## Logging in to the NYU HPC Clusters

### Quick Links

<table>
<thead>
<tr>
<th>HPC Home</th>
<th>Getting an account</th>
<th>Gentle Introduction to using HPC</th>
<th>Getting started on Prince</th>
<th>Prince How-to Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Logging in</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mac / Linux</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clusters and Storage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prince (HPC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dumbo (Hadoop)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brooklyn (OpenStack)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dalma (NYU Abu Dhabi)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transferring data to/from the clusters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transferring data to/from Prince cluster using Globus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Submitting jobs with sbatch</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Available software</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Licensed Software Available on the HPC Cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The HPC clusters (Prince and Dumbo) are not directly visible to the internet (outside the NYU Network). If you are outside NYU's Network (off-campus) you must first login to a bastion host named gw.hpc.nyu.edu

The diagram below illustrates the login path.

NOTE: The clusters can still access the internet directly. This may be useful when copying data from servers outside the NYU Network - see: How to copy files to and from the HPC clusters.
NOTE: Alternatively, instead of login to the bastion host, you can use VPN to get inside NYU's network and access the HPC clusters directly. Instructions on how to install and use the VPN client are available here.

NOTE: You can't do anything on the bastion host, except ssh to the HPC clusters.

In a nutshell

- From within the NYU network; that is, from an on-campus location, or after you VPN inside NYU's network, you can login to the HPC clusters directly.
  - To login to the HPC cluster Prince, simply use (replace NYUNetID with your NetId):  
    
    ```
    ssh NYUNetID@prince.hpc.nyu.edu
    ```
  
  - To login in to the Hadoop cluster (Dumbo)
    
    ```
    ssh NYUNetID@dumbo.hpc.nyu.edu
    ```

- From an off-campus location (outside NYU-NET), logging in to the HPC clusters is a two-step process:
  a. First login to the bastion host, gw.hpc.nyu.edu. From a Mac or Linux workstation, this is a simple terminal command (replace NYUNetID with your NetId). Your password is the same password you use for NYU Home:
    
    ```
    ssh NYUNetID@gw.hpc.nyu.edu
    ```

    You can’t do anything on the bastion host, except ssh to the cluster

  b. Next login to the cluster. For Prince, this is done with:
    
    ```
    ssh prince.hpc.nyu.edu
    ```

    For Dumbo, this is done with:
    
    ```
    ssh dumbo.hpc.nyu.edu
    ```

The full story

You need to ensure your workstation has the necessary software and settings to connect to the clusters and to use graphical interfaces. Here are instructions for preparing your workstation and logging in from a Windows / Linux / Mac.

SSH tunneling for easier login and data transfer

The two-stage access can be inconvenient, especially when transferring files to and from the clusters. Secure direct access and file transfer is possible by setting up SSH tunneling from your workstation to the HPC clusters. We have instructions on setting this up for Windows / Linux / Mac workstations.

What can I do on the login node?

The login nodes (Prince and Dumbo) are for preparing, submitting and monitoring scripts, analyzing results, moving data around and code development and simple compilation. Login nodes are Not suitable for running computational workloads: for Prince use this batch system.

Compiling a large source codebase, especially with heavy use of optimization or -ipo (interprocedural optimization), can use much memory and CPU time. In such circumstances it is best to use the batch system for compilation too, perhaps via an interactive batch job. Click here for more info about interactive batch jobs.