



AI Server for NVIDIA HGX Rubin NVL8, Liquid Cooled

(Supermicro SYS-222GS-R2BT-ALC)

Contents

Product overview	2
Features and benefits.....	3
Prominent feature.....	3
Platform support	4
Product sustainability	5
Product specifications.....	5
System requirements	6
Ordering information	7
Warranty information.....	7
Cisco and Partner Services	7
Cisco Capital.....	8
Learn more.....	8

An NVIDIA HGX Rubin NVL8 8-GPU system in a 2U direct liquid-cooled chassis, sold and supported by Cisco and managed through Cisco Nexus One. Packing 2.3 TB of HBM4 and NVLink 6 into a quarter of the rack space of the previous generation air-cooled server.

Product overview

This server brings the NVIDIA Rubin generation to the Cisco AI compute portfolio in its densest form. The underlying hardware is the Supermicro GPU SuperServer SYS-222GS-R2BT-ALC, a 2U liquid-cooled NVIDIA HGX Rubin NVL8 system built on the NVIDIA MGX rack architecture with Supermicro blind-mate busbar and coolant manifold connections for tool-free rack integration. As a rack-integrated liquid-cooled system, it is offered as part of a complete rack with onsite integration service. Cisco handles the selling motion, support through Cisco CX Services, and (Future) lifecycle management through Cisco Intersight. Supermicro builds and integrates the hardware.



Figure 1. AI server for NVIDIA HGX Rubin NVL8, liquid-cooled (Supermicro SYS-222GS-R2BT-ALC)



The system pairs dual x86 processors with an NVIDIA HGX Rubin 8-GPU connected over sixth-generation NVLink at 3.6 TB/s per GPU, 28.8 TB/s across the baseboard. Eight Rubin GPUs carry 2.3 TB of HBM4 with 176 TB/s of aggregate memory bandwidth and deliver up to 400 petaflops of NVFP4 compute per system. Nine systems fit in one rack for 72 NVLink-connected Rubin GPUs. Direct liquid cooling is inherent to the platform; for information on any liquid-cooling components, or for a quote on liquid cooling products specifically, contact your Cisco sales specialist.

Figure 2. AI rack for NVIDIA HGX Rubin NVL8, liquid-cooled



Features and benefits

Table 1 summarizes the primary features and benefits of the platform

Table 1. Features and benefits

Feature	Benefit
NVIDIA HGX Rubin NVL8 baseboard with NVLink 6 at 3.6 TB/s per GPU	Twice the GPU-to-GPU bandwidth of Blackwell; eight Rubin GPUs work as one coherent pool
2.3 TB HBM4 at 176 TB/s aggregate bandwidth	Next-generation models and KV caches stay resident with roughly 3x the memory bandwidth of HGX B300
Up to 400 petaflops NVFP4 per system	Trains next-generation models with a fraction of the GPUs prior platforms required
2U direct liquid-cooled chassis on the NVIDIA MGX rack architecture	9 systems and 72 Rubin GPUs per rack; a quarter of the rack space of an 8U air-cooled node
Blind-mate busbar power and coolant manifold connections	Tool-free rack integration and serviceability at high sustained power
8x NVIDIA ConnectX-9 SuperNICs	Next-generation east/west scale-out, one SuperNIC port per GPU
Dual x86 processors (next-generation Intel Xeon)	Keeps data pipelines fed with a familiar x86 software stack
Delivered as an integrated, liquid-cooled rack	Arrives built and tested with onsite integration service included in the motion
Cisco Nexus One Management	One operations model across the Cisco compute and networking estate
Cisco CX Services support	Single support contract for a multi-vendor system

Prominent feature

Eight Rubin GPUs in 2U: generational density that air cooling cannot reach

One 2U NVIDIA HGX NVL8 node carries the same eight-GPU NVLink domain as an 8U air-cooled NVIDIA HGX B300 system, with 2.3 TB of HBM4 instead of HBM3e, three times the memory bandwidth, double the NVLink bandwidth, and roughly four times the NVFP4 training compute. Nine of them stack into a single rack for 72 NVIDIA Rubin GPUs where the previous generation fit far fewer.

That density only works with liquid cooling. Direct liquid cooling carries the compute heat load through blind-mate manifolds, and power arrives over a rack busbar rather than per-node cords, so the system can be sold as part of a complete integrated rack with onsite service rather than as a standalone box. Cooling infrastructure is scoped per facility: in-row and in-rack CDU options and a liquid-to-air sidecar exist for data centers without facility water. For information on any liquid-cooling components, or for a quote on liquid-cooling products specifically, contact your Cisco sales specialist.



Platform support

Cisco offers this Supermicro platform in a defined configuration, given in Table 2. Supermicro has not yet published the full public data sheet for this SKU; the fields in Table 2 reflect the announced 2U HGX Rubin NVL8 platform.

Table 2. Configuration selected by Cisco, SYS-222GS-R2BT-ALC (HGX Rubin NVL8 liquid-cooled)

Item	Selected configuration
PID	SYS-222GS-R2BT-ALC (Supermicro SKU)
Platform	2U direct liquid-cooled system on the NVIDIA MGX rack architecture; complete rack and onsite integration service are required
GPU	1x NVIDIA HGX Rubin NVL8 8-GPU baseboard
GPU memory	2.3 TB HBM4 (8x 288 GB), 176 TB/s aggregate bandwidth
GPU interconnect	NVIDIA NVLink 6, 3.6 TB/s per GPU, 28.8 TB/s per system
CPU	2x x86 processors (next-generation Intel® Xeon®)
East-west NICs	8x NVIDIA ConnectX-9 SuperNICs
Rack density	9 systems per rack, up to 72 Rubin GPUs
Liquid cooling	Direct-to-chip DLC with blind-mate manifolds; CDU and liquid-to-air sidecar options scoped with your Cisco sales specialist
Software management	Cisco Nexus One
Support	Cisco CX services and Supermicro Services
Availability	October 2026



Product sustainability

Information about Cisco’s Environmental, Social and Governance (ESG) initiatives and performance is provided in Cisco’s CSR and sustainability [reporting](#).

Table 4. Cisco Environmental Sustainability Information

Sustainability Topic		Reference
General	Information on product-material-content laws and regulations	Materials
	Information on electronic waste laws and regulations, including our products, batteries and packaging	WEEE Compliance
	Information on product takeback and resuse program	Cisco Takeback and Reuse Program
	Sustainability Inquiries	Contact: csr_inquiries@cisco.com
Material	Product packaging weight and materials	Contact: environment@cisco.com

Product specifications

Table 5 reflects the NVIDIA HGX R200 NVL8 platform and Supermicro announcements for the 2U liquid-cooled system. Supermicro has not yet published the full data sheet for this SKU; product specifications may change without notice.

Table 5. Product specifications, Supermicro GPU SuperServer SYS-222GS-R2BT-ALC

Category	Specification
Form factor	2U rackmount, direct-liquid-cooled, NVIDIA MGX rack architecture with blind-mate busbar power and coolant manifold connections
GPU	NVIDIA HGX NVL8 baseboard; 8x NVIDIA Rubin GPUs
GPU memory	288 GB HBM4 per GPU at 22 TB/s; 2.3 TB and 176 TB/s aggregate per system
GPU compute	Up to 400 petaflops NVFP4 per system (up to 50 petaflops FP4 inference per GPU), other precisions also supported
GPU interconnect	NVIDIA NVLink 6 with NVSwitch; 3.6 TB/s per GPU, 28.8 TB/s total



Category	Specification
Processor	Dual socket x86: next-generation Intel Xeon
Scale-out networking	8x NVIDIA ConnectX-9 SuperNICs; NVIDIA Quantum-X800 InfiniBand or Spectrum-X Ethernet fabrics
Rack integration	9 systems per rack for up to 72 NVLink-connected Rubin GPUs; sold as a complete integrated rack with onsite service
Cooling	Supermicro direct liquid cooling (DLC); in-rack CDU, in-row CDU, and liquid-to-air sidecar options for facilities without water loops
Power	Rack busbar power delivery through power shelves (rack-level configuration)
System management	Cisco Nexus One Management Software

System requirements

Table 6 lists what must be in place before installation. This is a rack-integrated liquid-cooled system; for information on any liquid-cooling components, or for a quote on liquid-cooling products, contact your Cisco sales specialist.

Table 6. System requirements

Requirement	Description
Rack	Sold as part of a complete Supermicro-integrated rack with busbar power and coolant manifolds; standalone deployment in an existing rack is not offered
Power	Facility feeds are sized for the configured rack; per-rack power is set by GPU count and CPU selection and is confirmed at time of configuration
Liquid cooling	Facility water loop with CDU, or liquid-to-air sidecar for facilities without water; scope components and quotes with your Cisco sales specialist
Scale-out fabric	NVIDIA Quantum-X800 InfiniBand or Spectrum-X Ethernet infrastructure for multi-node deployments
Management	Cisco Nexus One

Ordering information

The base platform is a Supermicro SKU offered through Cisco as part of an integrated rack. Liquid-cooling components (coolant distribution units, manifolds, and liquid-to-air sidecars) are quoted separately: contact your Cisco sales specialist for information on any liquid-cooling components or for quotes on specific liquid-cooling products. To order, contact your Cisco account representative or visit the Cisco Ordering Home Page.

Table 7. Ordering information

Part number	Product description
SRS-RUBIN-NVL8-M0	Supermicro 48U DLC Rack, NVIDIA Rubin NVL8 222GS, No In-Rack CDU
SRS-RUBIN-NVL8-M1	Supermicro 48U DLC Rack, NVIDIA Rubin NVL8 222GS, with In-Rack CDU
LCS-SCDU-1K3LR001	1.8MW In-Row CDU, L2L, N+1 Pumps
LCM-SMLCPM3-3Y	Liquid Cooling Maintenance Service, SMLCPM3, 3-Year
LCS-SCDU-200AR001	200kW L2A Sidecar CDU, N+1 Pumps
SVC-NVSTDSUP-3Y	NVIDIA Enterprise Support Entitlement, 3Yr

Warranty information

Warranty terms for this Supermicro platform offered by Cisco are standard Supermicro warranty terms. Support is delivered through Cisco CX services under the applicable service contract; contact your Cisco sales specialist for more details.

Cisco and Partner Services

Cisco CX services covers this platform from planning through deployment and ongoing operations, with a single Cisco support contract even though the hardware is Supermicro-built. For more information, visit <https://www.cisco.com/go/services>.

Cisco Capital

Flexible payment solutions to help you achieve your objectives

Cisco Capital makes it easier to get the right technology to achieve your objectives, enable business transformation and help you stay competitive. We can help you reduce the total cost of ownership, conserve capital, and accelerate growth. In more than 100 countries, our flexible payment solutions can help you acquire hardware, software, services and complementary third-party equipment in easy, predictable payments. [Learn more.](#)

Learn more

For more on Cisco AI infrastructure, visit <https://www.cisco.com/go/ai>. Full platform specifications will be available on the Supermicro product page for the SYS-222GS-R2BT-ALC once published.

Learn more about Cisco and Supermicro partnership:

<https://www.cisco.com/site/us/en/solutions/global-partners/supermicro/index.html>.