

NEX Tips for New Users

Welcome to NEX. This document provides a collection of NAS Knowledge Base articles and short tips to address many of the questions asked by new NEX users. This guide is intended for NEX Science Users with access to the NEX high performance computing resources. For information about using the NEX web portal, please refer to the Help Center for the NEX web portal at <https://nex.nasa.gov/nex/help/> .

Please take a few minutes to read through this guide and the Knowledge Base and NEX articles referenced below. If you have additional questions, please submit them via email to support@nas.nasa.gov, and include “NEX” in the subject line of your email.

1) General guidance and the solution to most problems:

Please be sure to use the NAS Knowledge Base (KB). It’s a fantastic resource!

<http://www.nas.nasa.gov/hecc/support/kb/>

For general information about NEX data and software resources, please see:

<https://nex.nasa.gov/nex/projects/1304/>

2) Logging in

NAS users 2-factor authentication. Please see the NAS KB overview first:

http://www.nas.nasa.gov/hecc/support/kb/Two-Step-Login-Publickey+SecurID_231.html

The general login process involves the following two simple steps:

- 1) ssh from localhost to sfe1, sfe2, sfe3, or sfe4, for example:

```
$ ssh username@sfe1.nas.nasa.gov
```

- 2) The sfe machines are only used for access to NAS, and no work can be done on these machines. Once you login, you will need to ssh from an sfe to the machine where you intend to work (see #3 below)

To save time, it is recommended that you set-up SSH passthrough to avoid having to enter your password every time you connect to machines from within the NAS enclave:

http://www.nas.nasa.gov/hecc/support/kb/Setting-Up-SSH-Passthrough_232.html

3) Overview of NEX and NAS systems

a) Bridge nodes and pfe nodes (for small tasks and testing)

bridge1, bridge2, bridge3, bridge4, pfe

http://www.nas.nasa.gov/hecc/support/kb/Pleiades-Front-End-Usage-Guidelines_181.html

http://www.nas.nasa.gov/hecc/support/kb/news/Training-Materials-for-Using-the-Sandy-Bridge-Nodes-are-Available-Online_62.html

Your home filesystem on these machines stores 10GB, and should only be used to store code, and not large datasets.

http://www.nas.nasa.gov/hecc/support/kb/Pleiades-Home-Filesystem_227.html

b) Pleiades (for large computing jobs)

A few tips for submitting jobs on Pleiades:

qsub must be used to submit large jobs to Pleiades

http://www.nas.nasa.gov/hecc/support/kb/Commonly-Used-QSUB-Options-in-PBS-Scripts-or-in-the-QSUB-Command-Line_175.html

-q *queue_name*

Defines the destination of the job. The common possibilities for *queue_name* on Pleiades include *normal*, *debug*, *long*, and *low* AND *DEVEL*. The common possibilities for *queue_name* on Endeavour include *e-normal*, *e-long*, and *e-debug*.

Queue Structure:

http://www.nas.nasa.gov/hecc/support/kb/Queue-Structure_187.html

Queue name	NCPUs/ max/def	Time/ max/def	pr
low	--/ 8	04:00/ 00:30	-10
normal	--/ 8	08:00/ 01:00	0
long	--/ 8	120:00/ 01:00	0
debug	--/ 8	02:00/ 00:30	15
devel	--/ 1	02:00/	-- 149

Resource Request Examples:

http://www.nas.nasa.gov/hecc/support/kb/Resources-Request-Examples_188.html

Preparing to Run on Ivy Bridge Nodes:
http://www.nas.nasa.gov/hecc/support/kb/Preparing-to-Run-on-Pleiades-Ivy-Bridge-Nodes_446.html

For Ivy Bridge

```
#PBS -lselect=12:ncpus=20:mpiprocs=20:model=ivy
```

For Sandy Bridge

```
#PBS -lselect=15:ncpus=16:mpiprocs=16:model=san
```

For Westmere

```
#PBS -lselect=20:ncpus=12:mpiprocs=12:model=wes
```

For Nehalem

```
#PBS -lselect=30:ncpus=8:mpiprocs=8:model=neh
```

c) lfe1 (for large file transfer and storage, tape back-up of data)

lfe1, lfe2

d) NEX sandboxes (for medium jobs and testing of workflows; a special account must be requested for use of these machines)

lnxsrv105

4) NEX data

a) NEX data resources are shared in /nex/datapool/. This is a read-only filesystem. For additional information on the contents of many of these datasets, please see:

<https://nex.nasa.gov/nex/resources/127/>

b) Project / user datasets: NEX contains many datasets that are owned by individual projects and users. Datasets that are not part of the NEX datapool are considered to be a work in progress associated with the project or user that created them. To obtain access to these datasets, please contact the project team or PI.

c) Directory structure and naming conventions: Naming conventions are determined by the project that created each dataset. Guidance for standard naming conventions for new datasets is currently being developed.

d) Data transfers: For information on copying data to/from NEX and NAS resources, please see:

http://www.nas.nasa.gov/hecc/support/kb/File-Transfer-Overview_140.html

For small data transfers within the NAS enclave (e.g., from pfe to lfe), or to copy data back to your host machine, scp should work fine. For example:

```
# To scp 'testfile.txt' from the 'oneuser' home directory on 'hostname.org' to bridge3
bridge3> scp oneuser@hostname.org:~/testfile.txt /nobackup/oneuser/
```

```
# To scp 'testfile.txt' from bridge3 to the 'oneuser' home directory on 'hostname.org'
bridge3>scp testfile.txt oneuser@hostname.org:~/
```

```
# You can also retrieve files from a URL using wget
bridge3> wget "http://sampleserver.nasa.gov/data/data.tar"
```

For large data transfers, please use the **Shift** file tool:

http://www.nas.nasa.gov/hecc/support/kb/Use-Shift-for-Reliable-Local-and-Remote-File-Transfers_300.html

A few common problems:

- i) Be sure files have valid rwx permissions before transferring them from your local machine. This can be especially problematic when transferring files from a Windows machine, and files transferred to NAS may arrive with the permissions deleted, making them unusable.
 - ii) It can be difficult to copy files from NAS to a local host that has a dynamic IP, and configuration settings may need to be changed for your local host or firewall. If your scp or shift syntax looks correct, but file transfer is still not working, check with your local system administrator for assistance or contact the NAS help line.
- e) Tape back-up: For information on copying data to tape back-up, please see:

http://www.nas.nasa.gov/hecc/support/kb/Archiving-Data-Overview_148.html

For information and tips on using the Data Migration Facility (DMF) commands, please see: http://www.nas.nasa.gov/hecc/support/kb/Data-Migration-Facility-%28DMF%29-Commands_250.html

5) User workspace

a) Disk allocation: Each user has a default disk allocation of 500GB on /nobackupp[1-***] (e.g., /nobackupp1). The /nobackupp[1-***] dirs are not accessible on the sandbox machine(s). To request an increase of up to 2TB on the /nobackupp[1-***] dirs, send an email to support@nas.nasa.gov (include “NEX” in the subject line). Requests will be approved on a case-by-case basis, and will consider available NEX disk resources and current demand from all NEX users.

For users working on the sandbox machine(s), your workspace is located under the directory /nobackup/*username*/

b) CPU allocation: At present, all NEX users share a CPU quota. Since many NEX related projects are more data intensive than CPU intensive, CPU constraints have not yet been become an issue. Your CPU request should have been specified as part of your NEX application. Please be careful not to exceed your requested allocation without prior approval. If you are unsure, please check with support@nas.nasa.gov (include “NEX” in the subject line).

c) Loading modules and using software tools: Loading modules on NAS resources can be challenging. To streamline the set-up process for NEX users, we have added a module tree on Pleiades that includes many tools and libraries that are commonly used by the Earth science community. To access this module tree, one might add the following to their .bashrc or .cshrc file:

If you use CSH:

```
setenv MODULEPATH ${MODULEPATH}:/nex/modules/files
```

If you use BASH:

```
export MODULEPATH=$MODULEPATH:/nex/modules/files
```

Once the above variables are set you should be able use "module avail" to see the available modules and “module load” to import them if desired. These modules are provisional and are not fully supported by the nas support team so please keep this in mind if you use these. For more information, please see the following articles on the NEX website and the NAS knowledge base:

https://nex.nasa.gov/nex/projects/1217/wiki/nex_loadable_modules/
http://www.nas.nasa.gov/hecc/support/kb/Modules_115.html

For general information about using bash, please see:

http://www.hypexr.org/bash_tutorial.php

For general information about using the C shell, please see:

<http://www.cs.duke.edu/csl/docs/csh.html>

d) What tools are available and where are they located?

Please see the article below for a list of the tools available on NEX:

<https://nex.nasa.gov/nex/resources/125/>

e) Can I run a database on NEX?

Within a user’s workspace, users may run SQLite. At present, the only shared database that can be run on NEX is SQLite. For an example of an SQLite implementation, please see:

https://nex.nasa.gov/nex/projects/1217/wiki/landsat_sqlite_on_bridge_noodles/version/58/

6) Best practices and tips

a) Start with small tests: Please be sure to carefully test and debug your code using small data subsets and short-runs. While the NEX/NAS resources are powerful, they are a shared resource used by many projects. Poor management of memory and disk i/o can still overload the hardware.

b) Be very careful with memory management: Different machines on NAS have different amounts of memory. Exceeding the memory limits described in the article below can significantly slow a node or machine down.

http://www.nas.nasa.gov/hecc/support/kb/Pleiades-Configuration-Details_77.html

c) Minimize disk access: As a general rule, try to read large blocks of data at once into memory, and limit high frequency disk access to iteratively read small blocks of data.

d) Set the striping when creating large datasets: You will need to manually set the stripe for large files (generally >250GB).

http://www.nas.nasa.gov/hecc/support/kb/Lustre-Basics_224.html

d) Disks can fail: /nobackupp means what it says. Anything stored in a /nobackupp[1-*] directory will not be automatically backed-up. All code, as well as critical datasets, should be regularly backed up to tape on lou, or copied back to your local machine.

e) Ask for help (but first check the NEX website and NAS Knowledge Base)! There are always new problems to solve, but chances are that other users have encountered some version of the problem, especially if it has to do with loading modules or compiling libraries. Please check the NEX website and NAS Knowledge Base first, but if you can't find an answer, don't hesitate to ask for help.