DATASHEET

Vultr Cloud GPU: Accelerated by NVIDIA A100

Unleashing the full potential of
high-performance computing and AI

VULTR.COM

VULTR

Vultr Cloud GPU: Accelerated by NVIDIA A100

Unleashing the full potential of high-performance computing and AI

The NVIDIA A100 Tensor Core GPU on Vultr brings accelerated computing to the cloud. This state-of-the-art GPU delivers breakthrough performance for accelerating high-performance computing (HPC) and AI workloads. With Vultr Cloud GPU, accelerated by NVIDIA, you can now harness the power of the NVIDIA HGX A100 and A100 PCIe to supercharge your applications, providing you with a competitive advantage in today's rapidly evolving digital landscape.

Why it's important right now

The NVIDIA A100 GPU has emerged as a crucial component in industries such as AI and machine learning (ML), primarily due to its impressive computational power and ability to accelerate complex workloads. This cutting-edge GPU is built on the NVIDIA Ampere architecture, offering exceptional performance for training large-scale neural networks, conducting advanced data analytics, and supporting high-performance computing tasks. Vultr Cloud GPU enables businesses and researchers to access and deploy AI and machine learning solutions with ease, minimizing upfront investments in hardware and reducing time-to-market. By offering such an accessible platform, Vultr Cloud GPU, accelerated by NVIDIA, fosters innovation and growth in industries reliant on advanced computational capabilities, ultimately driving progress in AI and ML applications.

Use cases

Deep learning and AI

The NVIDIA A100 GPU is specifically designed to accelerate deep learning applications, such as natural language processing, computer vision, and recommendation systems. With the Vultr Cloud GPU, accelerated by NVIDIA, you can easily deploy and scale your AI-powered applications, driving innovation and competitive advantage.

Data analytics

The NVIDIA A100 GPU's parallel processing capabilities make it ideal for handling large-scale data analytics workloads. By leveraging Vultr Cloud GPU, accelerated by NVIDIA, you can process massive datasets and glean valuable insights faster than ever before.

Scientific simulations

The high-performance computing capabilities of the NVIDIA A100 GPU enable you to run complex scientific simulations and models at unprecedented speeds. With Vultr's infrastructure, researchers and scientists can unlock discoveries and drive progress in their respective fields.

Key benefits

Exceptional performance: The NVIDIA A100 GPU is built on the Ampere architecture, delivering industry-leading performance for various HPC and AI workloads. With 3rd generation NVIDIA Tensor Cores and Multi-Instance GPU technology, the NVIDIA A100 provides unmatched acceleration for diverse applications, including deep learning, data analytics, and scientific simulations.

Improved AI training and inference: The NVIDIA A100 GPU features Tensor Cores with Tensor Float (TF32) and automatic mixed precision, accelerating both training and inference of AI models, resulting in faster and more accurate results.

Scalable infrastructure: Vultr's flexible infrastructure allows you to scale your GPU resources up or down seamlessly, depending on your needs. This adaptability ensures you can optimize your resource usage and control costs while still benefiting from the NVIDIA A100's powerful computing capabilities.

Broad application support: The NVIDIA A100 is compatible with a wide range of deep learning frameworks, data analytics tools, and scientific simulation software, allowing users to easily leverage their preferred tools and libraries.

Global data center network: With Vultr's extensive global data center network, you can deploy your GPU resources close to your users, reducing latency and improving user experiences. This global presence ensures that your applications can perform at their best – wherever your users are.

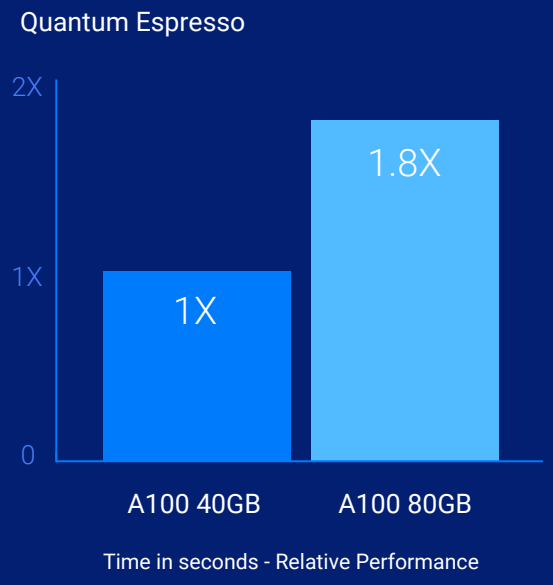
Conclusion

The NVIDIA HGX A100 and A100 PCIe GPU deliver unparalleled performance for HPC and AI, while Vultr's innovative platform makes them accessible to innovators and businesses alike, all while keeping cloud infrastructure costs low by minimizing upfront costs and enabling flexible resource management.

Learn more about
Vultr Cloud GPU

Contact us at vultr.com to get started. →

Up to 1.8X Higher Performance for HPC Applications



Specifications

NVIDIA A100 SXM	
FP64	9.7 TFLOPS
FP64 Tensor Core	19.5 TFLOPS
FP32	19.5 TFLOPS
TF32 Tensor Core	312 TFLOPS*
BFLOAT16 Tensor Core	624 TFLOPS*
FP16 Tensor Core	624 TFLOPS*
INT8 Tensor Core	1,248 TOPS*
GPU Memory	80GB HBM2e
GPU Memory Bandwidth	2,039 GB/s
Decoders	5 NVDEC 5 JPEG
Interconnect	NVLink: 600GB/s PCIe Gen5: 64GB/s

** Shown with sparsity.*