

One server, not four.

Five reasons your next inference deployment should be one server. Run every model your teams need in parallel, one per GPU, on AMD Instinct™ MI350P.

30 GPUs

MI350P on one host

4.32 TB

HBM3E on one server

30 models

in parallel,
one per GPU

0 hops

network hops
between GPUs



LIQID GPU chassis
4 to 30 MI350P over
PCIe Gen 5

One host
Dual-socket AMD
EPYC™ 9005

One model per GPU
No network between
GPUs

Five Reasons to Run One Server

When a job outgrows a server, the standard answer is to buy another. Here is what changes when every model runs on one.

- Every model gets its own GPU**
144 GB of HBM3E per MI350P holds a full 70B model at FP8, or gpt-oss-120b, on one card. Up to 30 GPUs on one AMD EPYC™ host means up to 30 models running in parallel.
30× MI350P · 4.32 TB HBM3E · 69 PFLOPS FP8 on one host

- There is no network between your GPUs**
Every GPU stays in one server on a Liquid PCIe Gen 5 fabric, peer-to-peer. There is no inter-node fabric to design, buy, power, tune or troubleshoot.
No NICs, no switch ports, no cluster to build

- One of everything, operationally**
One host, one OS image, one license footprint and one thing to patch, instead of four of each. Configurations are fixed and tested, from 4 to 30 GPUs.
Available now, from 4 to 30 GPUs

- It fits the building you already have**
MI350P is a standard dual-slot PCIe card at 600 W, passively air-cooled. A full 30-GPU system draws about 22 kW. No liquid cooling and no rack redesign.
~3.14 PFLOPS per kW · ~195 GB HBM per kW

- Grow by the GPU, not by the server**
Add GPUs and the next job's models to the same host: 10 GPUs for a team, 20 for a department, 30 for the enterprise.
Linear scaling measured from 1 to 30 GPUs

Bring us three numbers: the jobs you run, the model for each, and your peak concurrent users.

Built for
Enterprise AI teams · Regulated industries · Public sector · Neoclouds and AI service providers

Proven on
AMD Instinct™ MI350P · AMD EPYC™ 9005 · ROCm™ · 200+ Liquid deployments

Size Your System

sales@liquid.com · +1 303.500.1551