RIGHT-SIZING 3D DESIGN RESOURCES WITH NVIDIA OMNIVERSE ENTERPRISE

Supermicro workstations and servers power workloads for small, large, on-premises or remote teams
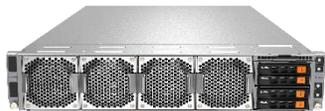
TABLE OF CONTENTS

SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

Executive Summary

3D design environments provide powerful visualization and simulation capability for many industries. Projects in media production, oil and gas exploration, medical imaging, electronic design automation (EDA) and mechanical computer-aided design and engineering (CAD/CAE) need advanced image rendering for design and analysis.

For teams in these and other industries, NVIDIA is transforming real-time collaboration around 3D designs with NVIDIA Omniverse™ Enterprise. It provides photorealistic 3D rendering and simulation in a workflow maximizing rendering iterations, streamlining review cycles, and enabling work from anywhere, on-premises or remote.

Supermicro is innovating solutions for Omniverse Enterprise deployments on a range of workstation and server hardware, right-sizing 3D design resources for each team. Three deployment configuration examples show a 6-user single workstation combining all needed capability, a pooled GPU 78-user rack-level deployment, and a new jointly developed OVX reference platform handling 64 heavy users or 256 light users in a rack-level solution with Ethernet or InfiniBand connectivity.

Reshaping 3D Design for Scalability

3D design acceleration began in the "box-on-desk" deployment model, with applications running on a workstation. Scaling a design team meant adding a designer and investing in another 3D graphics workstation for them. Workflows kicked off rendering jobs over lunch or at the end-of-day for overnight runs, followed by changes and more rendering. This 1:1 coupling between designer and workstation has limited scalability and made remote work difficult.

NVIDIA Omniverse Enterprise changes the 3D design scalability and remote work equation. It enables collaboration for creative teams on enterprise 3D visualization projects like never before. This new technology reduces costs and waste by enabling a simultaneous workflow, maximizes creativity by allowing new iterations without opportunity cost, and accelerates time to production by eliminating import-export workflows. Highlights of the Omniverse architecture:

- Rendering builds on advanced **NVIDIA RTX**™ accelerated real-time ray-tracing technology and adds dynamic allocation of graphics processing unit (GPU) cores, even from separate machines, for rendering jobs.
- Collaborative scene composition is enabled by Pixar's **Universal Scene Description** (USD), an extensible, open-source file framework that allows the interchange of 3D assets.
- **NVIDIA Omniverse**™ **Nucleus** database and collaboration engine manage data such as geometry, lights, materials, textures, and more for defining a virtual "world" in and around a design, including its environment.
- **NVIDIA Omniverse**™ **Connect** allows 3rd party client applications access to the Nucleus database using publish-subscribe messaging, letting many designers author live changes that everyone connected sees in real-time.
- **NVIDIA Omniverse**™ **Kit** is a toolkit for lightweight plugins building applications and microservices.
- **NVIDIA Omniverse**™ **Create** enables world-building, using advanced USD workflows such as layers, variants, instancing, animation caches, and more.
- **NVIDIA Omniverse**™ **View** lets non-technical users review 3D content interactively in full, photorealistic fidelity.
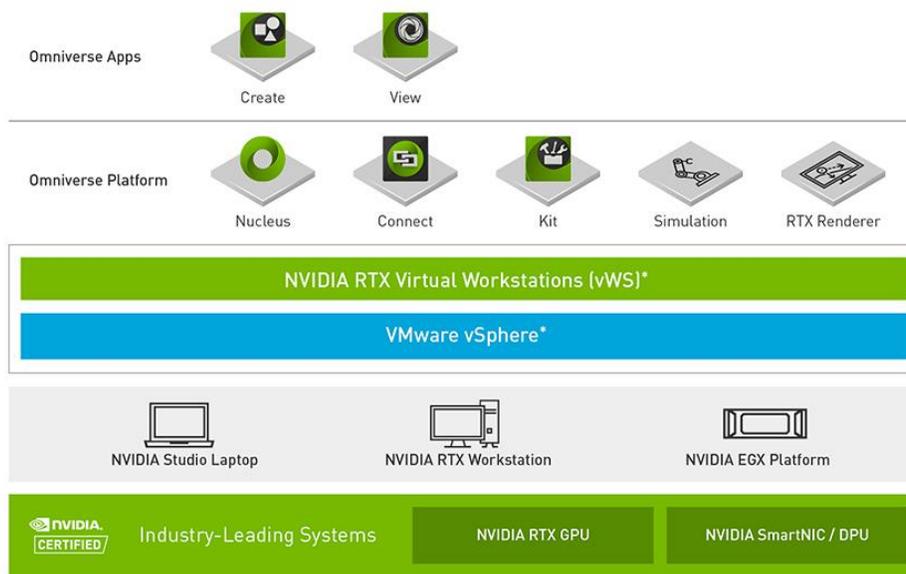
Figure 1 - NVIDIA Omniverse Software Stack - Image Courtesy of NVIDIA

Using Supermicro hardware platforms, 3D design teams can select a hardware topology and mix and match configurations for an Omniverse Enterprise solution best fitting their needs. For example, "creators" can tap into the full capability needed for design, while "reviewers" can run on lighter hardware platforms at a reduced software licensing cost. In addition, deployments are open to various workstations, deskside, and rack-scale servers right-sized for teams, projects, and workloads.

March 2022

Engineered Solutions for 3D Design Teams

Supermicro's server technology is proven through years of deployments in enterprise, data center, cloud computing, 5G telecom infrastructure, high-performance computing (HPC), edge computing, and other IT applications. US-based engineering and manufacturing teams provide innovative features, and ISO 9001-certified quality customers expect. These teams design and produce Supermicro motherboards, power supplies, enclosures and racks, and other components, increasing supply chain control. The firmware allows unlocked peripheral integration at the factory or by customers. Supermicro selects the right technology for delivering better price/performance with improved energy efficiency through its partnerships.

A key enabler for 3D design is NVIDIA's graphics processing unit (GPU) technology. Supermicro currently offers over 60 workstation and server configurations with installed NVIDIA graphics cards such as the NVIDIA RTX™ A6000 or NVIDIA A40. With the debut of NVIDIA Omniverse Enterprise, Supermicro is fine-tuning new workstation and server configurations matching hardware allocation to user performance profiles, maximizing resource usage.

Flexing GPU muscle for a variety of workloads

Supermicro excels at creating denser NVIDIA RTX GPU-based configurations targeting light to heavy Omniverse Enterprise workloads. The combination of cutting-edge AMD or Intel chipsets, fast PCI Express slots, Supermicro-designed high efficiency 80 Plus Titanium power supplies, and thermally simulated system cooling designs enable multi-GPU workstations and servers. In addition, high-speed Ethernet switching and IPMI 2.0 remote management help connect and manage deployments. Some possible use cases for various Supermicro hardware configurations:

- A 1:1 model, with a single GPU in an entry-level workstation for each designer, still works for light 3D design workloads
- Small 3D design teams can render on-demand using a more powerful workstation with several GPU cards installed
- Multi-location teams can use one medium-sized server with multiple GPU cards, connecting remotely over Ethernet
- A heavy media entertainment workload can deploy a rack-scale solution with up to 64 pooled GPUs from several servers

Expertise for Omniverse Enterprise deployments

Omniverse Enterprise is rapidly growing and being adopted quickly by customers as a new and exciting solution. NVIDIA featured Omniverse Enterprise during its CES 2022 presentation. Supermicro, as an NVIDIA ecosystem partner, keeps pace with new GPU applications, including artificial intelligence (AI) and on-demand 3D rendering. Initial deployments of the use cases above built vital Supermicro expertise with Omniverse Enterprise, and Supermicro can help deploy more solutions now. A typical Omniverse Enterprise engagement with Supermicro looks like this:

- A Supermicro solution architect interviews customer teams about roles, workloads, and deployment options
- Supermicro technology enablement teams look at specific requirements and configurations for workstations/servers
- Supermicro product teams install Omniverse Enterprise at the factory and launch clients on the customer's premises

During these engagements, the goal is always to provide efficient compute power where a team needs it for their workflow. For example, a solution architect can help size a platform for a Nucleus database and its Large File Transfer (LFT) technology, allocating disk storage and memory-based cache. Another example includes a rack-mount Ethernet switch for remote connections – each remote 4K display needs at least 1Gbit of Ethernet bandwidth to run at 60fps without latency.

Supermicro is the first solution provider to partner with NVIDIA on an OVX reference platform for low user count and high-intensity use cases. No other hardware vendor has yet deployed Omniverse Enterprise in rack-scale, high user counts solutions. Offering a wide variety of options, Supermicro can address more scenarios for Omniverse Enterprise deployments with results customers can count on – like three described next.

March 2022

Three Deployment Configuration Examples

For small teams of up to six users, one Supermicro A+ SuperWorkstation hosts a Nucleus server, an NVIDIA RTX™ Virtual Workstation (vWS), and VMware virtualization software.

A 64-core AMD RYZEN™ Threadripper™ Pro 3995WX processor handles threading in this configuration. Virtualization storage comes from 4x mirrored, striped 3.84TB m.2 SSDs, while 2x 7.68TB U.2 SSDs provide a Nucleus file store. Two NVIDIA RTX A6000 GPUs handle rendering. Dual-port 25GbE handles virtual connections for remote users.

Resource allocation creates a dedicated workspace for each user on the workstation, plus threads and storage for the Nucleus server and the vWS licensing server.

NVIDIA Omniverse Enterprise Nucleus Server VM: 32 vCPU threads, 52GB DRAM, 6TB Gen4 NVMe mirror NVIDIA RTX vWS license server VM: 4 vCPU threads, 8GB DRAM, 20GB SSD
6x User VM, each with: 10 vCPU threads, 32GB DRAM, 16GB vGPUs frame buffer, 960GB Gen4 NVMe
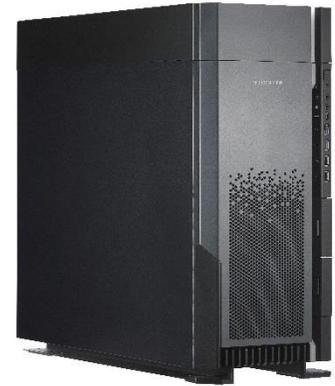
Image 1 - AS -5014A-TT Workstation

For a flexible GPU pool handling up to 78 users, a Supermicro rack-scale solution includes a standalone Nucleus server, 13 rendering servers, and a 100GbE switch.

The Omniverse Enterprise Nucleus Server is a 1114S-WN10RT, running an AMD EPYC™ 73F3 with 16 cores at 3.5GHz and 128GB DDR4-3200 DRAM. Storage is 6x 3.84TB U.3 SSDs. Two dual-port 100GbE connections feed the Ethernet switch for smooth database access.

Thirteen 2U-2 node user servers are either 2114GT-DNR (Intel® Xeon® W-3365 with 32 cores at 2.7GHz) or 210GP-DNR (AMD RYZEN™ Threadripper™ Pro 3975WX with 32 cores at 3.5GHz), each with 256GB DDR4-3200 DRAM, two 1.92TB U.3 SSDs, and two NVIDIA A40 GPUs. Dual-port 25GbE handles virtual connections.

Resource allocation creates a dedicated workspace for each user.
78x User VM, each with 16 vCPU threads, 80GB DRAM, 16GB vGPU frame buffer, 960GB SSD

Image 2 – 13 x 2U-2Node Servers

Developed jointly with NVIDIA, the SYS-420GP-TNR OVX reference platform targets heavy rendering or simulation requirements with GPU passthrough capability.

A 4U server houses eight NVIDIA A40 GPUs, each dedicated to a user or virtualized for multiple users. Processing is two Intel® Xeon® SP 8362 with 32 cores at 2.8GHz, backed by 1TB of DDR4 ECC DRAM. Storage is two 960GB NVMe boot drives and two 7368TB NVMe storage drives. In addition, two Mellanox ConnectX®-6 VPI cards provide Ethernet/InfiniBand connectivity at up to 200 GB/sec.

Eight Supermicro SYS-420GP-TNR servers can be combined with a 220U-TNR standalone Nucleus server and an Ethernet/InfiniBand switch in a rack-scale solution for mixed resource deployments. In a heavy usage scenario, each user gets a full NVIDIA A40 GPU for a total of 64 users per rack. The NVIDIA A40 GPUs can be virtualized with allocated CPU threads and storage for up to 256 users in a rack for medium and light resource usage.

Image 3 - 8 x SYS-420GP-TNR

March 2022

Bringing In the Right Skills

NVIDIA Omniverse Enterprise breaks out of the 1:1 model where a 3D design workstation dedicated to every single user was the only option. It relieves the hassle of planning workflows around scheduled rendering iterations. Instead, it enables on-demand rendering from workstations or servers available to all users in real-time. It also allows teams to staff and scale with flexibility. Teams can grab design and project talent with the right skills based anywhere with high-speed internet access.

Supermicro is a natural fit as a solution provider for Omniverse Enterprise deployments. With an in-depth understanding of architecting deployments from small teams to large media entertainment projects, Supermicro can match a team's use case to right-sized solutions saving energy and reducing the total cost of ownership. These solutions continue to evolve, as seen in a new, jointly developed reference platform:

"Supermicro's success in Omniverse Enterprise deployments ranging from small design teams to large, geographically distributed project teams caught our attention," says Bob Pette, Vice President, Professional Visualization at NVIDIA. "We created the NVIDIA OVX SYS-420GP-TNR reference platform working together for a proven configuration that ships quickly, getting customers up and running on Omniverse Enterprise faster."

The range of solutions Supermicro brings to 3D design and simulation teams – workstations, servers, and rack-scale – is formidable. It comes from years of experience supplying IT applications. With access to advanced parts from NVIDIA, AMD, Intel, and many other providers, plus in-house design capability for motherboards, power supplies, enclosures, and more, Supermicro can draw configurations together with the quality, deliverability, and performance needed.

Teams turning to Omniverse Enterprise should focus on revolutionizing their 3D content creation and simulation pipelines. Having wrong-sized hardware can get in the way of success. Worry-free projects start with a call to Supermicro, tapping into Omniverse Enterprise capability and deployment expertise no other solution provider offers.

For more information on Supermicro solutions for NVIDIA Omniverse Enterprise, please visit

https://www.supermicro.com/en/products/nvidia-omniverse

March 2022