UpStation

Design **UPS-DSN-20261103-BKCU**
Prepared 2026-11-03 · prices as of 2026-11-03
Prepared for **Example SaaS Co** · —

Budgetary design and estimate — not a quote

4× 8-GPU RTX PRO 6000 inference farm (32 GPUs), Texas colo

4× 8x NVIDIA RTX PRO 6000 Blackwell Server Edition (32 GPUs, PCIe class). Networks: front end 25G on 2 switches; oob 1G on 1 switch. 1 rack at ≤ 40 kW, 34.33 kW provisioned, in Texas colo. Every figure is computed from the ai.upstation.io catalog with the arithmetic shown below.

$724,604.56	**$7,724.00/mo**	**$817,292.56**	**$1,002,668.56**	**$1,188,044.56**
hardware (CapEx)	colocation, 34.33 kW	year one	3-year total	5-year total

Delivery: 5 weeks, set by 8x NVIDIA RTX PRO 6000 Blackwell Server Edition, Intel — SYS-422GA-NRT. **Colo:** Texas at $225/kW/month × 34.33 kW, 36-month term.

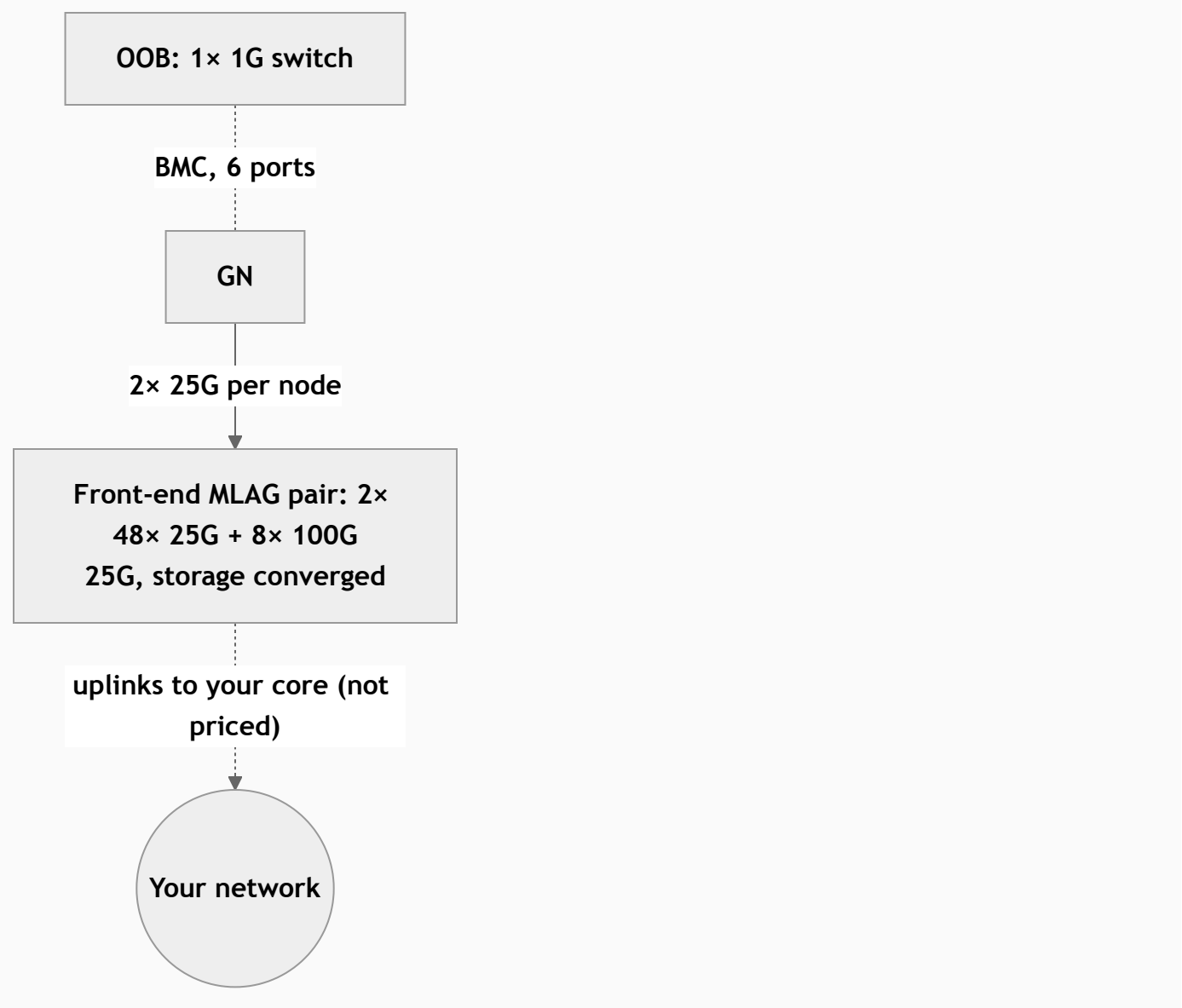
Requirements

Workload	inference
GPUs	32
Rack limit	40 kW
Colo	Texas
Timeline	within 6 weeks

How it was designed

1. GPUs → nodes: ceil(32 ÷ 8) = 4 × 8x NVIDIA RTX PRO 6000 Blackwell Server Edition, Intel (32 GPUs; nodes ship fully populated).
2. Class: PCIe — each GPU works alone (inference). Does the model at its precision fit one GPU's memory, and can each GPU work alone? Yes: PCIe. No: HGX. Distributed training is never served by PCIe servers.
3. Front-end: 2 ports per host, one to each switch of an MLAG pair of 48× 25G + 8× 100G switches; 4 ports per switch of 48 usable → 1 pair(s).
4. OOB: 4 hosts + 2 switches = 6 × 1G ports → 1 × 48× 1G RJ45 + 4× 10G.
5. GPU nodes per 40 kW rack = min(floor(40 U ÷ 4 U), floor(40 ÷ 8.47 kW)) = 4 (power binds).
6. Compute 1: 19 U, 34.33 kW provisioned → 1 × 415V 3-ph 60A circuit pair (A+B), 80% of each feed; 19 power cords per feed.

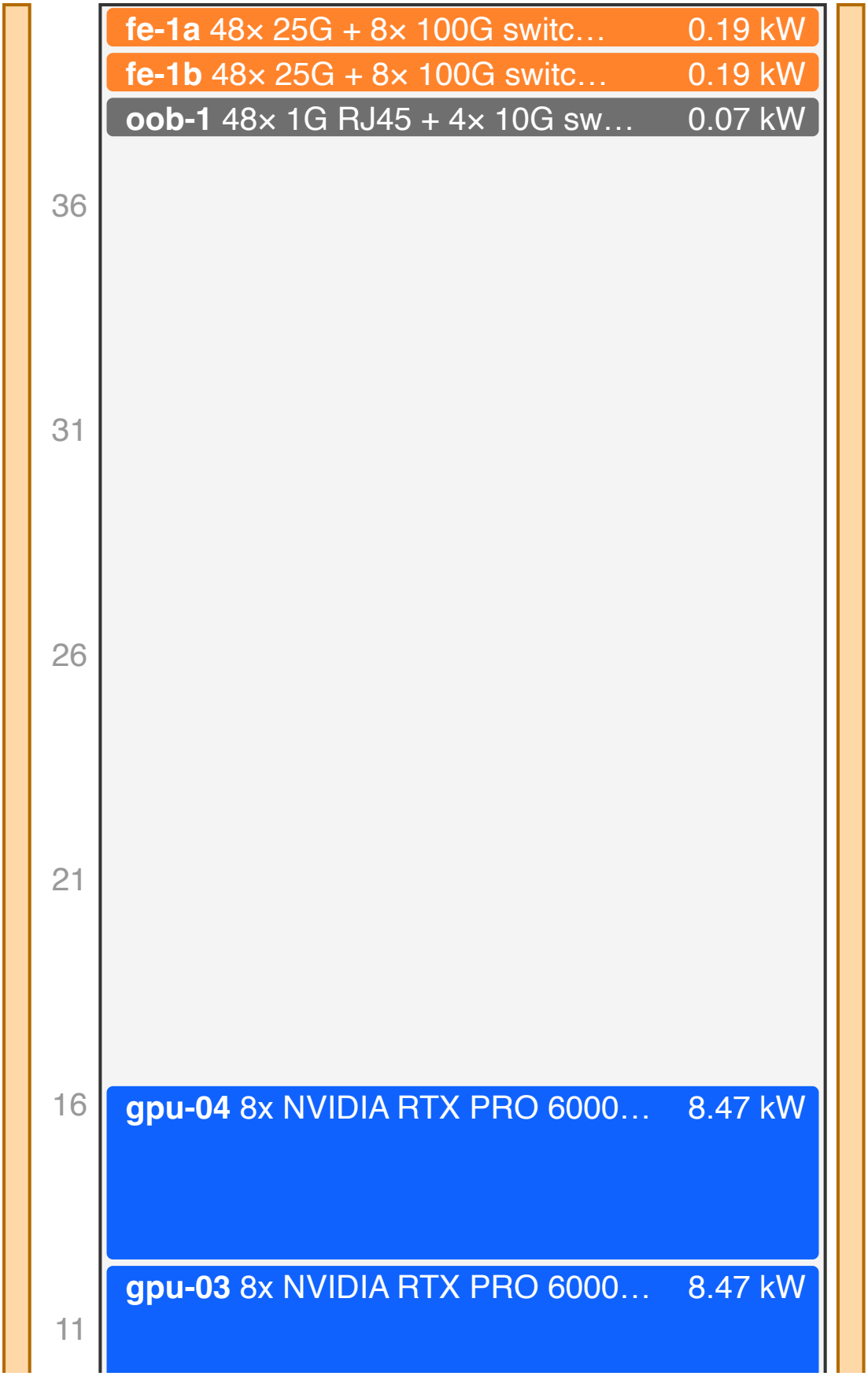
Network topology

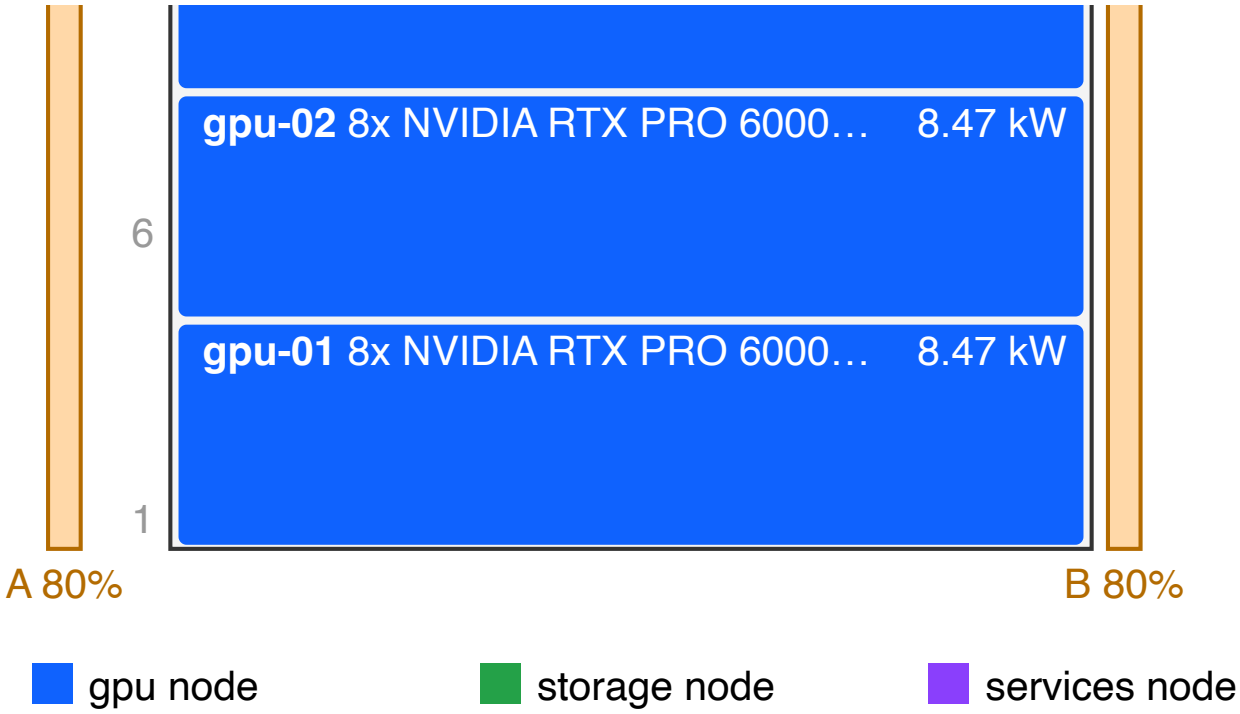


Rack elevation and power

Compute 1

19/40 U · 34.33 kW · 1× 415V 3-ph 60A A+B





Rack	Space	Provisioned	Circuits	Devices
Compute 1	19/40 U	34.33 kW	1× 415V 3-ph 60A (A+B)	4× 8x NVIDIA RTX PRO 6000 Blackwell Server Edition, 2× 48× 25G + 8× 100G switch — front-end pair, MLAG, 1× 48× 1G RJ45 + 4× 10G switch — OOB

Bill of materials

Item	Qty	Unit (USD)	Extended	Lead time	Source
GPU servers					
8x NVIDIA RTX PRO 6000 Blackwell Server Edition, Intel — SYS-422GA-NRT	4	177,430.31	709,721.24	In stock (barebone and RTX PRO 6000 cards, 2026-11-03); conf	/p/UPS-GPU-RTXPRO-8
Front-end network					
48× 25G + 8× 100G switch (front-end pair, MLAG) — FS S5470-48Y, PicOS	2	6,049.00	12,098.00	97 in stock (global warehouse, 2026-11-03). ~1 week	/p/UPS-NET-25G-48
+1x 25G 2x SFP28 CX6 NIC (AOC-631102AN-ADAT) (listed option, one per GPU node)	4	344.08	1,376.32		/p/UPS-GPU-RTXPRO-8
Front-end network cabling					
25G 2 m passive DAC — GPU nodes to the pair, 2 per node	8	39.00	312.00		/switches
100G 2 m passive DAC — peer link	2	54.00	108.00		/switches
Out-of-band					
48× 1G RJ45 + 4× 10G switch (OOB) — FS S3410-48TS, PicOS	1	989.00	989.00	593 in stock (global warehouse, 2026-11-03). ~1 week	/p/UPS-NET-1G-48
Hardware total			**$724,604.56**		

Assumptions

- GPU server: 8x NVIDIA RTX PRO 6000 Blackwell Server Edition, Intel (UPS-GPU-RTXPRO-8), the default PCIe line (proposed)

Not included

- Racks, PDUs, whips and cabling labour
- Uplinks to the buyer's core network
- Software (cluster manager, storage software), unless listed
- Taxes, shipping and duties
- Installation and on-site services

Next step. Book the intro (https://calendly.com/alecglover/intro-call) or email alec@upstation.io with the design ID; UpStation validates the design and introduces the supplier that fits.

Prepared by UpStation from public supplier prices published at ai.upstation.io (prices as of 2026-11-03). Budgetary only: the supplier confirms price, availability and lead time on order. UpStation does not resell; buyers transact directly with the supplier. alec@upstation.io