



QUADRO FOR EDUCATION & RESEARCH

6-2019

EDUCATION ENVIRONMENT CHALLENGES



Research problems, student projects more complex than ever, data sets & compute requirements growing exponentially



Educational and research software toolchains & workloads increasing in size and complexity



Demand for skilled data scientists & AI expertise requires students to hit the ground running after graduation



IT budgets not growing to match technology infrastructure demands

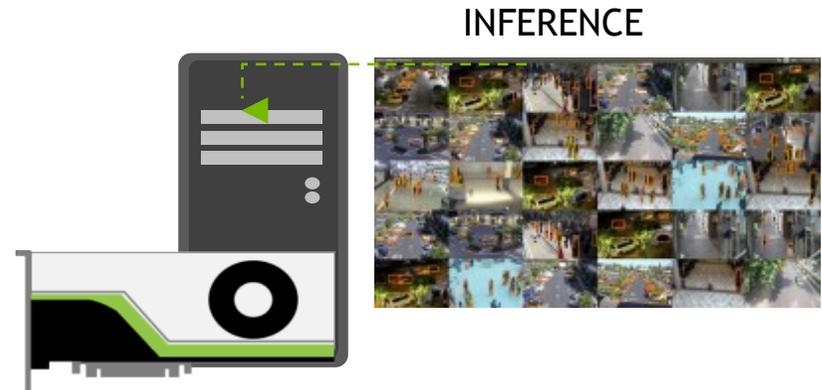
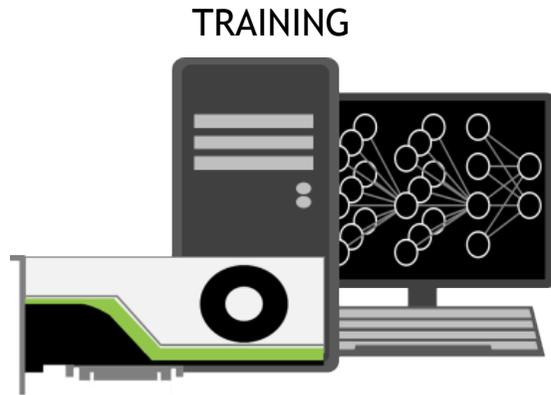
QUADRO HER VALUE PROPOSITION

Only NVIDIA Quadro solutions provide:

- Compute power, GPU memory capacity, and scalability required for today's demanding education & research projects
- Researcher and student access to professional level hardware & software - same used by professionals around the world
- Enterprise level hardware & software support - maximize uptime and minimize IT support requirements
- Server ready GPU solutions



QUADRO FOR DEEP LEARNING



Quadro value for DL training:

- Large memory, scalability for large training data sets, model optimization & validation
- Scalability with NVLink to speed up training tasks, provide development platform for server deployments

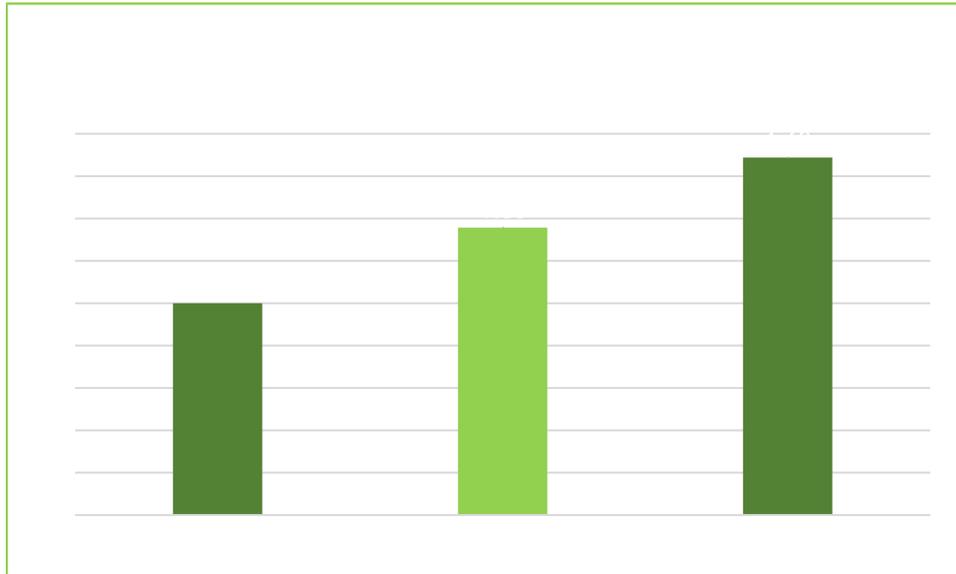
Quadro value for DL inferencing:

- Fast inferencing performance for the largest networks & datasets
- Support for encode/decode of multiple simultaneous video streams
 - Consumer cards limited to 2 simultaneous streams

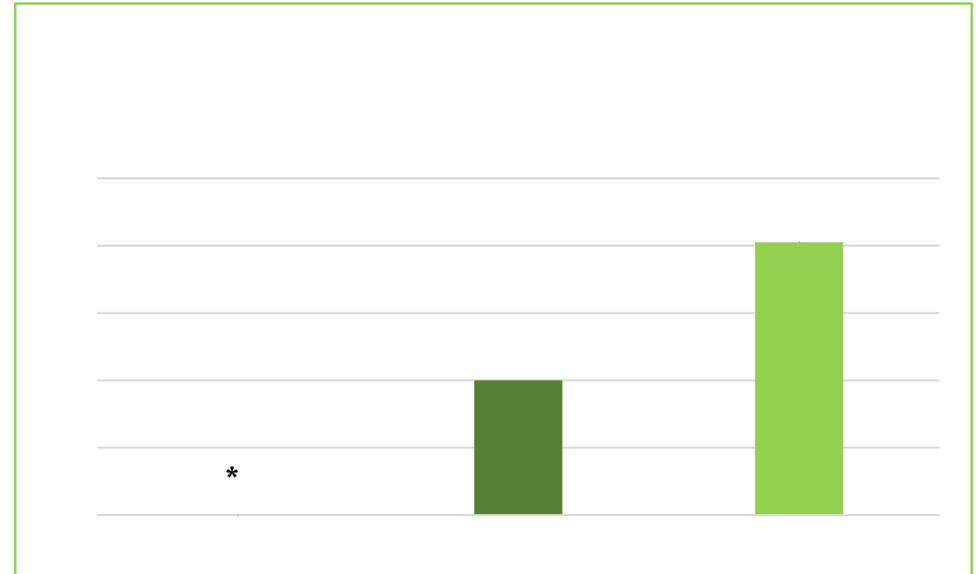
Quadro solutions are validated, tested, and used daily in demanding professional DL environments

DEEP LEARNING TRAINING PERFORMANCE

Quadro Provides Memory & Performance Required for DL Training



Quadro enables larger batch sizes to accelerate DL training

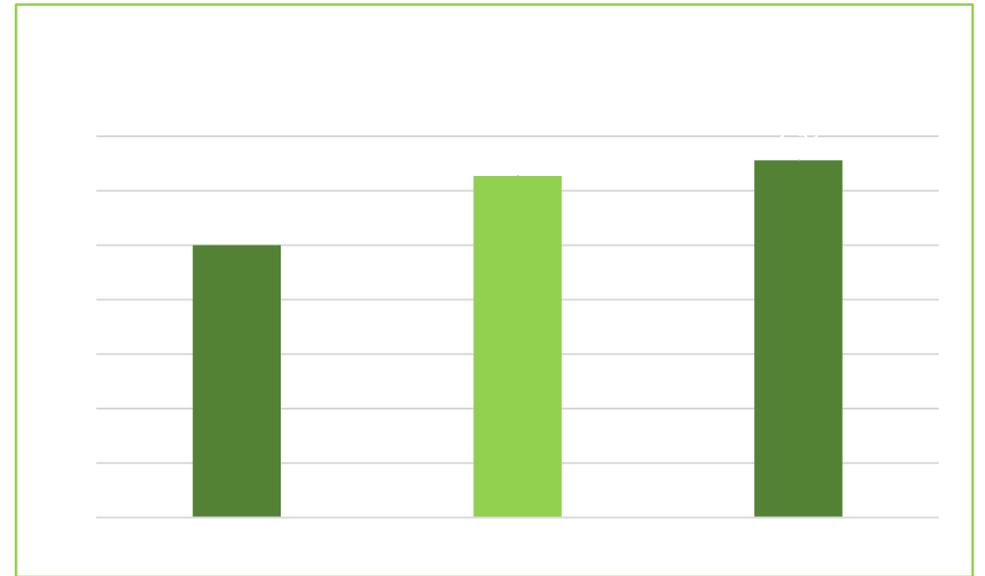
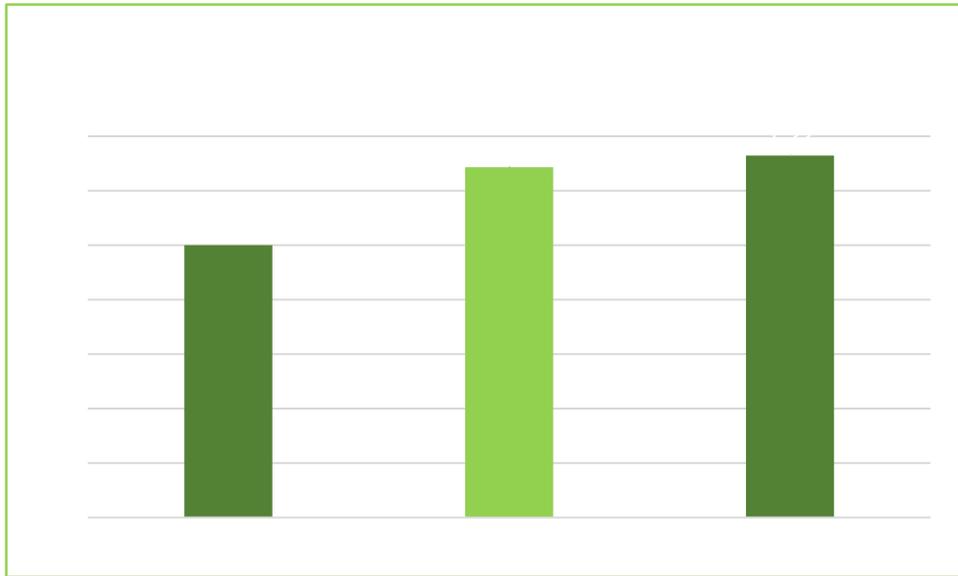


*Did not run due to insufficient memory

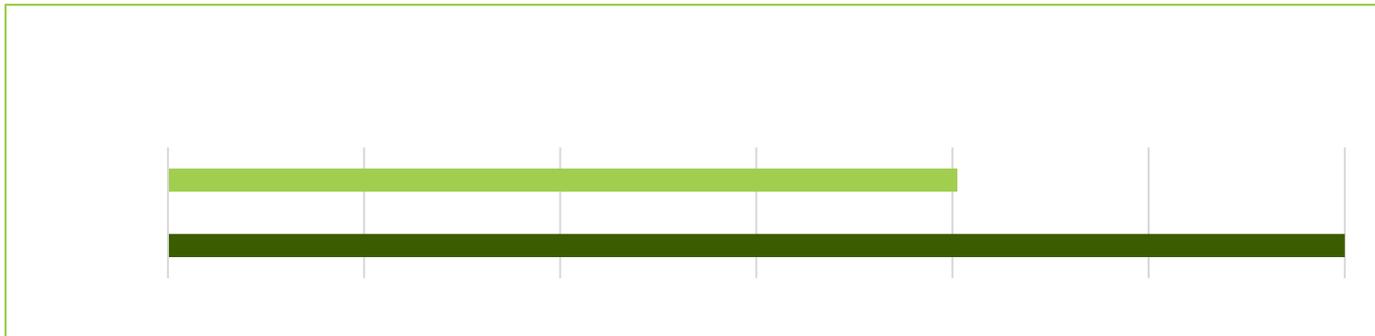
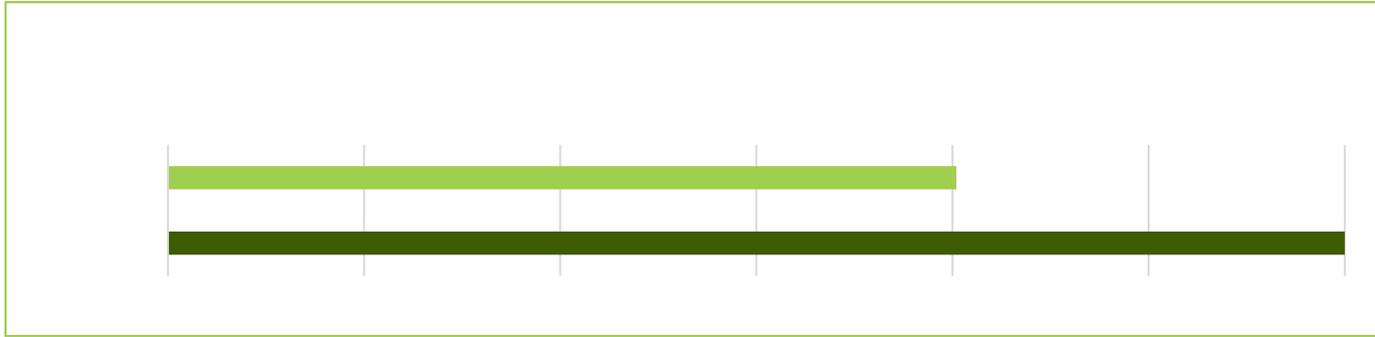
Large training datasets such as natural language processing, machine translations, etc., don't fit into consumer graphics memory

DEEP LEARNING INFERENCE PERFORMANCE

Deep Networks with Many Layers can Take Advantage of Quadro Memory



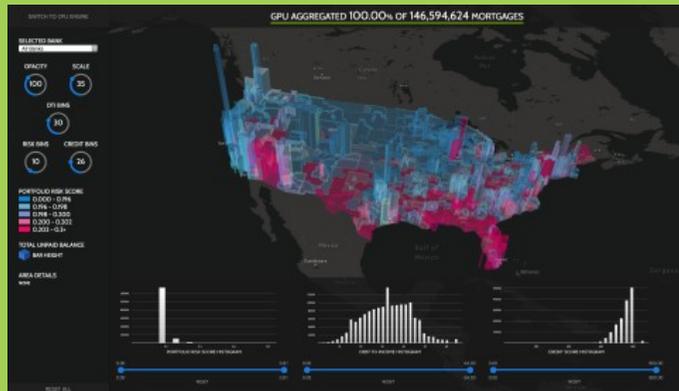
QUADRO - FASTER TIME TO SOLUTION*



*based on RTX 6000 ResNet-50 performance 33% faster than 2080Ti

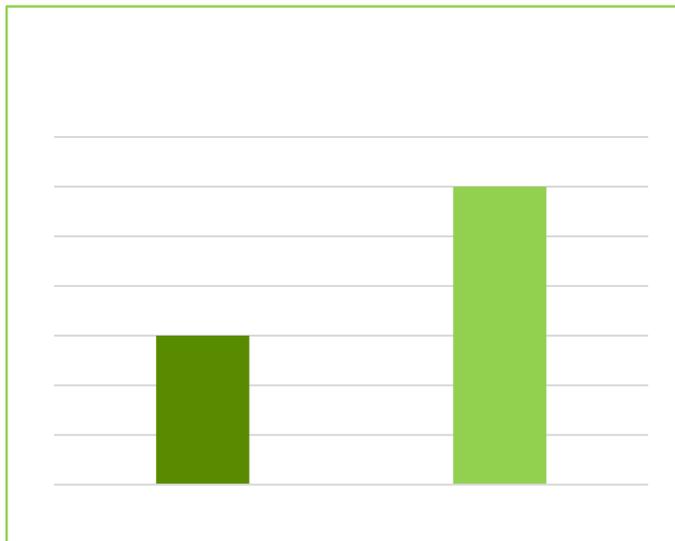
QUADRO REQUIRED FOR THE LARGEST WORKLOADS

Data Science Data Sets Require GPU Memory Only Available with Quadro



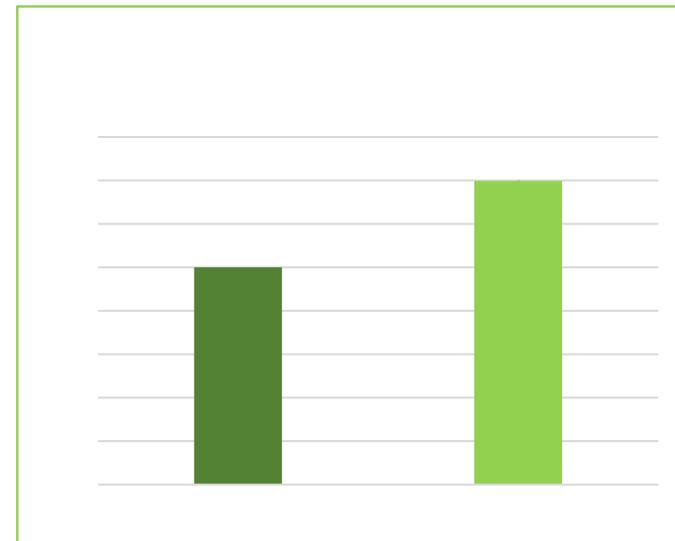
Data Science Workload Example:

Sample data set using Home Mortgage data in the US for 2016. A single GeForce RTX 2080Ti can only load 3 months worth of data. A single Quadro RTX 6000 GPU can load 6 months of data at a time. Two RTX 6000's can load the entire years worth of data in the combined GPU memory (dual 2080Ti's can only load 6 months of data).*



Quadro value for data science

- Large Quadro memory lets data scientists process more data to improve model training and accuracy.
- Quadro performance completes tasks faster



QUADRO FOR HPC

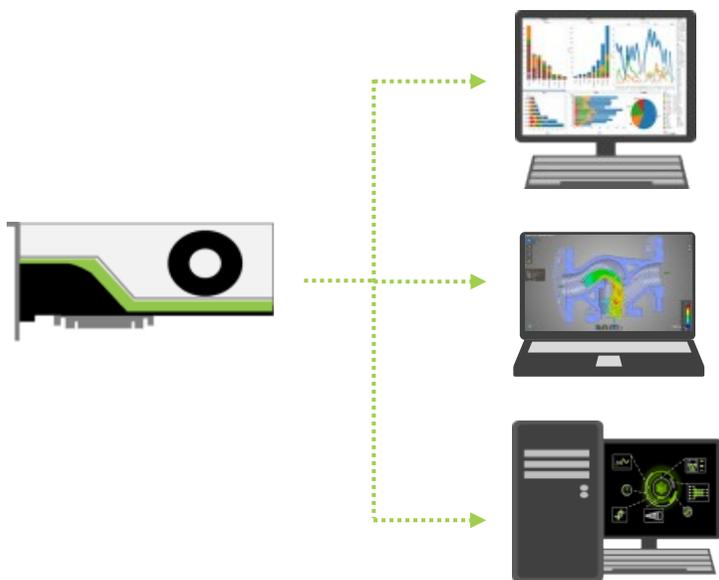


Quadro value for HPC:

- Larger GPU memory for the largest data sets and compute tasks
- Memory and compute power enables multi application workflows from compute to visualization
- Quadro is NGC Ready, tested and validated for NVIDIA NGC HPC software

QUADRO FOR VIRTUALIZATION

Quadro vDWS - Virtual Quadro Anywhere



Quadro value for virtualization:

- Get the features, performance of Quadro anywhere you need it.
- Partition a single RTX 6000 into multiple high performance virtual workstations
- Quadro vDWS only available on Tesla, Quadro RTX 6000/8000 GPUs

Learn more about Quadro vDWS solutions:

<https://www.nvidia.com/quadro-vdws/>

NVIDIA NGC SUPPORT SERVICES

Minimize Downtime And Maximize System Utilization

Support Coverage

- NGC DL & ML containers
- NVIDIA drivers
- NV-docker
- CUDA



Support by NVIDIA's subject matter experts



24x7 portal, phone and email access to create support cases

Live support during local, regional business hours for technical assistance

Availability

- Exclusively for NGC-Ready workstations
- Service agreement between NVIDIA & customer
- Purchase from OEM/System vendor



GeForce RTX 2080Ti



Quadro RTX 6000

NVIDIA Desktop System Use Case	PC-based HPC, AI/DL/ML workloads, small to medium size data sets, up to 22GB memory w/NVLink	Workstation-based HPC, AI/DL/ML workloads, largest data sets with up to 48GB memory w/NVLink, pre-installed data science software stack.
GPU Memory	11GB	24GB
ECC Memory	No	Yes - error free compute
NVLink support	2-way NVLink, 3 & 4 slot bridges	2-way NVLink, 2 & 3 slot bridges - fits into a wide variety of workstation chassis
Cooling Solution	Dual Axial Cooling	Blower Active Fan - ideal cooling solution for multi-GPU configurations
Pre-installed Data Science Software Stack from OEMs	No	Yes - up and running data science workloads within minutes
Nvidia support	N/A	Enterprise level pre/post sales support
OEM support	N/A	Enterprise level support
NGC Ready	No	Yes - validated and tested for running NVIDIA NGC software
NVIDIA NGC Support Services Option	N/A	Yes - support provided directly from NVIDIA
Quadro Advantage	<input type="checkbox"/>	Work with largest data sets to accelerate HPC, AI, ML workflows. NVIDIA and OEMs provide enterprise level testing, validation, and support. Additional NVIDIA NGC Support Services availability. GPU Cooling & NVLink solutions support widest range of system configurations.

QUADRO FOR HIGHER EDUCATION & RESEARCH

Quadro Brings Real Value

- Enterprise grade hardware, enterprise level software & hardware testing, validation, & support - ready for workstation & server deployments
- Performance & scalability - large GPU memory for large datasets & workloads today and room for growth in the future
- Maximum flexibility with the option to provide multiple virtual Quadro workstations (Quadro vDWS) from a single Quadro GPU



