Vultr offers a global cloud infrastructure platform through our partnership with Supermicro

Industry-leading cloud GPU, accelerated by AMD, and cloud compute at a global scale without the complexity or cost of hyperscale cloud providers

Competitive advantage: Enterprise cloud efficiency at a fraction of the cost

Transparent, predictable pricing with bandwidth included, delivering the best price-to-performance — with end-to-end infrastructure support for high-performance and AI workloads.

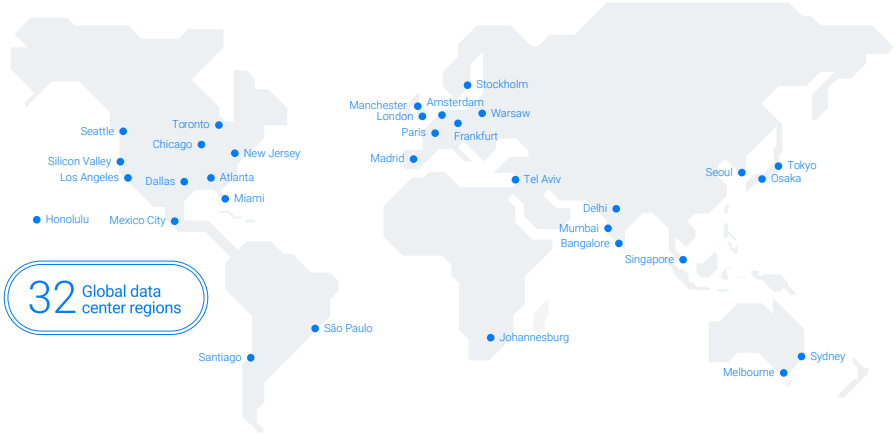
- 32 global cloud data center regions reaching 90% of the world's population in 2-40ms
- Global availability of full-stack cloud GPU – accelerated by AMD – and cloud compute
- 2M+ customers served in 186 countries
- Data residency, on demand sovereign cloud and compliance-ready infrastructure, including GDPR and HIPAA compliant
- Easy to use and API first –no certification needed - without vendor lock-in
- Composable, best-of-breed ecosystem of cloud services

30-50%
Lower cost than hyperscalers

20-30%
Lower costs than other independent cloud providers

Full native AI stack

Compute, storage, networking, and orchestration — not just GPUs like other neo cloud



Instantly deploy high-performance, composable cloud infrastructure worldwide

Cloud GPU

Enterprise-grade supercompute GPU clusters accelerated by AMD Instinct™ MI325X and MI300X.

PLUS: Fully-managed Kubernetes and databases

Cloud Compute

Spin up virtual machines in general purpose or optimized configurations in under 60 seconds.

Kubernetes

Deploy and scale containerized apps with a fully managed, CNCF-certified service compatible with the Cluster API.

Bare Metal

Stay in complete control of your environment with dedicated high-performance, single-tenant servers.

Cloud Storage

Harness secure, flexible, and scalable block, file system, and object storage or managed databases.

Networking

Deploy across the globe with the advanced Vultr CDN, and scale seamlessly with load balancers for greater performance and stability.

> "At ConfidentialMind, we specialize in deploying generative AI systems like semantic search, RAG, and AI agents across on-premises and cloud environments. Vultr Cloud GPUs, including Supermicro hardware and AMD Instinct™ MI300X, provide the computational power we need for running AI models. Vultr Kubernetes Engine streamlines deployment and scaling, enabling our customers to build AI solutions with data security and operational efficiency."
>
> Severi Tikkala, CTO at ConfidentialMind

How Vultr and Supermicro build smarter infrastructure

Vultr, Supermicro, and AMD deliver high-performance compute, AI, storage, and networking solutions engineered for scalability, efficiency, and modern workload demands.

Supermicro's modular building-block architecture accelerates the adoption of emerging technologies. Its flexible system design enables Vultr to deploy the latest infrastructure more rapidly—reducing time-to-market and giving customers earlier access to innovation for AI, cloud-native, and edge environments.

System-level power design enhancements—including improved airflow, optimized thermals, and platform-level tuning—enable higher performance per watt. This results in more compute capacity with greater energy efficiency, supporting cost-effective scaling for demanding workloads.

Backed by more than 30 in-house engineering teams and decades of expertise, Supermicro's solutions are proven at scale.

From the edge to the cloud to AI superclusters, Vultr leverages this foundation to offer highly reliable, cost-efficient infrastructure for AI training, inference, high-performance computing, and cloud-native applications.

Who we power: AI innovators, and HPC leaders

Vultr and Supermicro offer a ready-to-deploy, flexible infrastructure platform optimized for performance, energy efficiency, and fast scaling. Together with AMD, we deliver next-generation compute and AI capabilities, combining powerful GPUs, efficient server designs, and global cloud reach — giving customers a reliable, modern foundation to grow and innovate.

What we bring to you: Scalable, flexible, high-performance cloud

Vultr and Supermicro offer a ready-to-deploy infrastructure platform designed for performance, energy efficiency, and rapid scalability. In partnership with AMD, we deliver next-generation compute and AI capabilities—combining powerful GPUs, advanced server designs, and a global cloud footprint. This provides customers with a reliable, modern foundation to scale operations and drive innovation.

Enterprise-grade performance: Our platform is built with enterprise reliability in mind. Supermicro's engineering expertise, combined with Vultr's high-performance cloud infrastructure, ensures consistent uptime, low-latency connectivity, and robust support for demanding workloads such as AI training, real-time inference, and HPC.

Fast scalability: Whether scaling from pilot to production or expanding globally, our modular infrastructure enables fast deployment and seamless growth. Customers can provision resources on demand across Vultr's worldwide network—accelerating time-to-value and supporting evolving business needs.

Cost-optimized infrastructure: By leveraging energy-efficient system designs and eliminating unnecessary overhead, we deliver high performance at a lower total cost of ownership. Vultr's transparent pricing and flexible consumption model ensure cost predictability without compromising on capability.

Energy efficiency at scale: Optimized airflow, thermal tuning, and power-efficient components help reduce environmental impact while maintaining high compute density. This green, high-performance infrastructure supports sustainability goals without sacrificing operational efficiency.

Learn more about the Supermicro Portal and Vultr Cloud GPU, or start a Vultr GPU Proof of Concept today.

Contact us at vultr.com to get started.