



AMD Radeon PRO W7900

Also known as: W7900 · Radeon PRO W7900 48GB

The Radeon PRO W7900 is AMD's 2022-generation RDNA 3 workstation card with 48 GB of ECC GDDR6 on a 384-bit bus, 96 compute units, and a 295 W triple-slot actively cooled board on PCIe 4.0 x16. Its draw for AI work is capacity: 48 GB on one card at about \$4,000, with no HBM or fabric. It lacks the FP8 matrix support of the newer R9700, so it is an FP16 and INT8 card. AMD's own page carries a footnote that the W6000 and W7000 series are not designed or recommended for datacenter use, and GPU Smith treats it as a workstation and developer part.

48 GB GDDR6 ECC Memory · 384-bit	864 GB/s Memory bandwidth · peak	123 TFLOPS FP16 matrix · peak
123 TOPS INT8 matrix · INT4 245 TOPS	295 W Board power · 2x 8-pin, 750 W PSU recommended	Triple slot, active Form factor · PCIe 4.0 x16, 11 in

COMPUTE		01/05
Architecture	RDNA 3 · TSMC 5nm GCD, 6nm MCD	
Compute units	96	
Stream processors	6,144	
AI accelerators	192	
Ray accelerators	96	
FP32 / FP16 vector	61.3 TFLOPS	
FP16 matrix	123 TFLOPS	
INT8 matrix	123 TOPS	
INT4 matrix	245 TOPS	
Transistors	57.7 billion	

MEMORY02/05	
Capacity	48 GB GDDR6 · ECC supported
Bus width	384-bit
Bandwidth	864 GB/s · peak
Infinity Cache	96 MB
INTERCONNECT & FORM FACTOR03/05	
Host bus	PCIe 4.0 x16 · 3.0 backward compatible
Dimensions	11 in (280 mm) long · full height, triple slot
Display outputs	3x DisplayPort 2.1, 1x Enhanced Mini DisplayPort 2.1
POWER & THERMAL04/05	
Total board power	295 W
Power connectors	2x 8-pin
Minimum PSU	750 W · AMD recommendation
Cooling	Active
SOFTWARE & PLATFORM05/05	
OS support	Windows 10/11 64-bit, Linux x86-64
Compute stack	ROCm · check AMD's release notes for W7900 support on your distribution
Datacenter use	Not recommended by AMD · AMD footnote GD-239
Reference street price	About \$3,999 · reported, August 2026 tracker low
PROCUREMENT	
Typical lead time A 2022-generation part sold through retail and workstation channels; no lead time published. Stock varies by region (reported). Pricing About \$3,999 reported as the US low and 12-month median from a price tracker as of August 2026. AMD's original launch MSRP was not confirmed in the sources we checked. ■ Triple-slot width limits multi-card density; a 4-card build needs a purpose-built chassis. ■ Two 8-pin connectors are easier to power than 12V-2x6, which helps in older workstations. ■ At roughly three times the R9700 price for 16 GB more memory, it only makes sense where 48 GB on one card is the requirement.	

FIELD NOTES

- 48 GB on a single card avoids multi-GPU splits for many 30B-class models at 8-bit and lets you hold a quantised 70B model with a small context.
- No FP8 matrix support in AMD's published specs. Throughput on FP8 checkpoints will not benefit the way it does on the R9700; plan for FP16, INT8 or INT4.
- Memory bandwidth (864 GB/s) beats the R9700's 640 GB/s, which matters for decode-bound inference; the R9700 wins on matrix compute.
- AMD states this series is not designed for datacenter use. For rack deployment, use an Instinct part or a vendor-qualified server GPU.

QUESTIONS WE GET ON THIS PART

How much memory does the W7900 have?

48 GB of GDDR6 with ECC on a 384-bit bus, 864 GB/s peak bandwidth.

How much does the Radeon PRO W7900 cost?

About \$3,999 in the US as of August 2026, per a price tracker (reported).

W7900 vs R9700 for LLM inference?

The W7900 has 48 GB against 32 GB and higher bandwidth. The R9700 is the newer RDNA 4 card with FP8 support and, in Phoronix testing, about 1.47x the W7900 performance, at about a third of the price.

Is the W7900 supported in datacenters?

AMD's page says W6000 and W7000 series cards are not designed or recommended for datacenter use.

REFERENCES

- [1] <https://www.amd.com/en/products/graphics/workstations/radeon-pro/w7900.html>
- [2] <https://gpusmex.com/us/gpus/radeon-pro-w7900>
- [3] <https://www.phoronix.com/review/amd-radeon-ai-pro-r9700>