Session 1: Everything is Somewhere

Navigating the directory structure
Moving files around

Session 2: The computer as a nut - using Unix commands

Summary
Unix cheat sheet

Everything is somewhere, and everything happens somewhere. When you are using a Unix system (including Linux systems and HPC clusters), there are two kinds of somewhere:

1. What host (computer) am I on?
2. What filesystem and directory am I in?
The host determines what CPU and memory resources you have available, with whom you are sharing them and which filesystems (and therefore files and programs) are accessible.

The filesystem and directory (location within the filesystem) determine what files and programs are immediately visible, to you and to any program you run.

In fact, the same is true of a typical workstation, but on a workstation you have one host and one filesystem, shared with nobody, and the details of your location within the filesystem are hidden by the Windows or OSX GUI.

The diagram below is a simple map of some key places you move between when using an NYU HPC system (Tip: click the image for a closer view):

- From your workstation, you can access the Internet, and through it, the bastion host hpc.nyu.edu
  
The term bastion host refers to a server which stands between a private network - in this case the NYU HPC network - and a public network - the Internet. It protects the private network from attack from the outside. As HPC users, the significance of the bastion host is that we cannot directly access the HPC clusters.

Exercise
Try it. Open a terminal on your workstation and attempt to log in to Mercer with:

```bash
$ ssh -l NetID mercer.es.its.nyu.edu
```

If you are using Windows, you'll do this via PuTTY instead of a terminal. Click here to expand the instructions for PuTTY

The following page comes from the HPC Wiki. Replace "hpc.nyu.edu" with "mercer.es.its.nyu.edu" for this exercise.

Logging In - the easy, primitive way

When you installed PuTTY you will have added to your Desktop and also the the Start Menu and icon labeled “putty”. Double-clicking this icon will launch PuTTY, with the dialog initially in the “Session” category as in the screenshot below:

If this is not your first time through here, and you already saved a setup previously, you can select the saved session and hit “Load”. This will do the following steps for you so you can just hit “Open”. Otherwise, read on...

Enter the as Host Name "hpc.nyu.edu", leave the Port at its default value of 22 and select the "SSH" Protocol. Don't hit "Open" just yet!

Eventually, if not now, you will probably want to use some graphical applications. Enable this by selecting Connection -> SSH -> X11 from the Category menu on the left:
Check the "Enable X11 forwarding" box. Again, don't hit "Open" just yet...

You can save these settings to use again next time. Return to the "Session" category and give this setup a name in the "Saved Sessions" box (perhaps NYU HPC). Now click "Save".

Finally, hit "Open". You will get a command terminal for hpc.nyu.edu. This is a deliberately restricted environment from which you can do little other than to log in to one of the NYU clusters, which are named as follows:

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Host name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercer</td>
<td>mercer.es.its.nyu.edu</td>
</tr>
<tr>
<td>Dumbo</td>
<td>dumbo.es.its.nyu.edu</td>
</tr>
</tbody>
</table>

Therefore, log in using one of the following commands, according to which cluster you wish to use:

```
$ ssh -X NetID@mercer.es.its.nyu.edu
$ ssh -X NetID@dumbo.es.its.nyu.edu
```

In the boxes below and elsewhere in this wiki, the symbol "$" at the beginning of a line represents the command prompt - don't type the "$", type only the remainder of the line following the "$". Also wherever NetID appears, replace it with your NYU NetID.

On Mercer you may notice that you are now on a host named "login-0-0" or "login-0-3" or something similar. These clusters use multiple login nodes and which one you get depends how busy each is at the time. The login nodes are configured identically and see the same filesystems, so the specific node you are logged in to is not important.

Unless you have the SSH tunnel (set up in the tutorial prequel) running, this command will sit there making no response. Kill it by pressing Ctrl-C.

- From hpc.nyu.edu you can access the Mercer login nodes.

### Exercise

Log in again, this time via hpc.nyu.edu:

```
$ ssh -l NetID hpc.nyu.edu
$ ssh mercer
```

Again, Windows users will use PuTTY for the first step and for the second, ssh from the terminal PuTTY will start. Here's that Wiki page again...

### Logging In - the easy, primitive way

This method is useful as a first-time login, to ensure you are able to connect. For ongoing use however, we recommend setting up and using SSH tunneling instead.

When you installed PuTTY you will have added to your Desktop and also the the Start Menu and icon labeled "putty". Double-clicking this icon will launch PuTTY, with the dialog initially in the "Session" category as in the screenshot below:

If this is not your first time through here, and you already saved a setup previously, you can select the saved session and hit "Load". This will do the following steps for you so you can just hit "Open". Otherwise, read on...
Enter the as Host Name “hpc.nyu.edu”, leave the Port at its default value of 22 and select the “SSH” Protocol. Don’t hit “Open” just yet!

Eventually, if not now, you will probably want to use some graphical applications. Enable this by selecting Connection -> SSH -> X11 from the Category menu on the left:

Check the “Enable X11 forwarding” box. Again, don’t hit “Open” just yet...

You can save these settings to use again next time. Return to the “Session” category and give this setup a name in the “Saved Sessions” box (perhaps NYU HPC”). Now click “Save”.

Finally, hit “Open”. You will get a command terminal for hpc.nyu.edu. This is a deliberately restricted environment from which you can do little other than to log in to one of the NYU clusters, which are named as follows:

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Host name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercer</td>
<td>mercer.es.its.nyu.edu</td>
</tr>
<tr>
<td>Dumbo</td>
<td>dumbo.es.its.nyu.edu</td>
</tr>
</tbody>
</table>

Therefore, log in using one of the following commands, according to which cluster you wish to use:

```
$ ssh -X NetID@mercer.es.its.nyu.edu
$ ssh -X NetID@dumbo.es.its.nyu.edu
```

In the boxes below and elsewhere in this wiki, the symbol "$" at the beginning of a line represents the command prompt - don’t type the "$", type only the remainder of the line following the "$". Also wherever NetID appears, replace it with your NYU NetID.

On Mercer you may notice that you are now on a host named "login-0-0" or "login-0-3" or something similar. These clusters use multiple login nodes and which one you get depends how busy each is at the time. The login nodes are configured identically and see the same filesystems, so the specific node you are logged in to is not important.

Therefore, you can find which host you are on with the hostname command:

```
$ hostname
login-0-1.local
```

Mercer currently has four login nodes, so the hostname you see may not be the same as the one above. The login nodes are configured to be equivalent so for most purposes it makes no difference which specific login node you land on.

Mercer has a number of filesystems *mounted* (attached and visible). The diagram above only shows a few important ones. Some things to note are:

- `/home`, `/work` and `/scratch` are connected to the login nodes and also to the compute nodes
- `/archive` is connected to the login nodes, but not to the compute nodes
  
  So processes running on the compute nodes cannot see `/archive`

- The connection from `/scratch` to the compute nodes is faster - indicating faster - than the other connections on the system
  
  On NYU HPC clusters, `/scratch` is configured to provide fast I/O for jobs running on the compute nodes

- `/home` and `/scratch` are separate filesystems, in separate places.
  
  Moving files around within `/home` or within `/scratch` is trivial and fast, but to move files between `/home` and `/scratch`, the file must travel from one filesystem to the other over a network. Expect this to take longer!
Finally

We go to a host with `ssh ..` and to leave a host, type "`exit`".

(actually, "`exit`" leaves the shell - we'll discuss shells in session 2 - but by leaving the shell that `ssh` started upon your arrival, you will also leave the host)

Next: Navigating the directory structure