Rebellions RebelPOD (32-node)

256x Rebel100 | REBEL | pod

OVERVIEW

524	36.9 TB	1228.8 TB/s	224.0 kW
FP8 PFLOPS Dense	Aggregate Memory	Aggregate Bandwidth	Aggregate Power

PERFORMANCE METRICS

Precision	Peak (PFLOPS)	Per GPU (TFLOPS)	TFLOPS/W
FP16	262.1	1024.0	–
FP8	524.3	2048.0	–

All figures are dense. Vendors commonly headline the sparse number, which is twice the dense one.

CONFIGURATION

Accelerator	Rebel100
Count	256
Memory per GPU	144 GB
Bandwidth per GPU	4.8 TB/s

POWER & EFFICIENCY

Aggregate Power	224.0 kW
TDP per GPU	600 W
FP8 Efficiency	2.34 TFLOPS/W
Aggregate Bandwidth	1228.8 TB/s

SYSTEM DETAILS

Vendor: Rebellions
Family: Rebel
Architecture: REBEL
Form Factor: pod
Release Year: —
Process Node: 4nm

INTERCONNECT

Backend fabric 400G QSFP112-DD; frontend network 25G SFP28; optional 100G storage network

NOTES

The 32-node RebelPOD is eight NPU racks plus one management rack, holding 32 RebelServers and 256 Rebel100 cards. Rebellions rates it at 524.3 PFLOPS FP8 and 262.1 PFLOPS FP16, with 36.86 TB of HBM3E at 1,228.8 TB/s and 48 TB of DDR5 across 64 AMD EPYC 9355 processors totalling 2,048 cores and 4,096 threads. The NPU racks draw up to 224 kW and the whole pod is air-cooled, which is unusual at this scale. Rebellions also offers 64, 128 and 256-node configurations but publishes specifications only for this 32-node one, so the larger sizes are deliberately not listed here. All figures are dense, per Rebellions' own labelling of the Rebel100 compute; no sparse rate is published.