



Table of Contents

Why Supermicro Ultra System for VDI

2 Introduction

Key Benefits of a High Graphics Performance VDI

3 High Graphics Performance VDI Solutions

High Graphics Performance VDI Use Cases

GPU Options For High Graphic Performance VDI

NVIDIA Virtual GPU Software*

Configure Your Supermicro Ultra High Performance VDI

6 Supermicro 1U and 2U Systems + NVIDIA GPUs

Ultra 1U 1029U VDI with High Graphic Performance

Ultra 2U 2029U VDI with High Graphic Performance

Ultra 2U 2029U All-Flash NVMe with 2x NVIDIA[°] Tesla[°] GPUs VDI

9 Summary

Super Micro Computer, Inc. 980 Rock Avenue San Jose, CA 95131 USA www.supermicro.com

White Paper

High Performance GPU Accelerated Virtual Desktop Infrastructure Solutions with Supermicro Ultra SuperServers



Why Supermicro Ultra System for VDI

Supermicro Ultra rackmount servers are engineered to offer unrivaled performance, flexibility, scalability and serviceability and therefore ideally suited to CPU and memory intensive workloads.

Ultra systems feature dual sockets, 24 DIMM slots, and support high performance GPUs with up to 32GB video memory. The 1U systems can support up to 10 hot-swap 2.5" drive bays, while the 2U systems can support up to 24 drives to accommodate a wide range of application requirements.

Supermicro Ultra systems feature best cooling design for up to 165W CPU TDP with 56 cores/112 threads when GPU is installed, which can support up to 112 virtual machines with a 1:1 vCPU to physical thread ratio.

Introduction

VDI has become increasingly popular with the broadened availability of mobile broadband, higher bandwidth, more compute power in modern mobile devices, higher display resolutions and a trend of BYOD within enterprises.

Supermicro has developed a line of fully validated server solutions featuring high performance GPUs from NVIDIA, along with NVIDIA virtual GPU (vGPU) software, to address the rapidly growing Virtual Desktop Infrastructure (VDI) market. These solutions include a Supermicro Ultra, FatTwin[™] and 4U Tower SuperServers configured with dual Intel[®] Xeon[®] CPUs, up to 3TB DDR4 memory and NVIDIA[®] vGPU software with Tesla[®] GPUs based on the NVIDIA[®] Volta, Pascal[™], or Maxwell[™] architectures that deliver best-in-class vGPU software with virtual desktop experience to users anytime and anywhere.

NVIDIA vGPU technology brings the power of NVIDIA^{*} GPUs to virtual machines with an immersive user experience for anyone, from knowledge workers to engineers and designers. NVIDIA^{*} vGPU technology meets the increasing performance demands of modern apps and Windows 10, bringing graphics and compute acceleration to the data center, enabling IT to centralize apps and data, and deliver virtual desktops that offer improved management security and productivity.

Businesses can now deliver cost-effective high graphics performance with greater scalability, and transform traditional workflows and liberating users from the confines of PC's and offices, to enable real-time collaboration from anywhere, anytime, on any device.



Certified Hypervisors:

- VMware ESXi 5.5 U2 or newer
- Citrix XenServer 6.2 or newer

Key Benefits of a High Graphics Performance VDI

Simplified Deployment

Allocate GPU resources precisely during the assessment and proof-of-concept phase and ensure an optimized user experience from day one.

- **Multiple Monitor** Support multiple displays for up to four 4K monitors* per virtual machine, improving productivity and user experience.
- Native Performance Experience

With NVIDIA[®] vGPU software and Tesla[®] GPUs, virtual desktops can perform like native PCs to every user.

• Linux OS Applications

Deliver like-native performance of virtual desktops to software developers and EDA designers that rely on a Linux-based development environment.

• Optimized for Microsoft Windows 10

Windows 10 is Microsoft's most graphically intensive operating system to date, with common productivity applications like Office 365, Skype and YouTube requiring GPU acceleration to function properly, so many general purpose VDI users require a GPU for better user experiences.

High Graphics Performance VDI Solutions

High Graphics Performance VDI Use Cases

| Target Market | Use Cases | Common Applications |
|---------------|---|--|
| Healthcare | For remotely viewing and editing very large and complex medical images (PACS) | EMR , PACS, Eclipse Medical Imaging General purpose VDI with multi-monitor support and Windows 10. |
| Manufacturing | For remotely viewing and editing very large 3D models and images | PTC Creo, Siemens PLM NX, Autodesk AutoCAD, Dassault Systèmes SOLIDWORKS and CATIA, ANSYS Discovery Live and Fluent. General-purpose VDI using virtualized design and creative apps such as Adobe Creative Suite Linux-based EDA developer apps such as Cadence and Synopsis |
| Government | For remotely viewing, analyzing, and editing very large 3D models and images | ESRI ArcGIS, Autodesk AutoCAD, ANSYS Discovery Live and Fluent, Adobe Creative Cloud, MATLAB. General-purpose VDI and common office productivity apps. |
| Finance | For running network heavy apps (Bloomberg, Reuters, Eikon) on multi-monitors. | Bloomberg, Reuters, and other electronic trading platforms. Many financial services firms have propriety applications that they are gradually updating to support Windows 10. |
| AEC | For remotely viewing and editing very large 3D project files and images | Autodesk 3ds Max, AutoCAD and Revit, Ansys Fluent, Adobe Creative Suite, Bentley MicroStation, AECOsim, Nemetschek Vectorworks, Allplan, Trimble SketchUp, GRAPHISOFT ARCHICAD, Newforma Project Center. |

* NVIDIA GRID vPC software supports up to two 4K displays or four QHD displays. NVIDIA Quadro Virtual Data Center Workstation software supports up to four 4K displays.



GPU Options For High Graphic Performance VDI

NVIDIA[®] vGPU software brings graphics and virtualization capabilities to NVIDIA[®] Tesla[®] data-center deployments, and is currently supported on the following Tesla[®] GPUs.

| Supported Tesla° GPU | M10 | T4 | V100 |
|-------------------------|---|---|---|
| Recommend vGPU S/W | GRID [®] vPC | Quadro vDWS | Quadro [®] vDWS |
| GPU | 4 Maxwell GPUs | 1 NVIDIA Turing GPU | 1 Volta GPU |
| CUDA Cores | 2,560 (640 per GPU) | 2,560 | 5,120 |
| Memory Size | 32 GB GDDR5 (8 GB per GPU) | 16 GB GDDR6 | 16 GB HBM2 32 GB HBM2 |
| Use Case | Knowledge workers using modern productivity apps and Windows 10 requiring maximum user density and best TCO | Entry-level to high- end 3D design and engineering work- flows with Quadro vDWS | Ultra-high-end ren- dering, simulation, 3D design |

Hardware Resource Requirements

- CPU
- Recommend more CPU cores for higher VM scalability for Knowledge worker
- Recommend higher clock CPU for high performance professional application
- High Memory Speed and Capacity to provide VM
- Graphic Application needs lots of memory.
- High speed storage to provide high speed throughput and IO
- Large video, rendering files read/write needs high speed throughput and IO drives.
- High speed network interface to provide systems connections
- At least 10GbE is require. VMware's VM provide 10GbE VM connections.
- High Performance GPU with virtualization and direct share capability
- Also need higher capacity of GPU memory to provide larger quantity of VM.

NVIDIA Virtual GPU Software*

NVIDIA vGPU software is available in three editions: *NVIDIA GRID*^{*} *Virtual PC (GRID vPC)*, *GRID Virtual Applications* (GRID vApp), and *NVIDIA Quadro*^{*} *Virtual Data Center Workstation* (Quadro vDWS). The table below shows some feature comparisons.

| | Workstations | Desktops | Applications | |
|-----------------------------------|---|---|--|--|
| vGPU SW Editions | Quadro vDWS | GRID vPC | GRID vApps | |
| | Engineers/Architects/Designers | Knowledge/Business Worker | | |
| Use Case | For professional graphics applica- tions mainstream and high-end designers. | For virtual desktops delivering standard PC applications, brows- er, and high definition video. | Use with Citrix XenApp or other RDSH solutions like VMware Horizon Apps. | |
| Desktop Virtualization | Yes | Yes | | |
| RDSH App Hosting ³ | Yes | Yes | Yes | |
| RDSH Desktop Hosting ³ | Yes | Yes | Yes | |
| Windows Guest OS | Yes | Yes | Yes | |
| Linux Guest OS ¹ | Yes | Yes | | |
| Maximum Displays | 4 | Four QHD, Two 4K ¹ | 12 | |
| Maximum Resolution | 4096x2160 (4K) | 4096x2160 (4K) | 1280x1024 | |
| Data Sheet | Quadro vDWS Brief | GRID vPC Brief | GRID vApps Brief | |

¹ Support starting with NVIDIA virtual GPU software Spring 2018 release (version 6.0)

² GRID vApps supports one 1280x1024 display from the GPU card. However, XenApp renders to an offscreen buffer, so it can support multiple software-rendered displays at higher resolutions.

³ With packaged GRID vApps license

* Please review the NVIDIA Virtual GPU Packaging, Pricing and Licensing.

Configure Your Supermicro Ultra High Performance VDI

| Customer Need | System | CPU | NVIDIA GPU | NVIDIA vGPU SW |
|---|---|--------------------------|---|---------------------------------|
| Standard VDI: For Knowledge worker using common office productivity apps and Windows 10. | Ultra 1U 1029U/6019U With up to 24x DIMM And up to 10x Storages | 2x up to 28C/56T CPUs | 1x NVIDIA Tesla M10 32 Users' with 1B Profile | 32 GRID vPC Licenses |
| Advanced VDI: For general purpose VDI with multi monitor support on Windows 10 and engineers who work in 2D design workload. | Ultra 2U 2029U/6029U With up to 24x DIMM And up to 24x Storages | 2x up to 28C/56T CPUs | 2x NVIDIA Tesla M10 32 Users ¹ with 2B Profile ² 2 x 4k or 4 x HD displays ² | 32 GRID vPC Licenses |
| Entry-Mid End VDI : For Engineers, designers, CAD/CAE users for remotely viewing and editing very large 3D models and images. | Ultra 2U 2029U/6029U With up to 24x DIMM And up to 24x Storages | 2x up to 3.5GHz CPUs | 4x NVIDIA Tesla T4 32 ~ 16 users¹ with 2Q ~ 4Q profile | 32 ~ 16 Quadro vDWS Licenses |
| High End VDI: For Architects, Engineers, and Designers for high-end rendering, 3Ddesign and engineering workflows. | Ultra 2U 2029U/6029U With up to 24x DIMM And up to 24x Storages | 2x up to 3.5GHz CPUs | 2x NVIDIA Tesla V100 8 ~ 4 Users ¹ with 8Q ~ 16Q Profile | 8 ~ 4 Quadro vDWS Licenses |

¹ Number of users are varies depending on the application and customer workload.

² Support starting with NVIDIA virtual GPU software Spring 2018 release (version 6.0)









Supermicro 1U and 2U Systems + NVIDIA GPUs

Ultra 1U 1029U VDI with High Graphic Performance



Dual-socket support up to 52 cores / 112 threads

| Technical Specifications | |
|--------------------------|---|
| Processor Support | Dual Intel[®] Xeon[®] Scalable processors Up to 165W CPU TDP with one or more GPUs |
| Memory Capacity | • 24 DIMM slots, up to 6TB DDR4-2933 ECC RDIMM/LRDIMM |
| PCI-E Expansion Slots | 2 PCI-E Gen 3.0 x16 (2 full height 10.5" length) 2 PCI-E Gen 3.0 x8 (1 low profile, 1 internal low profile) |
| Networking | -TR4: 4x 1GbE ports -TR4T: 4x 10Gbase-T ports -TRT: 2x 10Gbase-T ports -TRTP: 2x 10G SFP+ ports -TRTP2: 2x 1GbE and 2x 10G SFP+ ports -TR25M: 2x 25GbE ports |
| Drive Bays | 10 hot-swap 2.5" drives bays (Default 10 SATA3 ports) Optional 2 NVMe/SATA3 Hybrid ports Optional 8 SAS3 ports support via Add-on Card |
| Power Supply | Redundant 750W Platinum Level digital power supplies |

- The number of users supported varies depending on specific applications and workloads.

- Q-series virtual GPU types require a Quadro vDWS license.
- B-series virtual GPU types require a GRID Virtual PC license.
- For more information about NVIDIA vGPU technology, see <u>https://docs.nvidia.com/grid/latest/grid-vgpu-user-guide/index.html</u>







Ultra 2U 2029U VDI with High Graphic Performance



Dual-socket support up to 52 cores / 112 threads

| Technical Specifications | |
|--------------------------|--|
| Processor Support | Dual Intel[®] Xeon[®] Scalable processors Up to 165W CPU TDP with one or more GPUs |
| Memory Capacity | • 24 DIMM slots, up to 6TB DDR4-2933 ECC RDIMM/LRDIMM |
| PCI-E Expansion Slots | 8 PCI-E: 1 PCI-E 3.0 x16 (FH 10.5" L), 7 PCI-E 3.0 x8 (5 FH 10.5"L, 1 LP, 1 LP Internal) -TR4T: 7 PCI-E: 2 PCI-E 3.0 x16 (FH 10.5" L), 5 PCI-E 3.0 x8 (4 FH 10.5"L, 1 LP) |
| Networking | -TR4: 4x 1GbE ports -TR4T: 4x 10Gbase-T ports -TRT: 2x 10Gbase-T ports -TRTP: 2x 10G SFP+ ports -TR25M: 2x 25GbE ports |
| Drive Bays | 24 hot-swap 2.5" drives bays (default 14 SATA3 ports) Optional 24 SAS3 ports support via Add-on Card(s) Optional 4 NVMe ports (2 NVMe from CPU1 + 2 NVMe from CPU2) Optional 2 rear hot-swap 2.5" drive bay kit |
| Power Supply | Redundant 1000W Titanium Level digital power supplies |

- The number of users supported varies depending on specific applications and workloads.

- Q-series virtual GPU types require a Quadro vDWS license.
- B-series virtual GPU types require a GRID Virtual PC license.
- For more information about NVIDIA vGPU technology, see <u>https://docs.nvidia.com/grid/latest/grid-vgpu-user-guide/index.html</u>









Ultra 2U 2029U All-Flash NVMe with 2x NVIDIA® Tesla® GPUs VDI



Dual-socket support up to 52 cores / 112 threads Up to 24 hot-swap U.2 NVMe drives for I/O intensive applications

| Technical Specifications | |
|--------------------------|--|
| Processor Support | Dual Intel[®] Xeon[®] Scalable processors Up to 165W CPU TDP with one or more GPUs |
| Memory Capacity | • 24 DIMM slots, up to 6TB DDR4-2933 ECC RDIMM/LRDIMM |
| PCI-E Expansion Slots | 2 PCI-E Gen 3.0 x16 (1 full height 10.5" length) 1 PCI-E Gen 3.0 x8 (low profile) |
| Networking | • 4x 10GBase-T Ethernet ports |
| Drive Bays | 24 hot-swap 2.5" drives bays 12 NVMe ports (NVMe from CPU1) 8 NVMe ports + 4 NVMe/SATA3 Hybrid ports (NVMe from CPU2) 2 rear 2.5" hot-swap bays (default 2 SATA3 ports) |
| Power Supply | Redundant 1600W Titanium Level digital power supplies |

- The number of users supported varies depending on specific applications and workloads.

- Q-series virtual GPU types require a Quadro vDWS license.
- B-series virtual GPU types require a GRID Virtual PC license.
- For more information about NVIDIA vGPU technology, see <u>https://docs.nvidia.com/grid/latest/grid-vgpu-user-guide/index.html</u>

Summary

VDI has become increasingly popular with the broadened availability of mobile broadband, higher bandwidth, more compute power in modern mobile devices, higher display resolutions and a trend of BYOD within enterprises.

Supermicro is leading the industry on VDI and focus on this area to provide the best solution for VDI infrastructure.

Supermicro, together with NVIDIA, has developed a line of fully configured GPU server solutions to address the rapidly growing Virtual Desktop Infrastructure (VDI) market.

Supermicro and NVIDIA Virtual GPU Web Resources

Supermicro GPU Qualified Platform Matrix

Supermicro + NVIDIA vGPU Solution

NVIDIA vGPU Evaluation license for up to 90 days

NVIDIA vGPU Website

NVIDIA vGPU TCO Calculator



This page is intentionally left blank.

This page is intentionally left blank.

About Super Micro Computer, Inc.

Supermicro[®] (NASDAQ: SMCI), the leading innovator in high-performance, high-efficiency server technology is a premier provider of advanced server Building Block Solutions[®] for Data Center, Cloud Computing, Enterprise IT, Hadoop/Big Data, HPC and Embedded Systems worldwide. Supermicro is committed to protecting the environment through its "We Keep IT Green[®]" initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

Learn more at www.supermicro.com

No part of this document covered by copyright may be reproduced in any form or by any means — graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system — without prior written permission of the copyright owner.

Supermicro, the Supermicro logo, Building Block Solutions, We Keep IT Green, SuperServer, Twin, BigTwin, TwinPro, TwinPro², SuperDoctor are trademarks and/or registered trademarks of Super Micro Computer, Inc.

Intel, the Intel logo, and Xeon are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

All other brands names and trademarks are the property of their respective owners.

© Copyright Super Micro Computer, Inc. All rights reserved.

