VULTR

DATASHEET

Vultr Cloud GPU

Large-scale dedicated clusters and flexible on-demand VMs, accelerated by AMD and NVIDIA GPUs

VULTR.COM

VULTR

Vultr Cloud GPU

Unleash the potential of flexible, large-scale high-performance compute, accelerated by AMD and NVIDIA GPUs.

Vultr Cloud GPU, accelerated by AMD and NVIDIA, stands as a beacon for next-gen GPU-accelerated infrastructure. Sidestepping the usual complications of driver setups and licensing, it offers users a direct conduit to the raw power of AMD and NVIDIA GPUs for any computational endeavor.

Why it's important right now

Accelerated computing is being rapidly adopted across industries and large-scale production deployments. Vultr Cloud GPU, accelerated by AMD and NVIDIA, enables breakthrough performance with fewer, more powerful servers, driving faster time to insights while saving money. The platform accelerates a broad array of workloads, from artificial intelligence (AI) training and inference to scientific computing and virtual desktop infrastructure (VDI) applications, with a diverse range of AMD and NVIDIA GPUs.

Power diverse workloads

AI training and inference

With AI models for conversational AI, natural language processing (NLP), and recommendation systems growing in size and complexity, training times have dramatically increased, resulting in lower productivity and higher costs. AMD and NVIDIA GPUs and a complete solution stack significantly accelerate your AI training on Vultr, resulting in faster model exploration, significant cost savings, and faster time to ROI.

High-performance computing

HPC, fused with AI and machine learning, is fueling the advancement of computational science, paving the way for breakthrough scientific discovery. AMD and NVIDIA GPUs accelerate large-scale HPC applications across a broad range of industries and domains, from weather forecasting and energy exploration to computational fluid dynamics and life sciences.

Graphics visualization

GPU-accelerated virtual workstations provide professional workstation performance for modern workloads — including real-time rendering, product design, engineering simulation, and content creation — with simplified management and the assurance that sensitive data is protected in the cloud and accessible from anywhere in the world.

Vultr Kubernetes Engine accelerated by AMD and NVIDIA GPUs

Create AMD- and NVIDIA-GPU accelerated Kubernetes clusters that will power your most resource-intensive workloads globally. This powerful combination empowers developers and innovators to build sophisticated AI and machine learning systems that can handle even the most complex challenges.

Vultr Bare Metal accelerated by AMD and NVIDIA GPUs

Vultr Bare Metal grants you access to the underlying physical servers. Featuring no virtualization layer and a single-tenant environment, these servers can power even the most resource-intensive workloads. Stay in full control of your environment with high-performance single-tenant dedicated servers, accelerated by AMD and NVIDIA GPUs.

Vultr Cloud GPU accelerated by AMD and NVIDIA

Vultr Cloud GPU virtual machines provision fully dedicated AMD and NVIDIA GPU resources that deploy in seconds. These instances can scale from fractions of a single GPU all the way up to full 8-GPU servers, ensuring that even the most demanding workloads are executed with precision and speed.

Key benefits

Global scalability

Vultr Cloud GPU, accelerated by AMD and NVIDIA, redefines agility in tech deployments. Whether you're a start-up or a global enterprise, you can initiate a cloud GPU instance tailored to your exact needs in seconds. As your requirements evolve — due to a growing user base or more intricate data processing tasks — you can effortlessly scale up or down, ensuring optimal resource allocation no matter where you operate.

Cost-effectiveness

Traditional computational setups can be financially prohibitive, particularly when maintaining hardware or updating to newer models. With Vultr Cloud GPU, accelerated by AMD and NVIDIA, businesses can dynamically allocate GPU resources, only paying for the precise power they need. It's a smarter, leaner approach to resource investment, eliminating wasted capacity and making budgeting more predictable.

Ease of use

Tapping into high-end computational power shouldn't be daunting. With pre-configured deployments and Vultr's streamlined control panel and API, even the most intricate AMD and NVIDIA GPU-driven workflows become straightforward. It's about offering high-tech prowess without the usual technical intricacies.

Reduce time-to-market with Vultr GPU Enabled Images

Vultr GPU Enabled Images offer finely tuned and integrated operating system and software environments, which instantly provision the full array of Vultr Cloud GPUs. The images are pre-configured for NVIDIA GPUs with the NVIDIA CUDA Toolkit, NVIDIA cuDNN, and NVIDIA Drivers, and for AMD GPUs with the ROCm™ open software ecosystem, for immediate deployment.

Deploy faster with Vultr Container Registry

From HPC to conversational AI to medical imaging to recommender systems and more, Vultr Container Registry offers ready-to-use containers — in one place — to speed up your application development and deployment process. All are accelerated by Vultr's preconfigured GPU Enabled Images, featuring AMD ROCm™ open software, NVIDIA drivers and CUDA Toolkit, and more.

Advanced computing with AMD and NVIDIA

High-performance computing reaches new heights with Vultr Cloud GPU, accelerated by AMD and NVIDIA. A seamless meld of Vultr's cloud capabilities and AMD and NVIDIA's industry-leading GPU tech ensures businesses are well-equipped to tackle current and forthcoming computational challenges. Whether your focus lies in Big Data, deep learning, or any GPU-driven task, Vultr guarantees a front-row seat in the tech revolution.

Vultr Cloud GPU, accelerated by AMD and NVIDIA

AMD Instinct™ MI325X Accelerator
Setting new standards in HPC, AI training, and deployments, with inference-optimized acceleration and efficiency

AMD Instinct™ MI300X Accelerator
Delivering powerful and efficient AI and HPC deployments with exceptional performance, and inference-optimized acceleration

NVIDIA HGX™ B200
Propelling a new era of accelerating computing and generative AI, integrating NVIDIA Blackwell GPUs with high-speed interconnect to power AI performance at scale

NVIDIA GH200 Grace Hopper™ Superchip
Delivering breakthrough acceleration of large-scale AI, model training and inference, and high-performance computing (HPC) applications

NVIDIA HGX H100 GPU
Delivering unprecedented acceleration to power the world's most advanced AI, data analytics, and HPC workloads

NVIDIA A100 GPU
Enabling researchers and scientists to combine simulation, data analytics, and AI to drive scientific progress

NVIDIA L40S GPU
Combining powerful AI compute with best-in-class graphics and media acceleration to power the next generation of data center workloads

NVIDIA A40 GPU
Combining professional graphics with powerful compute and AI, to meet today's design, creative, and scientific challenges

NVIDIA A16 GPU
Enabling virtual desktops and workstations with the power and performance to tackle any project from anywhere

Learn more about Vultr Cloud GPU, accelerated by AMD and NVIDIA, and contact us at vultr.com to get started.