



SUPERMICRO AND STRATUSCORE TEAM UP FOR GENERATIVE AI AND DESIGN WORKFLOW SOLUTION

Enterprise AI Becomes Available to Entire Organizations

TABLE OF CONTENTS

Executive Summary	1
Solution Highlights	2
For More Information	8
Additional Resources	8



SYS-751GE-TNRT

Executive Summary

Generative AI and Generative Design have emerged as critical considerations in various industries such as Media & Entertainment, Product Design, Manufacturing, Gaming, and AEC. Efficiently building and managing secure turnkey environments for these industries is paramount. To achieve a thriving Generative AI and Design environment, three primary challenges need to be addressed:

1. Adequate computing power: A powerful machine, workstation, or server is essential to handle complex computational requirements.
2. Data and workflow management: Effectively managing the data and image workflows associated with content generation is crucial.
3. Orchestration layer: An orchestration layer is needed to enable seamless communication and coordination between the

infrastructure and the workflow, yielding desired results.

Supermicro and StratusCore have collaborated to develop the industry's first turnkey SuperWorkstation designed specifically for Generative AI and Generative Design. This SuperWorkstation caters to the needs of DevOps administrators, managers, and other end users, empowering them to manage, control, and scale Generative Design workloads for both local and remote content creation teams.

The inclusion of StratusCore's no-code RAVEL Orchestrate™ further enhances the practical utilization of our AI SuperWorkstation within the content creation industries, allowing teams to quickly and efficiently commence their work. By managing Supermicro's Liquid Cooled AI SuperWorkstation's technical resources and other ecosystem components, this solution provides customers with a comprehensive and integrated approach.

This solution enables complex workflows and heavy workloads to be efficiently executed within a desk-sized space or easily integrated into existing infrastructures. Its small power footprint ensures a lower total cost of operation and ownership, making it a clear advantage for all stakeholders.

Today's DevOps customers, who require swift and efficient initiation of complex workflows with substantial workloads for their content creation teams, will find this solution highly beneficial. While the SuperWorkstation excels as a dedicated AI compute machine, it also offers the versatility of being transformed into a virtualized command center through RAVEL Orchestrate™. This solution allows administrators to deploy virtualized technical resources to local and remote content creative team members, showcasing unparalleled range and flexibility.

This solution effectively addresses the needs of modern content creation workflows by offering quick start capabilities, managing large workloads, and providing extensive utility through its SuperWorkstation and virtualized command center functionalities facilitated by RAVEL Orchestrate™.

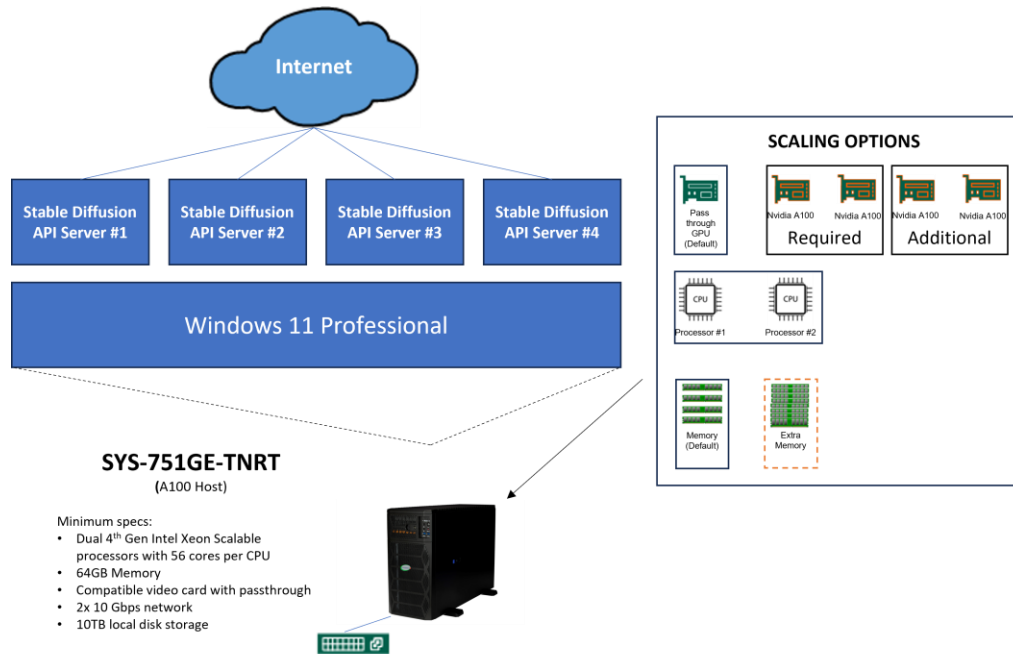
Solution Highlights

A deeper look at Supermicro's Secure Liquid Cooled SuperWorkstation integrated with RAVEL Orchestrate's SMART Assembly™:

Creators and intellectual property (IP) owners are increasingly interested in harnessing the potential of Generative AI and Generative Design technologies. They have a range of options at their disposal, including open-source technologies such as Stable Diffusion, Open AI, and Midjourney, as well as Adobe's Firefly, which offers a content authenticity model. Recognizing users' diverse needs and preferences, the SuperWorkstation solution ensures flexibility to support exploration and IP protection.

This collaborative solution allows creators and IP owners to freely explore the open-source technologies mentioned, empowering them to experiment and innovate with Generative AI and Generative Design. Simultaneously, it caters to those who prioritize safeguarding their intellectual property via engines such as Adobe's Firefly, which offers robust content authenticity measures. By providing support for both options, our SuperWorkstation solution ensures that users can choose the best approach to their objectives, whether exploration or protection of their valuable IP assets.

Use Case #1 –The Generative AI/Design Compute Working Server (GPU SuperServer)



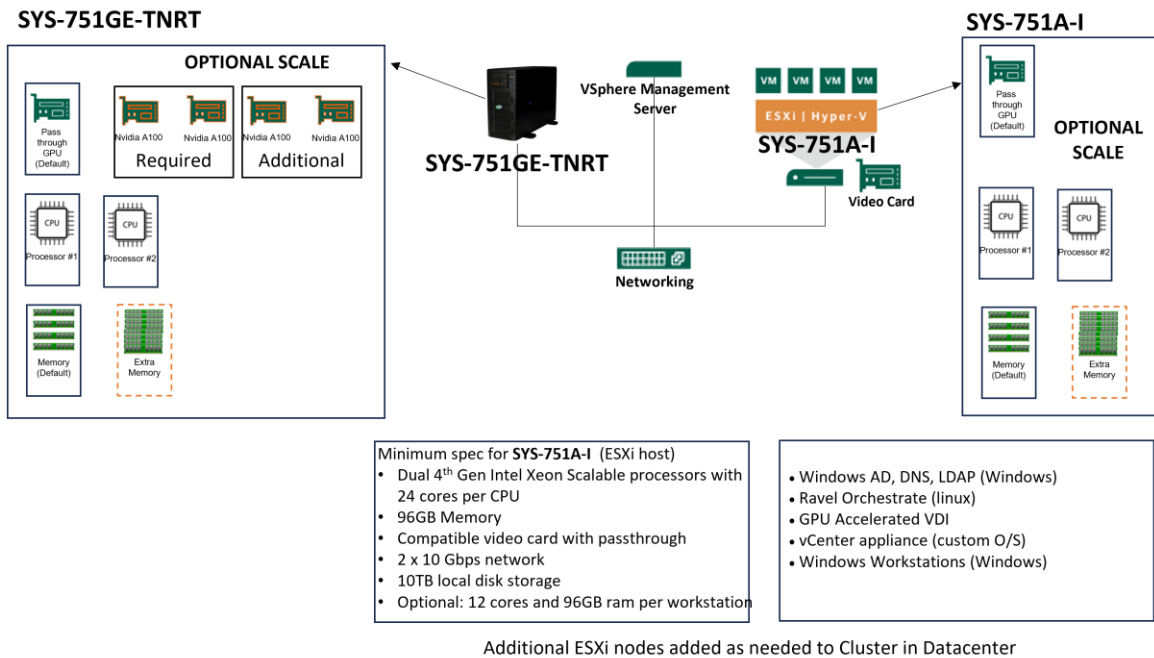
In this specific use case, an existing environment is already in place, whether on-premise, in the cloud, or a hybrid setup. The objective is to integrate the Liquid Cooled Workstation seamlessly with the existing systems. To accomplish this, two primary requirements need to be addressed:

- Validate the equipment and architecture outlined above as the optimized and approved schematic that will maximize the utilization of the Liquid Cooled AI Workstation.
- Deploy RAVEL Orchestrate™ to manage the associated infrastructure and environment efficiently.

By implementing RAVEL Orchestrate™, the Generative AI and Generative Design workflows can be effectively managed in conjunction with virtual workstations located either on-premise or in a data center. It ensures seamless integration with the associated infrastructure components, including license servers, storage area networks (SANs), network-attached storage (NAS), and more. Additionally, the flexibility of RAVEL Orchestrate™ extends to managing and enabling virtual environments for various content creation workflows.

This configuration provides comprehensive support for AI engine training, inference, 3D modeling, rendering, high-powered virtual workstations, as well as persistent servers and workstations. Moreover, it enables remote users to access and utilize the system efficiently, regardless of location.

Use Case #2: A clustered environment for Generative AI and Generative Design for SMBs or smaller installations with the ability to scale and collaborate with other team members.

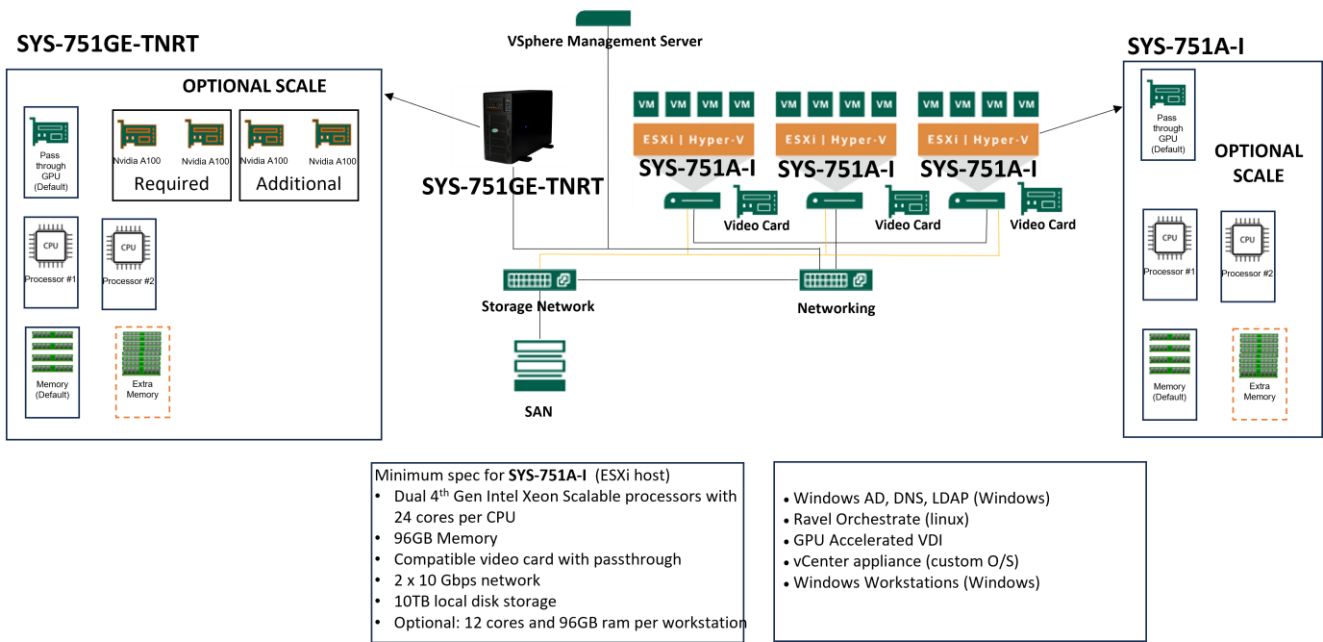


This solution is scalable in an established environment, whether on-premise, in the cloud, or in a hybrid setup, and this configuration provides a streamlined infrastructure solution. It focuses on deploying a compact and scalable system that efficiently supports GPU and AI/ML generative content production. This setup is particularly suitable for demonstration purposes, allowing for easy node expansion to enhance disaster recovery and ensure high availability in production scenarios.

RAVEL Orchestrate™ is critical in managing Generative AI and Generative Design workflows, seamlessly integrating with virtual workstations on-premise or in a data center. It effectively handles associated infrastructure components such as license servers, storage area networks (SANs), network-attached storage (NAS), and more. Furthermore, RAVEL Orchestrate™ exhibits the necessary flexibility to manage and enable virtual environments for any content creation workflow.

This configuration provides robust support for various tasks, including AI engine training, inference, 3D modeling, rendering, high-powered virtual workstations, and persistent servers and workstations. Moreover, it offers the advantage of facilitating remote user access, effectively allowing individuals to utilize the system regardless of physical location. Design achieving unprecedented peak performance.

General Use Case #3: A core cluster ready to scale immediately for medium and larger organizations and team members.



This configuration is intended for teams that have already established their environment, whether on-premise, in the cloud, or a hybrid setup and are prepared to leverage the power of GenAI/Design for their operations.

Designed to optimize medium to large-scale infrastructures, this configuration is adaptable to various established environments, including on-premise, cloud, and hybrid systems. Utilizing virtualized servers enables efficient GPU and AI/ML-driven content generation. This solution is particularly well-suited for production environments where high availability and disaster recovery mechanisms are already in place. It provides an optimal approach for deploying GPU-enabled workstations on a medium-to-large scale, regardless of whether the operating environment is cloud-based, hybrid, or on-premise.

RAVEL Orchestrator™ is the management tool for overseeing Generative AI and Generative Design workflows, seamlessly integrating with virtual workstations on-premise or in a data center. It effectively handles the associated infrastructure components, including license servers, storage area networks (SANs), network-attached storage (NAS), and more. Furthermore, RAVEL Orchestrator™ is versatile enough to manage and enable virtual environments for any content creation workflow.

This configuration offers comprehensive support for a wide range of tasks, such as AI engine training, inference, 3D modeling, rendering, high-powered virtual workstations, and persistent servers and workstations. Additionally, it facilitates remote access, allowing users to operate the system efficiently from anywhere.

Supermicro, in connection with StratusCore's RAVEL Orchestrator™, supports the following Generative AI and Design workflows:

AutoDesk, Inc	Adobe, Inc	AVID, Inc.	AI Generative Software	Knowledge Generative	Game Development
Media & Entertainment Collection including: Maya, 3DS Max, Motion Builder,	Creative Cloud Including: Photoshop, AfterEffects, Premiere Pro	NEXIS Storage	Stable Diffusion	Llama - Large Language Models (LLM)	Unity Unreal
Architecture, Engineering, and Construction: Collection including AutoCAD, Revit, Civil 3D	Generative AI & Design: Plug-in support includes Stability Diffusion,	File Transfer Service Media Central - Production Engine Media Index Production Services Cloud UX Ingest Asset Management	Stability-AI StableStudio Lamda	Stanford Alpaca	
Product Design & Manufacturing: Collection including Inventor, AutoCAD, Fusion360					

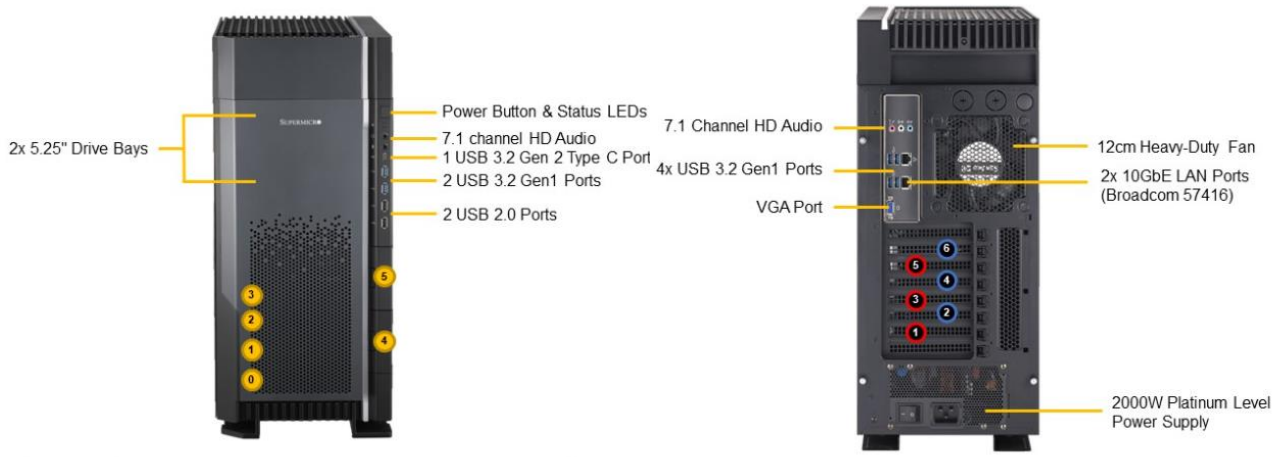
*Other uses cases, including Life Sciences and AI & Deep Learning, are also applicable

Supermicro Hardware Configurations:

Product Type	Supermicro Product	Processor	Memory	OS	Expansion	Network	Storage	Display	Power
A-100 box for Gen AI/Design	SYS-751GE-TNRT	Dual 4 th Gen Intel Xeon Scalable processors 8480+ - up to 112 Cores	64GB - Up to 1 TB DDR5 RAM	Windows 11 Pro	2 or 4 NVIDIA A100 80GB GPU	2x 10GB Base T Lan ports	1 -2x NVME for OS and Apps & 1 - 8x 4TB NVMe	PNY NVIDIA T1000	1+1 2200W Redundant 80 PLUS Titanium Level
ESXi Host	SYS-751A-I	Dual 4 th Gen Intel Xeon Scalable processors 8480+ - up to 112 Cores	256GB - Up to 1 TB RAM	VMWare ESXi v8 u3 with NVIDIA Drivers		2x 10GB Base T Lan ports	2x 6.7TB 2.5" Magnetic (OS and App installation) RAID Cluster	2x NVIDIA RTX 4000	2000W PS2 Multi-output 80PLUS Platinum Level
ESXi Host - Storage for	SSG-631E-E1CR16L	-	-	-	-	-	Storage SAN - iSCSI or FC connection	-	-

vSphere							. RAID 5 or better		
ESXi - Virtual workstations	-	16 Core - Up to 96 per compute resource	96GB of RAM	Windows 10 Pro, 11 Pro, Server 2019, or Server 2022	-	-	-	GPU Pass-through required for graphic acceleration	-

The Supermicro SYS-751A-I workstation is an air cooled system. Specifically, the layout is detailed below:



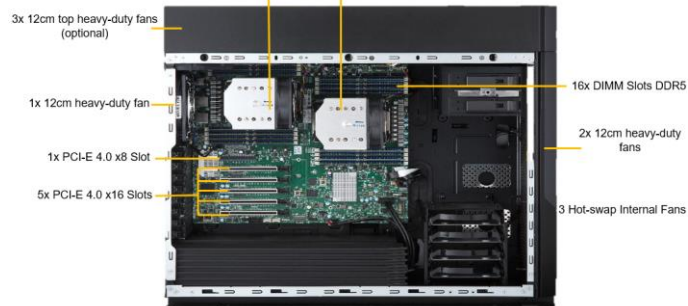
Drive Bay	Description
0 - 3	Default - 4 Hot-swap 3.5" SATA3 HDD/SSD drive bays (SAS3 supported with optional add on card and cables)
5 - 4	2 Hot-swap 2.5" SSD (optional)

Slot	Slot Description
1 - 5	PCIe 5.0 x16
6	PCIe 5.0 x8

SuperWorkstation SYS-751A-I

(Top View - System)

Dual Socket E (LGA-4677) Intel® 4th Gen Xeon® Scalable Processors



The Trial and the “Verify” Opportunity:

The Liquid Cooled AI SuperWorkstation powered by StratusCore’s RAVEL Orchestrate™ is now available for trial. The trial program is called Verify. StratusCore and Supermicro have collaboratively developed a comprehensive turn-key solution allowing end users to test and experience a sophisticated Generative AI and Generative Design workflow, executed on Supermicro's AI SuperWorkstation, and integrated with RAVEL Orchestrate™, a no-code SMART Assembly™ product, ensuring seamless workflow management. This demonstration leverages a combination of cutting-edge technologies, including:

1. Open source latent text-to-image and image-to-image diffusion models, utilizing Stable Diffusion, to generate high-quality content.
2. Adobe's Creative Tools, such as Adobe® Photoshop® and Adobe® Premiere® Pro , enable creative enhancements and content refinement.
3. Autodesk's Maya and 3DS Max provide robust 3D modeling and animation capabilities.
4. Virtualization technology for efficient resource allocation and management.
5. Secure access control and identity management to ensure data integrity and user authentication.

Sign-up for our VERIFY turnkey Proof Of Concept:

Select customers and partners are granted virtualized access to the SuperWorkstation trial, hosted at Supermicro's state-of-the-art lab in San Jose. Please note that availability for these trials is limited and provided on a first-come, first-serve basis. For further information and to sign up for this valuable opportunity to trial the solution, please visit the following link: <https://ravelinc.com/orchestrate-verify-supermicro/>

For More Information

For Integrators & Resellers

Interested in partnering to host your own Generative AI and Design POC, please get in touch with us at <https://ravelinc.com/partner/#> .

Additional Resources

To learn more about Supermicro’s Rack Scale AI Solutions, please visit:

- <https://www.supermicro.com/en/products/gpu>
- <https://www.supermicro.com/en/products/superworkstation?pro=cpu%3D2>

- <https://www.supermicro.com/en/support/resources/cpu-4th-gen-intel-xeon-scalable>

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.

Visit www.supermicro.com

STRATUSCORE / RAVEL

RAVEL Orchestrate's SMART Assembly™ feature plays a vital role in this demonstration by facilitating rapid deployments, scheduling, and management of various software images on the SuperWorkstation's virtual machines. This solution allows for seamless coordination and optimization of the workflow, ensuring the smooth execution of tasks within the Generative AI and Generative Design environment.

Visit www.ravelinc.com
