VULTR

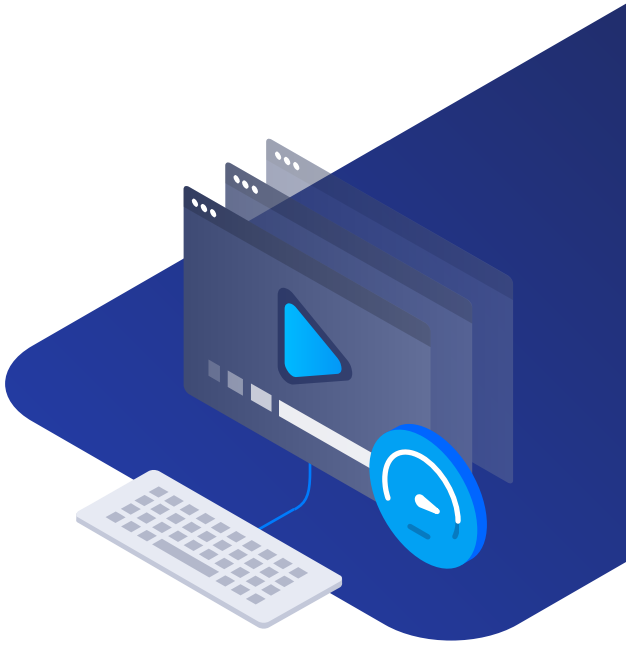
SOLUTION BRIEF

From Initial Renders to the Final Cut

Cloud GPUs can power the most complex video projects.

VULTR.COM

In the era of 4K and 8K videos, the demand for graphics processing power to handle sizable video files and related workloads continues to rise steadily. As a result, graphics processing units (GPUs) have become the standard architecture for media. Companies in the media industry are using GPUs for video rendering, visual effects (VFX), and even augmented reality applications. Engineering companies also use GPUs to render three-dimensional plans and blueprints for construction and manufacturing.



GPUs are the standard because they leverage hardware acceleration and compute unified design architecture (CUDA) for NVIDIA cards to reduce video processing timelines with the ideal blend of software and hardware. Traditionally users had to purchase these GPUs and store them on-premise or rent entire cloud GPUs.

When companies opt to purchase these GPUs for internal workloads, the overall cost for on-premises infrastructure may make these tools unaffordable, especially to smaller firms with capital constraints.

Cloud platforms have attempted to mitigate these challenges with products that support the flexible expansion of existing cloud infrastructure to accommodate even more computing, memory, and networking resources. The ability of a cloud platform to scale as demand grows is called hyperscaling and is used to support demanding big data operations and workflows.

The need for a global, efficient, and affordable cloud solution has inspired Vultr to innovate and create solutions built on the world's most powerful GPUs and offer them to users through a fractional, virtualized model at remarkably low costs.

Vultr Cloud GPUs bring video processing power to you

Organizations adopting GPU computing often have to choose between two infrastructures: cloud GPUs or on-premises GPUs. The case for cloud GPUs is getting stronger because as the processing demands for most organizations increase, the cost implication for on-premises cloud infrastructure rises in equal or greater proportion. Companies that opt for on-premises GPU implementations must account for hardware, personnel, and maintenance costs, which could increase steeply, especially for initial setups.

Cloud GPUs offer streamlined and managed solutions while optimizing cost and consumption. A key advantage of cloud GPUs is that they reduce the burden of complex setups and tuning. They can also be managed by your team anywhere in the world and are automatically maintained and upgraded. This may translate to lower adoption costs because your company incurs minimal or no infrastructure and personnel costs. However, the price is still often a barrier for companies that might otherwise use them.

Vultr solves the affordability problem with our Cloud GPU platform of fractional virtual NVIDIA GPUs at a fraction of the market rates.

Powerful cloud GPUs

Adding the NVIDIA A40 and NVIDIA A100 GPUs to Vultr Cloud GPU unlocks new opportunities for studios with video processing and augmentation workflows. Vultr Cloud GPU powered by the NVIDIA A40 provides optimized computing for professionals who handle processing-intensive tasks, such as visual effects rendering, computer-aided design (CAD), simulations, and AI modeling. The GPUs are optimized to handle enterprise-grade graphics software seamlessly, such as Cinema 4D, Blender, Maya, and Open Broadcaster Software (OBS), even when executing concurrent processes in distributed environments.

Vultr has lowered the adoption barrier to these technologies by simplifying the onboarding process for organizations while guaranteeing consistent performance regardless of workload. Cloud GPUs ensure global accessibility, which supports distributed (and remote) teams and projects.

Vultr has a simplified, robust, and intuitive control panel that promotes rapid configurations of its GPUs to enable engineers and designers to leverage the cloud NVIDIA GPUs. Also included is a developer-friendly application programming interface (API) and command-line interface (CLI), which diversify integration options with existing applications for further process streamlining and automation. Within minutes, you can have the GPUs provisioned and ready for processing, especially since the VM instances come with images for popular applications and operating systems ready for pre-installation. Images for most Linux distributions and Windows Server 2016, 2019, and 2022 are available, as well as one-click apps, including Anaconda, MLDev, NVIDIA NGC, and AIDungeon2.

Fractional GPUs

Traditionally, the leading cloud vendors have used a model where customers pay for entire GPUs, meaning small and large enterprises all have to adopt this model if they desire to have GPUs as part of their cloud infrastructure. The challenge is that cloud GPUs are very expensive to run, effectively raising the barrier to entry for small businesses. Even for large enterprises that can afford the exorbitant monthly fees, the traditional model elevates the risks of resource waste because the cost of running this infrastructure doesn't change even when usage is low.

Vultr's cutting-edge cloud GPU platform re-imagines the cloud GPU workflow with its advanced NVIDIA GPU virtualization platform. Vultr sets the standard for the delivery of fractional GPUs as a more efficient implementation of GPU-as-a-service through the virtualization of physical GPUs. Two of the world's leading GPUs, NVIDIA A100 and NVIDIA A40, power Vultr's virtual machines (VMs). By virtualizing these GPUs, customers can provision fractions of the processing time at a fraction of the standard cost while optimizing efficiency and performance.

Vultr pairs cloud VMs with a fraction of the physical GPU in the form of a virtual GPU (vGPU). On the user's side, the vGPU appears as a physical GPU, typically bundled with memory and computing power.

This model ensures optimal resource use and, most importantly, cost-effective processing of some of the most GPU-intensive AI workflows, including computer vision, natural language processing, and voice recognition.

GPUs for media

Today's streaming era shows that video data accounts for significant global Internet traffic. As content creation diversifies and expands, video and animation will continue to evolve, demanding computing resources that can handle intensive rendering or live video and animations. Even in traditional media and entertainment, such as film and television, production studios seek to increase entertainment value through high-quality content and VFX. As streaming becomes the default mode of content consumption, studios and creators are all working on upgrading their infrastructure to handle such intensive workloads in distributed environments.

Companies use GPUs to combine graphics computing systems with cutting-edge software, AI, and cloud technologies. Despite this immense potential in today's GPUs, organizations with intensive workflows and tight budgets may struggle to procure onpremises graphics infrastructure.

Cloud platforms such as Vultr offer GPUs as a service, enabling businesses to execute demanding graphics tasks such as real-time ray tracing, video transcoding, and programmable shading. This allows for the creation of modeling complex scenes that are visually appealing and as close to reality as possible and tackling previously difficult tasks such as real-time virtual productions and the generation of immersive three-dimensional worlds.

Not only do Vultr Cloud GPUs reduce setup and configuration time, but being on the cloud ensures customers can benefit from industry-standard security tools such as data encryption, security upgrades, network virtualization, and support when needed. This means even small companies can reap these benefits, leveling the playing field and allowing them to compete even with more established companies.

GPUs for engineering and manufacturing

GPU technologies have evolved to power more than just graphics rendering and AI workflows. Chip makers are stretching the bounds of GPU performance by developing graphics chips that can handle industrial designs for construction, manufacturing, graphics design, and engineering domains.

Futuristic GPUs can combine AI and engineering processes to achieve advanced and innovative techniques. For example, by applying an AI layer to ray tracing, new-generation GPUs can render directional light and shadows – processes that have significant utility in industrial design.

The refinements and efficiencies built into new GPU architectures have expanded the chips' computing power. Simulation workflows with long execution cycles now take a shorter time, enabling engineering teams to multiply their simulation scenarios and creating more opportunities for efficient and innovative designs.

Manufacturing is another industry that stands to benefit from high-performance GPUs. In the field of additive manufacturing (AM), for example, a common observation in the engineering community is that traditional computer-aided design may not work well as the primary design engineer when designing and manufacturing the next generation of 3D printers. This potential bottleneck has inspired firms to develop new design computation engines that leverage the power of GPUs to accelerate their performance and capacity to handle intensive AM tasks.

For 3D printers to produce sophisticated and precise designs, they rely on large high-resolution CAD files. The conventional approach to importing native CAD files into 3D printers and creating fabrications often takes hours without accelerated computing. Cloud GPUs and their support for accelerated computing engines present optimal environments for scaling the performance of such processes end-to-end so that customers can achieve all their manufacturing demands. GPUs specifically enable the processing of large manufacturing data to rapidly create 3D geometries that printers can produce in record time.

Cloud GPUs present innovation opportunities that industry-wide engineering teams can adopt. Engineering solution developers have grappled with the challenge of decoupled VM and GPU infrastructure, where the data is constantly transferred between storage and GPU for processing, creating process bottlenecks and hampering performance. A promising cloud solution for this problem is designing GPU architectures with data processing solutions wholly in the graphics chips. This approach reduces data exchange between different architectures and processes. GPUs could even power digital twins – virtual replicas of physical environments– as part of the 3D internet evolution, making the metaverse and other virtual worlds a reality.

Power complex video projects with Vultr

By bringing the power of NVIDIA GPUs to the cloud, Vultr makes these previously expensive resources accessible to teams with different graphics-heavy visual computing projects at reduced costs. These cloud GPUs are built to take on intensive workflows, including real-time video rendering, ray tracing, video transcoding, and visual effects rendering for media and entertainment.

Cloud GPUs are on course in manufacturing and engineering to revolutionize industrial processes such as additive manufacturing, 3D design and printing, and AI-based simulations. Customers can get fractional GPUs instead of expensive physical ones that on-premise implementations require. With fractional vGPUs, you pay for the resource you need with the option to scale when demand increases without the risk of resource wastage.

For top-notch resources at a fraction of the price of the big names, check out Vultr Cloud GPU.