NVIDIA MAXIMUS SYSTEM BUILDERS' GUIDE
MICROSOFT WINDOWS 7–64

DI-05471-001_v01 | August 2012
NVIDIA CONFIDENTIAL | Prepared and Provided Under NDA

User Instructions

DOCUMENT CHANGE HISTORY

DI-05471-001_v01

Version	Date	Authors	Description of Change
01	August 4, 2012	OO, TS	Initial release

TABLE OF CONTENTS

NVIDIA Maximus System Builders' Guide 5
- Audience 5
- Prerequisite Skills 6
- Contents 6
- Benefits of Maximus Technology 7

Installation 8
- Required Components to Enable Maximus Technology 8
  - Quadro Professional Graphics 9
  - Tesla C2075 9
  - NVIDIA Quadro Professional Graphics Driver 10
  - Implementing the Maximus Platform 10
  - Requirements for the Host Chassis 11
  - Using x16 Electrical PCIe Slots 11
- Requirements for Host Memory 13
- Requirements for Power Supplies and Auxiliary PCIe Cabling 14
- Connecting Display Monitors 16
- Using Scalable Link Interface (SLI) Ribbon Cables 16
- Quadro Driver 17
  - Prerequisites 17
  - Installation 17
  - Verifying System Installation 18
  - Checking Power Connections 19
  - Checking the Quadro Driver Version 20
  - Check GPU Recognition and Quadro Driver Installation 21
  - Checking Driver Recognition of Quadro and Tesla Cards 22
- Advanced Topics 23
  - Advanced Maximus Configurations 23
  - Tuning Applications 25
  - GPU-enabled Applications 25

Using the NVIDIA Maximus Configuration Utility 26

LIST OF FIGURES

Figure 1. NVIDIA Quadro Family for Maximus 9
Figure 2. NVIDIA Tesla C2075 10
Figure 3. PCIe x16 Slots in an HP Z800 Workstation 12
Figure 4. Typical Quadro 6000 and Tesla C2075 Combination 13
Figure 5. 6-pin-to-8-pin PCIe Power Adaptor 14
Figure 6. Power LEDs and PCIe Adapters 15
Figure 7. Workstation Configured with One Quadro 200 and two Tesla C2075 Cards 23
Figure 8. Maximus Configuration Utility Screen 24
Figure 9. NVIDIA Maximus Configuration Utility (MCU) Screen 26

LIST OF TABLES

Table 1. Tesla and Quadro GPU Specifications 11
Table 2. Power Consumption and Auxiliary Connector Requirements 14

NVIDIA MAXIMUS SYSTEM BUILDERS' GUIDE

! NVIDIA® Maximus™ platforms are available only through select OEMs. For a list of OEMs qualified by NVIDIA, refer to the Maximus web page at http://www.nvidia.com/maximus.

This document explains how to install, configure, and verify the primary hardware and software components of an NVIDIA Maximus platform for those who already have a platform from a qualified OEM (qualified OEMs listed at http://www.nvidia.com/maximus) and would like to upgrade their existing system.

This document *does not* cover vendor-specific GPU-enabled applications. For vendor-specific details about GPU-enabled applications, refer to the documentation provided by the software vendor.

AUDIENCE

This document is intended for professionals who are capable of installing cards on a workstation and making any system upgrades that may be necessary.

PREREQUISITE SKILLS

To complete the tasks described in this document, you should have an understanding of the following concepts and mastery of the skills:

- Personal computer hardware, terminology, and power connectivity
- Personal computer thermals and cooling
- Hardware and software combination troubleshooting
- Software driver installation
- Microsoft Windows 7 configuration and administration
- Microsoft Windows 7 device manager and driver installation

CONTENTS

The chapters in this document describe the required enabling components for a Maximus platform and how to enable Maximus technology and verify that your installation is correct.

The topics of the chapters are as follows:

- "Required Enabling Components for Maximus Technology" on page 9 identifies the components that you need to enable a system for Maximus technology.
- "Implementing the Maximus Platform" on page 10 describes the installation of the Quadro and Tesla cards, the Quadro driver and the system enhancements that may be required to support a Maximus-enabled system.
- "Verifying Your System Installation" on page 18 describes how to check that the Maximus technology is correctly installed and correctly recognized by the system.
- "Advanced Topics" on page 21 briefly discusses more advanced Maximus configurations, tuning of applications that use Maximus and a link to a list of GPU-enabled applications that use Maximus technology.
- "Appendix: Using the NVIDIA Maximus Configuration Utility" on page 22 describes the Maximus Configuration Utility, which you can use to check what GPU cards are installed on a system and to tune your Maximus configuration.

BENEFITS OF MAXIMUS TECHNOLOGY

For an overview of Maximus technology, its benefits, and how it is being used, go to http://www.nvidia.com/maximus.

INSTALLATION

This section identifies the components that are needed to enable a system for Maximus technology. It also describes the installation of the Quadro and Tesla cards, the Quadro driver, and the system-enhancements that may be required to support a Maximus-enabled system and provides a procedure to verify your installation.

REQUIRED COMPONENTS TO ENABLE MAXIMUS TECHNOLOGY

Following is a list of the critical components needed for a Maximus-enabled workstation:

- At least one NVIDIA Quadro 600, 2000, 4000, 5000, or 6000 graphics card (for computation and graphics). See *Quadro Professional Graphics* on page 9.
- At least one NVIDIA Tesla C2075 card (for computation). See *Tesla C2075* on page 9.
- A compatible professional NVIDIA software driver. See *NVIDIA Quadro Professional Graphics Driver* on page 10.
- Microsoft Windows 7—64 bit
- A host chassis from a qualified OEM with space for the Quadro and Tesla cards and the capacity to support the power demands.

The Maximus technology components are described in the following sections. System enhancements needed to support a Maximus-enabled workstation are described in *Implementing the Maximus Platform* on page 10.

Quadro Professional Graphics

The NVIDIA Quadro family of professional graphics solutions provides the visualization component of a Maximus-enabled workstation. A Quadro graphics card can also perform GPU computational processing, therefore providing another alternative for users to best utilize their system in demanding workflows. NVIDIA Maximus platform supports the following NVIDIA Quadro graphics cards (see Figure 1):

- Quadro 600
- Quadro 2000
- Quadro 4000
- Quadro 5000
- Quadro 6000

Note: Some retail Quadro cards may look different than those shown in the illustration. However, these cards are designed, built and supported by NVIDIA and have exactly the same functionality.

Figure 1. NVIDIA Quadro Family for Maximus

Tesla C2075

An NVIDIA Maximus platform requires the Tesla C2075. card This is the critical component of a Maximus platform.

The Tesla C2075 is the latest addition to the Tesla C series of compute companion processors, which are based on the Fermi architecture. Tesla C2075 features 448 NVIDIA CUDA™ cores and six GB of high-speed GDDR5 memory (see Figure 2).

Figure 2. NVIDIA Tesla C2075

NVIDIA Quadro Professional Graphics Driver

The NVIDIA Quadro Professional Driver is another critical component. Support for both the Quadro and Tesla cards are unified in this one NVIDIA driver.

The NVIDIA Quadro Professional driver enables runtime GPU processing tasks to operate among multiple GPUs in a system. The driver serves as the interface between professional GPU-enabled applications and the GPU hardware itself.

Implementing the Maximus Platform

As stated in the preceding chapter, a Maximus-enabled platform is a combination of an NVIDIA Quadro card, Tesla card, and software driver technology. To host this combination of GPUs, a capable chassis is required along with whatever other workflow-specific components are necessary. The recommendations in this chapter serve only as a guide and can cover only the components noted here. It is he system builder's responsibility to verify that proper power, cooling, thermals, and connectivity requirements are met for the total system.

! Disconnect all mains power before performing any system hardware work.

Requirements for the Host Chassis

You must ensure that the host chassis satisfies the requirements to host your particular configuration of Quadro and Tesla cards. For example, the host chassis must have sufficient physical space for each card and satisfy the cooling and power requirements for your card configuration.

Table 1 outlines the specifications of each supported Tesla and Quadro GPU.

Table 1. Tesla and Quadro GPU Specifications

Model	Onboard Memory	PCIe Slot Width	Max Power Consumption	PCIe Slot Type	Auxiliary Power Requirement
Tesla C2075	6GB	Double	225W	x16 Gen2	Yes
Quadro 6000	6GB	Double	204W	x16 Gen2	Yes
Quadro 5000	2.5GB	Double	152W	x16 Gen2	Yes
Quadro 4000	2GB	Single	142W	x16 Gen2	Yes
Quadro 2000	1GB	Single	62W	x16 Gen2	No
Quadro 600	1GB	Single	40W	x16 Gen2	No

A Maximus-enabled workstation contains at least one NVIDIA Quadro card and at least one NVIDIA Tesla C2075 card. A typical configuration is one Quadro 6000 and one Tesla C2075 with six GB of memory on each board. This configuration provides maximum memory and application flexibility.

Using x16 Electrical PCIe Slots

Peripheral Component Internet Express (PCIe) is a data and physical interconnect standard for attaching peripherals (such as graphics cards or other I/O devices) to the main data bus of a host computer system board. The Quadro and Tesla cards conform to the PCIe standard and are specifically built for PCIe x16 electrical slots.

There are several types of PCIe slots in modern workstations. Typicaly, high-end workstations from major OEMs such as HP, Dell, and Lenovo feature two or three PCIe x16 slots. Some specialized workstations may contain more PCIe x16 slots for larger configurations.

Slots on a workstation system board may appear physically similar, but can be electrically different in that the pin slots may not be connected to the data bus. Consult the documentation for your particular system board to identify the x16 electrical PCIe slots. For example, use slots that are labeled SLOT2 PCIe2 x16 (x16). Such slots are x16 mechanical and x16 electrical.

Try to avoid using slots that are labeled SLOT4 PCIe2 x16 (x8) for example. Such slots are x16 mechanical and x8 electrical. Due to reduced bandwidth on these slots, you may not be able to achieve full performance for your application.

The PCIe x16 slot often has a single spring-like plastic tab (commonly green) at one end. Take care when using these tabs because they can break easily. Typically, PCIe slots are labeled on the system board for convenient identification. Figure 3 shows an HP Z800 workstation system board; the PCIe x16 slots are shown outlined in red.

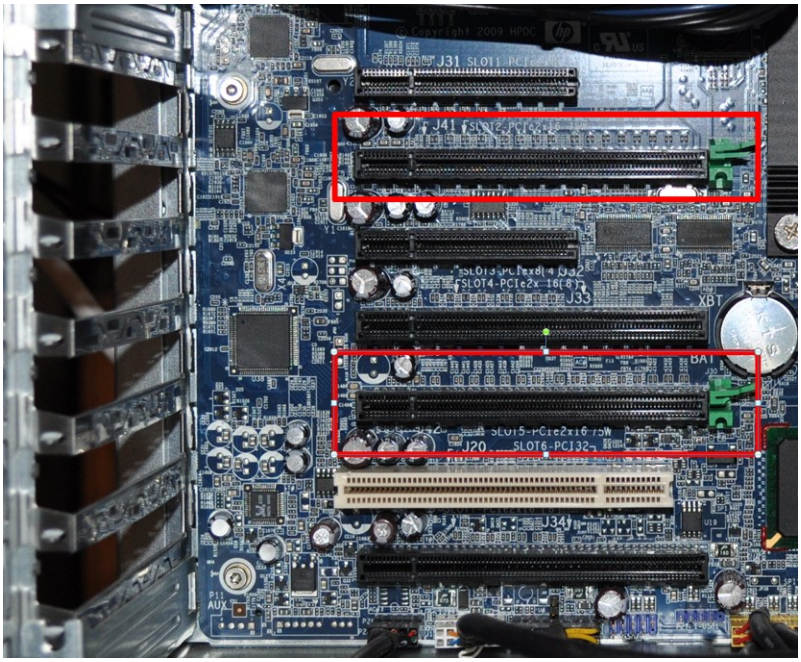


Figure 3. PCIe x16 Slots in an HP Z800 Workstation

Figure 4 shows a typical Quadro 6000 and Tesla C2075 combination. The Quadro card in the first x16 electrical PCIe slot should be installed in the slot nearest to the system board I/O connections (this card drives the displays for visualization). The Tesla board should be installed in the next available x16 electrical PCIe slot nearest the Quadro card.

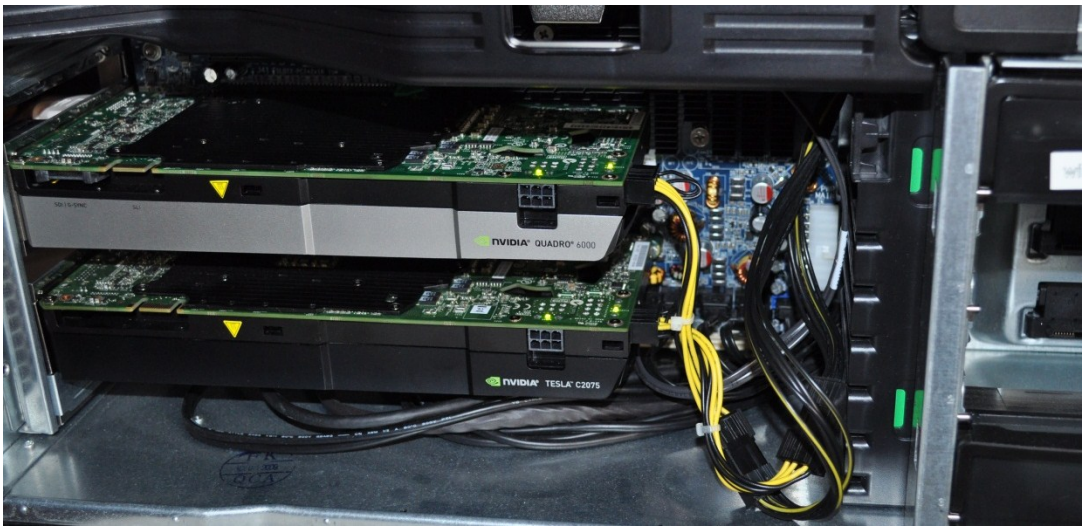


Figure 4. Typical Quadro 6000 and Tesla C2075 Combination

REQUIREMENTS FOR HOST MEMORY

For best results, ensure that your Maximus-enabled system has approximately three times as much system memory as aggregate onboard GPU memory. For example; if a system has one Quadro 6000 (six GB) and one Tesla C2075 (six GB) installed, the host system requires 36 GB (3 x 12 GB = 36 GB) of system memory to accommodate GPU operations. Note that this recommendation is intended for general guidance and will change by application; for system configurations recommended by NVIDIA, go to http://nvidia.com/maximus.

REQUIREMENTS FOR POWER SUPPLIES AND AUXILIARY PCIE CABLING

The NVIDIA Tesla card and many of the Quadro cards require auxiliary power within the host chassis. A suitable power supply is necessary to maintain system integrity under computational load. Many power supply wiring harnesses do not provide four PCIe 8-pin connectors and therefore may require 6-pin to 8-pin adaptors. Ensure that your system can deliver the necessary wattage and auxiliary power connectors for all cards in the system. If the supplied power is not adequate, the cards will not function properly. Table 2 specifies the power requirements and power connector options for the NVIDIA Tesla and Quadro cards.

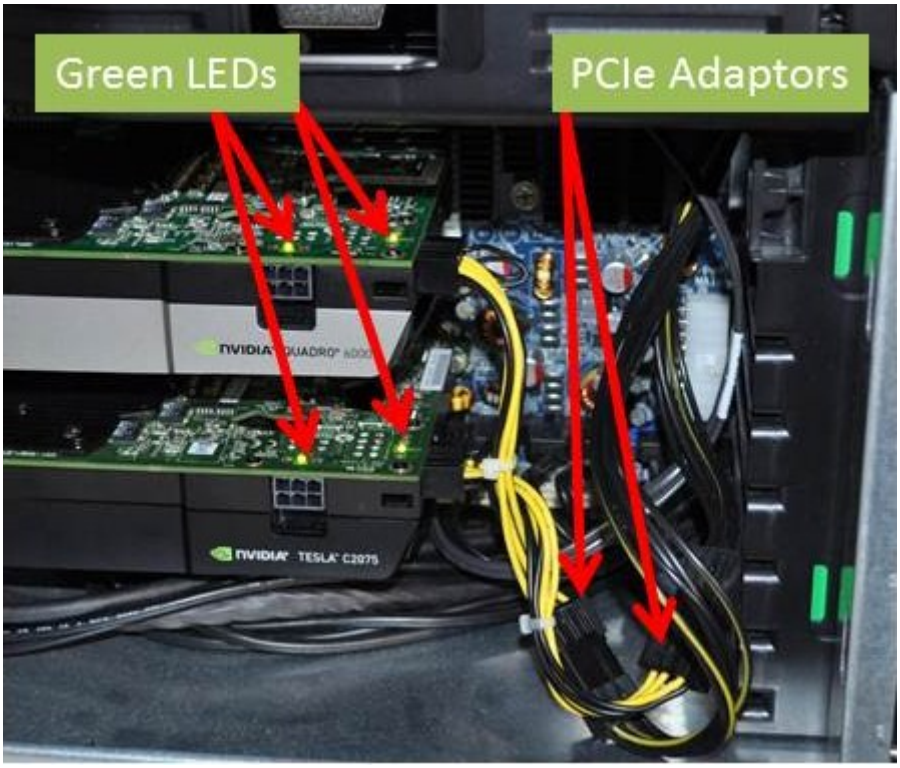
Table 2. Power Consumption and Auxiliary Connector Requirements

Model	Maximum Power Consumption	Required Auxiliary PCIe Power Connectors
Tesla C2075	225W	1/ea. 8-pin OR 2/ea. 6-pin
Quadro 6000	204W	1/ea. 8-pin OR 2/ea. 6-pin
Quadro 5000	152W	1/ea. 6-pin
Quadro 4000	142W	1/ea. 6-pin
Quadro 2000	62W	n/a
Quadro 600	40W	n/a

Power adaptors can be obtained to connect 6-pin PCIe chassis power cables to 8-pin device receptacles from several online retailers. Figure 5 shows one type of 6-pin-to-8-pin PCIe power adaptor.

Figure 5. 6-pin-to-8-pin PCIe Power Adaptor

Be sure to check that the two green LEDs on the Tesla and higher-end Quadro cards (if any) are illuminated when you power on the host system. If illuminated, the power connections are correct. The callouts in Figure 6 highlight the LEDs on the Tesla and Quadro cards and the PCIe adaptors.



Notes: Some NVIDIA Quadro boards, such as the Quadro 4000 and the Quadro 5000, may not have LEDs. If the auxiliary power cabling is not correct, the following message will display at boot time: **`Please power down and connect the PCIe power cable(s) for this graphics card.`**

Figure 6. Power LEDs and PCIe Adapters

CONNECTING DISPLAY MONITORS

! In a Maximus-enabled system, do not connect a monitor to a Tesla card.

In a Maximus-enabled system, connect a monitor to a *Quadro* card only. NVIDIA Quadro graphics cards provide several options for display connections.

In scalable visualization configurations (for an overview go to http://www.nvidia.com/object/quadro_svs_overview.html), dedicate Quadro cards to graphics and Tesla cards to computation. Intermixing graphics and compute tasks may yield undesirable frame rate and display synchronization artifacts.

USING SCALABLE LINK INTERFACE (SLI) RIBBON CABLES

! Do not connect a Quadro card and a Tesla card with an SLI ribbon cable.

SLI ribbon cables are not required to enable a system for multiple GPU computational operations. In general, use SLI ribbon cables to connect *two Quadro* graphics cards for scalable visualization or other frame-rate sensitive applications.

QUADRO DRIVER

Prerequisites

- Make sure all applications are closed before installing/updating the driver.
- Ensure the Quadro and Tesla cards are already installed.

Installation

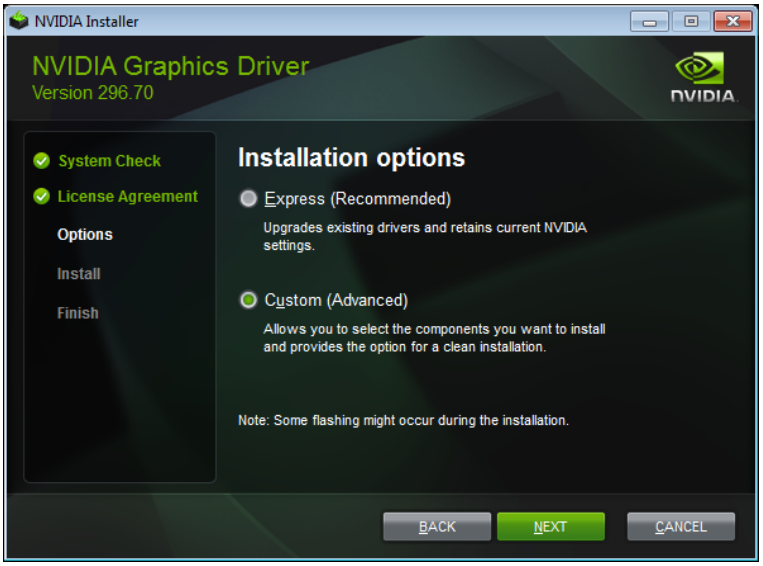
Install the NVIDIA Professional Quadro driver R275.89 or greater (within the R275 family). This is the recommended mainstream driver for enabling Maximus technology. You can download this driver from the NVIDIA website at http://www.nvidia.com.

> **!** You cannot use a Tesla driver to enable a Maximus-enabled workstation.

For specific applications, it is recommended to use ISV-certified drivers. Go to http://www.nvidia.com/page/partner_certified_drivers.html for a list of drivers.

Use the following procedure to install the graphics driver:

1. Download the NVIDIA Professional Quadro Driver R275.89 or greater (R275 family) from the NVIDIA website at http://www.nvidia.com and run the installer package
2. Accept the license agreement.
3. Select the Custom (Advanced) radio button under Installation options.
4. Click NEXT to begin installation.



5. Select Perform a clean installation checkbox on the Custom installation menu.

6. Click NEXT.

The screen(s) may go blank, flash, and/or show varying resolutions as the display driver software is unloaded and updated. This is normal as the display driver, operating system, and monitor(s) negotiate screen modes. It may be necessary to reboot the system.

Verifying System Installation

Use the following procedure to verify the system installation:

1. Ensure all Quadro and Tesla cards are properly powered. For more information, see Checking Power Connections on page 18.

   Hint: If the cards have green LEDs, they should be illuminated. If the Quadro cards do not have LEDs illuminated, check your display for system messages.

2. Check that the correct driver is installed. For more information, see *Checking the Quadro Driver Version*" on page 20.

   Hint: Go to NVIDIA Control Panel→System Information to check the driver version. In a standard configuration, a Quadro card is shown.

3. Check to make sure Microsoft Windows 7 recognizes the Quadro and Tesla cards as display adapters. For more information, see *Check GPU Recognition and Quadro Driver Installation* on page 21.

   Hint: Check for alerts; there should be none.

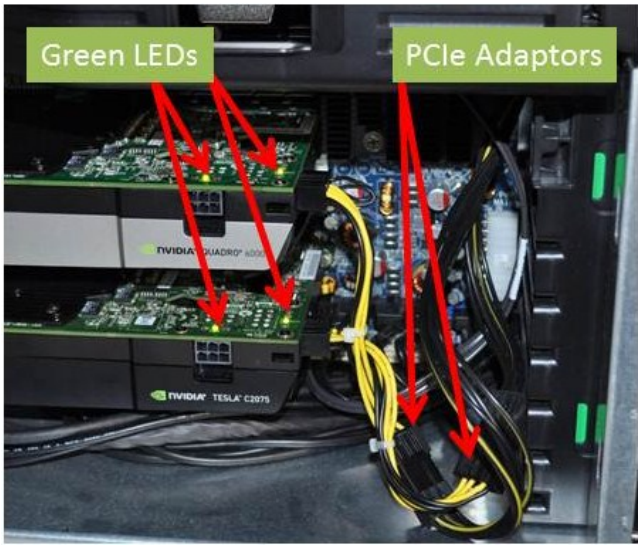
4. Check that all GPUs are physically installed in the system from the Maximus Configuration Utility (MCU). For more information, see *Checking Driver Recognition of Quadro and Tesla Cards* on page 22.

   Hint: All GPUs installed on the system should be listed in the MCU window.

Checking Power Connections

To ensure all Quadro and Tesla cards are properly powered, perform the following system checks:

1. Power on the system.
2. Check to make sure all the cards have green LEDs illuminated.



3. Check your display for system messages if the Quadro cards do not have LEDs illuminated,.

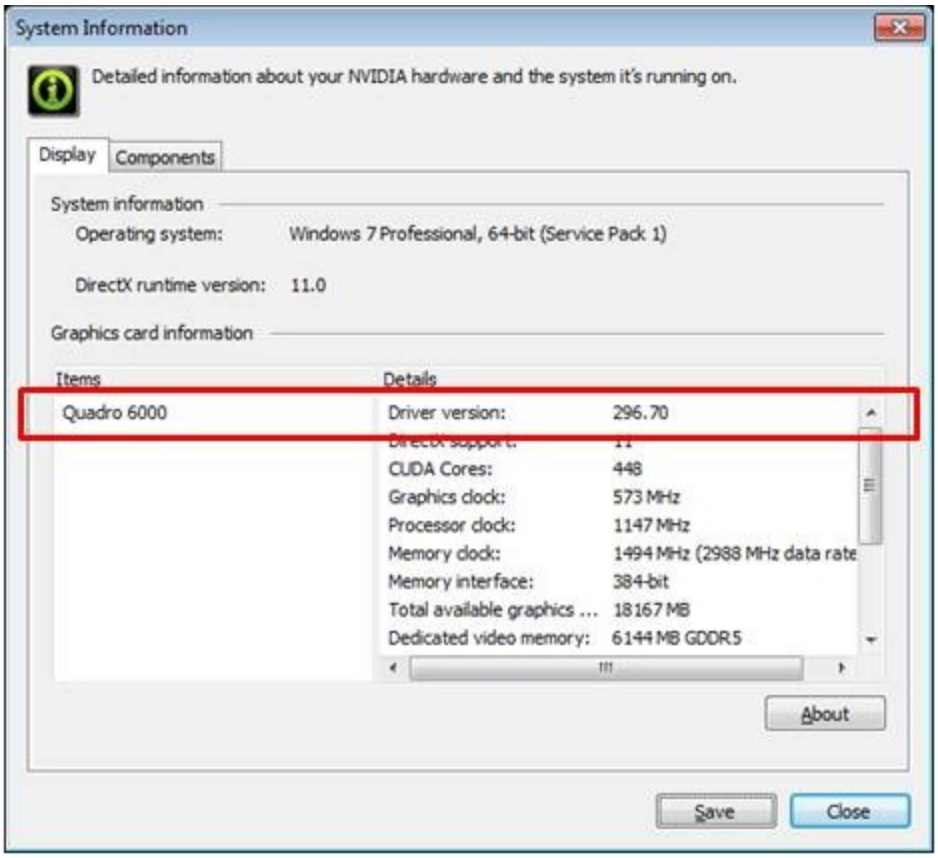
   Example: The displayed message **`PLEASE POWER DOWN AND CONNECT THE PCIe POWER CABLE(S) FOR THIS GRAPHICS CARD`** indicates that the correct power was not provided for an NVIDIA card. Check the power connections to all NVIDIA cards.

4. Check that all power cables are connected at both ends (card side and host-system side) if the green LEDs are not illuminated or a system message is displayed.

Checking the Quadro Driver Version

Use the following procedure to check the NVIDIA Quadro driver version displayed in the NVIDIA Control Panel:

1. Open the System Information menu from the NVIDIA Control Panel.
2. Check the driver version on the System Information menu. If the driver version is older than 275.89, reinstall the driver (see *Installation* on page 17).

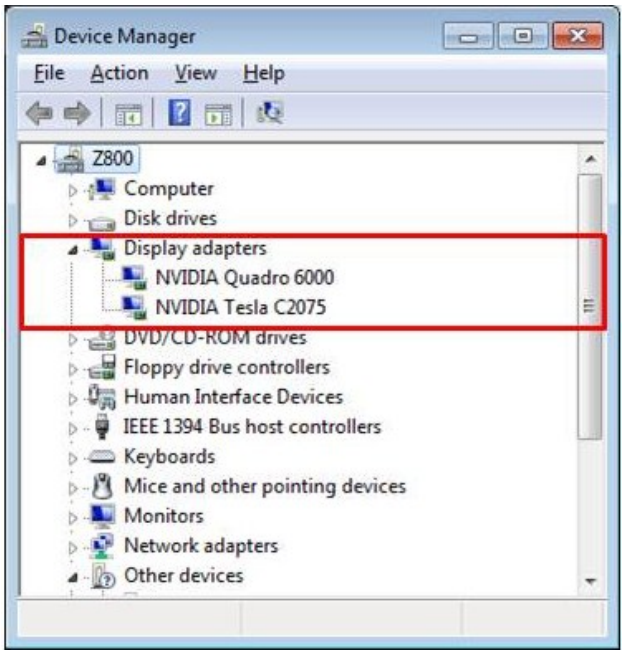


Note: In a standard configuration, the Tesla card is not displayed in the **NVIDIA Control Panel**. If displayed, no action is required by the user.

Check GPU Recognition and Quadro Driver Installation

Use the following procedure to check that the Quadro and Tesla cards are recognized by the operating system and the Quadro driver is correctly installed:

1. Open the Windows 7 Device Manager and check that the Quadro and Tesla cards are recognized as display adapters.



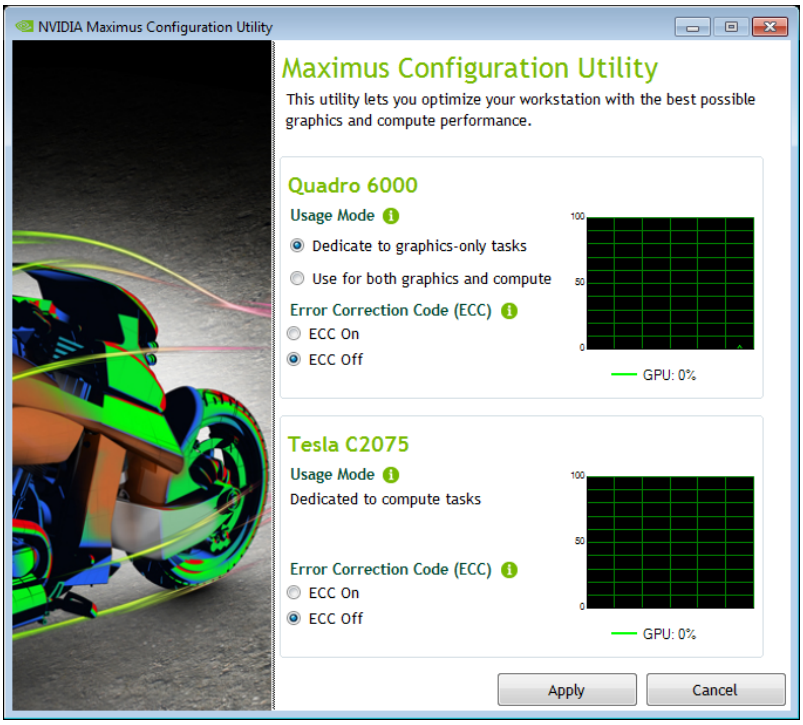
2. Check the driver versions, chassis slots, and power connections to the Quadro and Tesla cards if alerts are shown.
3. Check that the Quadro driver is correctly installed.

> Microsoft Windows Update will not necessarily provide the proper NVIDIA Quadro driver version. Please do not use the Microsoft Windows Update to upgrade your Quadro driver.

Checking Driver Recognition of Quadro and Tesla Cards

Use the Maximus Configuration Utility (MCU) to verify that all the GPUs that are physically installed in your system are listed in the MCU dialog window. The MCU is described on page 24.

1. Open the NVIDIA Maximus Configuration Utility.



2. Check that all Quadro and Tesla cards installed in your system are listed in the dialog window.

ADVANCED TOPICS

This section contains information about advanced Maximus configurations and advanced tuning of the GPUs in a Maximus system.

Advanced Maximus Configurations

The upper limit of GPUs that can be installed in a single workstation is often limited by the physical space and power consumption

It is possible to build Maximus-enabled systems with multiple Quadro cards and multiple Tesla cards. Some OEMs offer them as standard configurations. The main principles still apply to system implementation, verification, and configuration, as described in this document.

Figure 7 shows a workstation configuration with one Quadro 2000 card and two Tesla C2075 cards, where the Quadro 2000 is installed in the PCI Express x16 mechanical, x8 electrical slot of the workstation. This configuration was selected because the workflow and application requires less graphics visualization and more numerical computation.

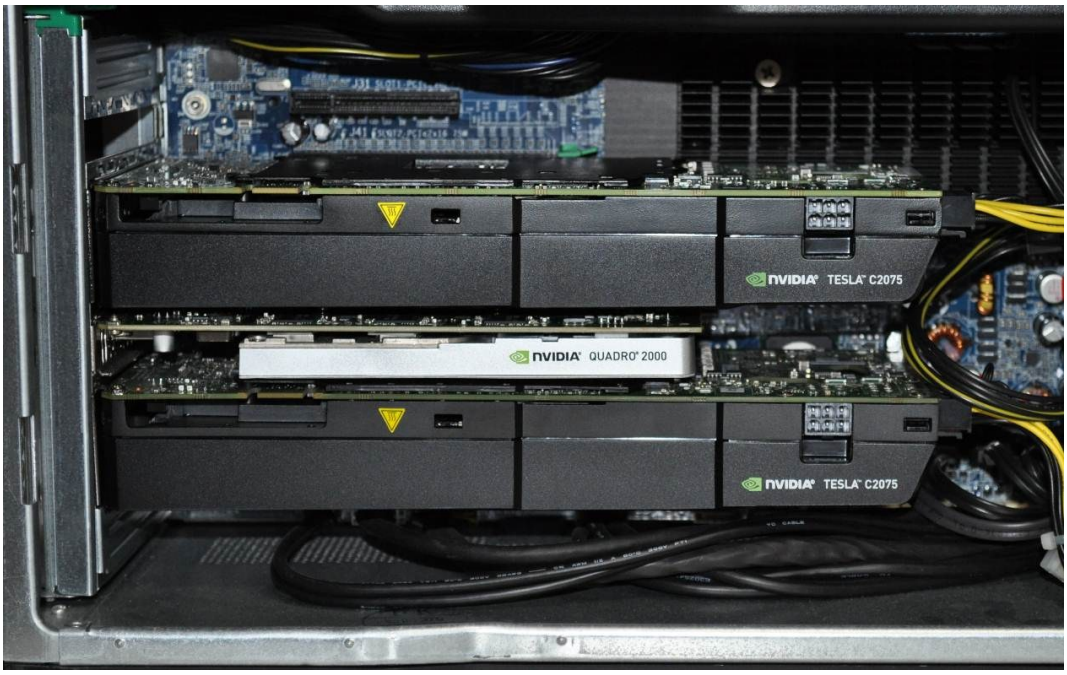


Figure 7. Workstation Configured with One Quadro 200 and two Tesla C2075 Cards

The workflow and application requires less graphics visualization and more numerical computation. When the Maximus Configuration Utility is run on this system, all Quadro and Tesla cards in the system will be listed (SEE Figure 8).

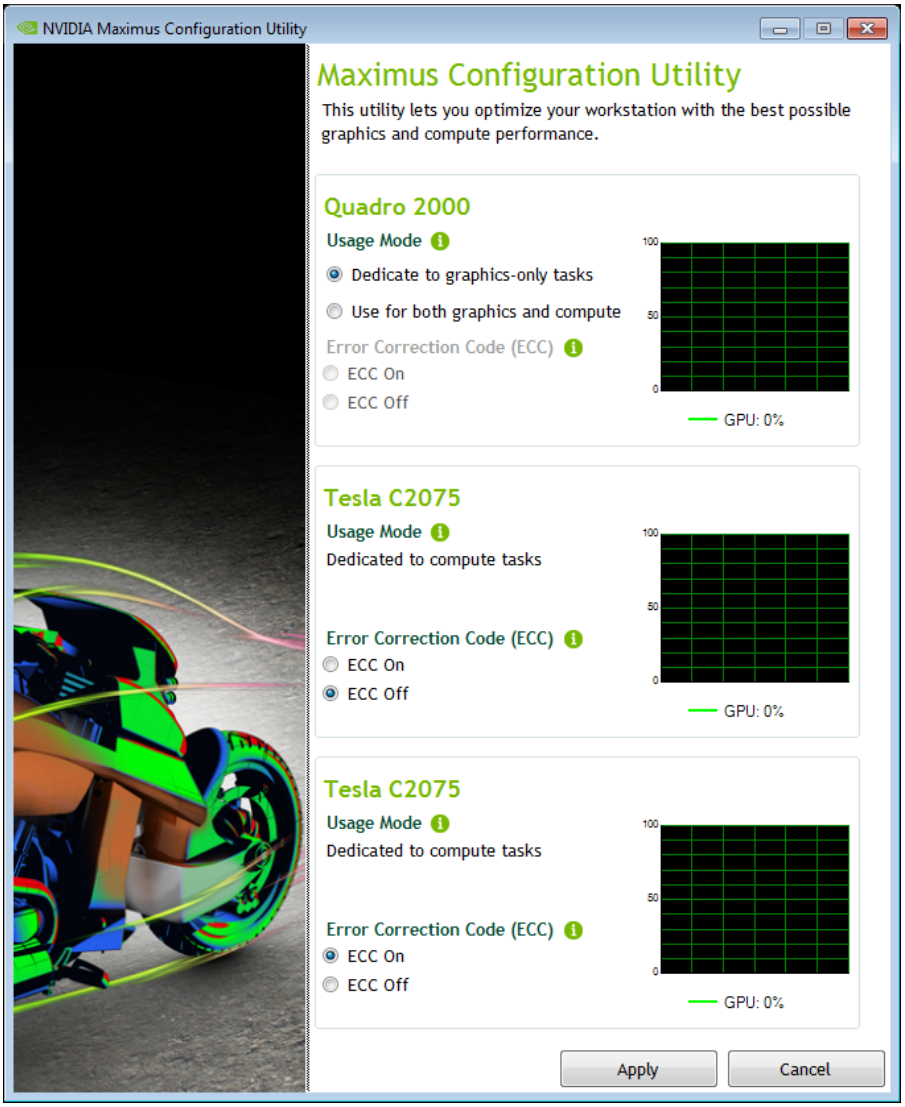


Figure 8. Maximus Configuration Utility Screen

Tuning Applications

A wide variety of GPU-enabled applications can leverage the power of an NVIDIA Maximus-enabled workstation. There are many different methods by which software vendors can author their applications to leverage the power. The methods are beyond the scope of this document; however, there are tuning adjustments that can be made to a Maximus-enabled system to accommodate certain software applications. Consult the software vendors if there are questions regarding specific tuning needs or contact your NVIDIA representative.

The MCU is a useful tool to monitor GPU activity.

GPU-enabled Applications

The availability of GPU-enabled applications is continuously expanding. Many software vendors are making use of not only one, but multiple GPUs in a system. These applications enable greater productivity and return on investment for users and their organizations. The benefits, which can be substantial, are dependent on workflow and organization.

An ever-growing list of GPU-enabled applications are listed on the following NVIDIA websites:

- http://www.nvidia.com/object/gpu-accelerated-applications.html
- http://www.nvidia.com/object/vertical_solutions.html

USING THE NVIDIA MAXIMUS CONFIGURATION UTILITY

There are several third-party GPU monitoring software tools available. These tools are not provided by NVIDIA and are not officially supported.

The NVIDIA Maximus Configuration Utility (MCU) is a separate graphical software utility that provides convenient GPU processing controls (Figure 9). The MCU provides GPU memory and utilization monitors for all supported GPUs in a Maximus-enabled system.

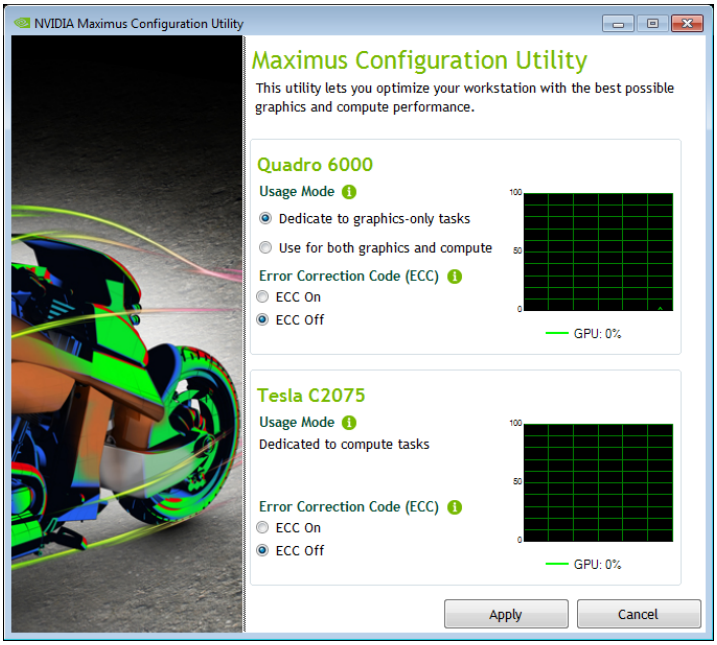


Figure 9. NVIDIA Maximus Configuration Utility (MCU) Screen

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.

HDMI

HDMI, the HDMI logo, and High-Definition Multimedia Interface are trademarks or registered trademarks of HDMI Licensing LLC.

ROVI Compliance Statement

NVIDIA Products that support Rovi Corporation's Revision 7.1.L1 Anti-Copy Process (ACP) encoding technology can only be sold or distributed to buyers with a valid and existing authorization from ROVI to purchase and incorporate the device into buyer's products.

This device is protected by U.S. patent numbers 6,516,132; 5,583,936; 6,836,549; 7,050,698; and 7,492,896 and other intellectual property rights. The use of ROVI Corporation's copy protection technology in the device must be authorized by ROVI Corporation and is intended for home and other limited pay-per-view uses only, unless otherwise authorized in writing by ROVI Corporation. Reverse engineering or disassembly is prohibited.

OpenCL

OpenCL is a trademark of Apple Inc. used under license to the Khronos Group Inc.

Trademarks

NVIDIA, the NVIDIA logo, Quadro, Tesla, and CUDA are trademarks and/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

© 2012 NVIDIA Corporation. All rights reserved.