

Remote File Transfer Commands

Category: File Transfers

Columbia Phase Out:

As of Feb. 8, 2013, the Columbia21 node has been taken offline as part of the Columbia phase out process. Columbia22-24 are still available. If your script requires a specific node, please make the appropriate changes in order to ensure the success of your job.

Summary: Use the file transfer commands `scp`, `bbftp`, `bbscp`, or `shiftc` when the source and destination are located on different hosts.

Use the following file transfer commands to transfer files either between NAS high-end computing hosts, or between a NAS host and a remote host such as your local desktop system.

Except for bbFTP, the basic syntax is: copy-command [options]...source...destination.

scp command (with/without HPN-SSH patch)

Secure Copy Protocol (SCP), based on the Secure Shell Protocol (SSH), is a means of securely transferring files between a local and a remote host. Both your authentication information (such as password or passcode) and your data are encrypted.

Normal scp (without the HPN-SSH patch)

The most widely used `scp` program is from OpenSSH.

Where is scp installed at NAS?

A copy of `scp` from OpenSSH without the patch is available on the Pleiades front ends (PFEs) and bridge nodes, all Columbia nodes, Lou, and the secure front ends (SFE).

Use `scp` on Columbia, Pleiades, Lou, or your local host to push files into or pull files out of the the DMZ files servers (DMZFS).

Do you need it installed on your local host?

If you have a version of SSH installed on your local host, `scp` is most likely already installed there.

When to use it?

Typically, `scp` is used to transfer small files within NAS (<< 5 GB) or offsite (<< 1 GB) that take a reasonable amount of time to complete.

Examples

"Outbound" means the file transfer command is initiated on a NAS host such as Columbia, Pleiades, or Lou, whether the file is being pushed or pulled. "Inbound" means the commands are initiated on your local host.

Omit `local_username@` and `nas_username@` in the examples below if your local username and NAS username are identical. Substitute your own filenames for the dummy parameter "foo." These examples assume you already know [how log into the NAS enclave](#).

For outbound transfer:

```
lou2% scp local_username@your_localhost.domain:foo ./foo2
```

For inbound two-step transfer:

```
your_localhost% scp foo nas_username@sfe1.nas.nasa.gov:foo2  
sfe1% scp foo2 lou2:
```

For inbound one-step transfer if [SSH-passthrough](#) has been set up correctly:

```
your_localhost% scp foo nas_username@lou2.nas.nasa.gov:foo2
```

To transfer files through the DMZFS, initiate the `scp` command from either a NAS host or your local host, not from the DMZFS. For example:

```
your_localhost% scp foo nas_username@dmzfs1.nas.nasa.gov:foo2  
pfe20% scp dmzfs1:foo2 .
```

In the last example above, the dot (`.`) means your current directory. If no path is given after the remote host specifier (for example, `dmzfs1.nas.nasa.gov:`), the file is copied to/from the home directory of the remote host.

Performance

Within the NAS secure enclave, depending on source and destination hosts and other factors, the performance range will be 40-100 MB/sec.

If your data will compress well, consider enabling compression by adding `-C` to your `scp` command line.

We recommend upgrading to OpenSSH version 5.0 or newer.

In cases where OpenSSH 5.0 or a newer version does not yield satisfactory performance, consider applying the HPN-SSH patch to your OpenSSH, detailed below.

HPN-SSH enabled scp

HPN-SSH is a patch for OpenSSH designed to eliminate a network throughput bottleneck that typically occurs in an SSH session over long-distance and high-bandwidth networks (that is, when the bandwidth-delay product is high). The bottleneck is eliminated by allowing internal flow control buffers to be defined and grow at runtime, rather than statically defining them, as OpenSSH does. The resulting performance increase can range from 10x to more than 50x, depending on the cipher used and host tuning.

HPN-enabled SSH is fully interoperable with other SSH servers and clients. HPN clients will be able to download faster from non-HPN servers, and HPN servers will be able to receive uploads faster from non-HPN clients. However, the host receiving the data must have a properly tuned TCP/IP stack.

Where is HPN-SSH installed at NAS?

- The client version of OpenSSH with the HPN patch is available on cfe2, c[21-24], and Lou
- On SUP, both the OpenSSH client and server have been patched with the latest HPN
- The OpenSSH server with HPN patch is installed on dmzfs1-hpn and dmzfs2-hpn

On cfe2, c[21-24], and Lou, the HPN-patched SSH programs are purposely named as `hpn-ssh`, `hpn-scp`, and `hpn-sftp` to distinguish them from the default non-HPN versions (`ssh`, `scp`, and `sftp`) of OpenSSH.

Do you need it installed on your local host?

To get good performance, an HPN-SSH server must be installed on your local system if data is to be received on your local system. Ask your local network staff for help to see if an HPN-SSH patch is needed.

Typical installation procedure:

1. Download [the latest HPN-SSH patch](#) (`openssh-x.xpx-hpnxxvx.diff.gz`) from Pittsburgh Supercomputing Center
2. Download [the latest OpenSSH source](#) (`openssh-x.xpx.tar.gz`)
3. Uncompress and untar the above distributions
4. Move the file `openssh-x.xpx-hpnxxvx.diff` to the directory `openssh-x.xpx`
5. Patch: `% cat ../openssh-x.xpx-hpnxxvx.diff | patch -p1`
6. Configure [OPTIONS]: `% ./configure --prefix`
7. Make [OPTIONS]: `% make`
8. `% make install-nosysconf`
9. Validate, as shown in the example box, below
`% ssh -v`
`OpenSSH_x.xpx-hpnxxvx`

Examples

```
lou2% hpn-scp -c arcfour -o TCPRecvBufPoll=yes source destination
your_localhost% scp -c arcfour -o TCPRecvBufPoll=yes source destination
```

Notes:

- **arcfour** (RC4) is a more CPU-efficient 128-bit cipher; you can also choose **NONE** for cipher so that there is no encryption of data
- Enabling **TCPRecvBufPoll** allows for the TCP receive buffer to be polled throughout the duration of the connection

Performance

With an HPN-SSH enabled `scp`, you can expect good performance for transferring large files to remote sites over a long distance with high-latency connections. Improvement over non-patched `scp` versions older than 4.7 may be 10x to 50x.

bbftp command

bbFTP is a high-performance remote file transfer protocol that supports parallel TCP streams for data transfers. Basically, it splits a single file in several pieces and sends them through parallel streams. The whole file is then rebuilt on the remote site. bbFTP also allows dynamically adjustable TCP/IP window sizes instead of a statically defined window size used by normal `scp`. In addition, it provides a secure control channel over SSH and allows data to be transferred in cleartext to reduce overhead in unnecessary encryption. These characteristics allow bbFTP to achieve transfers that are faster than with normal `scp`.

We recommend using `bbftp` in place of `scp` large data transfers over long distances.

Where is it installed at NAS?

Both the bbFTP server (`bbftpd`) and client (`bbftp`) are installed on all Columbia hosts, Lou, Pleiades front-end and bridge nodes, and the SUP.

For DMZFS, only the bbFTP server (`bbftpd`) is installed. Issue the `bbftp` command from Columbia, Pleiades, Lou, or your local host (if bbFTP client has been installed) to push files into DMZFS or pull files out of DMZFS.

Do you need it installed on your local host?

If you want to initiate `bbftp` from your local host, you must download and install the client version of bbFTP on your local host. If you want to initiate `bbftp` from a NAS system and transfer files from/to your local host, download and install the server version of bbFTP on your local host.

When to use it?

Consider using bbFTP when transferring large files (> 1 GB) offsite. Be sure to use multiple streams to get better transfer rates.

Example

bbFTP is like a non-interactive FTP, and the syntax can be complicated.

```
your_localhost% bbftp -u nas_username -e 'setnbstream 8; get filename' -E 'bbftpd -s -m
```

(Note that the above command may appear to be broken into two lines. When you use it, enter it all on one line.)

Performance

bbFTP typically transfers data 10-20 times faster than normal `scp`.

If you are not getting good performance, check with your network administrator to see if performance tuning is needed on your system. See the article bbFTP for more instructions on installing and using bbFTP.

bbscp

bbSCP, written in Perl by Greg Matthews at NAS, is a bbFTP wrapper that provides an scp-like command-line interface. It assembles the proper command-line for bbFTP and then executes `bbftp` to perform the transfers. bbSCP is designed and tested for bbFTP version 3.2.0.

bbSCP only encrypts usernames and passwords, it does *not* encrypt the data being transferred.

Where is it installed at NAS?

bbSCP is installed on all Columbia hosts, Lou, and Pleiades front-end and bridge nodes under `/usr/local/bin`.

Do you need it installed on your local host?

If you want to initiate `bbscp` from your local host, you need to:

- [Download and install bbftp-client-3.2.0](#) on your local host
- [Download bbSCP version 1.0.6](#) (also attached at the end of this article) and install it on your local host

When to use it?

Use the bbSCP script when you want the bbFTP functionality and performance but with scp-like syntax. It can be used to transfer files within NAS enclave or between NAS and a remote site.

Example

```
your_localhost% bbscp foo nas_username@lou2.nas.nasa.gov:
```

Performance

The performance of bbSCP is the same as that of bbFTP.

See [The bbscp Script](#) for more information (man page, performance tuning, test and verification).

shiftc

Shift (**shiftc**) is a NAS-developed tool for performing automated local and remote file transfers. Shift utilizes a variety of underlying file transports to achieve maximum performance for files of any size on any file system.

Where is it installed at NAS?

Shift is installed on cfe2, Lou, the Pleiades front-end nodes (PFEs), and the bridge nodes.

Do you need it installed on your local host?

For transfers between your local host and NAS systems, you must install the SUP client as discussed in [Use Shift for Reliable Local and Remote File Transfers](#).

When to use it?

Shift can be used as a drop-in replacement for **scp** or bbSCP between any enclave systems. For transfers between your local host and NAS systems, the transfer must be initiated from your local host with **shiftc** invoked via the [SUP client](#) (that is, using the command **sup shiftc**). If an encrypted transfer is required, use the **shiftc --encrypt** option.

Example

```
bridge2% shiftc /nobackupp2/username/foo lou:  
your_localhost% sup shiftc pfe:foo .
```

Performance

Shift uses the highest performing file transport that is available on both sides of the transfer, and is optimal for the sizes of the files being transferred. This means that Shift will be as fast as bbFTP for large transfers and faster than bbFTP for small and mixed transfers.

For more information, see [File Transfer: Overview](#) and [Use Shift for Reliable Local and Remote File Transfers](#).

Article ID: 142

Last updated: 14 Feb, 2013

Data Storage & Transfer -> File Transfers -> Remote File Transfer Commands

<http://www.nas.nasa.gov/hecc/support/kb/entry/142/?ajax=1>