HPC Data Transfer Nodes

Closely attached to the Prince cluster, the data-transfer nodes (DTNs) are (currently) two nodes optimized for transferring data between file systems as well as the NYU-net at large and machines beyond NYU. The DTNs have 10-Gb/s Ethernet connections to the Internet and FDR Infiniband connections to the filesystems exported from the HPC cluster. These filesystems include /home, /scratch, /archive, /beegfs (including), and /rw (i.e., Research Workspace). A valid HPC account is required to use the DTNs and the services they host.

Accessing the data-transfer nodes

The data-transfer nodes can be accessed in a variety of ways:

- From the public Internet or NYU-net: use the hostname dtn.hpc.nyu.edu
- From the Prince cluster (e.g., the login nodes): the hostname can be shortened to dtn

For example, to log in to a DTN from the Prince cluster, to carry out some copy operation, and to log back out, you can use a command sequence like:

```bash
ssh dtn
ergnc ...
logout
```

- Via specific tools like Globus: see below

Note that the allowed connection methods are restricted and generally limited to those described on this page.

Available data-transfer tools

Globus

Globus is the recommended tool to use for large-volume data transfers. It features automatic performance tuning and automatic retries in cases of file-transfer failures. Data-transfer tasks can be submitted via a web portal. The Globus service will take care of the rest, to make sure files are copied efficiently, reliably, and securely. Globus is also a tool for you to share data with collaborators, for whom you only need to provide the email addresses.

We have done some comparisons between the endpoints "nyu#prince" and "nyu#hpc". The latter performs a bit better. We are retiring the endpoint "nyu#prince".

The web portal is located at [https://www.globus.org/](https://www.globus.org/). Please select the organization "New York University" to login, and choose "nyu#hpc" for NYU HPC as one endpoint. The other endpoint should be on your workstation or other institution's server. You may follow the detailed instruction on [How To Log In and Transfer Files](https://www.globus.org/), prepared by the Globus team.

**scp and rsync**

Sometimes these two tools are convenient for transferring small files. Using the DTNs does not require to set up an SSH tunnel; use the hostname dtn.hpc.nyu.edu for one-step copying. See below for examples of commands invoked on the command line on a laptop running a Unix-like operating system:

```bash
scp HMLHWBGX7_n01_HK16.fastq.gz wd35@dtn.hpc.nyu.edu:/scratch/wd35/
```

```bash
rsync -av HMLHWBGX7_n01_HK16.fastq.gz wd35@dtn.hpc.nyu.edu:/scratch/wd35/
```

In particular, rsync can also be used on the DTNs to copy directories recursively between filesystems, e.g. (assuming that you are logged in to a DTN),
rsync -av /scratch/username/project1 /rw/sharename/

where username would be your user name, project1 a directory to be copied to the Research Workspace, and sharename the name of a share on the Research Workspace (either your NetID or the name of a project you're a member of).

rclone

Transferring files between the HPC Prince Cluster and Google Drive

FDT

http://monalisa.cern.ch/FDT/documentation_security.html