



NVIDIA Vera Rubin Superchip

Rubin

Superchip

2026

OVERVIEW

260.0 FP32 TFLOPS	576 GB VRAM HBM4	44.0 TB/s Memory Bandwidth	— TDP
----------------------	---------------------	-------------------------------	----------

PERFORMANCE METRICS

Precision	Peak TFLOPS	Bits	Type
FP32_TC	800.0	—	Tensor Core
FP6	35000.0	—	Standard
NVFP4	100000.0	—	Standard
INT8	500.0	8	Standard
FP8	35000.0	8	Standard
FP16	8000.0	16	Standard
BF16	8000.0	16	Standard
TF32	4000.0	32	Standard
FP32	260.0	32	Standard
FP64_TC	400.0	64	Tensor Core
FP64	67.0	64	Standard

MEMORY SPECIFICATIONS

Capacity	576 GB
Type	HBM4
Bandwidth	44.0 TB/s
Interface Width	—

POWER & EFFICIENCY

TDP	—
Est. Max Power	—
FP32 Efficiency	—
FP16 Efficiency	—

HARDWARE DETAILS

Vendor: NVIDIA	Architecture: Rubin	Process Node: —
Form Factor: Superchip	Launch Year: 2026	Variant: Superchip

INTERCONNECT

GPU-to-GPU: NVLink-C2C	Bandwidth: 1.8 TB/s
------------------------	---------------------

ENTERPRISE FEATURES

Sparsity: Yes