, SPA-3, SPB-1 and OOB-B face fan walls (1,050 mm, shared in time between switch withdrawal and fan-wall service).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Network cabinet elevations: SPA-1, CVS-A, OOB-A, SPB-3

Scale 1:20

Public Edition P2

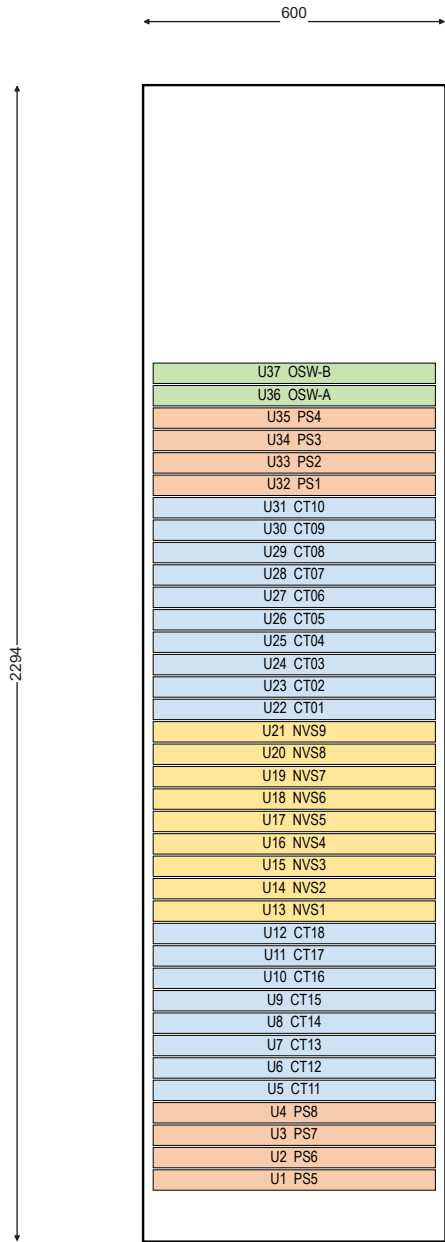
2026-10-05

Units: mm

N-402-EN

Sheet 20 of 30

GPU cabinet (GB300 NVL72, OEM, 600 × 1200 × 2294 mm): compute-tray NIC OSFP and DPU QSFP112 ports at the front; two in-rack OEM SN2201 management switches. Arrangement as the OEM reference for cabinet R043.



R043 front (OEM arrangement)

Network interfaces per GPU cabinet

Interface	Per cabinet	Connects to	Module / cable
CX-8 NIC OSFP 800G (2 × 400G)	72 (18 trays × 4)	Plane-A CL-A and plane-B CL-B leaves (N-102-EN)	MMA4Z00-NS-FLT; OM4 jumper via TL
BlueField-3 QSFP112 400G	36 (18 × 2)	NS-A and NS-B leaves (N-111-EN)	MMA1Z00-NS400 family*; OM4 jumper via TL
OEM SN2201 100G uplink	4 (2 switches × 2)	OOB aggregation OA-A / OA-B (N-121-EN)	MMA1B00-C100D; OM4 MPO-12 UPC
Tray / NVLink / power-shelf 1G management	80	In-rack OEM SN2201 (GPU cabinet TDP)	In-cabinet copper (OEM)

Notes

- GPU cabinet, trays, NVLink switch trays and in-rack SN2201 are OEM equipment; their power is inside the GPU cabinet TDP and is not counted in network load.
- GPU-side optical modules (MMA4Z00-NS-FLT 6,480 and MMA1Z00-NS400 family 3,240 in total): whether they are included in the GPU cabinet TDP is to be confirmed by the OEM (CONFIRM); they are not counted in network load.
- * BF3-side 400G SR4 QSFP112 multimode module of the MMA1Z00-NS400 family; exact OPN and shell variant for the B3240 cage HOLD (OEM/NVIDIA confirmation).
- Front: 1,200 mm tray-withdrawal zone in the cold aisle (N-311-EN). Rear: 1,000 mm access zone to the CDU front.
- Jumpers from the cabinet top drop from TL (bottom Z3550) and are dressed to the front-panel NIC and DPU ports.
- Selected media: optical NIC/DPU uplinks to the leaves (N-901-EN). Copper remains for cabinet-local 1G management.

Port map of R043 trays (from the reference port table): NIC r → leaf CL-x-LG15-Rr at the port shown (same port on rails 1–4); DPU ports to the NS leaves

Tray	NIC1–4 → CL-A-LG15-R1..R4 port	NIC1–4 → CL-B-LG15-R1..R4 port	DPU → NS-A leaf port	DPU → NS-B leaf port
CT01	P1/1	P1/1	NS-A-LG15:P1/1	NS-B-LG15:P1/1
CT02	P1/2	P1/2	NS-A-LG15:P1/2	NS-B-LG15:P1/2
CT03	P2/1	P2/1	NS-A-LG15:P2/1	NS-B-LG15:P2/1
CT04	P2/2	P2/2	NS-A-LG15:P2/2	NS-B-LG15:P2/2
CT05	P3/1	P3/1	NS-A-LG15:P3/1	NS-B-LG15:P3/1
CT06	P3/2	P3/2	NS-A-LG15:P3/2	NS-B-LG15:P3/2
CT07	P4/1	P4/1	NS-A-LG15:P4/1	NS-B-LG15:P4/1
CT08	P4/2	P4/2	NS-A-LG15:P4/2	NS-B-LG15:P4/2
CT09	P5/1	P5/1	NS-A-LG15:P5/1	NS-B-LG15:P5/1
CT10	P5/2	P5/2	NS-A-LG15:P5/2	NS-B-LG15:P5/2
CT11	P6/1	P6/1	NS-A-LG15:P6/1	NS-B-LG15:P6/1
CT12	P6/2	P6/2	NS-A-LG15:P6/2	NS-B-LG15:P6/2
CT13	P7/1	P7/1	NS-A-LG15:P7/1	NS-B-LG15:P7/1
CT14	P7/2	P7/2	NS-A-LG15:P7/2	NS-B-LG15:P7/2
CT15	P8/1	P8/1	NS-A-LG15:P8/1	NS-B-LG15:P8/1
CT16	P8/2	P8/2	NS-A-LG15:P8/2	NS-B-LG15:P8/2
CT17	P9/1	P9/1	NS-A-LG15:P9/1	NS-B-LG15:P9/1
CT18	P9/2	P9/2	NS-A-LG15:P9/2	NS-B-LG15:P9/2

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

GPU cabinet network interfaces (OEM reference)

Scale 1:15

Public Edition P2

2026-10-05

Units: mm

N-403-EN

Sheet 21 of 30

Network cabinets: position, existing source pairing, equipment, U use, PDUs and PSU inlets, input power (planning = typical × 1.25), heat to air, airflow at 15 K, mass and support-point load (equal sharing). Switch mass FACT; empty cabinet 180 kg CONFIRM.

Cabinet	Plane	X0–X1	UPS A / B (existing)	Switches	U used / free	PDUs	PSU inlets	Typical kW	Planning kW	Heat to air kW	Airflow m³/h	Mass kg	Support kN* (×1.5)	Front zone
NETA-CU01	A	2,355–3,155	UPS7-01 / UPS7-02	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU01	B	6,755–7,555	UPS7-02 / UPS7-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU02	A	7,555–8,355	UPS7-01 / UPS7-02	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU02	B	11,955–12,755	UPS7-02 / UPS7-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-01
NETA-CU03	A	14,445–15,245	UPS7-02 / UPS7-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-01
NETB-CU03	B	18,845–19,645	UPS7-03 / UPS7-04	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU04	A	19,645–20,445	UPS7-03 / UPS7-04	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU04	B	24,045–24,845	UPS7-04 / UPS7-05	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-02
NETA-CU05	A	26,645–27,445	UPS7-04 / UPS7-05	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-02
NETB-CU05	B	31,045–31,845	UPS7-05 / UPS7-06	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU06	A	31,845–32,645	UPS7-05 / UPS7-06	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU06	B	36,245–37,045	UPS7-06 / UPS7-01	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-03
NETA-CU07	A	38,745–39,545	UPS7-06 / UPS7-01	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-03
NETB-CU07	B	43,145–43,945	UPS7-05 / UPS7-06	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-04
OOB-A	central	45,495–46,295	UPS7-03 / UPS7-04	5	9 / 37	2	10	2.52	3.15	3.18	653	282	0.69 (1.04)	1050 to FW-04
CVS-A	central	46,295–47,095	UPS7-01 / UPS7-02	8	19 / 27	4	30	14.15	17.68	17.74	3,649	476	1.17 (1.75)	2450 aisle
SPA-1	central	47,095–47,895	UPS7-01 / UPS7-02	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	2450 aisle
SPA-2	central	47,895–48,695	UPS7-03 / UPS7-04	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	2450 aisle
SPA-3	central	48,695–49,495	UPS7-05 / UPS7-06	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	1050 to FW-05
SPB-1	central	51,145–51,945	UPS8-01 / UPS8-02	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	1050 to FW-05

Cabinet	Equipment composition	Cabinet	Equipment composition
NETA-CU01	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	NETA-CU06	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU01	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	NETB-CU06	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU02	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	NETA-CU07	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU02	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	NETB-CU07	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU03	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	OOB-A	SN2201 × 1; SN4600C × 4
NETB-CU03	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	CVS-A	4U border MPO panel × 1; SN2201 × 1; SN5610 × 7
NETA-CU04	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	SPA-1	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18
NETB-CU04	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	SPA-2	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18
NETA-CU05	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	SPA-3	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18
NETB-CU05	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10	SPB-1	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18

* Support kN: cabinet mass × 9.81 m/s² ÷ 4, with the weight shared equally by four support points (static gravity load). It is not the worst unequal support reaction and not a rail, joist or floor capacity. The value in brackets applies an illustrative factor of 1.5 for unequal sharing; it is not a calculated reaction. Support reactions and capacity: structural engineer (CONFIRM).

The CVS-cabinet 4U MPO panel is the border panel; its ports are unused (not reserved) and storage use is a candidate only.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Network cabinet schedule (1 of 2)

Scale NTS

Public Edition P2

2026-10-05

Units: mm

N-411-EN

Sheet 22 of 30

Network cabinets: position, existing source pairing, equipment, U use, PDUs and PSU inlets, input power (planning = typical × 1.25), heat to air, airflow at 15 K, mass and support-point load (equal sharing). Switch mass FACT; empty cabinet 180 kg CONFIRM.

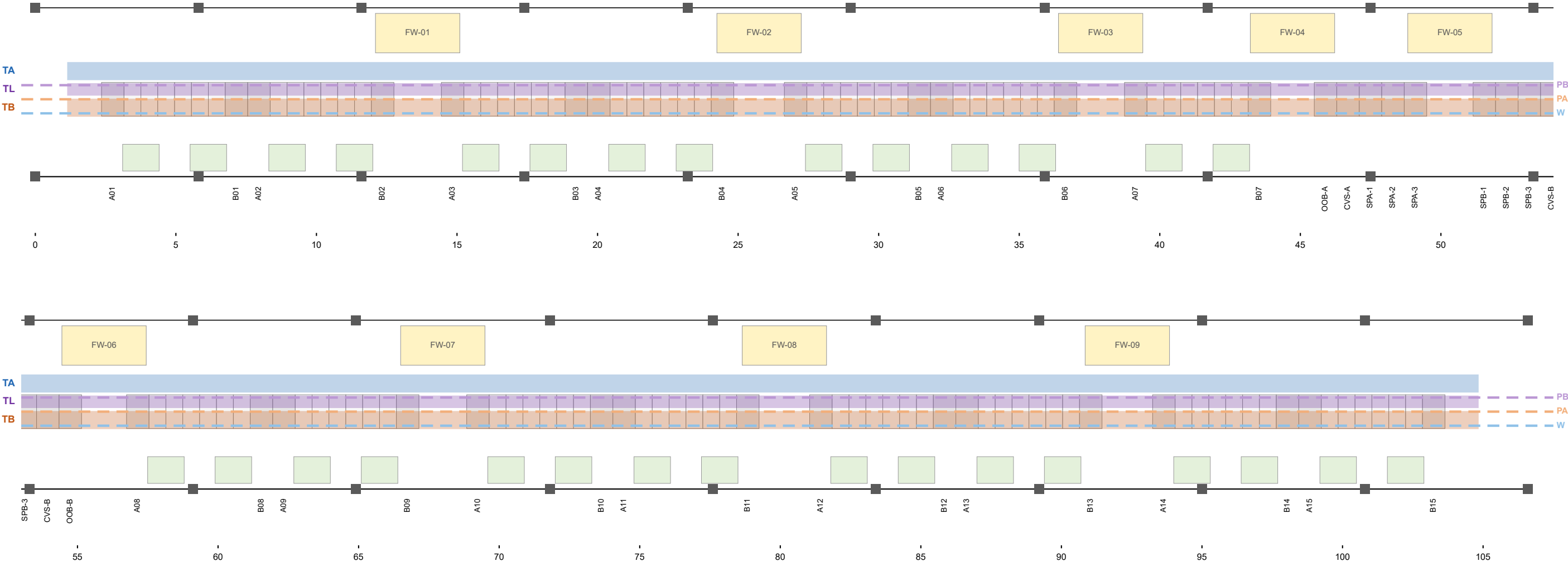
Cabinet	Plane	X0–X1	UPS A / B (existing)	Switches	U used / free	PDUs	PSU inlets	Typical kW	Planning kW	Heat to air kW	Airflow m³/h	Mass kg	Support kN* (×1.5)	Front zone
SPB-2	central	51,945–52,745	UPS8-03 / UPS8-04	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	2450 aisle
SPB-3	central	52,745–53,545	UPS8-05 / UPS8-06	19	46 / 0	4	74	36.22	45.27	45.33	9,323	794	1.95 (2.92)	2450 aisle
CVS-B	central	53,545–54,345	UPS8-07 / UPS8-01	8	19 / 27	4	30	14.15	17.68	17.74	3,649	476	1.17 (1.75)	2450 aisle
OOB-B	central	54,345–55,145	UPS8-03 / UPS8-04	5	9 / 37	2	10	2.52	3.15	3.18	653	282	0.69 (1.04)	1050 to FW-06
NETA-CU08	A	56,745–57,545	UPS8-01 / UPS8-02	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-06
NETB-CU08	B	61,145–61,945	UPS8-02 / UPS8-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU09	A	61,945–62,745	UPS8-01 / UPS8-02	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU09	B	66,345–67,145	UPS8-02 / UPS8-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-07
NETA-CU10	A	68,845–69,645	UPS8-02 / UPS8-03	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-07
NETB-CU10	B	73,245–74,045	UPS8-03 / UPS8-04	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU11	A	74,045–74,845	UPS8-03 / UPS8-04	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU11	B	78,445–79,245	UPS8-04 / UPS8-05	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-08
NETA-CU12	A	81,045–81,845	UPS8-04 / UPS8-05	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-08
NETB-CU12	B	85,445–86,245	UPS8-05 / UPS8-06	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU13	A	86,245–87,045	UPS8-05 / UPS8-06	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU13	B	90,645–91,445	UPS8-06 / UPS8-07	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-09
NETA-CU14	A	93,245–94,045	UPS8-06 / UPS8-07	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	1050 to FW-09
NETB-CU14	B	97,645–98,445	UPS8-07 / UPS8-01	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETA-CU15	A	98,445–99,245	UPS8-07 / UPS8-01	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle
NETB-CU15	B	102,845–103,645	UPS8-01 / UPS8-02	11	28 / 18	4	42	18.58	23.23	23.29	4,789	572	1.4 (2.11)	2450 aisle

Cabinet	Equipment composition
SPB-2	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18
SPB-3	1U horizontal manager × 9; SN2201 × 1; SN5610 × 18
CVS-B	4U border MPO panel × 1; SN2201 × 1; SN5610 × 7
OOB-B	SN2201 × 1; SN4600C × 4
NETA-CU08	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU08	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU09	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU09	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU10	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU10	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10

Cabinet	Equipment composition
NETA-CU11	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU11	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU12	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU12	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU13	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU13	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU14	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU14	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETA-CU15	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10
NETB-CU15	7U MPO-12 APC 72/U × 1; SN2201 × 1; SN5610 × 10

* Support kN: cabinet mass × 9.81 m/s² ÷ 4, with the weight shared equally by four support points (static gravity load). It is not the worst unequal support reaction and not a rail, joist or floor capacity. The value in brackets applies an illustrative factor of 1.5 for unequal sharing; it is not a calculated reaction. Support reactions and capacity: structural engineer (CONFIRM).
The CVS-cabinet 4U MPO panel is the border panel; its ports are unused (not reserved) and storage use is a candidate only.

Fibre trays TL 400×100 (bottom Z3550) and TA / TB 600×200 (bottom Z3900) above the cabinet row, and under-floor wet / power-A / power-B channels (Y2250 / 2750 / 3250). Plane-A trunks use TA only; plane-B trunks use TB only.



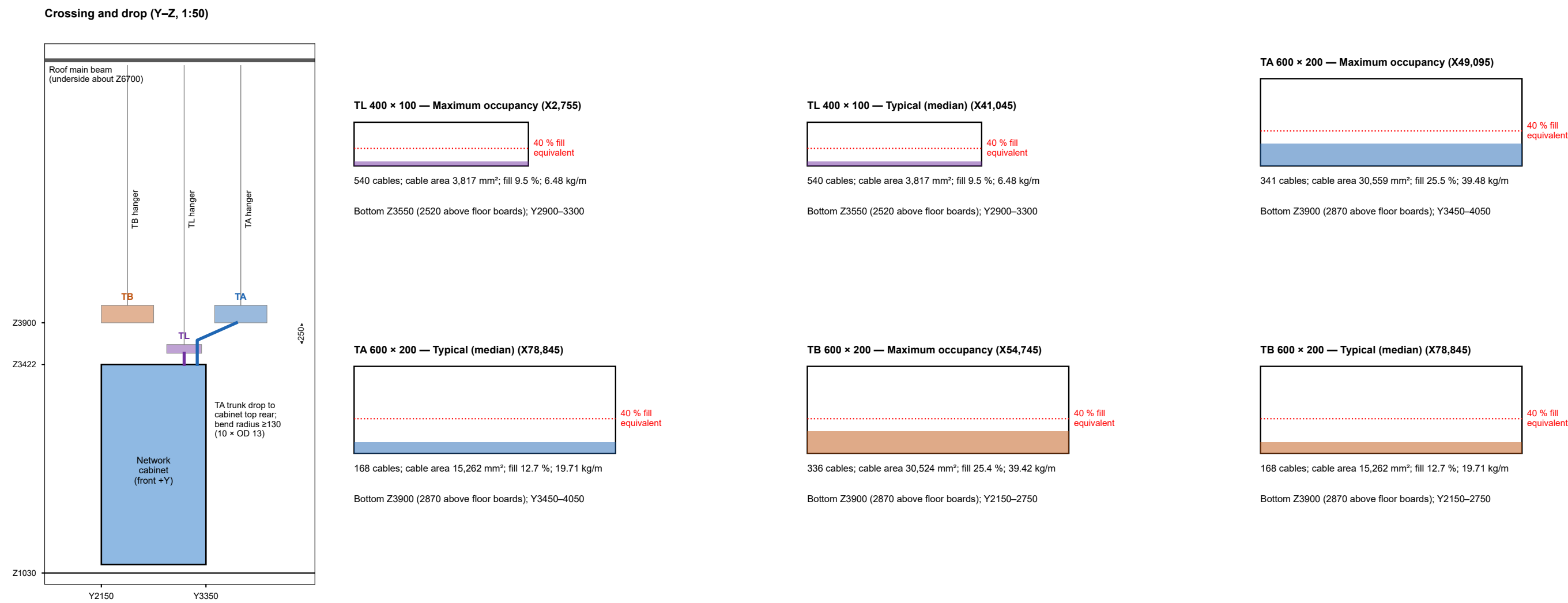
Tray labels at left; under-floor channels at right: W = wet (Y2250, facility context), PA = power A (Y2750), PB = power B (Y3250). Network cabinet IDs below the wall line (A01 = NETA-CU01). X in m.

Tray / channel	Size	Y range	Level	Carries
TL	400 × 100	Y2900–3300	bottom Z3550	NIC and DPU jumpers of both planes (common route)
TA	600 × 200	Y3450–4050	bottom Z3900	Plane-A trunks (leaf → SPA, NS → CVS-A) and plane-A OOB 100G uplinks
TB	600 × 200	Y2150–2750	bottom Z3900	Plane-B trunks (leaf → SPB, NS → CVS-B) and plane-B OOB 100G uplinks
Power A / B	under floor	Y2750 / Y3250	below Z1030	PDU A1/A2 and B1/B2 branch circuits (existing supply input)
Wet	under floor	Y2250	below Z1030	CDU secondary piping (facility context)

Notes

- The three fibre trays are separated in Y and do not cross in plan; TA / TB bottoms (Z3900) are 250 mm above the TL top (Z3650).
- Fibre is above the cabinet tops (Z ≥3550); power and water are below the floor boards (Z <1030): vertical separation.
- Tray hangers bear on the roof purlins; hanger, seismic bracing and purlin capacity: CONFIRM (structural engineer).
- Penetrations of the partition for power branches: position CONFIRM (facility).

Tray sections at the maximum-occupancy and typical positions (cable counts at the actual X position); crossing and drop detail in the Y–Z plane. Filled area is drawn as an equivalent depth (cable cross-section area ÷ tray width).



- Notes**
- Crossing: TA (Y3450–4050), TB (Y2150–2750) and TL (Y2900–3300) are side by side in Y and do not cross in plan. TA / TB bottoms Z3900 are 250 mm above the TL top Z3650; trunk drops pass over TL into the cabinet.
 - Cable units (N-503-EN): MPO-12 jumper OD 3.0 mm (DESIGN); trunk 18 × MPO-12 / 216 fibres, OD 13 mm envelope (DESIGN; product CONFIRM).
 - Fill is the cross-section area ratio with a 40 % limit (DESIGN; engineer of record / deploying organisation CONFIRM). Sections drawn at 1:10.
 - Supports at 1.5 m spacing (DESIGN); tray self-weight 15 kg/m (CONFIRM); maximum cable load 39.48 kg/m in TA at X49095.
 - Hangers, seismic bracing and roof purlins: capacity to be confirmed by the structural engineer (CONFIRM).

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Tray sections: maximum and typical occupancy,
crossing and drop

Scale 1:10 / 1:50

Public Edition P2

2026-10-05

Units: mm

N-502-EN

Sheet 25 of 30

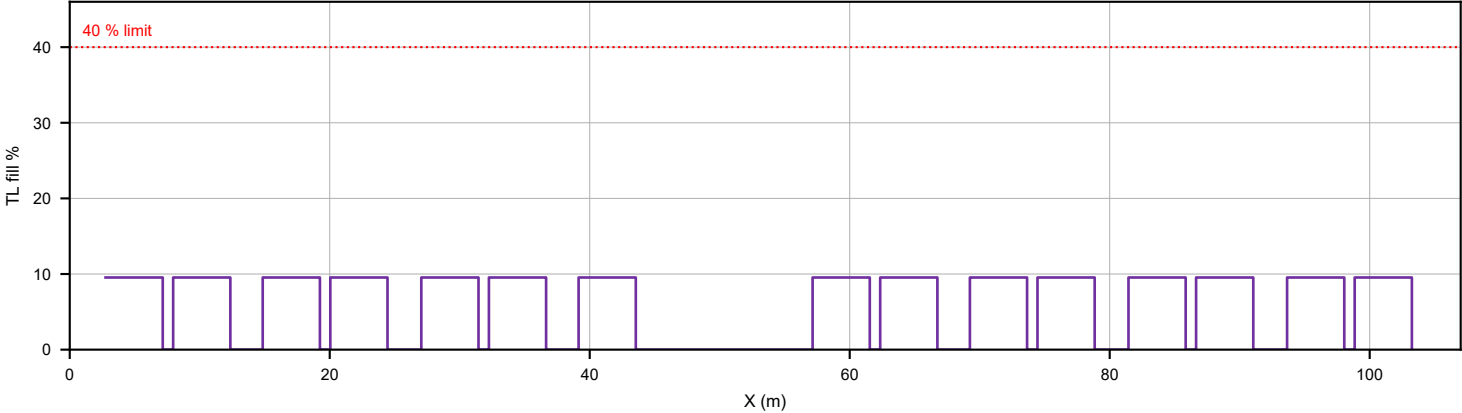
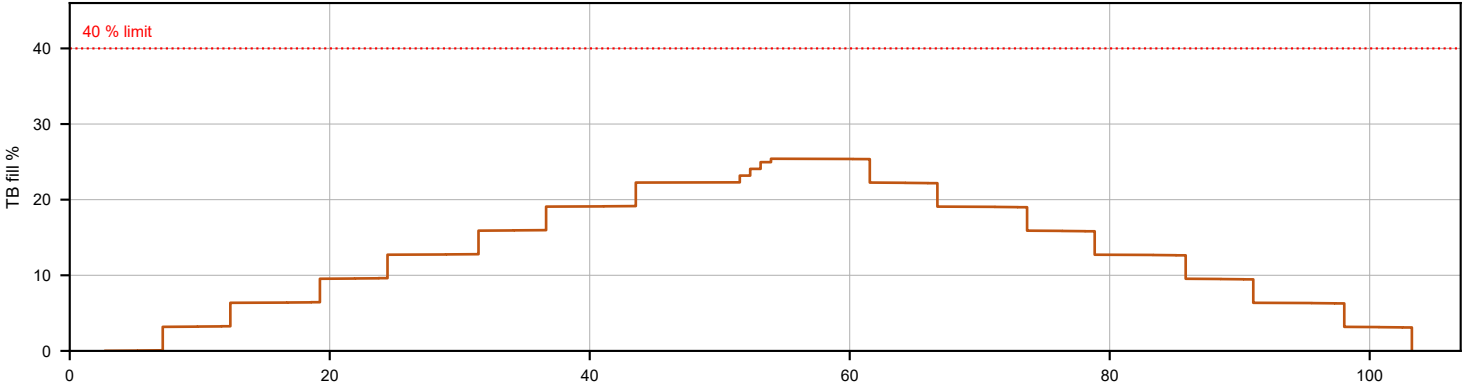
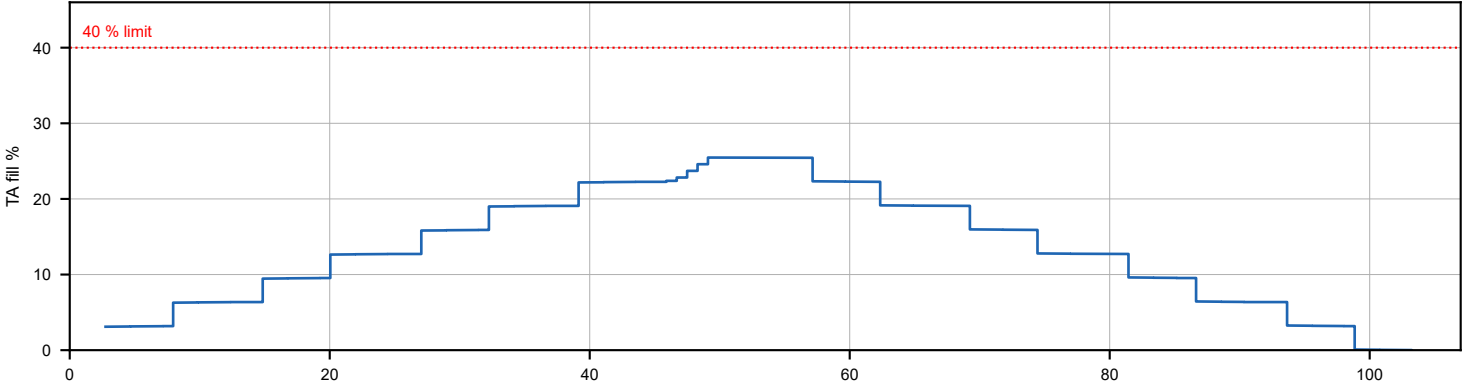
Cable counting units, trunk quantities by destination and tray fill profile along X (values from the reference trunk and tray-profile tables).

Unit	Definition	Status
Jumper, OC-1 to OC-4	MPO-12 APC jumper, OD 3.0 mm (NIC, DPU and 400G/800G module channels)	DESIGN
Jumper / leg, OC-5 OOB 100G	OM4 MPO-12 UPC; every module-facing end UPC female; current direct cords UPC at both ends	DESIGN; product CONFIRM
Trunk	18 × MPO-12 legs / 216 fibres, OD 13.0 mm envelope	DESIGN; product CONFIRM
Static bend radius	10 × OD (trunk 13 mm → 130 mm)	DESIGN
Installation bend radius	20 × OD	DESIGN
Tray fill limit	40 % of cross-section area	DESIGN; EOR / deploying organisation CONFIRM
End slack	0.5 m per end	DESIGN

Tray	Destination	MPO-12 legs	Trunks	Fibres
TA	CS-A	840	60	10,080
TA	SPA-1	2,160	120	25,920
TA	SPA-2	2,160	120	25,920
TA	SPA-3	2,160	120	25,920
TB	CS-B	840	60	10,080
TB	SPB-1	2,160	120	25,920
TB	SPB-2	2,160	120	25,920
TB	SPB-3	2,160	120	25,920
Total		14,640	840	175,680

Notes

- 1. Trunks = Σ roundup(legs per leaf–destination pair ÷ 18).
- 2. OOB 100G (OC-5) cords: MMA1B00-C100D ports are MPO UPC; module-facing ends UPC female; any later intermediate mated pair needs complementary pin/hole gender and compatible polish and key. OC-1 to OC-4 remain APC.
- 3. TL carries 540 jumpers (9.5 % fill) within each leaf-group section; its profile drops to zero between sections and across the central zone.
- 4. Maximum fill: TA 25.5 % at X49095 and TB 25.4 % at X54745 (entry to the central zone).



K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Cable units, trunk quantities and tray fill

Scale NTS

Public Edition P2

2026-10-05

Units: mm

N-503-EN

Sheet 26 of 30

Failure domains and service boundaries: physical endpoint loss, network reachability and host / job recovery are distinguished. Monthly 99.999 % is a design target to be verified by commissioning and operating measurements.

Level	Event	NIC: host power lost	NIC: both planes unreachable	NIC: one plane unreachable	BF3 tray: host lost	BF3 tray: one port	BF3 tray: no path	OOB racks: host lost	Min. leaf uplink fraction
Cable	One NIC→leaf jumper broken	0	0	1	0	0	0	0	1.0
Cable	One leaf→spine link broken	0	0	0	0	0	0	0	0.9815
Port / module	One leaf port cage (optical module) failed	0	0	2	0	0	0	0	1.0
Port / NIC	One CX-8 NIC (both half-ports share one cage) failed	0	1	0	0	0	0	0	1.0
Device	One compute leaf failed	0	0	54	0	0	0	0	1.0
Device	One compute spine failed	0	0	0	0	0	0	0	0.9815
Device	One north–south leaf failed	0	0	0	0	54	0	0	1.0
Cabinet	One plane-A network cabinet lost (NETA-CU08)	0	0	432	0	108	0	0	1.0
Cabinet	Spine cabinet SPA-1 lost	0	0	0	0	0	0	0	0.6667
Plane	Entire plane A lost (all A leaf/spine/NS leaf/CVS)	0	0	6480	0	1620	0	0	1.0
Tray (whole run)	All links in tray TA lost (whole trunk run damaged)	0	0	6480	0	1620	0	0	0.0
Tray (whole run)	All links in tray TB lost (whole trunk run damaged)	0	0	6480	0	1620	0	0	0.0
Tray (local cut)	TA cut at X30000 (unit CU-05, cabinet R029 X29845–30445); links passing the cut lost	0	0	2160	0	540	0	0	0.0
Tray (local cut)	TB cut at X30000 (unit CU-05, cabinet R029 X29845–30445); links passing the cut lost	0	0	1728	0	432	0	0	0.0
Tray (whole run)	All jumpers in the TL section of unit CU-08 lost	0	432	0	0	0	108	0	1.0
Tray (local cut)	TL cut at X59945 (between R045 and R046)	0	0	432	0	108	0	0	1.0
Generator segment (common)	Generator segment SEG-A lost (after battery depletion; risk case)	1008	0	288	252	1368	0	14	0.6667
Generator segment (common)	Generator segment SEG-B lost (after battery depletion; risk case)	1008	0	720	252	180	0	14	0.6667
Generator segment (common)	Generator segment SEG-C lost (after battery depletion; risk case)	1008	0	720	252	180	0	14	0.6667
Generator segment (common)	Generator segment SEG-D lost (after battery depletion; risk case)	1512	0	216	378	54	0	21	0.6667
Single UPS path	Any one of 13 UPS outputs lost (reference connectivity mapping, each path; maximum shown)	0	0	0	0	0	0	0	1.0

Shaded rows: common causes that leave NIC endpoints unreachable on both planes or DPU trays with no path.

Service boundaries: what a loss means for each service object

Service object	Meaning of loss	Recovery mechanism	Duration status
Physical NIC endpoint (6,480)	NIC / cage failure or host power loss	Replace NIC / restore power (maintenance, hours)	Physical endpoint loss; service accounting follows the network measurement scope defined by the deploying organisation
Network reachability (per plane)	Endpoint has no path in that plane	Other plane carries traffic; ECMP / BGP convergence	CONFIRM: convergence time by IST
Host service	Host loses network service	Host multi-NIC / multi-plane failover	CONFIRM: host failover by test
Job continuity	Training / inference job interrupted	Checkpoint restart	CONFIRM: workload property

Commissioning (IST) measurements: separate list, not tied to the service-object rows above

Commissioning (IST) measurement	Method	Record
Single link / port failure	Pull NIC and leaf–spine links per plane	Recovery time, loss, convergence
Single leaf / spine power-off	Remove power from an A or B device	Affected endpoints (compare with table), recovery
Single UPS path loss	Each of the 13 paths in turn	No network device reset; PSU transfer; PDU phase current
Network cabinet inlet temperature	Top / middle / bottom of each front door	Maximum inlet ≤35 °C
Job-level test	Representative job under the faults above	Interruption / restart time against monthly allowance

1. Monthly allowance for 99.999 %: 24.192 / 25.056 / 25.920 / 26.784 s for 28 / 29 / 30 / 31-day months. It is a design target, not a measured result. Monthly availability is measured per the supplement measurement rule (Section IS4.5): union of overlapping impaired intervals, stated measurement resolution and clock uncertainty, and no exclusion of common-cause events by definition. It is not derived from equipment MTBF or from soak-test duration.
2. Whole-run loss = every link in that tray; local cut = links passing the given X position. "One plane unreachable" does not mean zero interruption; recovery time is measured at IST. Topological reachability is not job continuity.
3. Common causes that remain in the reference design: one CX-8 NIC uses one OSFP cage for both planes; TL carries the jumpers of both planes (TL section loss in CU-08: 432 NICs unreachable on both planes, 108 DPU trays with no path, hosts powered). Generator-segment loss is a risk case after battery depletion.
4. The TA local cut at X30000 (2,160 NICs one plane, 540 BF3 trays one port) is represented in the supplement by fault case FT-16: logical isolation of the affected links at the switch side, without unplugging cords or touching the tray; isolation set and acceptance basis HOLD. No physical test has been executed.

K&K Data Service Inc. kkdatasvc.com/lab/downloads inquiry@kkdatasvc.com	K&K Data GPU Hall — Network Reference Design PUBLIC EDITION P2 Reference design for study and evaluation — not for construction	Failure domains, service boundaries and commissioning measurements Scale NTS	Public Edition P2 2026-10-05 Units: mm	N-801-EN Sheet 27 of 30
---	--	---	---	--

Cable, media and module summary for all 39,066 link records of the reference port table (grouped as in the reference link summary); both half-ports of a cage use the same module type and medium; label formats.

Class	Module A	Module B	Medium	Tray	Links	Calculated m	Panels
NS leaf → CVS	MMA4Z00-NS	MMA4Z00-NS	OM4 trunk (fan-out legs at spine)	TA	560	16.17–44.00	1
NS leaf → CVS	MMA4Z00-NS	MMA4Z00-NS	OM4 trunk (fan-out legs at spine)	TB	560	16.62–44.35	1
NS leaf → CVS	MMS4X00-NS	MMS4X00-NS	OS2 trunk (fan-out legs at spine)	TA	280	47.36–61.40	1
NS leaf → CVS	MMS4X00-NS	MMS4X00-NS	OS2 trunk (fan-out legs at spine)	TB	280	46.12–58.95	1
DPU → NS leaf	MMA1Z00-NS400 family*	MMA4Z00-NS	OM4 MPO-12 APC jumper	TL	3,240	6.09–10.34	0
Compute leaf → spine	MMA4Z00-NS	MMA4Z00-NS	OM4 trunk (fan-out legs at spine)	TA	4,320	14.45–43.23	1
Compute leaf → spine	MMA4Z00-NS	MMA4Z00-NS	OM4 trunk (fan-out legs at spine)	TB	4,320	14.80–43.68	1
Compute leaf → spine	MMS4X00-NS	MMS4X00-NS	OS2 trunk (fan-out legs at spine)	TA	2,160	43.95–60.33	1
Compute leaf → spine	MMS4X00-NS	MMS4X00-NS	OS2 trunk (fan-out legs at spine)	TB	2,160	44.70–61.08	1
Network cabinet 1G management	—	—	Cat6A patch (in-cabinet)	in-cabinet	510	1.50–1.50	0
Network cabinet 1G management	—	—	Cat6A patch (in-cabinet PDU management)	in-cabinet	76	1.50–1.50	0
NIC → compute leaf	MMA4Z00-NS-FLT	MMA4Z00-NS	OM4 MPO-12 APC jumper	TL	12,960	5.20–10.07	0
GPU rack in-rack 1G (OEM)	—	—	Cat6A patch (in-cabinet)	in-cabinet	7,200	2.00–2.00	0
OOB 100G uplink	MMA1B00-C100D	MMA1B00-C100D	OM4 MPO-12 UPC jumper / trunk leg	TA	220	6.51–63.95	0
OOB 100G uplink	MMA1B00-C100D	MMA1B00-C100D	OM4 MPO-12 UPC jumper / trunk leg	TB	220	6.91–58.99	0

Module	Cages	Mixed cages
MMA1B00-C100D	880	0
MMA1Z00-NS400 family*	3,240	0
MMA4Z00-NS	17,860	0
MMA4Z00-NS-FLT	6,480	0
MMS4X00-NS	4,880	0

Module	Description	Location	Cages	Counted in
MMA1B00-C100D	100G SR4 QSFP28, 100 m on OM4; MPO UPC	OOB 100G uplinks (both ends)	880	Network device power
MMA1Z00-NS400 family*	400G SR4 QSFP112, 50 m; exact OPN / shell HOLD	BlueField-3 (DPU) side	3,240	GPU rack side — inclusion in GPU rack TDP to be confirmed by OEM
MMA4Z00-NS	800G 2×SR4 OSFP, 50 m (reach stated with two patch panels)	Switch side, multimode	17,860	Network device power
MMA4Z00-NS-FLT	800G 2×SR4 OSFP flat-top, 50 m	GPU compute-tray NIC side	6,480	GPU rack side — inclusion in GPU rack TDP to be confirmed by OEM
MMS4X00-NS	800G 2×DR4 OSFP, 100 m single-mode (reach stated with two patch panels)	Switch side, single-mode	4,880	Network device power

Notes

- Label format (identical label at both ends, including the far end): L016201 CL-A-LG01-R1:P33/1 > SP-A-01:P1/1.
- Port = device ID + port (CL-A-LG15-R1:P1/1); cage = device ID + cage (CL-A-LG15-R1:P1); cabinet = NETA-CU08 (top, front and rear doors); PDU = NETA-CU08-PDU-A1; trunk = TRK-TA-CL-A-LG01-R1-SPA-1 (18 legs, each leg carries its L number).
- Optical module quantities are set by the design (by cage).
- Calculated m: calculated routed channel lengths of the reference layout model (DESIGN), not installed measurements; installed lengths and OLTS losses replace them in the as-built record.
- OOB 100G uplinks (OC-5): MMA1B00-C100D ports are MPO UPC; every module-facing cord end is MPO-12 UPC female and the current direct cords are UPC at both ends. OC-1 to OC-4 use MPO-12 APC. Intermediate mated pairs need complementary pin/hole gender and compatible polish and key; per-strand pin map HOLD.
- * BF3-side 400G SR4 QSFP112 multimode module of the MMA1Z00-NS400 family; exact OPN and shell variant for the B3240 cage HOLD (OEM/NVIDIA confirmation).
- Selected media: NIC and DPU uplinks to the leaves are optical. Each CX-8 is an 800G flat-top OSFP host port split into two 400G links to switches in two different cabinets (NETA and NETB). No validated copper part for this host-800G → 2 × switch-400G direction is listed; direct lengths would be 4.27–7.05 m across up to five GPU cabinets. This means no adoptable validated part, not physical impossibility.
- BF3: a listed part exists for the switch-800G → 2 × host-400G direction, but 690 of 1,620 trays exceed 5 m and 660 routes cross the GPU tray-withdrawal zone; DPU uplinks therefore also use fibre (one medium, one maintenance method, one tray).
- Copper remains for cabinet-local 1G management (Cat6A). If the OEM certifies a copper or active-copper part for the host-800G → 2 × 400G direction, units within its certified length and outside the tray-withdrawal zone can be revised together with the module, port and tray-occupancy tables.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Cable, media and optical module schedule; labelling

Scale NTS

Public Edition P2

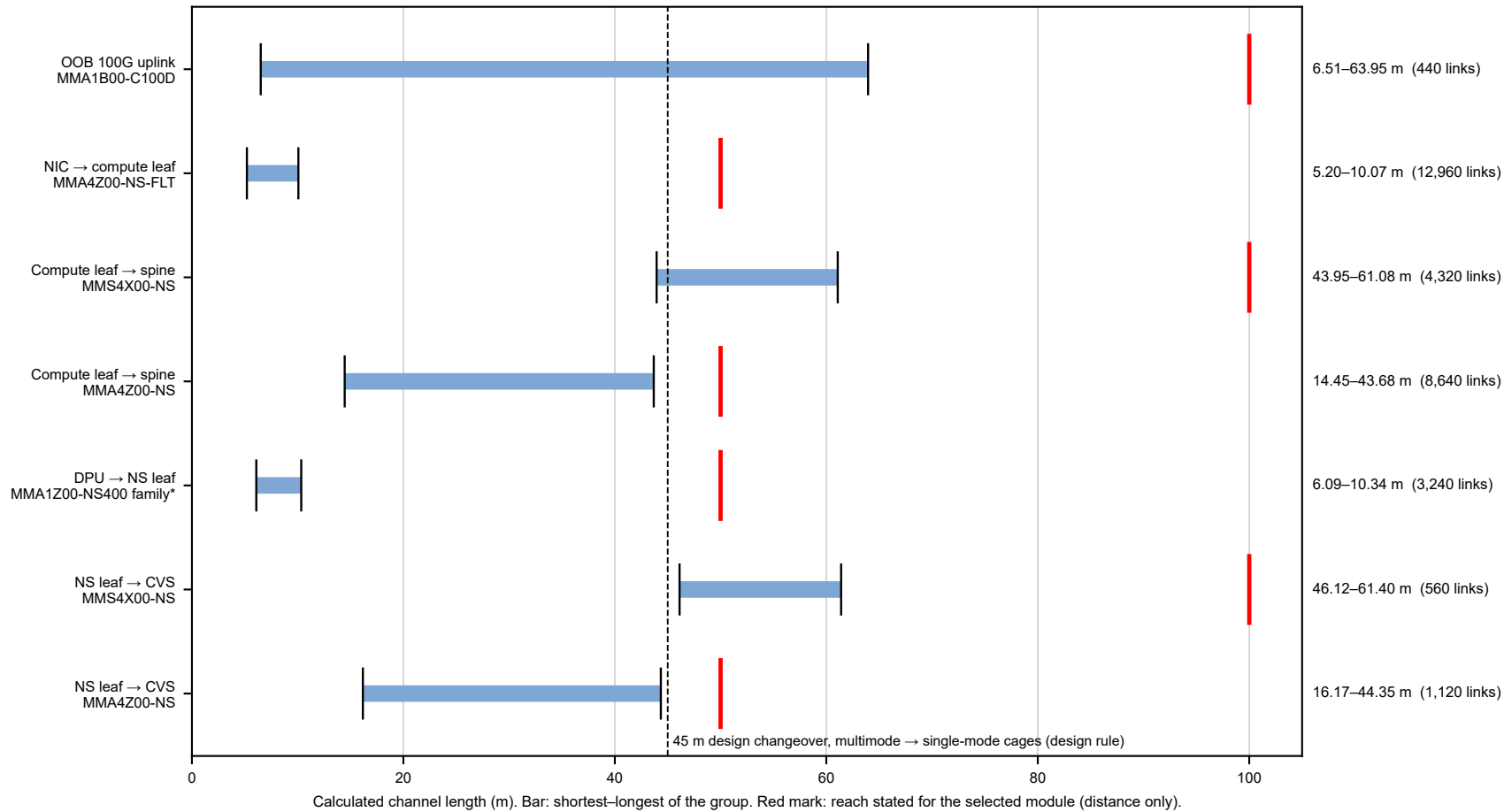
2026-10-05

Units: mm

N-901-EN

Sheet 28 of 30

Calculated channel lengths (reference layout model, DESIGN) of each link group against the reach stated for the selected module. Bars show the shortest and longest calculated channel of each group, not a distribution. Loss allowance, component maxima and test limits remain HOLD.



Longest calculated channels (from the reference port table)

Link	End A	End B	Module	Calculated m	Panels	Reach m	Reach – length m
L038250	NOOB-NETB-CU15:U1	OA-A-4:P30	MMA1B00-C100D	63.95	0	100	36.05
L038219	R090-OSW-A:U1	OA-A-3:P59	MMA1B00-C100D	63.43	0	100	36.57
L038220	R090-OSW-B:U1	OA-A-3:P60	MMA1B00-C100D	63.38	0	100	36.62
L038217	R089-OSW-A:U1	OA-A-3:P57	MMA1B00-C100D	62.83	0	100	37.17
L038218	R089-OSW-B:U1	OA-A-3:P58	MMA1B00-C100D	62.78	0	100	37.22
L038215	R088-OSW-A:U1	OA-A-3:P55	MMA1B00-C100D	62.23	0	100	37.77
L038216	R088-OSW-B:U1	OA-A-3:P56	MMA1B00-C100D	62.18	0	100	37.82
L038213	R087-OSW-A:U1	OA-A-3:P53	MMA1B00-C100D	61.63	0	100	38.37
L038214	R087-OSW-B:U1	OA-A-3:P54	MMA1B00-C100D	61.58	0	100	38.42
L023465	NS-A-LG29:P33/1	CS-A-1:P57/1	MMS4X00-NS	61.40	1	100	38.60
L023466	NS-A-LG29:P33/2	CS-A-1:P57/2	MMS4X00-NS	61.40	1	100	38.60
L023467	NS-A-LG29:P34/1	CS-A-1:P58/1	MMS4X00-NS	61.40	1	100	38.60

Module	Manufacturer statement (FACT)	Design use	Longest calculated m
MMA4Z00-NS / -FLT	2 × 400G SR4, 50 m; stated with two patch panels	45 m design changeover (design rule)	44.35
MMA1Z00-NS400 family*	400G SR4 QSFP112, 50 m (page retrieved for MMA1Z00-NS400-T)	45 m design changeover (design rule)	10.34
MMS4X00-NS	2 × 400G DR4 single-mode, 100 m; two patch panels	Cages whose longer half exceeds 45 m	43.95–61.40
MMA1B00-C100D	100G SR4, 100 m on OM4; MPO UPC	OOB 100G uplinks, checked against 100 m	63.95

- Notes**
- Distance check: all 31,280 optical links are within the reach stated for the selected module; channels have at most one patch panel. Insertion loss is counted by mated pairs: a pass-through adapter panel contributes one mated pair, a cassette with connectors on both faces contributes two. The manufacturer's reach statement with two patch panels does not define a channel loss allowance.
 - 45 m is the design changeover length between multimode and single-mode cages (manufacturer reach 50 m on OM4, stated with two patch panels); it is a design rule, not a manufacturer limit or a measured value. The longest 400G/800G multimode channel is 44.35 m.
 - OM4 100 m applies only to the selected 100G SR4 module (MMA1B00-C100D) on the OOB uplinks; it is not a general OM4 reach.
 - A cage is single-mode when its longer half-port exceeds 45 m, because both halves of a cage use the same module; 68 single-mode links are themselves shorter than 45 m (shortest 43.95 m).
 - Channel lengths are calculated routed lengths from the reference layout model (DESIGN), including the design allowances of the reference model; they are not installed measurements. Installed lengths and OLTs losses are recorded per channel at installation and replace the calculated values in the as-built record.
 - The insertion-loss estimates of the reference link table (OM4 3.0 dB/km, OS2 0.5 dB/km, MPO 0.35 dB per mated pair) are prior reference estimates only. They are not maximum specifications of selected products and do not define a channel allowance; the approved allowance, component maxima and test limits come from the selected module and cabling vendors (CONFIRM). Allowance, component maxima and instrument uncertainty: HOLD (supplement Section IS3.3).
 - * BF3-side 400G SR4 QSFP112 multimode module of the MMA1Z00-NS400 family; exact OPN and shell variant for the B3240 cage HOLD (OEM/NVIDIA confirmation).
 - Ranges and link counts per group from the reference link summary.

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Channel length against optical module reach

Scale NTS

Public Edition P2

2026-10-05

Units: mm

N-902-EN

Sheet 29 of 30

Network equipment interface requirements for the existing facility systems: device quantities, power inputs, heat and airflow, mass, cabinet and rail envelope, pathway loads and external network interfaces. Facility capability is an input and is not verified here.

Device	Qty	Typical W / device	Planning W / device	PSU rating W	Note
SN5610	422	1,656–2,006	2,070–2,508	2,000	Leaf / spine / NS leaf / CVS; 4 PSUs, ≥2 operating (FACT)
SN4600C	8	566–616	708–770	1,100	OOB aggregation; 2 PSUs
SN2201	40	103	129	not stated	Cabinet management; 2 PSUs

Demand item	Value	Condition	Status
Network input power, typical / planning	808.0 / 1,010.0 kW	Planning = typical × 1.25 (no published SN5610 maximum)	DESIGN
Heat to room air (planning, incl. PDU losses)	1,012.4 kW	Front-to-rear (C2P) airflow; all to room air	DESIGN
Airflow required at 15 K rise	208,212 m³/h	Sum of the 40 cabinet values	DESIGN
Rack PDUs	156 (32 A: 132; 63 A: 24)	3-phase 4-wire wye 400Y/230 V (400 V line-to-line, 230 V line-to-neutral), metered, C13/C19 outlets; outlet voltage per the selected PDU product (CONFIRM)	DESIGN
PSU inlets	1,784	C14–C15 cords; device range 200–240 V (SN5610 FACT)	DESIGN
Largest cabinet input (planning)	45.27 kW (SPA-1)	Spine cabinets SPA/SPB	DESIGN
Heaviest cabinet	794 kg (SPA-1)	Weight shared equally by 4 support points: 794 kg × 9.81 m/s² ÷ 4 ≈ 1.95 kN each (static gravity load); not the worst unequal reaction and not a rail, joist or floor capacity. 2.92 kN applies an illustrative ×1.5; it is not a calculated reaction	CONFIRM (structure)
Switch inlet air	≤35 °C	SN5610 manufacturer inlet limit; front-to-rear airflow	FACT

Demand per network cabinet type (per cabinet)

Cabinet type	Qty	Typical kW	Planning kW	Heat kW	Airflow m³/h	Mass kg	PDUs	PSU inlets
Unit cabinet NETA/NETB	30	18.58	23.23	23.29	4,789	572	4	42
Compute spine SPA/SPB	6	36.22	45.27	45.33	9,323	794	4	74
North–south spine CVS	2	14.15	17.68	17.74	3,649	476	4	30
OOB aggregation	2	2.52	3.15	3.18	653	282	2	10

Notes

- Network load counts the switches in network cabinets (power per installed optics) and PDU losses only. GPU-rack OEM SN2201 and GPU-side modules are in the GPU cabinet TDP or are to be confirmed by the OEM.
- Each network cabinet PDU pair A1/A2 and B1/B2 is fed from two different existing UPS paths; supply capacity, protection and spare circuits of the existing system are inputs, and no spare capacity is claimed.
- Heat is room-air heat served by the existing diffuse cooling; fan-wall and chilled-water capability is an input and is not verified by this package. Unused ports and free U are reported, not reserved; no expansion, storage or border capacity is implied.

Interface	Requirement	Status
Cabinet	800 × 1200 × 2294 mm, 48U (46U usable); cabinet top Z3422 on rails top Z1128; empty cabinet 180 kg	DESIGN / CONFIRM
Rails	NVIDIA 930-9SKIT-00L0-00A, posts 600–800 mm (750 used); front post 100 behind door	FACT / DESIGN
Front access	Switch withdrawal 1,000; at fan walls the 1,050 zone is shared in time with fan-wall service 850	DESIGN
Rear access	1,000 zone to CDU front or 1,950 to wall zone; PSU and fan replacement from the rear	DESIGN
GPU cabinet front	Tray withdrawal 1,200 (Y3350–4550); no fan wall faces GPU cabinets	DESIGN
Fibre trays	TL 400×100 Z3550; TA/TB 600×200 Z3900; cable load max 39.48 kg/m + tray 15 kg/m	DESIGN / CONFIRM
Tray supports	Hangers from roof purlins, 1.5 m spacing; seismic bracing and purlin capacity	CONFIRM (structure)
Power branches	Under-floor channels A (Y2750) and B (Y3250); partition penetrations	CONFIRM (facility)
North–south	EXT-IF-NS-A/B: 8 unused (not reserved) 400G per CVS, 7 CVS per side (56 per side, 112 in total) via the 4U border MPO panel	Deploying organisation defines upstream
Management	EXT-IF-OOB: unused 100G ports on OA-A/B-1..4	Deploying organisation defines upstream
Doors	Current door positions not supplied; check against fan walls and front zones	CONFIRM

Network cabinets with the 1,050 mm shared front zone

Fan wall	Network cabinets in front
FW-01	NETB-CU02, NETA-CU03
FW-02	NETB-CU04, NETA-CU05
FW-03	NETB-CU06, NETA-CU07
FW-04	NETB-CU07, OOB-A
FW-05	SPA-3, SPB-1
FW-06	OOB-B, NETA-CU08
FW-07	NETB-CU09, NETA-CU10
FW-08	NETB-CU11, NETA-CU12
FW-09	NETB-CU13, NETA-CU14

K&K Data Service Inc.

kkdatasvc.com/lab/downloads

inquiry@kkdatasvc.com

K&K Data GPU Hall — Network Reference Design

PUBLIC EDITION P2

Reference design for study and evaluation — not for construction

Network equipment interface requirements

Scale NTS

Public Edition P2

2026-10-05

Units: mm

N-951-EN

Sheet 30 of 30

Part III — Implementation Supplement

IS1 Scope, baseline and basis

IS1.1 Scope

This supplement is the implementation part (Part III) of the K&K Data GPU Hall network reference design. It covers the storage network access and capacity model, optical channel components and verification, operating configuration and network acceptance, and the register of clarifications CR-01 to CR-10 with their locations in this edition. It is a public reference document for study and evaluation; it is not an equipment selection, procurement, construction release or service-level commitment, and it contains no executed test results.

The network baseline is the reference design of Parts I and II and is not changed: 90 GPU cabinets (GB300 NVL72, OEM rack envelope 600 × 1200 mm), 40 network cabinets, 120 compute leaves and 54 spines per plane, and 39,066 port connection records. This supplement uses only the baseline summary data (fabric summary, optics budget, failure domains, tray sections, cabinets and network device U positions); it does not expand or recheck the port-level connection table. Supply, ownership, operations and commercial boundaries are outside its scope. Transformer, UPS, branch-circuit, generator, airflow, cooling and structural calculations are also outside its scope; the network states equipment-side demands only.

IS1.2 Status marks

Table IS1-1 Status marks

Mark	Meaning	Use in this supplement
FACT	Fact from a manufacturer source document or a fixed project input	Source number SRC-xx stated; cited only within the product and version to which it applies
DESIGN	Reference design value, or a value calculated from it	Quantities, ports, calculated lengths, patch-panel counts; calculated lengths are not measurements
STUDY	Example or conditional calculated value	Scenario parameters, 168 h test duration, schematic times; not a project conclusion
HOLD	Input not yet available or configuration not selected	Confirmation method and effect stated; no compliance judgement is made for the dependent cut

The deploying organisation defines the final network solution. The conditional options in this supplement are optional configurations with their checking methods; they are not selected. Where an input marked HOLD is not available, the dependent cut is given as a formula and threshold only, without a judgement of compliance or non-compliance. All HOLD inputs are listed in the register of Brief Section 11 (§11).

IS1.3 Principal sources

Technical facts are cited from manufacturer source documents only. Each document is recorded with its URL and access date.

Table IS1-2 Manufacturer source documents (accessed 2026-10-03)

ID	Document	Version / date	URL	Use
SRC-01	NVIDIA MMA4Z00-NS 800Gb/s Twin-port OSFP 2x400Gb/s Multimode 2xSR4 50m (PDF)	Rev 1.6, Feb 2025	docs.nvidia.com/networking/display/mma4z00-ns-800gb-s-twin-port-osfp-2x400gb-s-multimode-2xsr4-50m.pdf	Multimode module reach, connector, polarity, APC/UPC, Tx/Rx specification notes
SRC-02	NVIDIA MMS4X00-NS 800Gbps Twin-port OSFP 2xDR4 100m (web pages)	Rev 1.5, Jul 2025	networking-docs.nvidia.com/800gmms4x00ns/	Single-mode module reach (two patch panels), DDMI temperature, recommended cables
SRC-03	NVIDIA MMA1Z00-NS400-T 400Gb/s QSFP112 Multimode (web pages)	Pages updated 2024-08-14 / 2025-11-18	networking-docs.nvidia.com/mma1z00ns400t/	BF3-side module reach and connector; ordering-number gap
SRC-04	NVIDIA MMA1B00-C100D 100GbE QSFP28 SR4 (web pages)	Pages updated 2025-07-23 / 2026-05-07	networking-docs.nvidia.com/mma1b00c100dspec/	OOB module 100 m OM4, UPC MPO

ID	Document	Version / date	URL	Use
SRC-05	NVIDIA DGX SuperPOD Reference Architecture, DGX GB300 (PDF)	PDF created 2025-11-07	docs.nvidia.com/pdf/dgx-spod-gb300-ra.pdf	Storage structure and Table 5 storage performance guidance (STUDY calibration only)
SRC-06	NVIDIA NCCL User Guide: Environment Variables; Communicators	NCCL 2.32.3	docs.nvidia.com/deeplearning/nccl/user-guide/docs/env.html	NCCL variable defaults, timeouts, communicator error handling
SRC-07	NVIDIA Mission Control North-South Network GB200/GB300 NVL72 Deployment Guide: Planning and Design	2.3.0	docs.nvidia.com/mission-control/docs/ext-north-south-network-gb200-deploy-guide/2.3.0/ext-north-south-net-plan-design.html	North-south EVPN reference, recommended NOS version, address-plan reference

Source gaps: for MMA1Z00-NS400 only the -T page (Ethernet, flat top) was retrieved, so the ordering number applicable to the B3240 is HOLD. None of the four module documents gives a channel insertion-loss budget. Switch, NIC and NOS release notes were not retrieved, and the corresponding configuration items are HOLD. SRC-05 addresses DGX GB300 with an InfiniBand compute fabric; its values are used only as STUDY calibration references.

IS2 Storage network access and capacity model

IS2.1 Existing north–south and out-of-band structure

Each compute tray has one BlueField-3 (B3240); its two 400G ports connect to the A-side and B-side north–south (NS) leaves [DESIGN]. Each side has 30 NS leaves (SN5610), each with $54 \times 400\text{G}$ downlinks and $28 \times 400\text{G}$ uplinks, an oversubscription of $54/28 = 1.929:1$; each NS leaf also has 46 unused ports ($128 - 54 - 28$). Each side has 7 CVS, each using 120 ports with 8 unused, giving 56 unused 400G ports per side, which reach EXT-IF-NS-A/B through the 4U border patch panel in the CVS cabinet.

Table IS2-1 North–south capacity values (per side, per direction)

Item	Expression	Value	Status
NS leaf downlinks	$54 \times 400\text{G}$	21.6 Tb/s per leaf	DESIGN
NS leaf uplinks (C2)	$28 \times 400\text{G}$	11.2 Tb/s per leaf	DESIGN
Oversubscription	$54 / 28$	1.929 : 1	DESIGN
NS aggregate uplinks (C3)	$30 \times 28 \times 400\text{G}$	336 Tb/s	DESIGN
CVS unused ports	7×8	56	DESIGN
CVS unused port capacity	$56 \times 400\text{G}$	22.4 Tb/s	DESIGN (potential, not reserved)
Client port (C1)	$1 \times 400\text{G}$ per tray per side	0.4 Tb/s	DESIGN
Storage system throughput (C5)	—	—	HOLD

The 336 Tb/s and 22.4 Tb/s values are one-direction link capacities. They are not measured storage throughput, and they do not mean that all storage traffic passes the border ports or that all unused ports can be used for storage. The unused ports are candidates for the upstream border, for storage and for other attachments at the same time; any assignment is calculated separately.

The compute fabric ($4 \times \text{CX-8}$ per tray, each 800G split into $2 \times 400\text{G}$ to planes A and B) is physically separate from the NS network. Collective-communication bandwidth (all-reduce and similar) belongs to the compute fabric and is not counted as storage demand; the calculator rejects such traffic as an input. Out-of-band management runs from the two OEM SN2201 in each GPU cabinet ($80 \times 1\text{G}$) over 100G uplinks to the 4 SN4600C on each side and on to EXT-IF-OOB. Each of the 40 network cabinets has one SN2201 (NOOB) with $2 \times 100\text{G}$ uplinks to OA-A-4 and OA-B-4; the 440 OC-5 links include these 80. Management and scheduler-observation traffic does not enter the storage cuts.

IS2.2 Storage access candidate paths

The candidate storage access paths and capacity cuts are shown in Figure IS2-1 (rotated landscape page); the candidates are listed in Table IS2-2.

Table IS2-2 Storage access candidates (none selected)

ID	Path	Capacity cuts	Port source	Gap and conditions	Status
SA-0	BF3 → NS leaf → CVS → EXT-IF-NS (current condition, no storage equipment)	C1, C2, C3	—	Equipment beyond the border defined by the deploying organisation	DESIGN
SA-1	BF3 → NS leaf → CVS → border patch panel → storage beyond the border (edge attachment)	C1 to C5	CVS unused ports k, sharing the 56 per side with the upstream border	$k_{\text{req}} \leq 56 - k_{\text{border}}$; storage system, ports and throughput HOLD	STUDY
SA-2	BF3 → NS leaf → CVS → dedicated storage leaf per side → storage (separate attachment inside the fabric)	C1 to C5, plus the storage-leaf oversubscription (separate HOLD cut)	Storage-leaf uplinks occupy the same CVS unused port pool	C4 threshold as SA-1; SA-2 does not enlarge the $56 - k_{\text{border}}$ port pool, so SA-2 is equally short when it is exceeded; storage-leaf model, quantity and location HOLD	STUDY
SA-3	BF3 → NS leaf → storage on the same leaf, or via CVS on another leaf	C1; C2 and C3 by physical direction (read and write share one physical direction); C4' per hosting leaf	NS leaf unused ports (46 per leaf at most, not taken from the CVS pool)	Location, power and cooling of storage equipment in the unit cabinets not established; uniform placement is a STUDY assumption; failure state depends on replica distribution (HOLD); topology gap	STUDY (gap)

None of the three candidates is selected, and none includes installed storage equipment. SA-1 and SA-2 both consume the 56 unused CVS ports in the current topology: SA-1 connects the border ports directly to the storage system or storage network beyond the border; SA-2 places a dedicated storage leaf on each side, whose uplinks occupy the same port pool, so SA-2 adds no CVS ports for storage. SA-3 uses the unused ports of the NS leaves and does not take ports from the CVS pool; it is a topology-gap option, and the location, power and cooling of storage equipment in the unit cabinets are not yet established. The GB300 SuperPOD reference architecture (SRC-05) uses a dedicated storage leaf attached to the Ethernet spine layer with 400G non-blocking links; this is a fact of that reference architecture, not a selected configuration of this design.

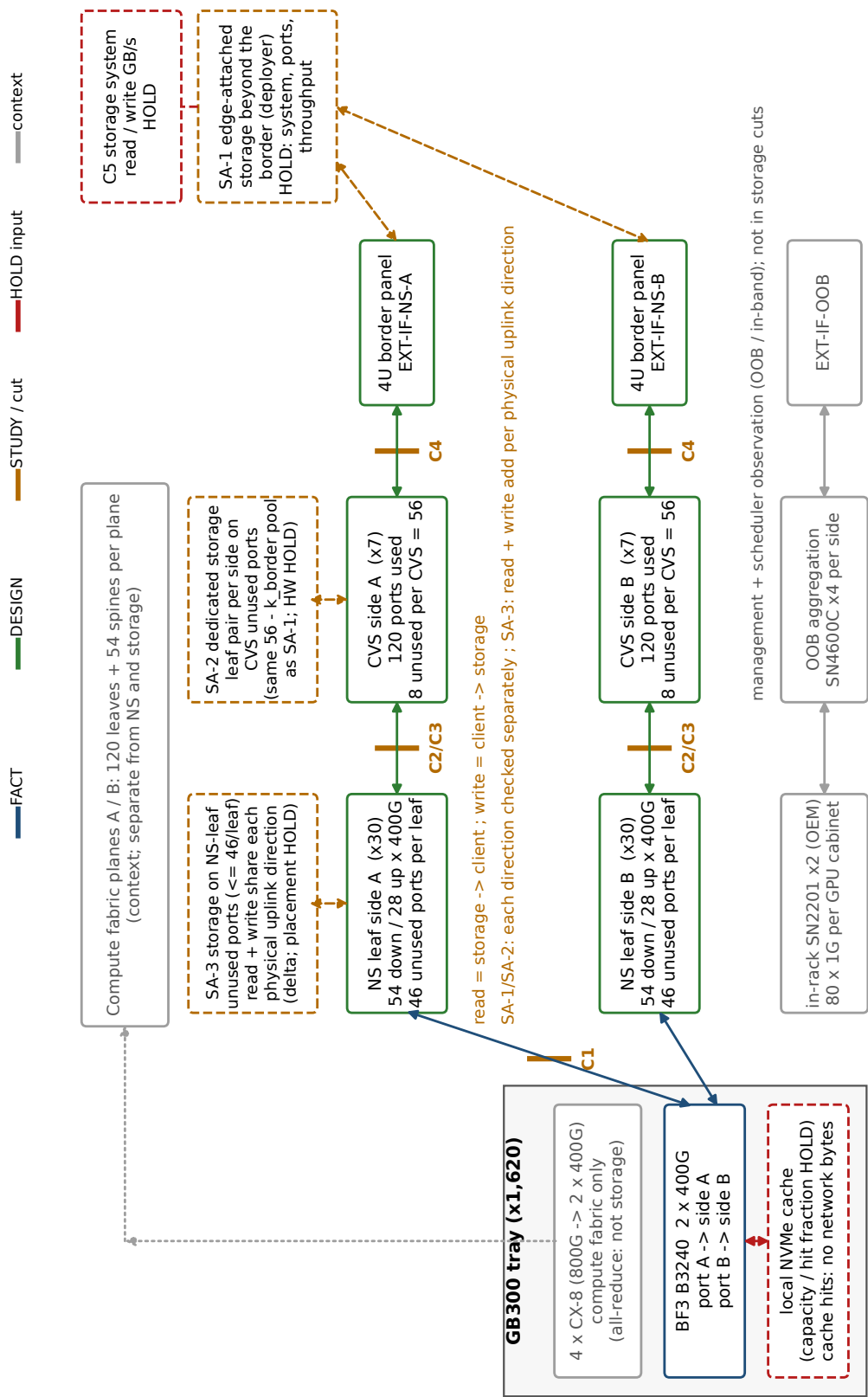


Figure IS2-1 Compute-tray DPU/NS/OOB current condition and the three storage access candidate paths (C1 to C5 are capacity cuts) (figure rotated to landscape)

IS2.3 Traffic classes and directions

Table IS2-3 Traffic classes

Class	Direction	Network path	Enters storage cuts	Model parameters
Dataset distribution / warm-up	Read	Storage → CVS → NS leaf → BF3 → local cache	Yes	U, F, h, T_deadline, concurrency
Local cache read	Within host	Local NVMe → GPU, not over the network	No	Hit fraction h (capacity HOLD)
Training read (cache miss)	Read	As dataset distribution	Yes	Rate or U/T, h
Checkpoint write	Write	BF3 → NS leaf → CVS → storage	Yes	Size, interval, deadline
Recovery read	Read	As dataset distribution; repeated reads of replicas counted in F	Yes	Size, F, deadline
In-band management	Both	Shares the BF3 ports	Reserved through r	r
Scheduler observation / telemetry	Both	Out-of-band or in-band management	Reserved through r, or not entered	—
Collective communication	Both	CX-8 compute fabric (separate planes)	No (input rejected)	—

A read (storage → client) occupies the CVS → leaf direction of the NS leaf uplinks and the switch → host direction at the BF3 port; a write (client → storage) occupies the opposite directions. The two directions of a full-duplex link are calculated separately; read and write are not added against one capacity.

IS2.4 Calculation method and units

Units: 1 TB = 10^{12} bytes; 1 GB/s = 10^9 bytes per second = 8 Gb/s = 0.008 Tb/s; 1 Tb/s = 10^{12} bits per second. The port capacity of 400 Gb/s is the one-direction line rate.

- Bytes transferred over the network: $W = U \times F \times (1 - h)$, where U is the unique data volume (TB), F the non-duplicated repeat-distribution factor (the same data fetched once by each of F clients) and h the local cache hit fraction (reads only).
- Payload rate: $R_{\text{payload}} = 8 \times W \times 10^{12} / T_{\text{deadline}}$ (bits per second), or a given aggregate rate (GB/s); line-rate demand $R_{\text{line}} = R_{\text{payload}} / \eta$, where η is the payload efficiency (protocol and frame overhead, HOLD).
- Side allocation: $R_s = R_{\text{line}} \times \sigma_s$ with $\sigma_A + \sigma_B = 1$, depending on how the storage client uses the two BF3 ports for multipath.
- On shared links (client port, NS leaf uplinks, NS aggregate uplinks) the usable capacity is multiplied by $(1 - r)$, where r is the share kept for in-band management and other traffic.
- Cuts: C1 client port 400G; C2 uplinks of one NS leaf $28 \times 400\text{G} = 11.2$ Tb/s; C3 NS aggregate uplinks per side 336 Tb/s; C4 ports assigned to storage $k \times 400\text{G}$; C5 storage system read/write throughput (HOLD).
- Client and leaf hot spots (C1, C2) are calculated by traffic group. Each flow states the number of trays covered, n, and a client-group label; the load of one client = the group's line rate $\times \sigma_s / n$. Groups do not dilute each other, and neither do read and write: one tray reading 100 GB/s is not averaged down because another 1,620 trays are writing. Flows of the same group are treated as the same set of trays and their loads are added. Where the overlap between different groups is unknown, the upper bound assumes all groups on the same trays (loads added) and the lower bound is the largest single-group load; where the tray counts of two groups add up to more than 1,620, the two groups must share trays and the lower bound is the sum of the two loads (pairwise pigeonhole). Shared trays that follow only from three or more groups are not included in the lower bound, so the lower bound stays low and a case that is in fact exceeded may be judged HOLD (conservative; it is never judged not exceeded in error). Where the groups are declared disjoint, the largest single-group load is used. C2 takes the worst placement of 54 trays on one leaf as the upper bound and the average of the side's demand over 30 leaves as the lower bound.
- Three judgements are used: "not exceeded" when the demand upper bound does not exceed the capacity; "exceeded" when the lower bound exceeds the capacity; HOLD when the capacity lies between the bounds or a required input is missing.
- Port demand and thresholds (C4): $k_{\text{req}} = \text{ceil}(R_s / 0.4 \text{ Tb/s})$. Where $k_{\text{req}} > k_{\text{assigned}}$, more ports must be assigned to storage; where $k_{\text{req}} > 56 - k_{\text{border}}$ (k_{border} being the ports kept for the upstream border), the CVS

unused port pool is short. SA-1 and SA-2 both draw ports from this pool (the SA-2 storage-leaf uplinks also come from these 56 ports), so above this threshold neither satisfies the demand in the current topology, and selecting or extending SA-2 does not remove the shortfall. The required topology gap is storage attachment ports outside this pool, for example SA-3 using the unused NS leaf ports (46 per leaf), or an added device tier whose uplinks are not taken from these 56 ports; neither is selected, and this supplement gives no port quantity for them. The oversubscription of the SA-2 storage leaf itself is a separate cut (model, quantity and downlink count HOLD), not the gap described above. Where the C2 or C3 utilisation exceeds 1, an NS uplink gap exists; where C1 exceeds 1, the host port is the limit. C5 depends on the storage system and is not judged before its throughput data is available.

- SA-3 physical directions: an NS leaf hosting storage sends read data upward (leaf → CVS) to clients on other leaves, and the client leaves receive it downward (CVS → leaf); checkpoint writes run in the opposite direction. Read and write therefore add in the same physical direction of the leaf uplinks, and logical read and write capacities checked separately cannot show that the distributed cut is satisfied. With uniform placement (storage and clients spread over 30 leaves, 1/30 of traffic local to the leaf; STUDY assumption): the aggregate demand per side in each physical direction = $(R + W) \times 29/30$; for one leaf, the upward direction $\leq R/H + \text{writes leaving the leaf's own clients}$, and the downward direction $\leq W/H + \text{reads entering the leaf's own clients}$ (H is the number of hosting leaves; local hits are not deducted, which is conservative); C4' checks R/H and W/H against $k_{\text{leaf}} \times 400\text{G}$ per hosting leaf. Unknown or non-uniform placement is HOLD; the SA-3 failure state depends on the distribution of storage replicas and is also HOLD.
- Input domain: rates and data volumes are finite and non-negative, deadlines finite and positive; σ_A and σ_B each lie between 0 and 1 and sum to 1; η lies in $(0, 1]$ and r in $[0, 1)$; tray counts are integers from 1 to 1,620 with no more than 54 trays per leaf; k_{border} does not exceed 56. Inputs outside the domain are rejected without any judgement. The baseline port assignment may not exceed $56 - k_{\text{border}}$ (SA-3: no more than 46 per leaf and no more than 30 hosting leaves); an assignment above $56 - k_{\text{border}}$ is accepted only as an explicitly marked topology-gap STUDY input, and the result is labelled as an unselected gap.
- Degraded states: when one side is lost, the lost side's demand moves to the surviving side with the transfer factor ρ , and the part not transferred, $(1 - \rho)$, is recorded as demand without a path; in each direction the check "served + no path = line-rate demand" is applied. The surviving side keeps its full capacity and is not halved automatically; ρ depends on the failover ability of the storage client and storage system (HOLD, tested in FT-14). When one CVS is lost, the side's aggregate uplinks fall to 6/7 and each leaf's uplinks to 24/28; the loss of storage ports needs an explicit assumption, otherwise it is HOLD. "Uniform, redirected on the same side" assumes the ports are spread evenly over the 7 CVS (the worst CVS holding $\text{ceil}(k/7)$) and the sessions move to the remaining ports of the same side; under "uniform, not redirected" that share is recorded as having no path. Ports concentrated on the lost CVS would cause a larger loss, which is outside this model. When one NS leaf is lost, only the 54 tray positions of that leaf in total are lost, not 54 for each client group. Because tray placement is unknown, the lost load is given as upper and lower bounds. The upper bound is the worst placement (with unknown group overlap, all groups on the same trays, each counted as $\min(n, 54)$; with declared disjoint groups or a single group, the 54 trays with the highest load). The lower bound is the load that must fall on that leaf, namely the part that the $29 \times 54 = 1,566$ tray positions of the other 29 leaves cannot hold (with unknown overlap, $\max(0, n - 1,566)$ for each group; with disjoint groups, the $\max(0, N - 1,566)$ trays with the lowest load). The failed side is judged on the upper bound of its remaining demand (side demand less the lower bound of the lost load), and the other side on the upper bound of the transferred demand ($\rho \times \text{upper bound of the lost load}$); each is compared with its lower bound to give not exceeded, exceeded or HOLD. Only the affected trays add the ρ share on the other side, and the client-port lower bound includes that share only for groups that are necessarily affected. The demand without a path, $(1 - \rho) \times \text{lost load}$, is also an interval, and demand conservation is checked at both placement extremes. The deadline stretch factor is given as upper and lower bounds; it is not given when any cut is HOLD or when demand without a path may exist. When all 1,620 trays belong to one group, the upper and lower bounds coincide (54/1,620) and the S12 result is unchanged. Where demand without a path exists, no deadline stretch factor is given, and no deadline is implied to be met.

The capacity calculator is covered by regression tests for unit conversion, saturation, direction, cache and repeat distribution, hot clients and client groups, SA-3 physical directions, the input domain, degraded states and demand conservation, the tray-position bounds for a lost NS leaf (including multi-group counter-examples), the per-bar data of Figure IS2-2, and HOLD gating.

IS2.5 Scenario parameters and results

Scenarios S01 to S12 are STUDY example values that show thresholds and sensitivity; they are not project workloads. S01 to S06 scale the aggregate read/write throughput per SU (8 DGX GB300 systems) of SRC-05 Table 5 linearly to 90 cabinets: Standard read 1,012.5 GB/s and write 506.25 GB/s; Enhanced read 3,150 GB/s (25.2 Tb/s) and write 1,575 GB/s (12.6 Tb/s). The linear scaling and its applicability to this design are assumptions. Common STUDY parameters: $\eta = 0.90$, $r = 0.10$, 1,620 concurrent trays, 54 trays on the worst leaf, $k_{\text{border}} = 16$ (ports per side kept for the border), $\sigma_A = \sigma_B = 0.5$. S13, S14 and S14d are hot-client examples; S06 and S06n compare two port assumptions for a lost CVS; S11f shows the HOLD of the SA-3 failure state. P00 is the project scenario, with all demand inputs HOLD.

Table IS2-4 Scenario inputs (STUDY examples; P00 is HOLD; flows without a stated tray count cover 1,620 trays)

Scenario	Description	Flows	Storage ports per side	Failure	Transfer / assumption	Status
S01	SRC-05 Standard scaled to 90 cabinets, concurrent read and write, SA-1, 16 ports per side	Training read 1,012.5 GB/s; checkpoint write 506.2 GB/s	16	normal	—	STUDY
S02	SRC-05 Enhanced scaled, SA-1, 16 ports per side	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	16	normal	—	STUDY
S03	SRC-05 Enhanced scaled, SA-1, 36 ports per side	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	36	normal	—	STUDY
S04	S03 with side A lost; all clients and storage move to side B	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	36	side_lost:A	$\rho = 1$	STUDY
S05	S03 with side A lost; no transfer (clients fixed to a side)	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	36	side_lost:A	$\rho = 0$	STUDY
S06	S03 with one side-A CVS lost; storage ports spread evenly over 7 CVS, sessions redirected to the remaining ports of the same side (assumption)	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	36	cvs_lost:A	Uniform, redirected on the same side	STUDY
S06n	S06, but sessions cannot be redirected to the remaining ports of the same side (assumption)	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	36	cvs_lost:A	Uniform, not redirected	STUDY
S07a	10 TB dataset distributed to the cache of every tray ($F = 1,620$), deadline 4 h	Dataset distribution $U = 10$ TB, $F = 1620$, $T = 14400$ s	16	normal	—	STUDY
S07b	10 TB dataset distributed in shards ($F = 1$), deadline 4 h	Dataset distribution $U = 10$ TB, $T = 14400$ s	16	normal	—	STUDY
S08	Synchronous checkpoint write of 50 TB, deadline 120 s	Checkpoint write $U = 50$ TB, $T = 120$ s	16	normal	—	STUDY
S09	Recovery read of 50 TB, 8 data-parallel replicas reading the same shards, deadline 300 s	Recovery read $U = 50$ TB, $F = 8$, $T = 300$ s	16	normal	—	STUDY
S10	Training read (Enhanced rate, cache hit 0.8) concurrent with the S08 write	Training read 3,150.0 GB/s, $h = 0.8$; checkpoint write $U = 50$ TB, $T = 120$ s	16	normal	—	STUDY
S11	Enhanced scaled, SA-3 distributed on unused NS leaf ports, 2 ports per leaf	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	2 per leaf (30 leaves, uniform)	normal	—	STUDY
S11f	S11 with side A lost (storage replica distribution unknown)	Training read 3,150.0 GB/s; checkpoint write 1,575.0 GB/s	2 per leaf (30 leaves, uniform)	side_lost:A	$\rho = 1$	STUDY
S12	S09 with one side-A NS leaf lost; affected trays move to side B	Recovery read $U = 50$ TB, $F = 8$, $T = 300$ s	16	ns_leaf_lost:A	$\rho = 1$	STUDY
S13	Hot client: 1 tray reads 100 GB/s while 1,620 trays write 1 GB/s in total	Recovery read 100.0 GB/s, 1 tray; checkpoint write 1.0 GB/s	16	normal	—	STUDY
S14	Two single-tray reads of 50 GB/s each, client-group relation unknown (worst case: same tray)	Recovery read 50.0 GB/s, 1 tray; recovery read 50.0 GB/s, 1 tray	16	normal	—	STUDY
S14d	S14 with the two groups declared to be different trays (disjoint)	Recovery read 50.0 GB/s, 1 tray; recovery read 50.0 GB/s, 1 tray	16	normal	Client groups disjoint	STUDY
P00	Project workload (all demand inputs HOLD)	Dataset distribution; training read; checkpoint write; recovery read	HOLD	normal	—	HOLD

Table IS2-5 Scenario results (worst side; C1 is the hottest client port; C5 storage system throughput is HOLD throughout and not judged)

Scenario	Direction	Payload Tb/s	Line rate Tb/s	Side: ports required / assigned	C1 utilisation	C4 utilisation	C2 utilisation	C3 utilisation	No path Tb/s	Threshold result
S01	Read	8.10	9.00	A: 12/16	0.008	0.703	0.015	0.015	0.00	C1 to C4 not exceeded at STUDY values
S01	Write	4.05	4.50	A: 6/16	0.004	0.352	0.007	0.007	0.00	C1 to C4 not exceeded at STUDY values
S02	Read	25.20	28.00	A: 35/16	0.024	2.188	0.046	0.046	0.00	Storage port demand 35 > assigned 16; C4 exceeded → port / topology gap
S02	Write	12.60	14.00	A: 18/16	0.012	1.094	0.023	0.023	0.00	Storage port demand 18 > assigned 16; C4 exceeded → port / topology gap
S03	Read	25.20	28.00	A: 35/36	0.024	0.972	0.046	0.046	0.00	C1 to C4 not exceeded at STUDY values
S03	Write	12.60	14.00	A: 18/36	0.012	0.486	0.023	0.023	0.00	C1 to C4 not exceeded at STUDY values
S04	Read	25.20	28.00	B: 70/36	0.048	1.944	0.093	0.093	0.00	Storage port demand 70 > assigned 36; demand 70 > available CVS unused ports 40 → topology gap outside this port pool needed (not selected); C4 exceeded → port / topology gap
S04	Write	12.60	14.00	B: 35/36	0.024	0.972	0.046	0.046	0.00	C1 to C4 not exceeded at STUDY values
S05	Read	25.20	28.00	B: 35/36	0.024	0.972	0.046	0.046	14.00	14.000 Tb/s line-rate demand without a path ($p = 0$)
S05	Write	12.60	14.00	B: 18/36	0.012	0.486	0.023	0.023	7.00	7.000 Tb/s line-rate demand without a path ($p = 0$)
S06	Read	25.20	28.00	A: 35/30	0.024	1.167	0.054	0.054	0.00	Storage port demand 35 > assigned 30; demand 35 > available CVS unused ports 34 → topology gap outside this port pool needed (not selected); C4 exceeded → port / topology gap
S06	Write	12.60	14.00	A: 18/30	0.012	0.583	0.027	0.027	0.00	C1 to C4 not exceeded at STUDY values
S06n	Read	25.20	28.00	A: 30/30	0.024	0.972	0.054	0.045	2.33	2.333 Tb/s line-rate demand without a path (uniform, not redirected; port share on the lost CVS)
S06n	Write	12.60	14.00	A: 15/30	0.012	0.486	0.027	0.023	1.17	1.167 Tb/s line-rate demand without a path (uniform, not redirected; port share on the lost CVS)
S07a	Read	9.00	10.00	A: 13/16	0.009	0.781	0.017	0.017	0.00	C1 to C4 not exceeded at STUDY values
S07b	Read	0.006	0.006	A: 1/16	0.000	0.000	0.000	0.000	0.00	C1 to C4 not exceeded at STUDY values
S08	Write	3.33	3.70	A: 5/16	0.003	0.289	0.006	0.006	0.00	C1 to C4 not exceeded at STUDY values
S09	Read	10.67	11.85	A: 15/16	0.010	0.926	0.020	0.020	0.00	C1 to C4 not exceeded at STUDY values
S10	Read	5.04	5.60	A: 7/16	0.005	0.437	0.009	0.009	0.00	C1 to C4 not exceeded at STUDY values
S10	Write	3.33	3.70	A: 5/16	0.003	0.289	0.006	0.006	0.00	C1 to C4 not exceeded at STUDY values
S11	Read	25.20	28.00	A: 2/2 (per leaf)	0.024	0.583	0.069	0.067	0.00	C1 to C4 not exceeded at STUDY values; C2/C3 are the physical directions shared by read and write (Table IS2-6)
S11	Write	12.60	14.00	A: 1/2 (per leaf)	0.012	0.292	0.069	0.067	0.00	C1 to C4 not exceeded at STUDY values; C2/C3 are the physical directions shared by read and write (Table IS2-6)
S11f	—	—	—	—	—	—	—	—	—	Not evaluated: storage replica distribution for SA-3 under failure is HOLD
S12	Read	10.67	11.85	B: 16/16	0.020	0.957	0.039	0.020	0.00	C1 to C4 not exceeded at STUDY values
S13	Read	0.80	0.89	A: 2/16	1.235	0.069	0.044	0.001	0.00	C1 hottest-client utilisation 1.235, exceeded
S13	Write	0.008	0.009	A: 1/16	0.000	0.001	0.000	0.000	0.00	C1 to C4 not exceeded at STUDY values
S14	Read	0.80	0.89	A: 2/16	1.235 (envelope, HOLD)	0.069	0.044	0.001	0.00	C1 hottest-client utilisation 1.235: envelope exceeded, lower bound not exceeded (client-group overlap unknown, HOLD)
S14d	Read	0.80	0.89	A: 2/16	0.617	0.069	0.044	0.001	0.00	C1 to C4 not exceeded at STUDY values
P00	—	—	—	—	—	—	—	—	—	Not evaluated: dataset.unique_tb, dataset.deadline_s, dataset.clients, train.unique_tb, train.deadline_s, train.clients are HOLD

Table IS2-6 SA-3 NS uplink cuts by physical direction (uniform placement is a STUDY assumption; read and write add in the same physical direction)

Scenario	Side	Logical read / write Tb/s	Physical cut	Demand Tb/s	Capacity Tb/s	Utilisation	Judgement
S11	A	14.00 / 7.00	Aggregate leaf → CVS	20.30	302.40	0.067	Not exceeded
S11	A		Aggregate CVS → leaf	20.30	302.40	0.067	Not exceeded
S11	A		One leaf, upward direction: hosted reads out + writes from the leaf's own clients	0.700	10.08	0.069	Not exceeded
S11	A		One leaf, downward direction: hosted writes in + reads to the leaf's own clients	0.700	10.08	0.069	Not exceeded
S11	B	14.00 / 7.00	Aggregate leaf → CVS	20.30	302.40	0.067	Not exceeded
S11	B		Aggregate CVS → leaf	20.30	302.40	0.067	Not exceeded
S11	B		One leaf, upward direction: hosted reads out + writes from the leaf's own clients	0.700	10.08	0.069	Not exceeded
S11	B		One leaf, downward direction: hosted writes in + reads to the leaf's own clients	0.700	10.08	0.069	Not exceeded

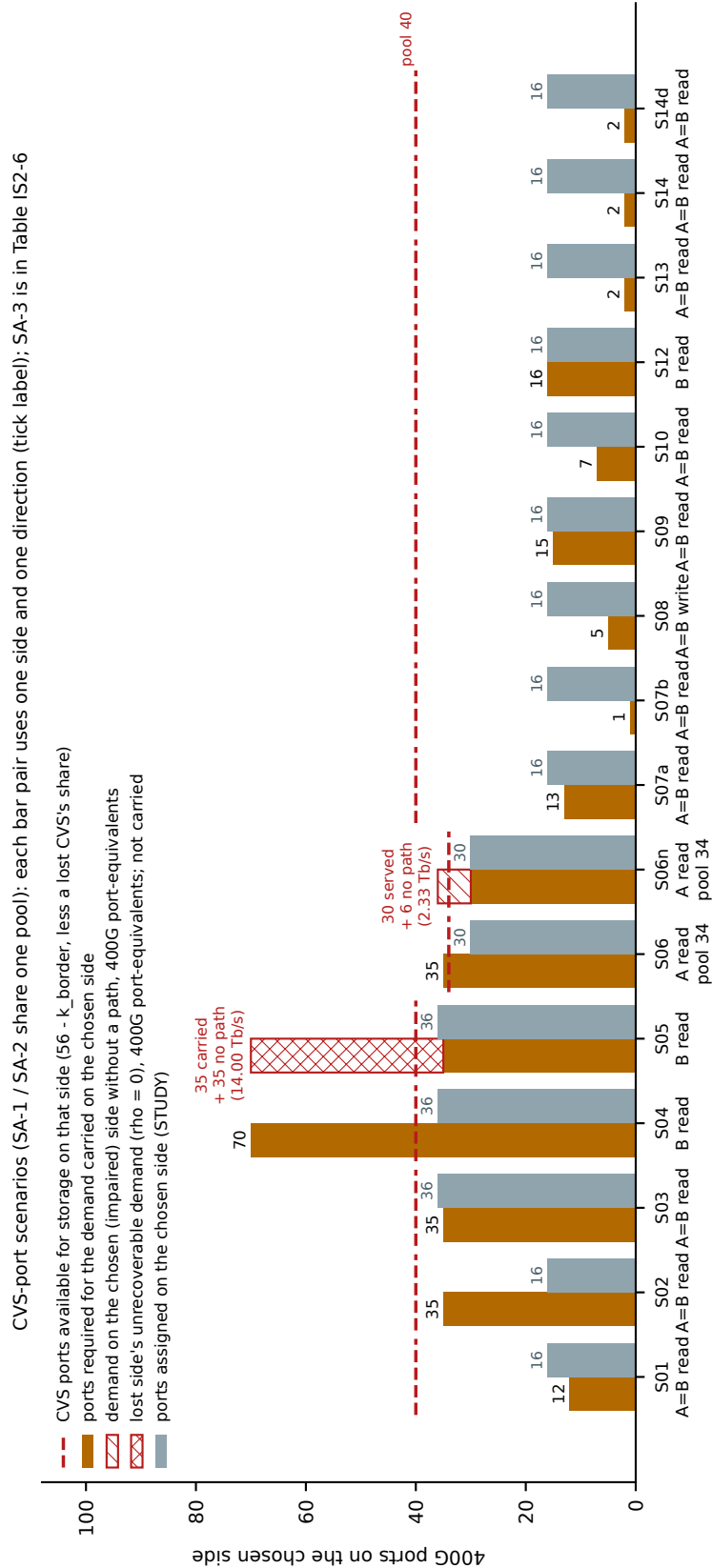


Figure IS2-2 CVS-port attachment scenarios (SA-1/SA-2): each bar pair uses one side and one direction (tick label) for ports required, no-path port-equivalents, ports assigned and the port pool available on that side (side-selection rule in IS2.6; SA-3 in Table IS2-6) (figure rotated to landscape)

IS2.6 Engineering meaning of the scenario results

- In the scenarios with clients spread over 1,620 trays (S01 to S12), the C1 to C3 utilisations are all below 0.1: the read demand scaled from SRC-05 Enhanced is about 15.6 Gb/s of payload per tray, far below one 400G port. The limits lie in the storage attachment ports (C4) and the storage system (C5). This conclusion does not apply where demand is concentrated on a few trays (see S13, S14).
- The difference between S02 and S03 comes from the port assignment: the Enhanced read line rate of 14.0 Tb/s per side needs 35 ports; with 16 ports the utilisation is 2.19, with 36 ports it is 0.97, which is close to saturation. The C4 capacity is the number of assigned ports $\times$ 400G and is not multiplied by $(1 - r)$, so the remaining headroom depends on the actual value of η and on the ports assigned, not on r .
- S04: with one side lost and full transfer, the surviving side needs 70 ports, more than the 40 ports left after the border deduction and more than the 56 unused ports per side. SA-1 and SA-2 share this port pool and neither can satisfy the demand; what is needed is a topology gap outside this port pool (not selected). In S05, without transfer, 14.0 Tb/s of read and 7.0 Tb/s of write demand have no path; the surviving-side utilisation covers only the served part and does not mean that service or deadlines are met. Together they show that the single-side capacity conclusion depends on the client transfer ability and cannot be replaced by halving the capacity.
- S06 and S06n: with one CVS lost and ports spread evenly, side A keeps 30 storage ports. If sessions can be redirected to the remaining ports of the same side, the read utilisation is 1.17, and after the border deduction the ports available on that side fall to 34, fewer than the 35 required. If they cannot be redirected, the share of the lost ports (read 2.33 Tb/s, write 1.17 Tb/s) has no path. Side-selection rule of Figure IS2-2: for the normal state and NS leaf loss, the side and direction with the highest C4 utilisation; for CVS loss, the impaired side; for side loss, the surviving side. Ports required, ports assigned, no-path port-equivalents and the port pool are all taken from that side and direction, and the demand of the healthy side is not added to the no-path demand of the failed side. In the figure, S06 and S06n are drawn for the read direction of the impaired side A: S06 requires 35 and has 30 assigned; for S06n the served part requires 30 ports, a further 6 port-equivalents have no path, and 30 are assigned; the available port pool is 34 for both. S05 is drawn for the read direction of the surviving side B: it carries a demand of 35 ports, and the unrecoverable 14.0 Tb/s of side A (35 port-equivalents) is shown separately with cross-hatching and is not counted as side-B demand. Both CVS-loss results depend on the HOLD port distribution and redirection ability.
- S07a and S07b: for the same 10 TB dataset, if every tray fetches its own copy ($F = 1,620$), 16,200 TB cross the network and 9.0 Tb/s of payload is needed within 4 hours; if the dataset is sharded and each tray fetches a different part ($F = 1$), only 0.006 Tb/s is needed. The distribution method has a larger effect on network demand than the link rate.
- S09: a 50 TB checkpoint read once by each of 8 data-parallel replicas needs 10.67 Tb/s of payload within 300 s; with 16 ports the utilisation is 0.93. S12 adds the loss of one NS leaf: side B then needs 16 ports at a utilisation of 0.96 (the 1,620 trays form one group, so the lost load is determinate). With several groups, only the 54 tray positions of one leaf are deducted. Example: two groups of 100 trays each, each reading 1,250 GB/s ($\eta = 1$, $r = 0$, $k_border = 0$, 12/40 ports on sides A/B, $\sigma = 0.5$, $\rho = 1$), with one side-A leaf lost. With the groups declared disjoint, side A keeps 7.3 to 10.0 Tb/s, all above 4.8 Tb/s, and the result is exceeded. With unknown overlap, side A keeps 4.6 to 10.0 Tb/s, spanning 4.8 Tb/s, and the result is HOLD, with no stretch factor given. This example is a calculator regression test, not a project scenario.
- S11 (SA-3, 2 storage ports per hosting leaf, uniform placement): the read utilisation per hosting leaf is 0.58. The leaf uplinks are calculated by physical direction: the aggregate per side carries $(14.0 + 7.0) \times 29/30 = 20.3$ Tb/s in each direction at a utilisation of 0.067, and one leaf carries about 0.70 Tb/s in each direction at a utilisation of 0.069 (Table IS2-6). The cost is that storage equipment must enter the unit cabinets, a gap condition that is not established; S11f shows that the SA-3 failure state is not judged until the replica distribution is known.
- S13: one tray reads at 100 GB/s while 1,620 trays write 1 GB/s in total. The hot tray needs a line rate of 0.444 Tb/s per side port against 0.36 Tb/s of usable capacity, a utilisation of 1.23, which is exceeded; the write traffic does not dilute this read demand. S14: two single-tray read groups of 50 GB/s each; if the two groups may be the same tray, the upper-bound utilisation is 1.23 and the lower bound 0.62, so the result is HOLD. S14d declares the two groups to be different trays, giving 0.62, not exceeded.

A project judgement requires the same calculator to be rerun once the HOLD inputs are available: dataset size and distribution method, local cache capacity and hit fraction, concurrent jobs and tray counts and the tray ranges of client groups, checkpoint size / interval / deadline, recovery-read method, storage system read/write throughput, client multipath and transfer factor, storage and border port assignment, port distribution and redirection ability when a CVS is lost, and SA-3 storage placement and replica distribution (U-01 to U-06, U-15; §11).

IS3 Optical channel components, specification and verification

IS3.1 Module facts

Table IS3-1 Optical module facts (NVIDIA product documents)

Ordering number	Form factor	Physical layer	Connector	Reach statement	Use (DESIGN)	Source	Status
MMA4Z00-NS	OSFP twin-port, finned top	2 x 400G SR4, 100G-PAM4/lane, 850 nm	2 x MPO-12/APC, 8 of 12 fibres used	50 m OM4 / 30 m OM3; 50 m assumes two patch panels	SN5610 multimode cages	SRC-01	FACT
MMA4Z00-NS-FLT	OSFP twin-port, flat top	as MMA4Z00-NS	2 x MPO-12/APC	50 m OM4 (as MMA4Z00-NS)	CX-8 NIC side (GB300 tray)	SRC-01	FACT (document lists the -FLT flat-top variant)
MMS4X00-NS	OSFP twin-port, finned top	2 x 400G DR4, 1311 nm, 100G-PAM4/lane	2 x MPO-12/APC, OS2	100 m; assumes two patch panels	SN5610 cages whose longer half-port exceeds 45 m	SRC-02	FACT
MMA1Z00-NS400 (doc retrieved: MMA1Z00-NS400-T)	QSFP112 single-port	400G SR4, 850 nm	MPO-12/APC	50 m OM4 / 30 m OM3; assumes two patch panels	BF3 B3240 side	SRC-03	-T page is FACT; ordering number applicable to the B3240 HOLD
MMA1B00-C100D	QSFP28	100GbE SR4, 850 nm, 25.78125 Gb/s/lane	MPO, un-angled (UPC)	100 m OM4 / 70 m OM3 at 100 Gb/s	OOB 100G uplinks (SN2201 to SN4600C)	SRC-04	FACT

The module optical ports carry alignment pins, so the cable connectors that mate with a module are unpinned (female); pinned connectors do not mate with a module and are used only to join two cable segments (SRC-01/02). Each intermediate mated pair therefore has one pinned and one unpinned side, with compatible key and end face. NVIDIA connects two modules directly with a Type-B crossover cable; where a channel includes patch panels, the whole channel keeps the end-to-end Tx → Rx crossover of every lane. APC and UPC connectors in green housings are not compatible (SRC-01/02). The MMA1B00-C100D has an un-angled UPC MPO (SRC-04), so in OC-5 the cable ends that mate with a module are UPC. The current OC-5 is a direct connection with no intermediate mated pair, and both ends of the cord mate with modules, so both ends are UPC. If an intermediate mated pair is added later, the end faces on the two sides of each mating face must match; UPC throughout OC-5 is the design choice of this supplement, while SRC-04 itself fixes only the module end face. The multimode reach statement of 50 m and the single-mode statement of 100 m both assume two patch panels in the link, but these reach statements are not a channel insertion-loss allowance.

IS3.2 Channel components and counting boundaries

The components, counting boundaries and reference planes of the five typical channel classes are shown in Figure IS3-1 (rotated landscape page); the class parameters are listed in Table IS3-2.

Table IS3-2 Channel classes (DESIGN, from the N-901-EN / N-902-EN summaries; 31,280 optical links in total)

Class	Name	End A	End B	Components	Patch panels / M	Calculated length m	Links	End face
OC-1	NIC → compute leaf	MMA4Z00-NS-FLT (CX-8)	MMA4Z00-NS (SN5610)	MPO-12/APC jumper, female–female, direct (TL tray)	0 / 0	5.20–10.07	12,960	APC
OC-2	DPU → NS leaf	MMA1Z00-NS400 (BF3)	MMA4Z00-NS (SN5610)	MPO-12/APC jumper, female–female, direct (TL tray)	0 / 0	6.09–10.34	3,240	APC
OC-3	Leaf → spine / NS leaf → CVS, multimode	MMA4Z00-NS	MMA4Z00-NS	Jumper → 7U unit patch panel → trunk, fan-out legs plugged directly into the spine/CSV modules	1 / 1	≤ 44.35	OC-3 + OC-4 together 14,640	APC
OC-4	Leaf → spine / NS leaf → CVS, single-mode	MMS4X00-NS	MMS4X00-NS	Jumper → 7U unit patch panel → trunk, fan-out legs plugged directly into the spine/CSV modules	1 / 1	43.95–61.40	OC-3 + OC-4 together 14,640	APC
OC-5	OOB 100G uplink	MMA1B00-C100D (SN2201)	MMA1B00-C100D (SN4600C)	MPO-12 jumper / trunk fan-out leg, no patch panel	0 / 0	≤ 63.95	440	Module-facing ends UPC (module FACT)

Counting boundaries: E is the interface between a module optical port and a cable connector; M is a mated pair of two connectors inside the channel. One patch panel is not necessarily one mated pair: a pass-through adapter is 1

pair, a cassette with connectors on both faces is 2 pairs, and splices are counted separately. The loss of the end interfaces E1/E2 is counted once: if the manufacturer includes it in the transmitter/receiver specification, it is not counted in the channel allowance; if the manufacturer's allowance includes the end interfaces, it is counted in the allowance. The assignment is to be confirmed in writing by the module and cable manufacturers (U-07). The per-strand pin and key order remains HOLD until the actual cable assemblies are selected (U-09). Lengths are calculated routed lengths of the reference model, not field measurements.

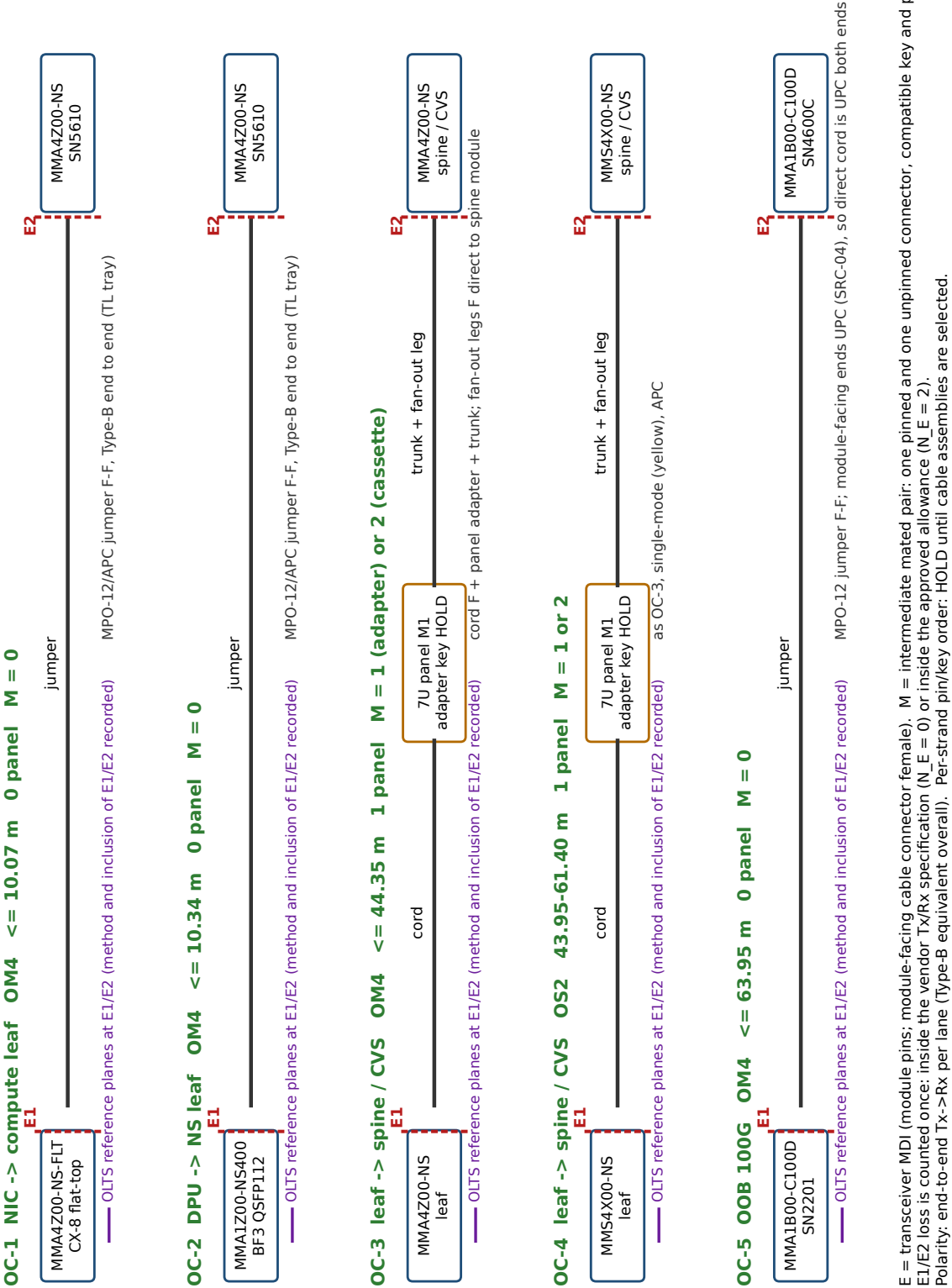


Figure IS3-1 Components of the five typical channel classes: end interfaces E1/E2, intermediate mated pairs M, patch panels, polarity / key / gender fields and OLTS reference planes (figure rotated to landscape)

IS3.3 Insertion-loss calculation method

Worst-case component loss: $L = \alpha \times \ell + (N_M + N_E) \times IL_M + N_S \times IL_S$, where α is the fibre attenuation (dB/km), ℓ the channel length (km), N_M the number of intermediate mated pairs, N_E the number of end interfaces counted in the allowance (0 or 2), and IL_M and IL_S the maximum mated-pair and splice losses of the selected products. Margin: $M = A_{ch} - L - U$, where A_{ch} is the channel allowance approved by the manufacturer and U the measurement or estimation uncertainty. Neither A_{ch} nor U is available, so no margin is calculated. From $M \geq 0$, the maximum permitted loss per interface is, for $N_M + N_E > 0$: $IL_M \leq (A_{ch} - U - \alpha \times \ell - N_S \times IL_S) / (N_M + N_E)$. Only with $N_E = 0$ and $N_S = 0$ does this simplify to $(A_{ch} - U - \alpha \times \ell) / N_M$. For $N_M + N_E = 0$ (direct jumper with the end interfaces inside the Tx/Rx specification; OC-1, OC-2 and OC-5 belong to this case on that basis), the allowance does not constrain IL_M , and mating quality is controlled by the end-face criteria and the Q1 inspection. A negative right-hand side means that the fibre and splices alone already exceed the allowance. All quantities in the formula come from the selected products and test method; A_{ch} and U are HOLD (U-07, U-08), and this supplement substitutes no values.

The channel allowance is not derived as average launch power minus receiver sensitivity: these are separate specification limits of the transmitter and the receiver (SRC-01 states that the minimum average launch power is informative only, and the receiver sensitivity is defined under stated TDECQ/SECQ conditions), and they do not form a link budget.

The table below gives the component loss of the longest channel of each class using prior reference estimates (OM4 3.0 dB/km, OS2 0.5 dB/km, 0.35 dB per mated pair), to show the differences caused by the counting basis and the patch-panel type; these coefficients are not maximum specifications of selected products.

Table IS3-3 Reference-estimate component loss (shows counting differences only; A_{ch} and U not available, margin not calculated)

Class	End-interface basis	Patch-panel type	Length m	N_M	N_E	Reference loss dB	A_ch	Margin
OC-1	End interfaces in Tx/Rx specification	No patch panel	10.07	0	0	0.030	HOLD	Not calculated
OC-1	End interfaces in channel allowance	No patch panel	10.07	0	2	0.730	HOLD	Not calculated
OC-2	End interfaces in Tx/Rx specification	No patch panel	10.34	0	0	0.031	HOLD	Not calculated
OC-2	End interfaces in channel allowance	No patch panel	10.34	0	2	0.731	HOLD	Not calculated
OC-3	End interfaces in Tx/Rx specification	Adapter	44.35	1	0	0.483	HOLD	Not calculated
OC-3	End interfaces in Tx/Rx specification	Cassette (+1 pair)	44.35	2	0	0.833	HOLD	Not calculated
OC-3	End interfaces in channel allowance	Adapter	44.35	1	2	1.183	HOLD	Not calculated
OC-3	End interfaces in channel allowance	Cassette (+1 pair)	44.35	2	2	1.533	HOLD	Not calculated
OC-4	End interfaces in Tx/Rx specification	Adapter	61.40	1	0	0.381	HOLD	Not calculated
OC-4	End interfaces in Tx/Rx specification	Cassette (+1 pair)	61.40	2	0	0.731	HOLD	Not calculated
OC-4	End interfaces in channel allowance	Adapter	61.40	1	2	1.081	HOLD	Not calculated
OC-4	End interfaces in channel allowance	Cassette (+1 pair)	61.40	2	2	1.431	HOLD	Not calculated
OC-5	End interfaces in Tx/Rx specification	No patch panel	63.95	0	0	0.192	HOLD	Not calculated
OC-5	End interfaces in channel allowance	No patch panel	63.95	0	2	0.892	HOLD	Not calculated

IS3.4 Product confirmation matrix

Table IS3-4 Product confirmation matrix (ordering numbers not supported by a document are not listed)

Component	Required attributes	Known fact / design value	Confirmation method	Evidence	Status
Switch-side multimode module	OSFP finned top, 2 × 400G SR4, MPO-12/APC × 2	MMA4Z00-NS (SRC-01, Rev 1.6)	Order confirmation and firmware version readback	Order list, readback record	FACT (model)
NIC-side multimode module	OSFP flat top, fits the CX-8 cage	MMA4Z00-NS-FLT (SRC-01 lists the flat-top variant)	OEM approval for the GB300 tray cage	OEM written confirmation	HOLD
BF3-side module	QSFP112 400G SR4, MPO-12/APC	Page retrieved for MMA1Z00-NS400-T (SRC-03)	OEM/NVIDIA confirm the ordering number and shell applicable to the B3240	Written confirmation	HOLD
Single-mode module	OSFP 2 × 400G DR4, OS2, MPO-12/APC	MMS4X00-NS 100 m (SRC-02, Rev 1.5)	Order confirmation	Order list	FACT (model)
OOB module	QSFP28 100G SR4, MPO UPC	MMA1B00-C100D (SRC-04)	Order confirmation	Order list	FACT (model)
Multimode jumpers OC-1/OC-2	OM4, MPO-12/APC female–female, Type-B, length steps by calculated length	NVIDIA recommends the MFP7E10 series, 3 to 50 m (SRC-01/03; candidate, not selected)	Manufacturer data sheet: maximum insertion loss, end-face geometry, polarity diagram	Data sheet, factory test report	HOLD
Single-mode OC-4 jumpers	OS2, MPO-12/APC, Type-B	NVIDIA recommends the MFP7E30 series, 1 to 100 m (SRC-02; candidate)	As above	As above	HOLD
Multimode trunk OC-3	18 × MPO-12 legs / 216 fibres, OD 13.0 mm envelope (DESIGN)	Product not selected	Manufacturer data sheet: fibre count, outer diameter, bend radius, leg length, insertion loss	Data sheet, factory test report	HOLD
Single-mode trunk OC-4	As above, OS2	Product not selected	As above	As above	HOLD
7U unit patch panel	72 ports per U (DESIGN, N-411-EN), APC adapters	Adapter or cassette not decided	Confirm the adapter key orientation (key-up/key-up or key-up/key-down) and the number of mated pairs	Data sheet, polarity diagram	HOLD
4U border patch panel	CVS cabinet, border / storage candidate	Port count follows the assignment of the deploying organisation	Deploying organisation confirms port use and end face	Interface table	HOLD
OOB jumper / trunk fan-out leg OC-5	OM4, MPO-12, female–female; module-facing ends UPC	Module end face UPC (SRC-04); a direct jumper is therefore UPC at both ends; UPC throughout is the design choice	Data sheet confirms end face, gender and polarity	Data sheet	HOLD
End-face inspection and cleaning tools	Suitable for MPO APC and UPC	—	End-face criteria required by the manufacturer	Inspection images and records	HOLD
OLTS instrument and reference cords	Multi-fibre MPO, 850 nm and 1310 nm, APC/UPC reference cords	—	Instrument uncertainty U and reference method	Calibration certificate, method record	HOLD

IS3.5 Sample coverage

Table IS3-5 Verification sample coverage rules

ID	Coverage criterion	Selection
CV-1	Longest calculated channel of each class	OC-1 10.07 m, OC-2 10.34 m, OC-3 44.35 m, OC-4 61.40 m, OC-5 63.95 m; then the next-longest channels in turn until both end units (CU-01, CU-15) are represented
CV-2	Most connector interfaces	Every OC-3/OC-4 patch-panel type actually supplied (adapter and any cassette used); the leg positions at both ends of the 18-leg trunks
CV-3	Worst measured loss	After installation, the N channels with the highest OLTS result per class and per wavelength (N is HOLD and follows the acceptance basis); all results within the uncertainty of the limit are included
CV-4	Both media and all end types	OM4 and OS2; OSFP finned top (switch), OSFP flat top (CX-8), QSFP112 (BF3), QSFP28 (OOB)
CV-5	Both half-ports of a twin-port cage	P/1 and P/2 of the same cage, including NIC cages whose two half-ports go to NETA and NETB respectively
CV-6	Module operating conditions	Module case temperature at the cage inlet with the highest measured temperature; adjacent cages fully populated; every firmware/version bundle; every module manufacturing-site code appearing in the delivery

IS3.6 Test stages and records

Table IS3-6 Test stages

Stage	Name	Scope	Records	Acceptance basis
Q1	Passive channel: compatibility, polarity, end face, OLTs	All installed channels (earlier design-stage requirement retained); end-face inspection before every mating	End-face type (APC/UPC, matching on both sides of each mating face), gender (module pinned, module-facing cable end unpinned; each intermediate mated pair one pinned and one unpinned side), key orientation (compatible key on each mated pair), end-to-end Tx → Rx correspondence per lane, end-face images and results, OLTs loss per fibre and per wavelength, reference method, launch/receive reference cords, instrument and uncertainty	Channel allowance and end-face criteria approved by the selected cable/module manufacturers (HOLD)
Q2	Physical layer and FEC with modules installed	Sampled per CV-1 to CV-6; all links read once at turn-up	Link state, speed, lane count, FEC mode at both ends, per-lane transmit and receive optical power, bias current, case temperature, voltage, pre-FEC bit error ratio or corrected codewords/symbols, uncorrectable codewords, link-down count	Module/switch/NIC manufacturer thresholds for the installed firmware version (HOLD); FEC correction within the manufacturer limits is normal operation
Q3	Sustained traffic	Samples per CV-1 to CV-6 run at line rate; STUDY sample duration 168 h	Time-stamped counter snapshots at fixed intervals: corrected/uncorrectable FEC, FCS/symbol errors, drops by cause, PFC sent/received per priority, ECN marks, CNP sent/received, link flaps, module resets, DDM trends	Rates normalised per hour and per bit; thresholds from the manufacturer and the acceptance basis of the deploying organisation (HOLD); 168 h is a STUDY test duration and does not demonstrate reliability or service life
Q4	Optical component faults and recovery	Module reseal, jumper disconnect/reconnect and administrative cage shutdown on designated samples only	Detection time, link re-establishment time, FEC/bit errors after reseal, cleaning records	FT test rows; manufacturer reseal guidance (HOLD)

Table IS3-7 Counter classification

Category	Counters	Distinction	Action on
Congestion	Buffer/queue drops, PFC pauses, ECN marks, CNP	No physical-layer counter growth on the same port in the same period	Traffic classes, buffer and congestion-control configuration
Physical signal	Uncorrectable codewords, FCS/symbol errors, link flaps, DDM outside the manufacturer range	Independent of load; may coexist with congestion	Channel (clean/reseat/replace), module, cage
FEC correction (normal)	Corrected codewords, pre-FEC bit error ratio	Normal on PAM4 links; trend and manufacturer limits apply	Monitoring; action on manufacturer thresholds or step changes

FEC correction counts are a normal operating phenomenon on PAM4 links and are not an automatic failure condition; thresholds and acceptance values are taken from the manufacturer documents of the installed version (U-14). 168 h is a recommended STUDY test duration, not an industry requirement, and it does not demonstrate service-life reliability or 99.999 % availability. The records separate congestion (buffer drops, PFC, ECN, CNP) from physical signal errors (uncorrectable codewords, FCS/symbol errors, link flaps); the two categories are counted separately for the same period.

IS4 Operating configuration and network acceptance

IS4.1 Configuration matrix

The configuration matrix is bound to vendor and version. All version combinations form one version bundle, and performance and fault test results are valid for that version bundle only. The "conditional options" in the table are not selected values; for unknown items the confirmation method and acceptance check are stated. The NOS is not yet decided, so no command lines are given, only configuration items and readback requirements.

Table IS4-1 Operating configuration matrix (conditional options not selected)

ID	Configuration item	Dependency	Conditional options and factual basis	Confirmation method	Acceptance check	Status
CFG-01	Version: CX-8 NIC firmware, driver and RDMA user-space libraries	Compatible combination released by the GB300 tray OEM	OEM-released combination, or NVIDIA network software stack release combination (one of the two, locked as a whole)	Version numbers and PSID in the written compatibility matrices of the OEM and NVIDIA	Readback of all trays matches the version bundle list item by item	HOLD
CFG-02	Version: BF3 B3240 firmware and operating mode (NIC/DPU mode, host network services)	North-south routing scheme; storage client	NIC mode; DPU mode with host-side network services	NVIDIA BF3 and Mission Control release notes	Mode readback; link and address readback on both side ports	HOLD
CFG-03	Version: switch NOS and version (SN5610 / SN4600C / SN2201)	Network operations practice of the deploying organisation	Cumulus Linux (SRC-07 recommends v5.14.0 for GB200/GB300 north-south; a fact of that guide only), or a NOS selected by the deploying organisation	After selection, vendor release notes and known-issue list	Version readback; zero difference between running configuration and template, or each difference registered	HOLD
CFG-04	Port: cage breakout and speed: 800G OSFP cage 2 × 400G; QSFP112 400G; QSFP28 100G	Reference port assignment	Breakout fixed by the design; configuration syntax follows NOS and NIC versions	Port-breakout support statement of the selected version	Readback of speed and lane count (4 × 100G-PAM4) for every half-port	DESIGN (syntax HOLD)
CFG-05	Port: FEC mode (same at both ends)	Module, switch and NIC versions	FEC mode specified by the vendors for 400G SR4 / DR4 and 100G SR4	Module and equipment vendor documents	FEC readback identical at both ends; Q2 counters readable	HOLD
CFG-06	Port: MTU: host IP MTU, RoCE path MTU and switch Ethernet frame size specified separately	If NS uses VXLAN, the underlay frame size must accommodate the encapsulation	Host and switch values set separately; switch frame size ≥ host IP MTU + layer-2 header (+ encapsulation)	MTU definitions in the NIC, storage and NOS documents	End-to-end maximum-size probe with DF set; RoCE active MTU readback; no fragmentation and no oversize-drop counts	HOLD
CFG-07	Routing: compute fabric L3 leaf-spine: routing protocol, ECMP hash fields, adaptive routing or packet spraying	NIC congestion control and out-of-order receive capability	Static ECMP (RoCEv2 UDP source port in the hash); vendor adaptive-routing scheme. SRC-06: NCCL_IB_ADAPTIVE_ROUTING is off by default on RoCE networks	Switch and NIC vendor version requirements for the selected scheme	Readback of 54 equal-cost uplinks per leaf; 53 after one is disconnected	HOLD
CFG-08	Routing: failure detection and convergence: link detection, protocol keepalive/BFD interval, route convergence	CFG-03, CFG-07	Values HOLD; checked in one table with the NCCL timeout (CFG-13)	Value ranges supported by the NOS	Detection and convergence times measured in the FT tests	HOLD
CFG-09	Routing: north-south: EVPN multihoming VXLAN (SRC-07 reference for GB200/GB300) or the scheme of the deploying organisation; border peering	Address plan; external routers	EVPN multihoming VXLAN; L3 scheme of the deploying organisation	Network design document of the deploying organisation	Both sides reachable from every BF3; border routing table readback	HOLD
CFG-10	QoS: traffic-class mapping table: compute RoCE, CNP, NCCL control, storage RoCE, TCP management, OOB → DSCP / priority / PFC / ECN / queue	Congestion-control scheme; storage vendor requirements	Filled in per the template of Table IS4-2; no external example values adopted	Mapping statements of the NIC, switch and storage vendors and the policy of the deploying organisation	Mapping readback identical at both ends; marking checked by packet capture	HOLD
CFG-11	QoS: PFC priority, headroom, ECN thresholds, CNP priority	CFG-10; calculated cable lengths (longest 61.40 m)	Values per the vendor headroom calculation method	Switch vendor buffer calculation document	PT-08 and Q3 counters	HOLD
CFG-12	QoS: NIC congestion-control algorithm and parameters	CFG-10, CFG-11	Vendor default; vendor Ethernet AI scheme	Vendor documents	PT-05, PT-08	HOLD

ID	Configuration item	Dependency	Conditional options and factual basis	Confirmation method	Acceptance check	Status
CFG-13	NCCL: NCCL version and network plugin; NCCL_IB_HCA, GID, TC, QP, timeout, CROSS_NIC, TOPO_FILE	CFG-10 (TC value derived from the selected DSCP), CFG-14	SRC-06 facts: GID_INDEX default AUTO (from 2.32u1); QPS_PER_CONNECTION default 1; TIMEOUT 20 × RETRY_CNT 7 about 30 s; CROSS_NIC default 2; ROCE_VERSION_NUM default 2	NCCL_DEBUG=INFO log records the actual devices, GID, TC and plugin	Log consistent with the template; timeout compared with the FT convergence times in one table	HOLD
CFG-14	Host: host-side presentation of the CX-8 800G after breakout: two independent RDMA devices or one merged device	OEM and NVIDIA drivers	Independent devices (each listed in NCCL_IB_HCA); merged device (NCCL_IB_MERGE_NICS behaviour applies)	OEM/NVIDIA documents and readback on the actual system	Device list readback; PT-04 with plane A, plane B and both	HOLD
CFG-15	Storage: storage client: RDMA, BF3 two-side multipath and failover method	Storage system selection	Both sides active; one side active with the other standby; fixed per side	Storage vendor client documents	Transfer behaviour measured in FT-14, consistent with the redirect parameter of the storage model	HOLD
CFG-16	Monitoring: counter set, sampling period, clock synchronisation accuracy	CFG-03; monitoring platform	Streaming telemetry; polling; period HOLD	Platform documents and measured clock synchronisation	Measured record of sampling period Δ and clock offset ε	HOLD
CFG-17	Out-of-band: SN2201 / SN4600C configuration, VLAN/VRF, address plan, external management interface	Management network of the deploying organisation	CIDR and VRF partitioning of SRC-07 used as reference only	Management network design of the deploying organisation	Every BMC and management port reachable; EXT-IF-OOB routing readback	HOLD
CFG-18	Control: version bundle and change control: all versions form one version bundle ID, and acceptance results are bound to the version bundle	All CFG rows	Design rule	—	After a change, the affected PT/FT rows are repeated and recorded	DESIGN

IS4.2 Traffic-class mapping template

The table below is a template to be filled in. DSCP, priority, PFC/ECN enablement, NCCL TC and similar values are all HOLD, and no example values from external articles are adopted. NCCL_IB_TC and NCCL_IB_FIFO_TC are converted from the selected marking values per the NIC vendor definition; NCCL_IB_GID_INDEX defaults to AUTO from NCCL 2.32u1 (SRC-06).

Table IS4-2 Traffic-class mapping template (to be filled in)

Traffic class	Endpoints	DSCP	Switch priority	PFC	ECN	Queue / buffer	NIC TC	NCCL variable	Source and version	Readback result
Compute RoCE data	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—
CNP	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—
NCCL control messages	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—
Storage RoCE	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—
TCP management and user	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—
OOB management	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	HOLD	—

IS4.3 Performance tests

Performance tests are run per version bundle, separately for plane A, plane B and both planes together, and combined with actual storage concurrency and representative collective communication. The acceptance basis is the manufacturer reference result for the same version bundle or the target value of the deploying organisation (HOLD).

PT-03 covers compute-fabric paths between CU-01 and CU-15. The single-mode cages it includes belong to the compute leaf → spine channel class, whose longest calculated length is 61.08 m (N-901-EN, TB); this value is the longest calculated length of the channel class, not a measured end-to-end path length from CU-01 to CU-15. The 61.40 m channel is an NS leaf → CVS channel (N-902-EN, NS-A-LG29 → CS-A-1); it is covered by CV-1 in Q1/Q2 and is noted when PT-06 records the NS channels used.

PT-06 and PT-08 are run with the actual workload of the deploying organisation (U-01, U-02), and the warm-up, checkpoint and recovery deadlines are carried into record fields: for each workload type, the data volume (TB), start and completion times, completion time and effective throughput (data volume ÷ completion time) are recorded and

compared with the corresponding deadline; the comparison rule follows the acceptance basis of U-12. Deadlines, data volumes and acceptance basis are HOLD, and this supplement makes no judgement of compliance.

Table IS4-3 Performance tests

ID	Test	Conditions	Records	Acceptance basis
PT-01	Point-to-point on one leaf	Two trays on the same rail in the same leaf group, plane A and plane B measured separately	RDMA one-way/bidirectional bandwidth, latency distribution, counters	HOLD
PT-02	Across leaf groups	Same rail in different units, via the spine	As above, plus ECMP path distribution	HOLD
PT-03	Far-end path	Between CU-01 and CU-15, including compute-fabric single-mode cages (longest calculated length of this class 61.08 m, N-901-EN); this is the longest calculated length of the compute-fabric channel class, not a measured end-to-end path from CU-01 to CU-15; the 61.40 m channel is an NS channel (NS leaf → CVS), covered by CV-1/Q2 and recorded in PT-06	As above, including Q2 optical parameters	HOLD
PT-04	Plane A / B / A+B	Two 400G half-ports per NIC: A only, B only, A+B together	Total bandwidth, balance between planes, host-side device selection	HOLD
PT-05	Collective communication	all-reduce, all-gather, reduce-scatter, all-to-all; scale: 1 leaf group, 4 units, all 1,620 trays; message-size sweep HOLD	Algorithm and bus bandwidth, tail latency, congestion counters	HOLD
PT-06	Storage concurrency	Read/write concurrency per the storage scenario table, run together with PT-05; dataset, checkpoint and recovery read per the actual workload of U-01/U-02	NS and compute fabric counted separately; storage throughput; job step time. For each workload type (warm-up/dataset distribution, checkpoint write, recovery read): dataset or checkpoint data volume (TB), start and completion times, completion time, effective throughput (data volume ÷ completion time), corresponding deadline (U-01/U-02, HOLD) and the comparison result per the acceptance basis of U-12; the NS channels used include the longest calculated channel of 61.40 m (NS-A-LG29 → CS-A-1)	HOLD
PT-07	Management and out-of-band	Management traffic and OOB reachability maintained throughout PT-05/PT-06	Probe loss and latency	HOLD
PT-08	Aggregation congestion	Many-to-one (restart fan-in) onto one leaf downlink; many-to-one onto a storage port group; recovery read per the actual workload of U-02	PFC, ECN, CNP, drop causes; recovery-read data volume (TB), start and completion times, completion time, effective throughput, recovery deadline (U-02, HOLD) and the comparison result per the acceptance basis of U-12	HOLD

IS4.4 Fault and recovery tests

Fault injection preferably uses logical isolation (administrative shutdown of ports or devices); trays are not cut and cabinet power is not disconnected. Physical unplugging is performed only on designated samples, with cleaning and end-face records. The affected-object counts are taken from the reference failure-domain table (N-801-EN) for comparison with the measurements. Each test records network path recovery, service throughput recovery and training-job recovery separately; one does not substitute for another. The NCCL default timeout `NCCL_IB_TIMEOUT = 20` with `NCCL_IB_RETRY_CNT = 7` amounts to about 30 s (SRC-06); route convergence times are compared in one table with the timeout settings in use to judge whether a link fault causes a job error. FT-05 isolates only the leaf downlinks of compute plane A and does not involve the NS side. FT-06 (both host-side CX-8 half-ports in one cage) and FT-15 (switch-side OSFP twin-port cage) are two different common causes: the first removes one NIC from both planes at once, the second removes one plane from each of two NICs.

FT-16 is a representative multi-link logical-isolation case that maps the N-801-EN event of tray TA cut at X30000 (CU-05): the reference failure-domain table gives 2,160 NICs with one plane unreachable and 540 BF3 trays affected on one side for that row; TB at the same position gives 1,728 / 432, for comparison only. The event affects all trunk links passing the cut point at once and is not the same event as FT-02 (single link), FT-05 (whole-plane

downlinks) or FT-07 (CU-08 TL section). The injection method is logical isolation of the affected links at the switch-side ports, without unplugging cords or touching the tray; the isolation set is derived from the reference connection table by cut position at test preparation and is HOLD. The affected-object counts are compared with 2,160 / 540, and the acceptance basis remains HOLD with the U-11 version bundle and the U-12 acceptance rules. This case is not a physical event and does not represent an executed test. Out-of-band aggregation faults are covered by FT-11.

Table IS4-4 Fault and recovery tests (not executed; affected objects from the reference failure-domain table)

ID	Event	Injection method	Affected objects	Impact category
FT-01	One NIC → leaf link	Administrative port shutdown at the switch side; jumper may be unplugged on designated samples	1 NIC unreachable on one plane	Path
FT-02	One leaf → spine link	Administrative port shutdown	Leaf uplinks 53/54 (0.9815)	Path / throughput
FT-03	One compute leaf	Isolate all ports or shut down per the vendor method	54 NICs unreachable on one plane	Path
FT-04	One compute spine	Isolate all ports	53/54 uplinks on every leaf	Throughput
FT-05	All leaf downlinks of compute plane A (one compute-fabric plane)	Logical isolation of the downlink ports of all 120 compute leaves of plane A, without power-off; the NS side is outside this test (single NS device faults in FT-08 to FT-10; effect of losing one NS side on storage in scenarios S04/S05 and FT-14)	6,480 NICs unreachable on one plane (compute-fabric part of the "entire plane A" row of the reference failure-domain table)	Path / throughput
FT-06	Shared host-side NIC/cage (both CX-8 half-ports in one cage)	Shut down the switch ports of both half-ports of that NIC at the same time	1 NIC unreachable on both planes (endpoint loss)	Endpoint
FT-07	TL route section CU-08	Logical isolation of all CU-08 TL jumpers at the switch-side ports, without touching the tray	432 NICs unreachable on both planes; 108 BF3 trays without a path	Endpoint / path
FT-08	One NS leaf	Isolate all ports	54 BF3 trays on one side	Path
FT-09	One CVS	Isolate all ports	NS uplinks of that side 6/7; storage ports on that CVS	Throughput
FT-10	One network cabinet NETA-CU08	Logical isolation of the 11 switches in the cabinet	432 NICs on one plane; 108 BF3 trays on one side	Path
FT-11	One OOB aggregation switch	Isolate all ports	Connected SN2201 move to the other side	Management
FT-12	Planned maintenance	One leaf and one spine in turn drained → upgraded → restored	As FT-03 / FT-04, planned	Maintenance
FT-13	Version rollback	Roll back to the previous version bundle	Affected devices	Maintenance
FT-14	Storage port group on one side	Logical isolation of the storage port group	Corresponding storage scenarios S04/S05	Storage
FT-15	Switch-side twin-port cage (one leaf OSFP cage/module)	Shut down both half-ports P/1 and P/2 of that leaf cage at the same time; module may be removed on designated samples	2 NICs each lose one plane and keep the other (reference failure-domain table, row CL-A-LG01-R1:P1)	Path
FT-16	Multiple links in a local TA route section (representative mapping: N-801-EN, TA cut at X30000)	Logical isolation of the affected links at the switch-side ports, without unplugging cords or touching the tray; isolation set derived from the reference connection table by cut position at test preparation (HOLD)	2,160 NICs unreachable on one plane; 540 BF3 trays on one side (reference failure-domain table, row TA@X30000); comparison: TB at the same position 1,728 / 432	Path / throughput

Table IS4-5 Evidence fields for every fault test

No.	Field
1	Version bundle ID
2	Preconditions and baseline (traffic, jobs, counter snapshot)
3	Injection time t ₀ (synchronised clock)
4	Detection time t _d
5	Route convergence time t _r
6	Throughput recovery time t _s (reaching baseline fraction x; x is HOLD)

No.	Field
7	Job result: continued / continued after stall / error / restarted
8	Job recovery time t_j
9	NCCL timeout settings and logs
10	Affected-object counts compared with the reference failure-domain table
11	Recovery actions and times
12	Hashes of counter and telemetry files

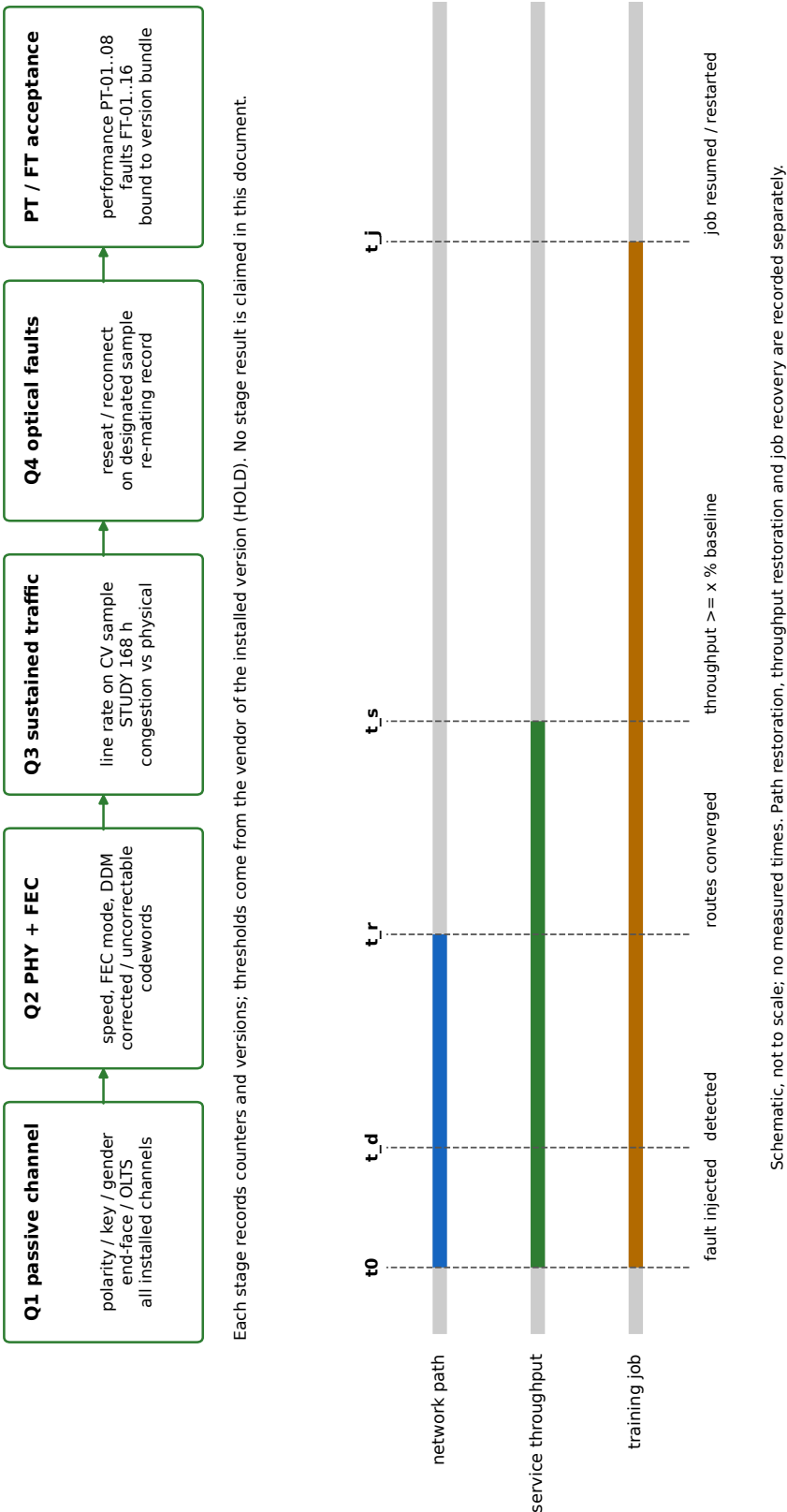


Figure IS4-1 Optical channel verification stages and fault-recovery timing: path, throughput and job measured separately (schematic, not measured) (figure rotated to landscape)

IS4.5 Monthly availability measurement

The 99.999 % calendar-month availability is a design target, not a measured achievement, and it is not derived from equipment MTBF or from a one-week soak test. The measurement proposed below is offered for adoption; the measurement object and accounting rules finally used are decided by the deploying organisation (HOLD, U-13).

- **Measurement object:** the primary indicator is the "compute fabric service unavailable" time, that is, periods in which the reference probe throughput is below the baseline fraction x , or in which any in-service endpoint has lost its paths in all planes (x is HOLD); the weighted values of M1 endpoint-plane path-unavailable seconds and M2 endpoint-unavailable seconds are reported as well.
- **Overlap:** overlapping or adjoining events within one indicator are merged as a union and are not added event by event.
- **Resolution and uncertainty:** probe period Δ and synchronisation error ϵ of the event time sources; the duration uncertainty of each merged interval is $\pm(\Delta + 2\epsilon)$. The monthly lower bound is the sum over merged intervals of $\max(0, d_i - \Delta - 2\epsilon)$; the upper bound is the sum over intervals of $(d_i + \Delta + 2\epsilon)$, not exceeding the length of the month; the tolerance is not applied only once to the whole month (see Table IS4-9). If the measured upper bound does not exceed the budget, the target is met at that resolution; if the lower bound exceeds the budget, it is not met; otherwise it cannot be determined at that resolution.
- **Missed detection:** interruptions shorter than Δ may not be detected. The measurement covers detected periods only and does not estimate missed events; "met" means only that the upper bound of the detected periods is within the budget, not a complete proof of compliance. Whether missed detections are checked with an independent probe of shorter period is decided by the deploying organisation (HOLD).
- **Common faults** are not excluded by definition in order to meet the target (single NIC/cage, TL section, planned maintenance and similar are all counted; whether planned maintenance is reported separately follows the rules of the deploying organisation, and results are reported both including and excluding it).

Table IS4-6 Monthly budget for 99.999 %

Days in month	Seconds in month	Permitted unavailable time s
28	2,419,200	24.192
29	2,505,600	25.056
30	2,592,000	25.920
31	2,678,400	26.784

Table IS4-7 Effect of resolution on the measurement (Δ and ϵ are STUDY examples)

Probe period Δ s	Clock error ϵ s	Uncertainty per merged interval s	Share of 30-day budget (per interval)
0.01	0.001	± 0.012	0.05%
0.1	0.01	± 0.12	0.46%
1	0.05	± 1.1	4.24%
5	0.5	± 6	23.15%

Table IS4-8 Measurement example (illustrative data, not measured)

Item	Value
Events (example)	E1 100–118 s; E2 110–124 s; E3 3,000–3,002.5 s
Sum event by event (not used)	34.5 s
Union	26.5 s (2 merged intervals)
Interval at $\Delta = 1$ s, $\epsilon = 0.05$ s	24.3 to 28.7 s
30-day budget	25.920 s
Judgement	Cannot be determined at this resolution (lower bound below budget, upper bound above budget)

Table IS4-9 Multi-interval measurement example: uncertainty accumulated per interval (illustrative data, not measured)	
Item	Value
Events (example)	10 non-overlapping interruptions of 2.0 s
Union	20.0 s (10 merged intervals)
Uncertainty per merged interval	± 1.10 s ($\Delta = 1$ s, $\epsilon = 0.05$ s)
Monthly interval (sum of per-interval bounds)	9.0 to 31.0 s
30-day budget	25.920 s
Judgement	Cannot be determined at this resolution
Comparison: $\pm(\Delta + 2\epsilon)$ applied once to the whole month (not used)	18.9 to 21.1 s, which would be misjudged as met

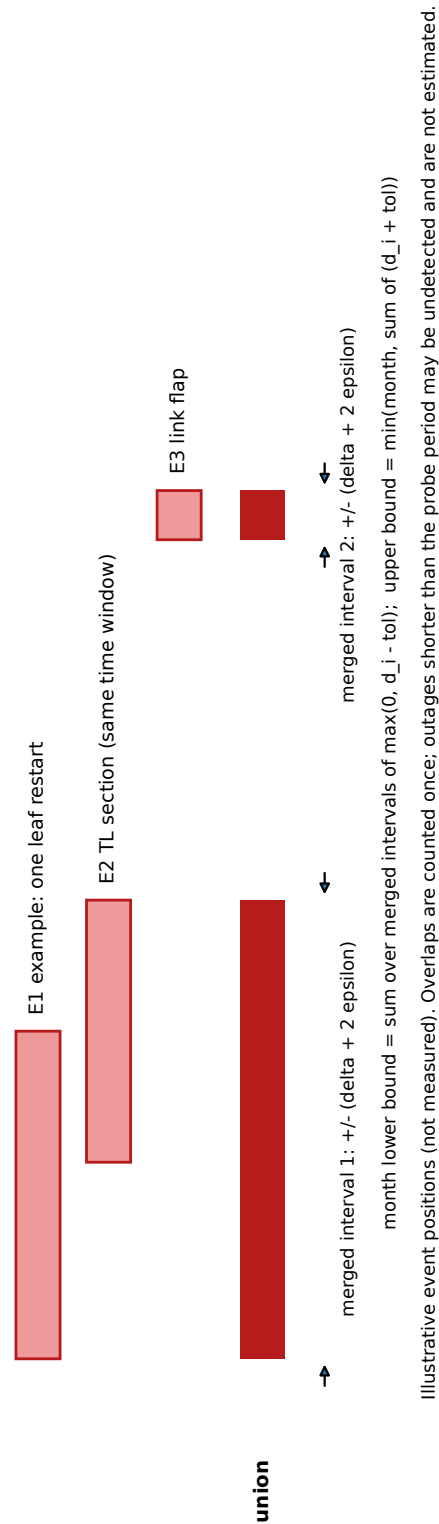


Figure IS4-2 Overlapping events merged as a union; resolution uncertainty applied to each merged interval, with monthly bounds summed per interval (illustrative event positions, not measured) (figure rotated to landscape)

IS5 Clarification register

The clarifications below are applied in the brief and drawing book of this edition. The table lists, for each item, the previous location and wording (as carried by Public Edition P1), the issue, the replacement text with its basis, and the locations in this edition; the previous and replacement texts are given in English so that they correspond directly. The CR-04 locations include N-402-EN, and the N-951-EN count basis is aligned to 56 per side, 112 in total; the CR-10 locations include Brief §6, N-403-EN and N-902-EN. CR-01 only clarifies the basis of the equal-sharing support-point load and adds no structural verification condition; for CR-02 the previous values are correct and only the line-to-line and line-to-neutral voltages are written out; CR-05 is an actual connector end-face compatibility correction. CR-10 adopts only the MMA1Z00-NS400 family name; the exact ordering number and shell variant remain HOLD. The existing common causes are unchanged, and no new physical design is selected.

Table IS5-1 Clarification register and locations in Public Edition P2

ID / topic	Previous location and wording	Issue	Replacement text and basis	Status and P2 location
CR-01 Support-point load statement	Brief §6 table "Heaviest cabinet"; N-951-EN; N-411-EN / N-412-EN column "Foot kN (×1.5)" Previous wording: 794 kg; 1.95 kN per foot (2.92 kN with ×1.5)	The value is the equally shared gravity load $794 \text{ kg} \times 9.81 \text{ m/s}^2 \div 4$; the equal-sharing assumption was not stated, so it could be read as the worst support reaction or as a floor or rail capacity. No load per square foot appears in the previous text; any derived document that states a load per square foot is changed to per support point as well.	794 kg cabinet mass (SPA-1). With the weight shared equally by four support points, $794 \text{ kg} \times 9.81 \text{ m/s}^2 \div 4 \approx 1.95 \text{ kN}$ per support point (static gravity load). This is not the worst unequal support reaction and not a rail, joist or floor capacity. The 2.92 kN value applies an illustrative factor of 1.5 for unequal sharing; it is not a calculated reaction. Support reactions and capacity: structural engineer (CONFIRM). Basis: arithmetic: $794 \times 9.80665 / 4 = 1,946.6 \text{ N}$	Applied in P2 Brief §6 table row Heaviest cabinet and the following paragraph; N-411-EN, N-412-EN column Support kN* and table note; N-951-EN row Heaviest cabinet
CR-02 Voltage notation	Brief §2.4, §6 table "Rack PDUs"; N-951-EN Previous wording: 3-phase 400Y/230 V	400Y/230 V is the wye three-phase four-wire notation: 400 V is the line-to-line voltage and 230 V the line-to-neutral voltage; the previous text did not write this out and could be misread.	3-phase 4-wire wye 400Y/230 V (400 V line-to-line, 230 V line-to-neutral); outlet voltage per the selected PDU product (CONFIRM). Basis: definition of the notation; network equipment PSU input range 200–240 V is an existing FACT	Applied in P2 Brief §2.4 equipment table, §6 row Rack PDUs; N-951-EN row Rack PDUs
CR-03 Calculated and measured lengths	Brief §5.1, §5.3; N-301-EN, N-303-EN, N-901-EN, N-902-EN Previous wording: Selected channel lengths are 5.20–10.07 m ... Longest channel 44.35 m / 61.40 m / 63.95 m	The lengths are calculated routed lengths of the reference layout model (DESIGN), not post-installation measurements; they must not be read as field FACT.	Channel lengths are calculated routed lengths from the reference layout model (DESIGN), including the design allowances of the reference model; they are not installed measurements. Installed lengths and OLTS losses are recorded per channel at installation and replace the calculated values in the as-built record. Basis: source of the reference-model link lengths	Applied in P2 Brief §5.1, §5.3; N-301-EN, N-303-EN table headers and notes; N-901-EN header Calculated m and note 4; N-902-EN title, table header and note 5
CR-04 Unused ports and expansion capacity	Brief §2.1, §2.2, §9; N-101-EN, N-111-EN, N-951-EN; N-411-EN/N-412-EN "4U border/storage MPO panel"; N-402-EN BP label and note 2; N-951-EN north–south row count basis Previous wording: 20 spare ports per leaf; 8 spare per spine; 56 spare 400G ports per side; 4U border/storage MPO panel	"Spare" is easily understood as reserved expansion or storage capacity; these ports are unused and not reserved, and any use for border, storage or other attachment requires its own capacity and port calculation.	unused (not reserved) ports: 20 per compute leaf, 8 per spine, 46 per NS leaf, 8 per CVS (56 per side, 112 in total). They are not dedicated expansion, storage or border capacity; any assignment is a configuration item of the deploying organisation checked with the storage/border capacity model (supplement Section IS2). The CVS-cabinet 4U MPO panel is the border panel; storage use is a candidate only. N-951-EN: EXT-IF-NS-A/B: 8 unused (not reserved) 400G per CVS, 7 CVS per side (56 per side, 112 in total). Basis: reference fabric summary; supplement storage model	Applied in P2 Brief §2.1, §2.2, §2.3, §9; N-101-EN, N-111-EN (56 per side, 112 in total), N-121-EN, N-402-EN labels and notes 2/3, N-411-EN/N-412-EN equipment and table notes, N-951-EN North–south row and note 3

ID / topic	Previous location and wording	Issue	Replacement text and basis	Status and P2 location
CR-05 OOB 100G connector end face	Brief §5.2 cable counting units; N-503-EN; N-901-EN class "OOB 100G uplink" Previous wording: jumper MPO-12 APC ... OOB 100G uplink: OM4 MPO-12 jumper / trunk leg	The MMA1B00-C100D optical port is an un-angled UPC MPO (SRC-04 FACT); APC and UPC in green housings are not compatible (SRC-01/SRC-02). The previous counting unit was written uniformly as jumper MPO-12 APC without separating the OOB channel, which conflicts with the OOB module end face. SRC-04 fixes only the module end face; the cable end that mates with a module must be UPC and unpinned. The current OC-5 is a direct connection without an intermediate mated pair and both ends mate with modules, so both jumper ends are UPC. If an intermediate mated pair is added later, each pair needs complementary pin/hole gender and compatible end face and key; UPC throughout is the design choice for this channel, not a general requirement of SRC-04.	OOB 100G channels connect MMA1B00-C100D modules (MPO UPC per NVIDIA product page). Every module-facing cord end is an MPO-12 UPC female connector; in the current direct assembly (no intermediate mated pair) both ends of each cord are therefore UPC. Any intermediate mated pair added later needs complementary pin/hole gender and compatible end-face polish and key on its two faces; UPC throughout is the design choice for these channels. Counting unit: MPO-12 UPC, OM4 (product CONFIRM). Basis: SRC-04; SRC-01/02	Applied in P2 (product condition CONFIRM) Brief §2.3, §5.2; N-121-EN description and note 4; N-403-EN interface table; N-503-EN counting-unit table and note; N-901-EN OOB media row and note 5
CR-06 Patch panels and mated-pair counting	Brief §5.3 "at most one patch panel per channel"; N-902-EN note 1 Previous wording: All 31,280 optical links are within the reach stated for the selected module, with at most one patch panel per channel.	One patch panel is not necessarily one mated pair: a pass-through adapter is 1 pair, a cassette with connectors on both faces is 2 pairs; the manufacturer's "two patch panels" reach statement is not a channel insertion-loss allowance.	All 31,280 optical links are within the reach stated for the selected module; channels have at most one patch panel. Insertion loss is counted by mated pairs: a pass-through adapter panel contributes one mated pair, a cassette with connectors on both faces contributes two. The manufacturer's reach statement with two patch panels does not define a channel loss allowance. Basis: reach statements of SRC-01/02/03; supplement Section IS3	Applied in P2 Brief §5.3; N-902-EN note 1
CR-07 Reference insertion-loss coefficients	Brief §5.3 last paragraph; N-902-EN header note Previous wording: The insertion-loss estimates in the link table (OM4 3.0 dB/km, OS2 0.5 dB/km, MPO 0.35 dB per mated pair) are reference values only.	It must be clear that these are not maximum specifications of selected products and cannot be used to derive a channel allowance.	... are prior reference estimates only. They are not maximum specifications of selected products and do not define a channel allowance; the approved allowance, component maxima and test limits come from the selected module and cabling vendors (CONFIRM). Basis: SRC-01 to SRC-04 give no channel insertion-loss budget	Applied in P2 Brief §5.3 last paragraph; N-902-EN note 6
CR-08 45 m design changeover length	Brief §5.3 table; N-902-EN Previous wording: 400G/800G multimode design threshold 45 m	45 m is the design rule for changing from multimode to single-mode cages; it is neither a manufacturer limit nor a measured value.	45 m is the design changeover length between multimode and single-mode cages (manufacturer reach 50 m on OM4, stated with two patch panels); it is a design rule, not a manufacturer limit or a measured value. Basis: SRC-01	Applied in P2 Brief §5.3 table and paragraph; N-902-EN 45 m line in the chart, module table and note 2

ID / topic	Previous location and wording	Issue	Replacement text and basis	Status and P2 location
CR-09 Availability measurement	Brief §8; N-801-EN note 1; N-000-EN note 5 Previous wording: The monthly availability figure of 99.999 % is a design target ...	The target statement is kept; a reference to the measurement rule is added so that availability is not inferred from equipment MTBF or a short soak test.	(append) Monthly availability is measured per the supplement measurement rule (Section IS4.5): union of overlapping impaired intervals, stated measurement resolution and clock uncertainty, and no exclusion of common-cause events by definition. It is not derived from equipment MTBF or from soak-test duration. Basis: supplement Section IS4	Applied in P2 Brief §8; N-000-EN note 5; N-801-EN note 1
CR-10 BF3-side 400G module ordering number	Brief §5.3; N-901-EN; reference optics budget (data unchanged); Brief §6, N-403-EN, N-902-EN Previous wording: MMA1Z00-NS400, 400G SR4 QSFP112, 50 m	The NVIDIA page retrieved under SRC-03 is MMA1Z00-NS400-T; the exact ordering number and shell variant applicable to the B3240 cage remain HOLD pending written OEM/NVIDIA confirmation.	BF3 side: 400G SR4 QSFP112 multimode module of the MMA1Z00-NS400 family (exact OPN and shell variant for the B3240 cage CONFIRM with OEM/NVIDIA). Basis: SRC-03	Applied in P2 (ordering number HOLD, U-10) Brief §5.3 table and paragraph, §6; N-403-EN interface table and notes 2/3; N-901-EN module column and note 6; N-902-EN chart labels, module table and note 7

Common causes retained:

- Both half-ports of one CX-8 share one cage: a NIC or cage failure removes that NIC from both planes at once (endpoint loss).
- The TL jumper tray carries the NIC and DPU jumpers of both planes: when the whole CU-08 section is lost, 432 NICs are unreachable on both planes and 108 BF3 trays have no path.
- In 27 of the 40 network cabinets both UPS paths belong to the same generator segment: loss of a generator segment after battery depletion is a risk case.

IS6 Register of pending inputs

The register of engineering inputs pending confirmation, U-01 to U-15, with the sections affected, the source of information and the confirmation method of each item, is consolidated together with the field-verification items FV-01 to FV-12 in Brief Section 11 (§11). Section references in that register beginning with "IS" refer to this supplement.