DUNE Computing Training December 2021 edition : Mission Setup

Objectives

- Get ready to do the tutorial
- Understand the authentication procedures
- Set up your environment for DUNE
- Do an exercise to help us check if all is good
- Get streaming and grid access

Requirements

You must be on the DUNE Collaboration member list and have a valid FNAL or CERN account. .

> ✏️ **Note**
>
> The instructions below are for FNAL accounts. If you do not have a valid FNAL account but a CERN one, go at the bottom of this page to the "Setup on CERN machines".

1. Kerberos business

If you already are a kerberos-aficionado, go to the next section. If not, we give you a little tour of it below.

What is it? Kerberos is a computer-network authentication protocol that works on the basis of tickets.

Why does FNAL use Kerberos? Fermilab uses Kerberos to implement strong authentication, so that no passwords go over the internet (if a hacker steals a ticket, it is only valid for a day).

How it works? Kerberos uses tickets to authenticate users. Tickets are made by the kinit command, which asks for your kerberos password (info on kerberos password here (https://fermi.servicenowservices.com/kb_view.do?sysparm_article=KB0011294)). The kinit command reads the password, encrypts it and sends it to the Key

Distribution Centre (KDC) at FNAL. The Kerberos configuration file, which lists the KDCs, is stored in a file named krb5.conf. On Linux and Mac, it is located here:

> **Code**
>
> ```
> /etc/krb5.conf
> ```

If you do not have it, create it. A FNAL template is available here (https://authentication.fnal.gov/krb5conf/) for each OS (Linux, Mac, Windows). More explanations on this config file are available here (https://fermi.servicenowservices.com/kb_view.do?sysparm_article=KB0011315) if you're curious.

To log in to a machine, you need to have a valid kerberos ticket. You don't need to do this every time you login, only when your ticket is expired. Kerberos tickets last for 26 hours. To create your ticket:

> **Bash**
>
> ```
> kinit -f username@FNAL.GOV
> ```

The -f option means your ticket is forwardable. A forwardable ticket is one that originated on computer A, but can be forwarded to computer B and will remain valide on computer B. Careful: you need to write FNAL.GOV in uppercase.

To know if you have a valid ticket, type:

> **Code**
>
> ```
> klist
> ```

Typical output of klist on macOS looks like this:

> **Output**
>
> ```
> Mac-124243:~ trj$ klist
> Credentials cache: API:F083D4AF-D901-41BE-99AF-92FE01838C3B
>         Principal: trj@FNAL.GOV
> Issued              Expires             Principal
> Jan 20 15:29:41 2021  Jan 21 17:29:38 2021  krbtgt/FNAL.GOV@FNAL.GOV
> ```

If your ticket has not expired yet but will soon, you can refresh it for another 26 hours by typing:

> **Bash**
>
> ```
> kinit -R
> ```

Refreshing a ticket can be done for up to one week after it was initially issued.

Running into issues? Users logging in from outside the Fermilab network may be behind Network Address Translation (NAT) routers. If so, you may need an "addressless" ticket. For this, add the option -A:

> **Bash**
>
> ```
> kinit -A -f username@FNAL.GOV
> ```

Some users have reported problems with the Kerberos utilities provided by Macports and Anaconda. Macintosh users should use the system-provided Kerberos utilities – like /usr/bin/kinit. Hint: use the command

Bash
```bash
which kinit
```

to report the path of the kinit that's first in your $PATH search list. If you have another, non-working one, a suggestion is to make an alias like this one:

Bash
```bash
alias mykinit="/usr/bin/kinit −A −f myusername@FNAL.GOV"
```

Then you can simply use:

Code
```
mykinit
```

2. ssh-in

What is it? SSH stands for Secure SHell. It is a safe protocol used for connecting to remote machines. The configuration is done in your local file in your home directory:

Bash
```bash
cat ~/.ssh/config
```

If it does not exist, create it. A minimum working example to connect to FNAL machines is as follow:

Output
```
Host *.fnal.gov
ForwardAgent yes
ForwardX11 yes
ForwardX11Trusted yes
GSSAPIAuthentication yes
GSSAPIDelegateCredentials yes
```

Now you can try to log into a machine at Fermilab. There are now 15 different machines you can login to: from dunegpvm01 to dunegpvm15 (gpvm stands for 'general purpose virtual machine' because these servers run on virtual machines and not dedicated hardware, others nodes which are indented for building code run on dedicated hardware). The dunegpvm machines run Scientific Linux Fermi 7 (SLF7). To know the load on the machines, use this monitoring link: dunegpvm status.

How to connect? The ssh command does the job. The -Y option turns on the xwindow protocol so that you can have graphical display and keyboard/mouse handling (quite useful). But if you have the line "ForwardX11Trusted yes" in your ssh config file, this will do the -Y option. For connecting to e.g. dunegpvm07, the command is:

> **Bash**
>
> ```
> ssh username@dunegpvmXX.fnal.gov
> ```

where XX is a number from 01 to 15. If you experience long delays in loading programs or graphical output, you can try connecting with VNC. More info: Using VNC Connections on the dunegpvms (https://wiki.dunescience.org/wiki/DUNE_Computing/Using_VNC_Connections_on_the_dunegpvms).

3. Get a clean shell

To run DUNE software, it is necessary to have a 'clean login'. What is meant by clean here? If you work on other experiment(s), you may have some environment variables defined (for NOvA, MINERvA, MicroBooNE). Theses may conflict with the DUNE environment ones.

Two ways to clean your shell once on a DUNE machine:

Cleanup option 1: Manual check and cleanup of your custom environment variables To list all of your environment variables:

> **Bash**
>
> ```
> env
> ```

The list may be long. If you want to know what is already setup (potentially conflicting later with DUNE), you can grep the environment variables for a specific experiment (here the -i option is 'case insensitive'). Here is an example to list all NOvA-specific variables:

> **Bash**
>
> ```
> env | grep -i nova
> ```

Here you can tweak your .bashrc/.profile file(s) to temporarily comment out those (the "export" commands that are setting custom environment variables).

Cleanup option 2: Backup your login scripts and go minimal at login (recommended)

A simpler solution would be to rename your login scripts (for instance .bashrc as .bashrc_save and/or .profile as .profile_bkp) so that your setup at login will be minimal and you will get the cleanest shell. For this to take into effect, you will need to exit and reconnect through ssh.

4. Setting up the DUNE software

To set up your environment with DUNE, the command is:

```bash
source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh
```

You should see in your terminal the following output:

```
Setting up larsoft UPS area... /cvmfs/larsoft.opensciencegrid.org/products/
Setting up DUNE UPS area... /cvmfs/dune.opensciencegrid.org/products/dune/
```

👁 How to make custom setup command with aliases 🔼

Not familiar with aliases? Read below.

To create unix custom commands for yourself, we use 'aliases':

```
alias my_custom_commmand='the_long_command_you_want_to_alias_in_a_shorter_custom_name'
```

For DUNE setup, you can type for instance:

```bash
alias dune_setup='source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh'
```

So next time you type:

```
dune_setup
```

Your terminal will execute the long command. This will work for your current session (if you disconnect, the alias won't exist anymore). You can store this in your (minimal) .bashrc or .profile if you want this alias to be available in all sessions. The alias will be defined but not executed. Only if you type the command `dune_setup` yourself.

5. Exercise! (it's easy)

This exercise will help organizers see if you reached this step or need help.

1) Start in your home area `cd ~` and create the file `.dune_presetup_202105.sh`. Files starting with a dot are hidden files and can be seen with the command `ls -la` (**a** for **a**ll files). Write in it the following:

Code

```
export DUNETPC_VERSION=v09_22_02
alias dune_setup='source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh'
```

When you start the training, you will have to source this file:

Bash

```
source ~/.dune_presetup_202105.sh
```

Then, to setup DUNE, use the created alias:

Bash

```
dune_setup
```

2) Create a working directory in the `dune/app` and `pnfs/dune` areas (these will be explained during the training):

Bash

```
mkdir -p /dune/app/users/${USER}
mkdir -p /pnfs/dune/scratch/users/${USER}
```

3) Print the date and add the output to a file named `my_first_login.txt`:

Bash

```
date >& /dune/app/users/${USER}/my_first_login.txt
```

4) With the above, we will check if you reach this point. However we want to tailor this tutorial to your preferences as much as possible. We will let you decide which animals you would like to see in future material, between: "puppy", "cat", "squirrel", "sloth", "unicorn pegasus llama" (or "prefer not to say" of course). Write your desired option on the second line of the file you just created above.

> ✏ **Note**
>
> If you experience difficulties, refer to the support options mentioned on Indico. Please mention in your message this is about the Setup step 5. Thanks!

6. Getting setup for streaming and grid access

In addition to your kerberos access, you need to be in the DUNE VO (Virtual Organization) to access to global DUNE resources. This is necessary in particular to stream data and submit jobs to the grid. If you are on the DUNE collaboration list and have a Fermilab ID you should have been added automatically to the DUNE VO.

To check if you are on the VO, two commands. The kx509 gets a certificate from your kerberos ticket. On a DUNE machine, type:

Bash

```
kx509
```

Output

```
Authorizing ...... authorized
Fetching certificate ..... fetched
Storing certificate in /tmp/x509up_u55793

Your certificate is valid until: Wed Jan 27 18:03:55 2021
```

To access the grid resources, you will need a proxy. More information on proxy is available here (https://cdcvs.fnal.gov/redmine/projects/sbndcode/wiki/Get_a_certificate_proxy). This is to be done once every 24 hours per login machine you're using to identify yourself:

Bash

```
export ROLE=Analysis
voms-proxy-init -rfc -noregen -voms=dune:/dune/Role=$ROLE -valid 120:00
```

Output

```
Your identity: /DC=org/DC=cilogon/C=US/O=Fermi National Accelerator Laboratory/OU=People/CN=Claire David/CN=UID:cdavid
Contacting  voms1.fnal.gov:15042 [/DC=org/DC=incommon/C=US/ST=Illinois/L=Batavia/O=Fermi Research Alliance/OU=Fermilab/CN=voms1.fnal.gov] "dune" Done
Creating proxy ................................ Done

Your proxy is valid until Mon Jan 25 18:09:25 2021
```

Report this by appending the output of `voms-proxy-info` to your first login file:

Bash

```
voms-proxy-info >> /dune/app/users/${USER}/my_first_login.txt
```

> ✏️ **Issues**
>
> If you have issues here, Please mention in your message it is the Step 6 of the setup. Thanks!

> **❶ Success**
>
> If you obtain the message starting with `Your proxy is valid until ...` Congratulations! You are ready to go!

Set up on CERN machines (Optional)

Caution: the following instructions are for those of you who do not have a valid FNAL account but have access to CERN machines.

1. Source the DUNE environment

CERN access is mainly for ProtoDUNE collaborators. If you have a valid CERN ID and access to lxplus via ssh, you can setup your environment for this tutorial as follow:

Bash

```bash
source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh
```

Output

```
Setting up larsoft UPS area... /cvmfs/larsoft.opensciencegrid.org/products/
Setting up DUNE UPS area... /cvmfs/dune.opensciencegrid.org/products/dune/
```

2. Access tutorial datasets

Normally, the datasets are accessible through the grid resource. But with your CERN account, you may not be part of the DUNE VO yet (more on this during the tutorial). We found a workaround: some datasets have been copied locally for you. You can check them here:

Bash

```bash
ls /afs/cern.ch/work/t/tjunk/public/may2021tutorialfiles/
```

Output

```
np04_raw_run005387_0019_dl5_reco_12900894_0_20181102T023521.root
PDSPProd4_protoDUNE_sp_reco_stage1_p1GeV_35ms_sce_datadriven_41094796_0_20210121T214555
Z.root
```

3. Notify us

You should be good to go. If however you are experiencing issues, please contact us as soon as possible. Be sure to mention "Setup on CERN machines" if that is the case, and we will do our best to assist you.

> ❗ **Success**
>
> If you can list the files above, you should be able to do most of the tutorial on LArSoft.

> ☑ **Warning**
>
> Connecting to CERN machines will not give you the best experience to understand storage spaces and data management. If you obtain a FNAL account in the future, you can however do the training through the recorded videos that will be made available after the event.