



NVIDIA RTX SERVER FOR VIRTUAL
WORKSTATION WITH TERADICI
CLOUD ACCESS SOFTWARE ON
PENGUIN RELION XE2114GT DESIGN
GUIDE

VERSION 1.0



TABLE OF CONTENTS

- Chapter 1. SOLUTION OVERVIEW 1**
- 1.1 RTX Server Overview 1
- Chapter 2. SOLUTION DETAIL..... 2**
- 2.1 Solution Configuration 3

Chapter 1.

SOLUTION OVERVIEW

NVIDIA RTX Server™ is a validated combination of qualified partner systems hosting NVIDIA® Quadro® RTX™ 6000 or RTX 8000 GPUs and Quadro® Virtual Data Center Workstation (Quadro vDWS) software to provide virtualized design workstations where artists and designers can create professional, photorealistic images for the Media & Entertainment, Architecture, Engineering & Construction (AEC), and Manufacturing & Design industries.

1.1 RTX SERVER OVERVIEW

In today's media rich world, there is a massive surge of content production. Designers and artists are looking for ways to create faster and more efficiently anywhere on the device of their choice. IT looks to enable this in a secure, easily managed, cost-effective and flexible way. NVIDIA RTX Server is an ideal platform to increase user productivity while increasing data center utilization and reducing cost.

The audience for this document includes, but is not limited to, sales engineers, field consultants, professional services, partner engineers, and IT Managers who wish to take advantage of a solution that is purpose built and optimized to deliver a virtualized design and creation workflow.

Chapter 2.

SOLUTION DETAIL

NVIDIA RTX Server for Virtual Workstation with Teradici Cloud Access Software on Penguin Relion XE2114GT is a reference design comprised of (a) NVIDIA Quadro RTX 8000 or Quadro RTX 6000 graphics cards; (b) VMware vSphere Hypervisor with vCenter; (c) NVIDIA Quadro vDWS software; (d) Penguin Relion XE2114GT server; (e) thin client[spec below in Table 1]; (f) Teradici Cloud Access Software with PCoIP® Ultra. This validated solution provides unprecedented graphics performance in a virtualized environment on professional applications at a fraction of the cost, space, and power consumption to individual local workstations.

NVIDIA Quadro RTX 8000 and RTX 6000, powered by the NVIDIA Turing™ architecture and the NVIDIA RTX platform, brings the most significant advancement in computer graphics in over a decade to professional workflows. Designers and artists can now wield the power of hardware-accelerated ray tracing, deep learning, and advanced shading to dramatically boost productivity and create amazing content faster than ever before.

NVIDIA Quadro vDWS enables delivery of the most powerful virtual workstations from the data center or cloud to any device, anywhere. It lets IT virtualize any application from the data center with a native workstation user experience, eliminate constrained workflows, and flexibly scale GPU resources to run multiple workloads. Efficiently centralize applications and data for a dramatically lower IT operating expense and focus IT resources on managing strategic projects versus individual workstations – all while enabling a more secure, work-from-anywhere environment with reduced threat of data loss or leakage.

VMware vSphere provides a powerful, flexible, and secure foundation for business agility that accelerates your digital transformation. With vSphere, you can support new workloads and use cases while keeping pace with the growing needs and complexity of your infrastructure. vSphere is the heart of a secure software defined data center (SDDC), securing applications, data, infrastructure, and access. Advanced security capabilities fully integrated into the hypervisor and powered by machine learning, provide better visibility, protection and faster response time for security incidents. vSphere helps you run, manage, connect and secure your applications in a common operating environment across the hybrid cloud.

Teradici Cloud Access Software with PCoIP® Ultra enhancements enables NVIDIA RTX servers to deliver highly interactive graphics to remote users without compromise. Leveraging NVIDIA NVENC encoding technology with security and performance benefits of the Teradici PCoIP protocol, graphic designers, artists, and CAD/CAM users can work productively over virtually any network.

As more organizations look to grow their compute capabilities, they look to GPU accelerators to do so. Deploying GPU's allows you to offload compute-intensive functions from CPUs, massively improving processing times. This is why workloads like artificial intelligence and deep learning, which need immense computing resources to process extreme amounts of data, benefit from GPU density. The Penguin Computing® Relion® XE2114GT server packs 4 GPUs in a 2U rackmount form factor, while not sacrificing CPU performance, with two Intel® Xeon® Scalable Processors. Choose the Relion XE2114GT server to power data-intensive workloads that need the computing power of GPU accelerators

2.1 SOLUTION CONFIGURATION

Table 1 outlines the system and virtual machine configuration recommended for Media and Entertainment industry designers. Depending on their workload size, frame buffer allocation can be changed by configuring the NVIDIA virtual GPU (vGPU) type on virtual machine.

Table 1: Designer virtual machine configuration

Component	Vender & Model	Details
System	Penguin Relion XE2114GT	<ul style="list-style-type: none"> • High Performance Dual Intel Xeon Gold 6126 processors: 2.6-3.7GHz; 12 Cores, or • Higher User Density Dual Intel Xeon Gold 6148 processors: 2.4-3.7GHz; 20 Cores • 1.5TB Memory (128GB recommended for each user), system memory size might change based on customer's deployment configuration • Network attached storage, SSD recommended for higher performance • 10GbE network ports
Graphics Hardware	4x Quadro RTX 8000 or RTX 6000	<ul style="list-style-type: none"> • RTX 8000 48GB frame buffer, 4,608 CUDA Cores, 72 RT Cores, 576 Tensor Cores, • RTX 6000 24GB frame buffer, 4,608 CUDA Cores, 72 RT Cores, 576 Tensor Cores,

Graphics Software	<p>Quadro vDWS Software</p> <p>(vGPU 8.0 or later)</p> <p>128GB memory / 16 vCPU cores</p> <p>vGPU frame buffer: 12GB / 16GB / 24GB</p>	<p>Examples</p> <ul style="list-style-type: none"> • 16GB FB per user GRID_RTX8000-16Q: 3 users, Max. 9 users supported • 12GB FB per user GRID_RTX8000-12Q: 4 users, Max. 12 users supported GRID_RTX6000-12Q: 2 users, Max. 6 users supported
Hypervisor	VMware vSphere 6.7U1 or later	Enterprise license
Client Device	Any Client	<ul style="list-style-type: none"> • Intel® Core™ i5-8250U or higher • DDR4 2 x SO-DIMM
Remote Access Software	Teradici Cloud Access Plus (subscription based)	<ul style="list-style-type: none"> • Windows Graphics Agent • Linux Graphics Agent • Software clients

Notice

ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE.

Information furnished is believed to be accurate and reliable. However, NVIDIA Corporation assumes no responsibility for the consequences of use of such information or for any infringement of patents or other rights of third parties that may result from its use. No license is granted by implication of otherwise under any patent rights of NVIDIA Corporation. Specifications mentioned in this publication are subject to change without notice. This publication supersedes and replaces all other information previously supplied. NVIDIA Corporation products are not authorized as critical components in life support devices or systems without express written approval of NVIDIA Corporation.

Trademarks

NVIDIA, the NVIDIA logo, and RTX Server, Turing, and Quadro are trademarks or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2019 NVIDIA Corporation. All rights reserved.