**Guide for Installing AI & Deep Learning Programs (Linux Ubuntu 16.04 – 17)**

**List of Programs we will be installing in this guide.**

1. **Nvidia Drivers (Latest Version from Nvidia Website)**
2. **Nvidia Cuda 9.1**
3. **Nvidia Digits**
4. **Nvidia cuDNN**
5. **Torch**
6. **Caffe**
7. **PyTorch**
8. **TorchVision**
9. **Tensor Flow**
10. **Keras**

**Pre-Installation:**

**Before using this guide make sure you have the following things installed on your system. This guide is applicable for all the Systems running Nvidia Graphics Card/GPU.**

**Make sure you have a fresh copy of Linux Ubuntu 16.04 or greater installed on your system before attempting to use this guide.**

**This guide was developed on a system running Core i7 8700K and GTX 1080Ti but can be used for any Nvidia GPU. I will provide all the steps as easy as possible along with explanations and screen shots.**

**All the commands use in this guide are all single line commands on terminals**

**Nvidia Drivers Installation:**

First, open up a terminal and run this command

**sudo su**

and login as root.

Now after your login its time to update the system. To update your system run this command.

**apt-get update && apt-get upgrade**

After you update the system. Its time to uninstall any older drivers or packages of Nvidia. As you know that Ubuntu by default comes with its own drivers for Nvidia which are on an older version as compare to the current latest drivers. And Cuda 9.1 supports R390 or 390.XX version of drivers.

To uninstall the old drivers please use these command below

**apt-get purge nvidia\***

**apt-get autoremove**

This command will remove your old Nvidia drivers and will allow you to install new drivers. Please note it is not necessary for you to have Nvidia drivers installed of if after running the commands you cannot find any packages for Nvidia its completely fine you can continue ahead.

Now we have the old drivers or installation removed we can go ahead and install the new drivers.

Download the new drivers from Nvidia Official Website. All the supported GPUs are mentioned on the Nvidia drivers page.

<http://www.nvidia.com/download/driverResults.aspx/132530/en-us>

After you download the drivers now its time for installation. The drivers you will download will be a **run** file so before installing we must do a few modifications to that file and install other dependencies for the drivers first.

Run this command.

**apt-get install build-essential gcc-multilib dkms**

The above command will install all the necessary file which we will need for the driver installation.

**Create Blacklist for Nouveau Driver**

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

**blacklist nouveau**

**options nouveau modeset=0**

To create a file open a terminal and type in

**gedit /etc/modprobe.d/blacklist-nouveau.conf**

and copy and paste the above 2 lines and click save.

Now execute **sudo update-initramfs -u** and reboot the computer.

### After you reboot you might notice a few changes on display and resolution. Don’t worry because we will be installing the drivers now.

### Stop lightdm/gdm/kdm

After the computer is rebooted. We need to stop the desktop manager before executing the runfile to install the driver. **lightdm** is the default desktop manager in Ubuntu. If GNOME or KDE desktop environment is used, installed desktop manager will then be **gdm** or **kdm**.

For Ubuntu 16.04, executing

**sudo service lightdm stop** (or use gdm or kdm instead of lightdm)

After you execute the above command the GUI (Graphical User Interface) Will be stopped. Now use this button combination to switch to CLI (Command Line) to install the drivers.

**Ctrl + Alt + F2**

After executing the above command login with your Ubuntu Username and Password which you setup while installing Ubuntu

**Executing the Runfile/Installing Drivers:**

After above batch of preparation, we can eventually start executing the runfile.

Now using the **cd** command navigate to the folder where you downloaded the Drivers. In my case it was in Downloads folder.

**cd ~/Downloads/**

Now type **ls** and it will display you the file which you have downloaded.

The downloaded file will not haver any permissions to run so we need to grant it the executable permission before we can install the drivers. Type in the command below.

**chmod +x (Name of your downloaded file)**

**e.g**

**chmod +x NVIDIA-Linux-x86\_64-390.48.run**

After you have executed the above command we just need to run the installer.

**sudo ./NVIDIA-Linux-x86\_64-390.48.run --dkms -s**

Wait for the driver installation to finish and then reboot your computer by just typing **reboot**  in the terminal.

After the reboot you will see the display back to normal and all the drivers will be installed.

**Check the Installation**

After a successful installation, **nvidia-smi** command will report all your CUDA-capable devices in the system.

### Common Errors and Solutions

1. ERROR: Unable to load the **'nvidia-drm'** kernel module.

* One probable reason is that the system is boot from UEFI but Secure Boot option is turned on in the BIOS setting. Turn it off and the problem will be solved.
* For any other errors please do contact us and let us know.

**Additional Notes:**

**nvidia-smi -pm 1** can enable the persistent mode, which will save some time from loading the driver. It will have significant effect on machines with more than 4 GPUs.

**nvidia-smi -e 0** can disable ECC on TESLA products, which will provide about 1/15 more video memory. Reboot is reqired for taking effect. nvidia-smi -e 1 can be used to enable ECC again.

**nvidia-smi -pl <some power value>** can be used for increasing or decrasing the TDP limit of the GPU. Increasing will encourage higher GPU Boost frequency, but is somehow DANGEROUS and HARMFUL to the GPU. Decreasing will help to same some power, which is useful for machines that does not have enough power supply and will shutdown unintendedly when pull all GPU to their maximum load.

**-i <GPUID>** can be added after above commands to specify individual GPU.

These commands can be added to **/etc/rc.local** for excuting at system boot.

**Nvidia Cuda 9.1 Installation (including Cuda Samples):**

After we have completed the above steps to install the drivers now its time for the Cuda 9.1 Installation. There is a documentation available on Nvidia website as well but it is complicated and does not explain everything well. So, in this guide I will break down the steps and make it as simple as possible.

One of the things which really confused me as well was that Nvidia provides run file or deb file for Cuda installation which comes with the drivers so why did we have to go through the above steps if they already have it in the Cuda installation file. The reason is the Cuda 9.1 requires R390 drivers and the version available in the installation is R387. Don’t know if Nvidia has made a mistake or what but when I attempted to install the drivers using the Cuda Run file it crashed the system and I was unable to get the display or anything working properly.

First of all lets download the CUDA 9.1 From Nvidia Website.

<https://developer.nvidia.com/cuda-downloads?target_os=Linux&target_arch=x86_64&target_distro=Ubuntu&target_version=1604&target_type=runfilelocal>

Please download all the files on that link and we will install them one by one. Make sure you download the RunFiles.

After download all the Run Files. You main Base Installer needs to be extracted first. The reason is that Nvidia Cuda 9.1 base installer comes with 3 files inside one is the Cuda Toolkit another are the samples and the last one are the old drivers. As I mentioned before we cannot use the old drivers because of the Cuda 9.1 support so if we try to install the Run file before extracting it will install all those 3 Run Files inside which can cause a system crash and driver override.

Now after downloading navigate to the installation file folder by **cd** command.

After you are inside the folder which contains the CUDA installation file. Run this command to extract the files.

**chmod +x cuda\_7.5.18\_linux.run**

**./cuda\_7.5.18\_linux.run --extract=$HOME** (It will extract the files inside the root folder **~** use the command

**cd ~**

to navigate inside the root folder and install the CUDA Tool Kit).

Now let’s run the installation file first.

**sudo ./cuda-linux.9.1.85-23083092.run**

Now just follow the instructions on terminal and accept everything. Same goes for installing sample.

**sudo ./cuda-samples.9.1.85-23083092-linux.run**

Please note do not install the drivers as it might crash the system that’s why we extracted the Run File.

After the installation is complete we will now setup Cuda for our work environment before we can start using it.

**Setting up the Environmental Variables:**

We will need to setup the environmental variable first so we can access the Cuda Toolkit and commands from anywhere in our terminal.

There are basically 2 ways either you can set them up temporarily or permanent. I will explain both ways.

**Temporary Method for Environment Variables:**

Cope and paste these commands one by one.

**sudo export PATH=/usr/local/cuda-9.1/bin${PATH:+:${PATH}}**

**sudo export LD\_LIBRARY\_PATH=/usr/local/cuda-9.1/lib64${LD\_LIBRARY\_PATH:+:${LD\_LIBRARY\_PATH}}**

Both the commands are single line commands so in copy & paste if you get a line break in the terminal please type the commands exactly mentioned above. The above method will setup the paths only temporary once you reboot the system it will be gone and you have too set it up again.

Now type **nvcc --help**  if this command works then it means Cuda has successfully installed and working.

**Permanent Method for Environment Variables:**

For permanent we will change and add the Paths inside the PATH file in /etc folder.

On your terminal type

**sudo gedit /etc/environment**

And hit enter. It will open up the file with the PATH. Now copy and paste this line at the end of the pat variable but before the “ marks and put the **:** sign after the /usr/local/games

**/usr/local/cuda-9.1/bin**

e.g

**PATH=”/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games:/usr/local/cuda-9.1/bin”**

And save the file. It will permanently set the path until and unless you remove it your self.

Now we are going to copy the library path. Library Path will be copied into the bashrc profile.On you terminal type

**echo "export LD\_LIBRARY\_PATH=$LD\_LIBRARY\_PATH:/usr/local/cuda-9.1/lib64" >> ~/.bashrc**

After following the above steps reboot the computer and all the paths will be setup for you. Now type **nvcc --help**  if this command works then it means Cuda has successfully installed and working.

You might need to install additional libraries while working with Image Processing.

**sudo apt-get install g++ freeglut3-dev build-essential libx11-dev libxmu-dev libxi-dev libglu1-mesa libglu1-mesa-dev**

After you install additional libraries a symbolic link can be broken. So, what you need to do is reinstall the driver which you downloaded in the first step. But before that reboot the computer and press **CTRL + ALT + F2**  to get a terminal screen and stop the lightdm service

**sudo service lightdm stop** (or use gdm or kdm instead of lightdm)

After that navigate to where you downloaded the Run File and reinstall drivers.

**sudo ./NVIDIA-Linux-x86\_64-390.48.run --dkms -s**

Wait for the driver installation to finish and then reboot your computer by just typing **reboot**  in the terminal.

After the reboot you will see the display back to normal and all the drivers will be installed. Nad the broken links will be fixed.

### [7.2.1. Install Persistence Daemon](http://docs.nvidia.com/cuda/cuda-installation-guide-linux/index.html" \l "install-persistenced)

The NVIDIA Persistence Daemon can be started as the root user by running:

**/usr/bin/nvidia-persistenced –verbose**

Add it to the /etc/rc.local to make it run on verytime when the system boots up.

**[7.1.2. POWER9 Setup](http://docs.nvidia.com/cuda/cuda-installation-guide-linux/index.html" \l "power9-setup) For Tesla GPUs only. Not for other Nvidia GPUs**

Because of the addition of new features specific to the NVIDIA POWER9 CUDA driver, there are some additional setup requirements in order for the driver to function properly. These additional steps are not handled by the installation of CUDA packages, and failure to ensure these extra requirements are met will result in a non-functional CUDA driver installation.

There are two changes that need to be made manually after installing the NVIDIA CUDA driver to ensure proper operation:

1. Create and enable a systemd service file or init script that runs the NVsIDIA Persistence Daemon as the first NVIDIA software during or at the end of the boot process. The following service file example is sufficient for most installations:

[Unit]

Description=NVIDIA Persistence Daemon

Wants=syslog.target

[Service]

Type=forking

PIDFile=/var/run/nvidia-persistenced/nvidia-persistenced.pid

Restart=always

ExecStart=/usr/bin/nvidia-persistenced --verbose

ExecStopPost=/bin/rm -rf /var/run/nvidia-persistenced

[Install]

WantedBy=multi-user.target

Copy the above text into the following file:

/usr/lib/systemd/system/nvidia-persistenced.service

And run the following command:

**$** sudo systemctl enable nvidia-persistenced

1. Disable a udev rule installed by default in some Linux distributions that cause hot-pluggable memory to be automatically onlined when it is physically probed. This behavior prevents NVIDIA software from bringing NVIDIA device memory online with non-default settings. This udev rule must be disabled in order for the NVIDIA CUDA driver to function properly on POWER9 systems.

On RedHat Enterprise Linux 7, this rule can be found in:

/lib/udev/rules.d/40-redhat.rules

On Ubuntu 17.04, this rule can be found in:

/lib/udev/rules.d/40-vm-hotadd.rules

The rule generally takes a form where it detects the addition of a memory block and changes the 'state' attribute to online. For example, in RHEL7, the rule looks like this:

SUBSYSTEM=="memory", ACTION=="add", PROGRAM="/bin/uname -p", RESULT!="s390\*", ATTR{state}=="offline", ATTR{state}="online"

This rule must be commented out, removed, or changed so that it does not apply to POWER9 NVIDIA systems.

You will need to reboot the system to initialize the above changes.