



DU-05347-001_v7.0 | March 2015

Installation and Verification on Linux Systems



TABLE OF CONTENTS

Chapter 1. Introduction	1
1.1. System Requirements	1
1.1.1. x86 32-bit Support	2
1.2. About This Document	3
Chapter 2. Pre-installation Actions	4
2.1. Verify You Have a CUDA-Capable GPU	4
2.2. Verify You Have a Supported Version of Linux	4
2.3. Verify the System Has gcc Installed	5
2.4. Choose an Installation Method	5
2.5. Download the NVIDIA CUDA Toolkit	6
2.6. Handle Conflicting Installation Methods	6
Chapter 3. Package Manager Installation	8
3.1. Overview	8
3.2. Redhat/CentOS	8
3.3. Fedora	9
3.4. SLES	10
3.5. OpenSUSE	11
3.6. Ubuntu	12
3.7. Additional Package Manager Capabilities	12
3.7.1. Available Packages	12
3.7.2. Package Upgrades	13
Chapter 4. Runfile Installation	14
4.1. Overview	14
4.2. Installation	14
4.3. Disabling Nouveau	16
4.3.1. Fedora	16
4.3.2. RHEL/CentOS	16
4.3.3. OpenSUSE	16
4.3.4. SLES	17
4.3.5. Ubuntu	17
4.4. Device Node Verification	17
4.5. Advanced Options	18
4.6. Uninstallation	19
Chapter 5. CUDA Cross-Platform Environment	21
5.1. CUDA Cross-Platform Installation	21
5.2. CUDA Cross-Platform Samples	22
TARGET_ARCH	22
TARGET_OS	22
TARGET_FS	22
Copying Libraries	23

5.3. Nsight Eclipse Edition	23
Chapter 6. Post-installation Actions	24
6.1. Mandatory Actions	24
6.1.1. Environment Setup	24
6.2. Recommended Actions	24
6.2.1. Install Writable Samples	24
6.2.2. Verify the Installation	25
6.2.2.1. Verify the Driver Version	25
6.2.2.2. Compiling the Examples	25
6.2.2.3. Running the Binaries	25
6.3. Optional Actions	27
6.3.1. Install Third-party Libraries	27
6.3.2. Install the source code for cuda-gdb	27
Chapter 7. Frequently Asked Questions	29
How do I install the Toolkit in a different location?	29
Why do I see "nvcc: No such file or directory" when I try to build a CUDA application?	29
Why do I see "error while loading shared libraries: <lib name="">: cannot open shared object file</lib>	:
No such file or directory" when I try to run a CUDA application that uses a CUDA library?.	. 30
How can I extract the contents of the installers?	30
How can I tell X to ignore a GPU for compute-only use?	30
Why doesn't the cuda-repo package install the CUDA Toolkit and Drivers?	31
How do I get CUDA to work with Bumblebee?	31
Chapter 8 Additional Considerations	32

Chapter 1. INTRODUCTION

CUDA[®] is a parallel computing platform and programming model invented by NVIDIA. It enables dramatic increases in computing performance by harnessing the power of the graphics processing unit (GPU).

CUDA was developed with several design goals in mind:

- Provide a small set of extensions to standard programming languages, like C, that enable a straightforward implementation of parallel algorithms. With CUDA C/C++, programmers can focus on the task of parallelization of the algorithms rather than spending time on their implementation.
- Support heterogeneous computation where applications use both the CPU and GPU. Serial portions of applications are run on the CPU, and parallel portions are offloaded to the GPU. As such, CUDA can be incrementally applied to existing applications. The CPU and GPU are treated as separate devices that have their own memory spaces. This configuration also allows simultaneous computation on the CPU and GPU without contention for memory resources.

CUDA-capable GPUs have hundreds of cores that can collectively run thousands of computing threads. These cores have shared resources including a register file and a shared memory. The on-chip shared memory allows parallel tasks running on these cores to share data without sending it over the system memory bus.

This guide will show you how to install and check the correct operation of the CUDA development tools.

1.1. System Requirements

To use CUDA on your system, you will need the following installed:

- CUDA-capable GPU
- A supported version of Linux with a gcc compiler and toolchain
- NVIDIA CUDA Toolkit (available at http://developer.nvidia.com/cuda-downloads)

The CUDA development environment relies on tight integration with the host development environment, including the host compiler and C runtime libraries, and

is therefore only supported on distribution versions that have been qualified for this CUDA Toolkit release.

Table 1 Native Linux Distribution Support in CUDA 7.0

Distribution	Kernel	GCC	GLIBC	ICC	PGI	XLC	
x86_64							
Fedora 21	3.17.4	4.9.2	2.20				
RHEL 7.x	3.10	4.8.2	2.17				
RHEL 6.x	2.6.32	4.4.7	2.12			ИО	
CentOS 7.x	3.10	4.8.2	2.17				
CentOS 6.x	2.6.32	4.4.7	2.12				
OpenSUSE 13.2	3.16.6	4.8.3	2.19		>=14.9		
OpenSUSE 13.1	3.11.6	4.8.1	2.18	15.0.0			
SLES 12	3.12.28	4.8.6	2.19				
SLES 11 SP3	3.0.76	4.3.4	2.11.3				
Ubuntu 14.10	3.16	4.9.1	2.19				
Ubuntu 14.04	3.13	4.8.2	2.19				
SteamOS 1.0-beta	3.10.11	4.7.2	2.17				
Ubuntu 12.04 (DEPRECATED)	3.2	4.6.3	2.15				
	ARMv7 (aarch32)						
Ubuntu 14.04	3.13	4.8.2	2.19	NO	NO	NO	
POWER8(*)							
Ubuntu 14.10	3.16	4.9.1	2.19	NO	NO	13.1.1	
Ubuntu 14.04	3.16	4.8.2	2.19	NO	NO	13.1.1	

Table 2 Cross-build Environment Linux Distribution Support in CUDA 7.0

Host Distribution	Targeting Architectures (Linux)				
x86_64	x86(**) ARMv7 POWER8				
Ubuntu 14.04	DEPRECATED	YES	YES		
Ubuntu 12.04	DEPRECATED	DEPRECATED	NO		
SteamOS 1.0-beta	DEPRECATED	NO	NO		

^(*) Only the Tesla K40m GPU is supported for CUDA 7.0 on POWER8.

1.1.1. x86 32-bit Support

Support for x86 32-bit applications on x86_64 Linux is limited to use with:

^(**) x86 support is limited. See the x86 32-bit Support section for details.

- GeForce GPUs with Kepler or higher architecture
- CUDA Driver
- CUDA Runtime (cudart)
- CUDA Math Library (math.h)
- CUDA C++ Compiler (nvcc)
- CUDA Development Tools

Support for this configuration is only available in the .run file installer.

1.2. About This Document

This document is intended for readers familiar with the Linux environment and the compilation of C programs from the command line. You do not need previous experience with CUDA or experience with parallel computation. Note: This guide covers installation only on systems with X Windows installed.



Many commands in this document might require *superuser* privileges. On most distributions of Linux, this will require you to log in as root. For systems that have enabled the sudo package, use the sudo prefix for all necessary commands.

Chapter 2. PRE-INSTALLATION ACTIONS

Some actions must be taken before the CUDA Toolkit and Driver can be installed on Linux:

- Verify the system has a CUDA-capable GPU.
- Verify the system is running a supported version of Linux.
- Verify the system has gcc installed.
- Download the NVIDIA CUDA Toolkit.
- ► Handle conflicting installation methods.



You can override the install-time prerequisite checks by running the installer with the -override flag. Remember that the prerequisites will still be required to use the NVIDIA CUDA Toolkit.

2.1. Verify You Have a CUDA-Capable GPU

To verify that your GPU is CUDA-capable, go to your distribution's equivalent of System Properties, or, from the command line, enter:

```
$ lspci | grep -i nvidia
```

If you do not see any settings, update the PCI hardware database that Linux maintains by entering update-pciids (generally found in /sbin) at the command line and rerun the previous lspci command.

If your graphics card is from NVIDIA and it is listed in http://developer.nvidia.com/cuda-gpus, your GPU is CUDA-capable.

The Release Notes for the CUDA Toolkit also contain a list of supported products.

2.2. Verify You Have a Supported Version of Linux

The CUDA Development Tools are only supported on some specific distributions of Linux. These are listed in the CUDA Toolkit release notes.

To determine which distribution and release number you're running, type the following at the command line:

```
$ uname -m && cat /etc/*release
```

You should see output similar to the following, modified for your particular system:

```
x86_64
Red Hat Enterprise Linux Workstation release 6.0 (Santiago)
```

The **x86_64** line indicates you are running on a 64-bit system. The remainder gives information about your distribution.

2.3. Verify the System Has gcc Installed

The **gcc** compiler is required for development using the CUDA Toolkit. It is not required for running CUDA applications. It is generally installed as part of the Linux installation, and in most cases the version of gcc installed with a supported version of Linux will work correctly.

To verify the version of gcc installed on your system, type the following on the command line:

```
$ gcc --version
```

If an error message displays, you need to install the *development tools* from your Linux distribution or obtain a version of **gcc** and its accompanying toolchain from the Web.

2.4. Choose an Installation Method

The CUDA Toolkit can be installed using either of two different installation mechanisms: distribution-specific packages, or a distribution-independent package. The distribution-independent package has the advantage of working across a wider set of Linux distributions, but does not update the distribution's native package management system. The distribution-specific packages interface with the distribution's native package management system. It is recommended to use the distribution-specific packages, where possible.



Distribution-specific packages and repositories are not provided for Redhat 5. For Redhat 5, the stand-alone installer must be used.



Standalone installers are not provided for architectures other than the x86 and x86_64 releases. For both native as well as cross development, the toolkit must be installed using the distribution-specific installer. See the CUDA Cross-Platform Installation section for more details.

2.5. Download the NVIDIA CUDA Toolkit

The NVIDIA CUDA Toolkit is available at http://developer.nvidia.com/cuda-downloads.

Choose the platform you are using and download the NVIDIA CUDA Toolkit

The CUDA Toolkit contains the CUDA driver and tools needed to create, build and run a CUDA application as well as libraries, header files, CUDA samples source code, and other resources.

Download Verification

The download can be verified by comparing the MD5 checksum posted at http://developer.nvidia.com/cuda-downloads/checksums with that of the downloaded file. If either of the checksums differ, the downloaded file is corrupt and needs to be downloaded again.

To calculate the MD5 checksum of the downloaded file, run the following:

\$ md5sum <file>

2.6. Handle Conflicting Installation Methods

Before installing CUDA, any previously installations that could conflict should be uninstalled. This will not affect systems which have not had CUDA installed previously, or systems where the installation method has been preserved (RPM/Deb vs. Runfile). See the following charts for specifics.

Table 3 CUDA Toolkit Installation Compatibility Matrix

		Installed Toolkit Version == X.Y Installed Toolkit Version		t Version != X.Y	
		RPM/Deb	run	RPM/Deb	run
Installing	RPM/Deb	No Action	Uninstall Run	No Action	No Action
Toolkit Version X.Y	run	Uninstall RPM/Deb	Uninstall Run	No Action	No Action

Table 4 NVIDIA Driver Installation Compatibility Matrix

		Installed Driver Version == X.Y		Installed Driver Version != X.Y	
		RPM/Deb run		RPM/Deb	run
Installing Driver	RPM/Deb	No Action	Uninstall Run	No Action	Uninstall Run
Version X.Y	run	Uninstall RPM/ Deb	No Action	Uninstall RPM/ Deb	No Action

Use the following command to uninstall a Toolkit runfile installation:

```
$ sudo /usr/local/cuda-X.Y/bin/uninstall_cuda_X.Y.pl
```

Use the following command to uninstall a Driver runfile installation:

```
$ sudo /usr/bin/nvidia-uninstall
```

Use the following commands to uninstall a RPM/Deb installation:

```
$ sudo apt-get --purge remove <package_name>
$ sudo yum remove <package_name>
                                                                                       # Ubuntu
```

Fedora/Redhat/CentOS

OpenSUSE/SLES

^{\$} sudo zypper remove <package_name>

Chapter 3. PACKAGE MANAGER INSTALLATION

3.1. Overview

The Package Manager installation interfaces with your system's package management system. When using RPM or Deb, the downloaded package is a repository package. Such a package only informs the package manager where to find the actual installation packages, but will not install them.

If those packages are available in an online repository, they will be automatically downloaded in a later step. Otherwise, the repository package also installs a local repository containing the installation packages on the system. Whether the repository is available online or installed locally, the installation procedure is identical and made of several steps.

Distribution-specific instructions detail how to install CUDA:

- Redhat/CentOS
- Fedora
- ► SLES
- OpenSUSE
- Ubuntu

Finally, some helpful package manager capabilities are detailed.

These instructions are for native development only. For cross-platform development, see the CUDA Cross-Platform Environment section.

3.2. Redhat/CentOS

- 1. Perform the pre-installation actions.
- 2. Satisfy DKMS dependency

The NVIDIA driver RPM packages depend on other external packages, such as DKMS and libvdpau. Those packages are only available on third-party repositories,

such as EPEL. Any such third-party repositories must be added to the package manager repository database before installing the NVIDIA driver RPM packages, or missing dependencies will prevent the installation from proceeding.

3. Address custom xorg.conf, if applicable

The driver relies on an automatically generated xorg.conf file at /etc/X11/xorg.conf. If a custom-built xorg.conf file is present, this functionality will be disabled and the driver may not work. You can try removing the existing xorg.conf file, or adding the contents of /etc/X11/xorg.conf.d/00-nvidia.conf to the xorg.conf file. The xorg.conf file will most likely need manual tweaking for systems with a non-trivial GPU configuration.

4. Install repository meta-data

\$ sudo rpm --install cuda-repo-<distro>-<version>.<architecture>.rpm

5. Clean Yum repository cache

\$ sudo yum clean expire-cache

6. Install CUDA

\$ sudo yum install cuda

If the i686 libvdpau package dependency fails to install, try using the following steps to fix the issue:

```
$ yumdownloader libvdpau.i686
$ sudo rpm -U --oldpackage libvdpau*.rpm
```

7. Add libcuda.so symbolic link, if necessary

The libcuda.so library is installed in the /usr/lib{,64}/nvidia directory. For pre-existing projects which use libcuda.so, it may be useful to add a symbolic link from libcuda.so in the /usr/lib{,64} directory.

8. Perform the post-installation actions.

3.3. Fedora

1. Perform the pre-installation actions.

2. Address custom xorg.conf, if applicable

The driver relies on an automatically generated xorg.conf file at /etc/X11/xorg.conf. If a custom-built xorg.conf file is present, this functionality will be disabled and the driver may not work. You can try removing the existing xorg.conf file, or adding the contents of /etc/X11/xorg.conf.d/00-nvidia.conf to the xorg.conf file. The xorg.conf file will most likely need manual tweaking for systems with a non-trivial GPU configuration.

3. Satisfy Akmods dependency

The NVIDIA driver RPM packages depend on the Akmods framework which is provided by the RPMFusion free repository. The RPMFusion free repository must be added to the package manager repository database before installing the NVIDIA driver RPM packages, or missing dependencies will prevent the installation from proceeding.

4. Install repository meta-data

```
$ sudo rpm --install cuda-repo-<distro>-<version>.<architecture>.rpm
```

5. Clean Yum repository cache

```
$ sudo yum clean expire-cache
```

6. Install CUDA

```
$ sudo yum install cuda
```

The CUDA driver installation may fail if the RPMFusion non-free repository is enabled. In this case, CUDA installations should temporarily disable the RPMFusion non-free repository:

```
$ sudo yum --disablerepo="rpmfusion-nonfree*" install cuda
```

If also installing the <code>gpu-deployment-kit</code> package, the <code>cuda</code> and <code>gpu-deployment-kit</code> packages should be either installed using separate instances of <code>yum</code>:

```
$ sudo yum install cuda
$ sudo yum install gpu-deployment-kit
```

Or, installed while also specifying the **cuda-drivers** package:

```
$ sudo yum install cuda cuda-drivers gpu-deployment-kit
```

If a system has installed both packages with the same instance of **yum**, some driver components may be missing. Such an installation can be corrected by running:

```
$ sudo yum install cuda-drivers
```

If the i686 libvdpau package dependency fails to install, try using the following steps to fix the issue:

```
$ yumdownloader libvdpau.i686
$ sudo rpm -U --oldpackage libvdpau*.rpm
```

7. Add libcuda.so symbolic link, if necessary

The libcuda.so library is installed in the /usr/lib{,64}/nvidia directory. For preexisting projects which use libcuda.so, it may be useful to add a symbolic link from libcuda.so in the /usr/lib{,64} directory.

8. Perform the post-installation actions.

3.4. SLES

- 1. Perform the pre-installation actions.
- 2. Install repository meta-data

```
$ sudo rpm --install cuda-repo-<distro>-<version>.<architecture>.rpm
```

3. Refresh Zypper repository cache

```
$ sudo zypper refresh
```

4. Install CUDA

```
$ sudo zypper install cuda
```

The driver is provided in multiple packages, nvidia-gfxG03-kmp-desktop, nvidia-gfxG03-kmp-default, nvidia-gfxG03-kmp-trace, and their Unified Memory variants. When installing cuda, the correct driver packages should also be specified. Without doing this, zypper will select packages that may not work on the system. Run the following to detect the flavor of kernel and install cuda with the appropriate driver packages:

5. Add the user to the video group

```
$ sudo usermod -a -G video <username>
```

6. Install CUDA Samples GL dependencies

The CUDA Samples package on SLES does not include dependencies on GL and X11 libraries as these are provided in the SLES SDK. These packages must be installed separately, depending on which samples you want to use.

7. Perform the post-installation actions.

3.5. OpenSUSE

- 1. Perform the pre-installation actions.
- 2. Install repository meta-data

```
$ sudo rpm --install cuda-repo-<distro>-<version>.<architecture>.rpm
```

3. Refresh Zypper repository cache

```
$ sudo zypper refresh
```

4. Install CUDA

```
$ sudo zypper install cuda
```

The driver is provided in multiple packages, nvidia-gfxG03-kmp-desktop, nvidia-gfxG03-kmp-default, nvidia-gfxG03-kmp-trace, and their Unified Memory variants. When installing cuda, the correct driver packages should also be specified. Without doing this, zypper will select packages that may not work on the system. Run the following to detect the flavor of kernel and install cuda with the appropriate driver packages:

5. Add the user to the video group

```
$ sudo usermod -a -G video <username>
```

6. Perform the post-installation actions.

3.6. Ubuntu

- 1. Perform the pre-installation actions.
- 2. Install repository meta-data



When using a proxy server with aptitude, ensure that wget is set up to use the same proxy settings before installing the cuda-repo package.

```
$ sudo dpkg -i cuda-repo-<distro> <version> <architecture>.deb
```

3. Update the Apt repository cache

```
$ sudo apt-get update
```

4. Install CUDA

```
$ sudo apt-get install cuda
```

5. Perform the post-installation actions.

3.7. Additional Package Manager Capabilities

Below are some additional capabilities of the package manager that users can take advantage of.

3.7.1. Available Packages

The recommended installation package is the **cuda** package. This package will install the full set of other CUDA packages required for native development and should cover most scenarios.

The **cuda** package installs all the available packages for native developments. That includes the compiler, the debugger, the profiler, the math libraries,... For x86_64 patforms, this also include NSight Eclipse Edition and the visual profiler It also includes the NVIDIA driver package.

On supported platforms, the cuda-cross-armhf and cuda-cross-ppc64el packages install all the packages required for cross-platform development to ARMv7 and POWER8, respectively. The libraries and header files of the target architecture's display driver package are also installed to enable the cross compilation of driver applications. The cuda-cross-<arch> packages do not install the native display driver.

The packages installed by the packages above can also be installed individually by specifying their names explicitly. The list of available packages be can obtained with:

```
$ yum --disablerepo="*" --enablerepo="cuda*" list available  # RedHat & Fedora
$ zypper packages -r cuda  # OpenSUSE & SLES
$ cat /var/lib/apt/lists/*cuda*Packages | grep "Package:"  # Ubuntu
```

3.7.2. Package Upgrades

The **cuda** package points to the latest stable release of the CUDA Toolkit. When a new version is available, use the following commands to upgrade the toolkit and driver:

The **cuda-cross-<arch>** packages can also be upgraded in the same manner.

The cuda-drivers package points to the latest driver release available in the CUDA repository. When a new version is available, use the following commands to upgrade the driver:

Some desktop environments, such as GNOME or KDE, will display an notification alert when new packages are available.

To avoid any automatic upgrade, and lock down the toolkit installation to the X.Y release, install the **cuda-X-Y** or **cuda-cross-<arch>-X-Y** package.

Side-by-side installations are supported. For instance, to install both the X.Y CUDA Toolkit and the X.Y+1 CUDA Toolkit, install the **cuda-X.Y** and **cuda-X.Y+1** packages.

Chapter 4. RUNFILE INSTALLATION

This section describes the installation and configuration of CUDA when using the standalone installer. The standalone installer is a ".run" file and is completely self-contained.

4.1. Overview

The Runfile installation installs the NVIDIA Driver, CUDA Toolkit, and CUDA Samples via an interactive text-based interface.

The installation steps are listed below. Distribution-specific instructions on disabling the Nouveau drivers as well as steps for verifying device node creation are also provided.

Finally, advanced options for the installer and uninstallation steps are detailed below.

The Runfile installation does not include support for cross-platform development. For cross-platform development, see the CUDA Cross-Platform Environment section.

4.2. Installation

- 1. Perform the pre-installation actions.
- 2. Disable the Nouveau drivers.
- 3. Reboot into text mode (runlevel 3).

This can usually be accomplished by adding the number "3" to the end of the system's kernel boot parameters.

Since the NVIDIA drivers are not yet installed, the text terminals may not display correctly. Temporarily adding "nomodeset" to the system's kernel boot parameters may fix this issue.

Consult your system's bootloader documentation for information on how to make the above boot parameter changes. The reboot is required to completely unload the Nouveau drivers and prevent the graphical interface from loading. The CUDA driver cannot be installed while the Nouveau drivers are loaded or while the graphical interface is active.

- 4. Verify that the Nouveau drivers are not loaded. If the Nouveau drivers are still loaded, consult your distribution's documentation to see if further steps are needed to disable Nouveau.
- 5. Run the installer and follow the on-screen prompts:

```
$ sudo sh cuda <version> linux.run
```

The installer will prompt for the following:

- EULA Acceptance
- CUDA Driver installation
- CUDA Toolkit installation, location, and /usr/local/cuda symbolic link
- CUDA Samples installation and location

The default installation locations for the toolkit and samples are:

Component	Default Installation Directory
CUDA Toolkit	/usr/local/cuda-7.0
CUDA Samples	\$(HOME)/NVIDIA_CUDA-7.0_Samples

The /usr/local/cuda symbolic link points to the location where the CUDA Toolkit was installed. This link allows projects to use the latest CUDA Toolkit without any configuration file update.

The installer must be executed with sufficient privileges to perform some actions. When the current privileges are insufficient to perform an action, the installer will ask for the user's password to attempt to install with root privileges. Actions that cause the installer to attempt to install with root privileges are:

- installing the CUDA Driver
- installing the CUDA Toolkit to a location the user does not have permission to write to
- installing the CUDA Samples to a location the user does not have permission to write to
- creating the /usr/local/cuda symbolic link

Running the installer with **sudo**, as shown above, will give permission to install to directories that require root permissions. Directories and files created while running the installer with **sudo** will have root ownership.

If installing the driver, the installer will also ask if the openGL libraries should be installed. If the GPU used for display is not an NVIDIA GPU, the NVIDIA openGL libraries should not be installed. Otherwise, the openGL libraries used by the graphics driver of the non-NVIDIA GPU will be overwritten and the GUI will not work. If performing a silent installation, the --no-opengl-libs option should be used to prevent the openGL libraries from being installed. See the Advanced Options section for more details.

If the GPU used for display is an NVIDIA GPU, the X server configuration file, /etc/X11/xorg.conf, may need to be modified. In some cases, nvidia-xconfig can be used to automatically generate a xorg.conf file that works for the system. For non-standard systems, such as those with more than one GPU, it is recommended to manually edit the xorg.conf file. Consult the xorg.conf documentation for more information.



Installing Mesa may overwrite the /usr/lib/libGL.so that was previously installed by the NVIDIA driver, so a reinstallation of the NVIDIA driver might be required after installing these libraries.

- 6. Reboot the system to reload the graphical interface.
- 7. Verify the device nodes are created properly.
- 8. Perform the post-installation actions.

4.3. Disabling Nouveau

To install the Display Driver, the Nouveau drivers must first be disabled. Each distribution of Linux has a different method for disabling Nouveau.

The Nouveau drivers are loaded if the following command prints anything:

\$ 1smod | grep nouveau

4.3.1. Fedora

1. Create a file at /usr/lib/modprobe.d/blacklist-nouveau.conf with the following contents:

```
blacklist nouveau
options nouveau modeset=0
```

2. Regenerate the kernel initramfs:

\$ sudo dracut --force

4.3.2. RHEL/CentOS

1. Create a file at /etc/modprobe.d/blacklist-nouveau.conf with the following contents:

```
blacklist nouveau
options nouveau modeset=0
```

2. Regenerate the kernel initramfs:

\$ sudo dracut --force

4.3.3. OpenSUSE

1. Create a file at /etc/modprobe.d/blacklist-nouveau.conf with the following contents:

```
blacklist nouveau
options nouveau modeset=0
```

2. Regenerate the kernel initrd:

\$ sudo /sbin/mkinitrd

4.3.4. SLES

No actions to disable Nouveau are required as Nouveau is not installed on SLES.

4.3.5. Ubuntu

1. Create a file at /etc/modprobe.d/blacklist-nouveau.conf with the following contents:

```
blacklist nouveau
options nouveau modeset=0
```

2. Regenerate the kernel initramfs:

\$ sudo update-initramfs -u

4.4. Device Node Verification

Check that the device files/dev/nvidia* exist and have the correct (0666) file permissions. These files are used by the CUDA Driver to communicate with the kernel-mode portion of the NVIDIA Driver. Applications that use the NVIDIA driver, such as a CUDA application or the X server (if any), will normally automatically create these files if they are missing using the *setuid* nvidia-modprobe tool that is bundled with the NVIDIA Driver. However, some systems disallow setuid binaries, so if these files do not

exist, you can create them manually either by running the command **nvidia-smi** as root at boot time or by using a startup script such as the one below:

```
#!/bin/bash
/sbin/modprobe nvidia
if [ "$?" -eq 0 ]; then
  # Count the number of NVIDIA controllers found.

NVDEVS=`lspci | grep -i NVIDIA`

N3D=`echo "$NVDEVS" | grep "3D controller" | wc -l`

NVGA=`echo "$NVDEVS" | grep "VGA compatible controller" | wc -l`
  N=\ensuremath{`expr}\ \$N3D + \$NVGA - 1\ensuremath{`}
  for i in `seq 0 $N`; do
    mknod -m 666 /dev/nvidia$i c 195 $i
  mknod -m 666 /dev/nvidiactl c 195 255
else
 exit 1
fi
/sbin/modprobe nvidia-uvm
if [ "$?" -eq 0 ]; then
  # Find out the major device number used by the nvidia-uvm driver
  D=`grep nvidia-uvm /proc/devices | awk '{print $1}'`
  mknod -m 666 /dev/nvidia-uvm c $D 0
else
 exit 1
fi
```

4.5. Advanced Options

Action	Options Used	Explanation
Silent Installation	silent	Required for any silent installation. Performs an installation with no further user-input and minimal commandline output based on the options provided below. Silent installations are useful for scripting the installation of CUDA. Using this option implies acceptance of the EULA. The following flags can be used to customize the actions taken during installation. At least one ofdriver,uninstall,toolkit, andsamples must be passed if running with non-root permissions.
	driver	Install the CUDA Driver.
	toolkit	Install the CUDA Toolkit.

Action	Options Used	Explanation
	toolkitpath= <path></path>	Install the CUDA Toolkit to the <path> directory. If not provided, the default path of /usr/local/cuda-7.0 is used.</path>
	samples	Install the CUDA Samples.
	samplespath= <path></path>	Install the CUDA Samples to the <path> directory. If not provided, the default path of \$(HOME)/NVIDIA_CUDA-7.0_Samples is used.</path>
	uninstall (DEPRECATED)	Unnstall the CUDA Driver.
Extraction	extract= <path></path>	Extracts the individual driver, toolkit, and samples runfile installers to <path>.</path>
Overriding Installation Checks	override	Ignores compiler, third-party library, and toolkit detection checks.
No OpenGL Libraries	no-opengl-libs	Prevents the driver installation from installing NVIDIA's GL libraries. Useful for systems where the display is driven by a non-NVIDIA GPU. In such systems, NVIDIA's GL libraries could prevent X from loading properly.
Overriding Kernel Source	kernel-source-path= <path></path>	Tells the driver installation to use <path> as the kernel source directory when building the NVIDIA kernel module. Required for systems where the kernel source is installed to a non-standard location.</path>
Custom Temporary Directory Selection	tmpdir= <path></path>	Performs any temporary actions within <path> instead of /tmp. Useful in cases where /tmp cannot be used (doesn't exist, is full, is mounted with 'noexec', etc.).</path>
Show Verbose Output	verbose	Prints all logging information to stdout as well as the log file.
Show Installer Options	help	Prints the list of command-line options to stdout.

4.6. Uninstallation

To uninstall the CUDA Toolkit, run the uninstallation script provided in the bin directory of the toolkit. By default, it is located in /usr/local/cuda-7.0/bin:

\$ sudo /usr/local/cuda-7.0/bin/uninstall_cuda_7.0.pl

To uninstall the NVIDIA Driver, run nvidia-uninstall:

\$ sudo /usr/bin/nvidia-uninstall

To enable the Nouveau drivers, remove the blacklist file created in the Disabling Nouveau section, and regenerate the kernel initramfs/initrd again as described in that section.

Chapter 5. CUDA CROSS-PLATFORM ENVIRONMENT

Cross-platform development is only supported on Ubuntu systems, and is only provided via the Package Manager installation process.

We recommend selecting Ubuntu 14.04 as your cross-platform development environment. This selection helps prevent host/target incompatibilities, such as GCC or GLIBC version mismatches.

5.1. CUDA Cross-Platform Installation

Some of the following steps may have already been performed as part of the native Ubuntu installation. Such steps can safely be skipped.

These steps should be performed on the x86_64 host system, rather than the target system. To install the native CUDA Toolkit on the target system, refer to the native Ubuntu installation section.

- 1. Perform the pre-installation actions.
- 2. Enable the foreign architecture

The foreign architecture must be enabled in order to install the cross-platform toolkit. To enable armhf or ppc64el as a foreign architecture, the following commands must be executed, replacing <arch> with the desired architecture:

```
$ sudo dpkg --add-architecture <arch>
$ sudo apt-get update
```

3. Install repository meta-data



When using a proxy server with aptitude, ensure that wget is set up to use the same proxy settings before installing the cuda-repo package.

```
$ sudo dpkg -i cuda-repo-<distro> <version> amd64.deb
```

4. Update the Apt repository cache

```
$ sudo apt-get update
```

5. Install the cross-platform CUDA Toolkit

To install the cross-platform toolkit for armhf or ppc64el, the following command must be executed, replacing <arch> with the desired architecture:

\$ sudo apt-get install cuda-cross-<arch>

6. Perform the post-installation actions.

5.2. CUDA Cross-Platform Samples

This section describes the options used to build cross-platform samples.

TARGET_ARCH=<arch> and TARGET_OS=<os> should be chosen based on the supported targets shown below. TARGET_FS=<path> can be used to point nvcc to libraries and headers used by the sample.

Table 5 Supported Target Arch/OS Combinations

		TARGET OS			
		linux	darwin	android	qnx
	x86_64	YES	YES	NO	NO
TARGET ARCH	armv7l	YES	NO	YES	YES
IARGET ARCIT	aarch64	NO	NO	YES	NO
	ppc64le	YES	NO	NO	NO

TARGET_ARCH

The target architecture must be specified when cross-compiling applications. If not specified, it defaults to the host architecture. Allowed architectures are:

- **x86 64** 64-bit x86 CPU architecture
- armv71 32-bit ARM CPU architecture, like that found on Jetson TK1
- aarch64 64-bit ARM CPU architecture, found on certain Android systems
- ppc641e 64-bit little-endian IBM POWER8 architecture

TARGET OS

The target OS must be specified when cross-compiling applications. If not specified, it defaults to the host OS. Allowed OSes are:

- ▶ linux for any Linux distributions
- darwin for Mac OS X
- android for any supported device running Android
- qnx for any supported device running QNX

TARGET FS

The most reliable method to cross-compile the CUDA Samples is to use the TARGET_FS variable. To do so, mount the target's filesystem on the host, say at /mnt/target. This is

typically done using **exportfs**. In cases where **exportfs** is unavailable, it is sufficient to copy the target's filesystem to /mnt/target. To cross-compile a sample, execute:

```
$ make TARGET_ARCH=<arch> TARGET_OS=<os> TARGET_FS=/mnt/target
```

Copying Libraries

If the TARGET_FS option is not available, the libraries used should be copied from the target system to the host system, say at <code>/opt/target/libs</code>. If the sample uses GL, the GL headers must also be copied, say at <code>/opt/target/include</code>. The linker must then be told where the libraries are with the <code>-rpath-link</code> and/or <code>-L</code> options. To ignore unresolved symbols from some libraries, use the <code>--unresolved-symbols</code> option as shown below. <code>SAMPLE_ENABLED</code> should be used to force the sample to build. For example, to cross-compile a sample which uses such libraries, execute:

5.3. Nsight Eclipse Edition

Nsight Eclipse Edition supports cross-platform development. See the Nsight Eclipse Edition Getting Started Guide for more details.

Chapter 6. POST-INSTALLATION ACTIONS

The post-installation actions must be manually performed. These actions are split into mandatory, recommended, and optional sections.

6.1. Mandatory Actions

Some actions must be taken after installing the CUDA Toolkit and Driver before they can be completely used.

6.1.1. Environment Setup

The PATH variable needs to include /usr/local/cuda-7.0/bin

The LD_LIBRARY_PATH variable needs to contain /usr/local/cuda-7.0/lib64 on a 64-bit system, and /usr/local/cuda-7.0/lib on a 32-bit system

► To change the environment variables for 64-bit operating systems:

```
$ export PATH=/usr/local/cuda-7.0/bin:$PATH
$ export LD_LIBRARY_PATH=/usr/local/cuda-7.0/lib64:$LD LIBRARY PATH
```

► To change the environment variables for 32-bit operating systems:

```
$ export PATH=/usr/local/cuda-7.0/bin:$PATH
$ export LD_LIBRARY_PATH=/usr/local/cuda-7.0/lib:$LD_LIBRARY_PATH
```

6.2. Recommended Actions

Other actions are recommended to verify the integrity of the installation.

6.2.1. Install Writable Samples

In order to modify, compile, and run the samples, the samples must be installed with write permissions. A convenience installation script is provided:

```
$ cuda-install-samples-7.0.sh <dir>
```

This script is installed with the cuda-samples-7-0 package. The cuda-samples-7-0 package installs only a read-only copy in /usr/local/cuda-7.0/samples.

6.2.2. Verify the Installation

Before continuing, it is important to verify that the CUDA toolkit can find and communicate correctly with the CUDA-capable hardware. To do this, you need to compile and run some of the included sample programs.



Ensure the PATH and LD_LIBRARY_PATH variables are set correctly.

6.2.2.1. Verify the Driver Version

If you installed the driver, verify that the correct version of it is loaded. If you did not install the driver, or are using an operating system where the driver is not loaded via a kernel module, such as L4T, skip this step.

When the driver is loaded, the driver version can be found by executing the command

\$ cat /proc/driver/nvidia/version

Note that this command will not work on an iGPU/dGPU system.

6.2.2.2. Compiling the Examples

The version of the CUDA Toolkit can be checked by running **nvcc -v** in a terminal window. The **nvcc** command runs the compiler driver that compiles CUDA programs. It calls the **gcc** compiler for C code and the NVIDIA PTX compiler for the CUDA code.

The NVIDIA CUDA Toolkit includes sample programs in source form. You should compile them by changing to ~/NVIDIA_CUDA-7.0_Samples and typing make. The resulting binaries will be placed under ~/NVIDIA_CUDA-7.0_Samples/bin.

6.2.2.3. Running the Binaries

After compilation, find and run **deviceQuery** under **~/NVIDIA_CUDA-7.0_Samples**. If the CUDA software is installed and configured correctly, the output for **deviceQuery** should look similar to that shown in Figure 1.

```
./deviceQuery Starting...

CUDA Device Query (Runtime API) version (CUDART static linking)

Detected 1 CUDA Capable device(s)

Device 0: "Tesla X20c"

CUDA Oriver Version / Runtime Version

CUDA Oriver Version / Runtime Version

CUDA Capability Major/Minor version number:

Total amount of global memory:

(13) Multiprocessors, (192) CUDA Cores/MP:

Memory Bus Width:

12 Cache Site:

130-bit 1320-bit 1320-b
```

Figure 1 Valid Results from deviceQuery CUDA Sample

The exact appearance and the output lines might be different on your system. The important outcomes are that a device was found (the first highlighted line), that the device matches the one on your system (the second highlighted line), and that the test passed (the final highlighted line).

If a CUDA-capable device and the CUDA Driver are installed but **deviceQuery** reports that no CUDA-capable devices are present, this likely means that the **/dev/nvidia*** files are missing or have the wrong permissions.

On systems where **SELinux** is enabled, you might need to temporarily disable this security feature to run **deviceQuery**. To do this, type:

```
$ setenforce 0
```

from the command line as the *superuser*.

Running the **bandwidthTest** program ensures that the system and the CUDA-capable device are able to communicate correctly. Its output is shown in Figure 2.

```
[CUDA Bandwidth Test] - Starting...
Running on...
 Device 0: Quadro K5000
 Quick Mode
 Host to Device Bandwidth, 1 Device(s)
 PINNED Memory Transfers
  Transfer Size (Bytes)

Bandwidth(MB/s)
                              5798.4
  33554432
 Device to Host Bandwidth, 1 Device(s)
 PINNED Memory Transfers
                           Bandwidth(MB/s)
6378.4
  Transfer Size (Bytes)
  33554432
 Device to Device Bandwidth, 1 Device(s)
 PINNED Memory Transfers
                             Bandwidth(MB/s)
   Transfer Size (Bytes)
   33554432
                               133606.8
Result = PASS
```

Figure 2 Valid Results from bandwidthTest CUDA Sample

Note that the measurements for your CUDA-capable device description will vary from system to system. The important point is that you obtain measurements, and that the second-to-last line (in Figure 2) confirms that all necessary tests passed.

Should the tests not pass, make sure you have a CUDA-capable NVIDIA GPU on your system and make sure it is properly installed.

If you run into difficulties with the link step (such as libraries not being found), consult the *Linux Release Notes* found in the **doc** folder in the CUDA Samples directory.

6.3. Optional Actions

Other options are not necessary to use the CUDA Toolkit, but are available to provide additional features.

6.3.1. Install Third-party Libraries

Some CUDA samples use third-party libraries which may not be installed by default on your system. These samples attempt to detect any required libraries when building. If a library is not detected, it waives itself and warns you which library is missing. To build and run these samples, you must install the missing libraries.

6.3.2. Install the source code for cuda-gdb

The cuda-gdb source is automatically installed with the runfile installation method.

To obtain a copy of the source code for cuda-gdb using the RPM and Debian installation methods, the **cuda-gdb-src** package must be installed.

The source code is installed as a tarball in the /usr/local/cuda-7.0/extras directory.

Chapter 7. FREQUENTLY ASKED QUESTIONS

How do I install the Toolkit in a different location?

The Runfile installation asks where you wish to install the Toolkit and the Samples during an interactive install. If installing using a non-interactive install, you can use the --toolkitpath and --samplespath parameters to change the install location:

The RPM packages don't support custom install locations though the package managers (Yum and Zypper), but it is possible to install the RPM packages in custom locations using rpm's --relocate parameter:

```
$ rpm --install --relocate /usr/local/cuda-7.0=/my/new/toolkit rpmpackage.rpm
```

You will need to install the packages in the correct dependency order; this task is normally taken care of by the package managers. For example, if package "foo" has a dependency on package "bar", you should install package "bar" first, and package "foo" second. You can check the dependencies of a RPM package as follows:

```
$ rpm -qRp rpmpackage.rpm
```

The Deb packages do not support custom install locations. It is however possible to extract the contents of the Deb packages and move the files to the desired install location. See here for more details on extracting Deb packages.

Why do I see "nvcc: No such file or directory" when I try to build a CUDA application?

Your PATH environment variable is not set up correctly. Ensure that your PATH includes the bin directory where you installed the Toolkit, usually /usr/local/cuda-7.0/bin.

```
$ export PATH=/usr/local/cuda-7.0/bin:$PATH
```

Why do I see "error while loading shared libraries: lib name>: cannot open shared object file: No such file or directory" when I try to run a CUDA application that uses a CUDA library?

Your LD_LIBRARY_PATH environment variable is not set up correctly. Ensure that your LD_LIBRARY_PATH includes the lib and/or lib64 directory where you installed the Toolkit, usually /usr/local/cuda-7.0/lib{,64}:

```
$ export LD LIBRARY PATH=/usr/local/cuda-7.0/lib:$LD LIBRARY PATH
```

How can I extract the contents of the installers?

The Runfile can be extracted into the standalone Toolkit, Samples and Driver Runfiles by using the --extract parameter. These standalone Runfiles can be further extracted by running:

```
$ ./runfile.run --tar mxvf
The RPM packages can be extracted by running:
```

```
$ rpm2cpio rpm_package.rpm | cpio -idmv
```

The Deb packages can be extracted by running:

```
$ dpkg-deb -x deb_package.deb output_dir
```

How can I tell X to ignore a GPU for compute-only use?

To make sure X doesn't use a certain GPU for display, you need to specify which **other** GPU to use for display. This is done by editing the xorg.conf file located at /etc/xorg/xorg.conf.

You will need to add a section that resembles the following to your xorg.conf file:

```
Section "Device"

Identifier "Device0"

Driver "driver_name"

VendorName "vendor_name"

BusID "bus_id"

EndSection
```

The exact details of what you will need to add differ on a case-by-case basis. For example, if you have two NVIDIA GPUs and you want the first GPU to be used for display, you would replace "driver_name" with "nvidia", "vendor_name" with "NVIDIA Corporation" and "bus_id" with the Bus ID of the GPU.

The Bus ID will resemble "PCI:00:02.0" and can be found by running lspci.

Why doesn't the cuda-repo package install the CUDA Toolkit and Drivers?

When using RPM or Deb, the downloaded package is a repository package. Such a package only informs the package manager where to find the actual installation packages, but will not install them.

See the Package Manager Installation section for more details.

How do I get CUDA to work with Bumblebee?

Bumblebee works by establishing the right runtime environment to execute the targeted application through the <code>optirun</code> command. To run, debug, or profile a CUDA application, <code>optirun</code> must set up the correct runtime environment, which includes, among other items, loading the appropriate kernel modules. The Bumblebee project should provide the information on how to proceed. Please see the Bumblebee Documentation for more information.

Chapter 8. <u>ADDITIONAL CONSIDERATIONS</u>

Now that you have CUDA-capable hardware and the NVIDIA CUDA Toolkit installed, you can examine and enjoy the numerous included programs. To begin using CUDA to accelerate the performance of your own applications, consult the *CUDA C Programming Guide*, located in /usr/local/cuda-7.0/doc.

A number of helpful development tools are included in the CUDA Toolkit to assist you as you develop your CUDA programs, such as NVIDIA[®] Nsight[™] Eclipse Edition, NVIDIA Visual Profiler, cuda-gdb, and cuda-memcheck.

For technical support on programming questions, consult and participate in the developer forums at http://developer.nvidia.com/cuda/.

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 and the NVIDIA logo 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

© 2009-2015 NVIDIA Corporation. All rights reserved.

