Accessing the Prince cluster

Quick Links

HPC Home
Getting an account
Gentle Introduction to using HPC
Getting started on Prince
Prince How-to Articles

Logging in
Windows
Mac / Linux

Clusters and Storage
Prince (HPC)
Dumbo (Hadoop)
Brooklyn (OpenStack)
<table>
<thead>
<tr>
<th>Dala (NYU Abu Dhabi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfering data to/from the clusters</td>
</tr>
<tr>
<td>Transfering data to/from Prince cluster using Globus</td>
</tr>
<tr>
<td>Submitting jobs with sbatch</td>
</tr>
<tr>
<td>Available software</td>
</tr>
<tr>
<td>Licensed Software Available on the HPC Cluster</td>
</tr>
<tr>
<td>Building Software</td>
</tr>
<tr>
<td>Slurm Tutorial</td>
</tr>
<tr>
<td>Tutorials</td>
</tr>
<tr>
<td>FAQs</td>
</tr>
<tr>
<td>Scratch Area Cleanup</td>
</tr>
<tr>
<td>Programming for Biologists</td>
</tr>
</tbody>
</table>
## Running jobs on the Prince Cluster

### Accessing the Prince Cluster

- From Windows workstation
- From Mac workstation

### Software and Environment Module

- Job script and resource request

### Introduction to job scheduling

- Submitting jobs with `sbatch`
- Requesting resources
The HPC clusters cannot be accessed directly from the Internet - if you are outside of the NYU network, access is a two-step process:

1. **Log in to hpc2.nyu.edu**
   - This is a "bastion host", providing a secure gateway between the NYU HPC clusters and the Internet. You will find that on hpc2.nyu.edu itself you can do very little save but to log in to an NYU HPC cluster.

2. **Log in to the cluster you wish to use**
   - This will put you on a "login" node for that cluster, from where you can manipulate files and initiate jobs. This is not the place for heavy compute work: you submit that as a job to the compute nodes via the batch queuing system. Instructions for using the batch system can be found [here](but you don't need to worry about that yet).

### What is a SSH tunnel

In computer networking, a computer decides what to do with an incoming network packet according to the "port" it arrived on. The port is simply a number attached to the packet. Certain ports are reserved for specific functions, for example packets arriving on port 22 are assumed to be intended for the SSH handler, so the computer passes those packets to SSH to interpret. Other port numbers are available to use for whatever you like, and as long as the same port is not used for different things on the same computer, everything works.
With SSH Tunneling, you will start an SSH session between your workstation and the bastion host hpc2.nyu.edu, and instruct that session to create a tunnel. Your workstation will make one end of the tunnel, at "localhost, port 8026" ("localhost" is the computer's name for itself, so packets arriving at your workstation port 8026 will be sent into the tunnel). The bastion host will make the other end of the tunnel, at "prince.hpc.nyu.edu, port 22", so anything coming through the tunnel will be forwarded to the normal SSH port (22) of Prince. The fact that your workstation cannot see Prince does not matter, it only needs to see its end of the tunnel.

The following diagram illustrates the process. It looks complex, but only requires 2 steps: the blue text shows what happens when you create the tunnel (step 1) and the green arrows indicate using the tunnel (step 2).

![Diagram showing SSH tunnel setup]

You only need to do step 1 once, and then you can use the tunnel (step 2) as many times as you like - for example, you might have two terminal sessions and a WinSCP session all using the same tunnel created with step 1.

In these instructions we are using port 8026. If it happens that another program on your computer is watch this port (which is fairly unlikely) then it won't work, and you'll need to choose a different port number, eg 9020, and substitute that throughout these instructions. 4-digit numbers starting with an 8 or a 9 are usually good ones to choose.