

AFRL
DoD Shared Resource Center
(DSRC)

Archive Server
User's Guide

January 21, 2010

Release Notes

Issued 21 January 2010

- Removed references to Eagle

Issued 08 April 2009

- Reviewed and updated to reflect change from AFRL MSRC to AFRL DSRC

Issued 20 January 2008

- Reviewed and updated to reflect change from ASC MSRC to AFRL MSRC

Issued 07 December 2007

- Removed references to SGI Origin 3900 (hpc11)
- Added references to SGI Altix 4700 (HAWK)

Issued 20 June 2007

- Removed references to Scientific Visualization workstations

Issued 1 October 2006

- Removed references to Compaq ES-45 (hpc09, hpc10).

Issued 1 June 2006

- Added references to HP XC (FALCON).

Issued 1 October 2005

- Removed references to Compaq ES-40 (hpc05)
- Added references to SGI Altix (EAGLE)

Issued 1 October 2004

- Removed references to IBM P3 (hpc04)
- Updated server type to SunFire 6800

Issued 1 October 2003

- Removed references to Compaq GS-320 (hpc08)

Issued 30 May 2003

- Added references to SGI Origin 3900 (hpc11)

Issued 8 April 1997

- Initial release

Table of Contents

1.0	Introduction.....	4
1.3	Document Conventions.....	4
2.0	AFRL DSRC Archive Server.....	4
3.0	Using the Archival Server.....	5
3.1	The Archive Command.....	5
3.1.2	Using the Archive Command in the Batch Environment.....	5
3.2	Creating a Directory.....	5
3.3	Linking to the Created Directory.....	5
3.4	Using Archive Files.....	6
3.5	Methods for Storing and Retrieving Data.....	6
3.5.1	cp and mv.....	6
3.5.2	rcp and scp.....	7
3.5.3	ftp.....	8
3.5.4	“archive” command.....	9
3.5.5	safeftp.....	10
3.6	Other Commands	10
3.6.1	Standard Long Listing (sls).....	10
3.6.2	Non-Kerberized rsh on the AFRL DSRC Application Servers.....	11
4.0	Customer Service.....	12
4.1	Consolidated Customer Assistance Center (CCAC).....	12
4.2	AFRL DSRC Support.....	12
4.3	AFRL DSRC Website.....	12

1.0 Introduction

This document provides an overview of the archival capabilities of the Air Force Research Laboratory (AFRL) Department of Defense Supercomputing Research Center (DSRC) and how they may be used. The AFRL DSRC is physically located in Building 676 of Area B at Wright-Patterson Air Force Base near Dayton, Ohio.

1.1 Assumed Background of the Reader

It is assumed that the reader of this guide has a firm grasp of the concepts required to use the Linux operating system and to program in either the C, C++, FORTRAN 77, FORTRAN 90 or FORTRAN 95 languages. Please see <http://>

1.3 Document Conventions

Unless specified otherwise, % represents the command line prompt. Anything after the % should be entered at the command line prompt in the user's terminal.

file-1 ... file-N (file names) and dir-1 ... dir-N (directory names) are the files and directories to be archived.

1.4 Additional Information

Much of the information presented in this document is available online through the man pages and is accessible while logged in by typing:

```
man {command name}
```

2.0 AFRL DSRC Archive Server

This facility provides long-term storage for user files and data backup files. The archival server uses the StorageTek Application Storage Manager (ASM) to move data files to tape when the files are inactive. The ASM is also used to move the files back to disk when a user needs to access the contents of the files. The file status information of the files that are migrated to tape remains on disk, which gives the appearance that the files are still stored on disk. The archival server is available from the AFRL DSRC High Performance Application Servers.

The archival server uses the StorageTek ASM to maintain the near-on-line storage system. ASM periodically migrates data files to tape, but retains the directory information about the files on disk. When a user attempts to access the data in a migrated file (e.g., attempting to copy the file), ASM will automatically copy the data file back to disk. The archival server provides large file-space capacity at the expense of longer access times (e.g., the time needed to read a file from tape) than from disk. Since the access time on this system can be much longer than the access time on other file systems and file servers, the archival server is best suited for long-term and permanent storage and the storage of large files.

Infrequently accessed files, large data files, and file archives of entire directories are best stored on the archival server. Users should not directly use files located on the archival server. Changes made by editing or appending to these files may not be saved if one of the Sun servers fails. Users should copy the files from the archival server to the system where the files are needed and copy the modified files back to the archival server. Possible destinations are the user's \$HOME directory, the global /scratch directory, the /workspace directory of an appropriate AFRL DSRC High Performance Server, or space on a local disk (such as a hard disk attached to the user's local workstation).

NOTE: The AFRL DSRC Archival Storage System is not a replacement for the normal backups made by the system administrators and should not be used as an extension of a user's disk allocation.

2.1 Hardware

The archival storage system for the AFRL DSRC (herein, denoted as the archival server) is currently configured with a total capacity of 500 terabytes of near-on-line storage. The system consists of a set of dual Sun SunFire 6800 cluster servers with on-line RAID storage attached to dual-combined Storage Technology Corporation (StorageTek) 9310 tape silos.

3.0 Using the Archival Server

Users cannot login directly to the archival server, but they are allowed to connect via the ftp, rcp, rsh (remote shell) and ssh/scp (secure shell/copy) commands. In addition, the locally installed utilities archive and safeftp provide other interfaces to the archival server. Each user is given an archive directory,¹ the user's archive directory, on the archival server which is network file system (NFS)-mounted on the AFRL DSRC Application Servers. Thus, users can directly access their archive directory from any AFRL DSRC Application Server.

3.1 The Archive Command

The archive system (**\$ARCHIVE_HOME**) is mounted on the interactive node of the system, however, to improve transfer speeds for batch jobs, **\$ARCHIVE_HOME** is not mounted to the batch nodes. To transfer files from **\$ARCHIVE_HOME**, users will have to use either un-kerberized rcp or the archive command. The archive command is a tool to the AFRL DSRC to help users with transferring files to and from **\$ARCHIVE_HOME**.

The AFRL DSRC startup scripts create the environment variables **ARC** and **msas** associated with the archival server. **ARC** contains the path of the user's archive directory. The user's archive directory path can be simply written as **\$ARCHIVE_HOME** as used in the examples in this User's Guide. **MSAS** contains the hostname of the server on which the user's archive directory is mounted. **\$ARCHIVE_HOST** is used where a machine hostname is required in this User's Guide. It is recommended that users use **\$ARCHIVE_HOST** and **\$ARCHIVE_HOME** for the archive server name and archive directory to avoid problems when these values change.

3.1.1 Basic Syntax for the Archive Command

The basic syntax for the archive command is:

```
% archive get [getopts] file1 [file2 ...]
% archive put [putopts] file1 [file2 ...]
```

More information on the archive command can be found on line via the archive man page (man archive).

3.1.2 Using the Archive Command in the Batch Environment

To transfer files in the batch environment, you will need to use un-kerberized rcp or the archive command, as shown below:

```
% /usr/bin/rcp ${msas}:/msas*/foo/bar
% $WRK/username archive get archive_filename local_filename
% archive put local_filename
```

3.2 Creating a Directory

Subdirectories are used to organize files within the user's archive directory. In the examples below, files are copied to and from subdirectories that must already exist.

To create a new subdirectory, use the mkdir command:

```
% cd $ARCHIVE_HOME
% mkdir dirname
```

or

```
% mkdir $ARCHIVE_HOME/dirname
```

3.3 Linking to the Created Directory

To create a link to the new directory in the user's current working directory, use the ln command:

```
% ln -s $ARCHIVE_HOME/dirname ./dirname
```

where \$ARCHIVE_HOME/dirname is the new subdirectory on the archival server. The user can access the subdirectory by referencing the link dirname.

Users can also put their data files directly in their \$ARCHIVE_HOME directory and create no subdirectories. This arrangement may become unwieldy as more files are stored in the user's archive directory.

3.4 Using Archive Files

Although rcp, scp and cp can transfer entire directories to and from the archival server, users are strongly advised to use archive files instead because they are much more efficient for storage and transfer.

3.4.1 Using Tar Files

The tar (tape archive) command creates and extracts from archive files. tar archive files, which store the directory structure as well as the file contents, are recommended for storing directories and collections of related files. For additional information regarding tar files, please see the man page by typing "man tar" at the command line.

In the below example, "tarfile" is the archive file created by the tar command.

```
% tar -cf tarfile file-1 ... file-N dir-1 ... dir-N
```

To extract the created tar file (named "tarfile" in this example), type the following:

```
% tar -xf tarfile
```

3.5 Methods for Storing and Retrieving Data

The user has three methods for copying files to and from the archival server. They are: cp and mv, rcp and scp, ftp.

In the examples below, the files are copied to and from the user's current working directory for convenience. This is not a requirement. Also, since rcp, scp, ftp and rsh begin in the user's archive home directory \$ARCHIVE_HOME on the archival server, the relative path dirname is equivalent to \$ARCHIVE_HOME/dirname. If a final "." is included, this is indicative of the current directory.

3.5.1 cp and mv

The cp command on an NFS-mounted filesystem is recoverable when a server fails, but it will be slower than the other methods. It can also be used in shell scripts. In the below examples, \$ARCHIVE_HOME/dirname is the destination directory on the archival server. Also, removing the file once it has been copied and moved will keep the

Using cp and mv to store directories or projects:

Step 1: Create a tarfile.

```
% tar -cf tarfile file-1 ... file-N dir-1 ... dir-N
```

Step 2: Copy the file for storage.

```
% cp file-1 ... file-N $ARCHIVE_HOME/dirname
```

Step 3: Move the file for storage.

```
% mv file-1 ... file-N $ARCHIVE_HOME/dirname
```

Step 4: Remove the original file from work directory, or from the home directory.

% rm tarfile

Using cp to retrieve migrated files and/or directories and untar them:

Step 1: Retrieve the files.

% cp \$ARCHIVE_HOME/dirname/file-1 ... \$ARCHIVE_HOME/dirname/file-N .

Step 2: Untar the files from the location they have been moved to.

% tar xf tarfile

Step 3: Remove the original tarfile from the work directory, or from the home directory.

% rm tarfile

3.5.2 rcp and scp

The rcp and scp commands is more complicated to use than cp. Only non-Kerberized rcp (see Section 2.3.3) can be used in batch the archival server and machines outside the AFRL DSRC. This requires that users use non-Kerberized rcp in any batch scripts.

The full rcp path names required to use rcp on the AFRL DSRC systems are:

Host Name	Path Name
Hawk	/usr/bin/rcp
Falcon	/usr/bin/rcp

Since rcp and scp have the same syntax, only rcp is used in the examples.

Using rcp to store files and/or directories (without creating a tar file):

% rcp file-1 ... file-N \${msas}:dirname

Using rcp to store directories or projects (with a tar file):

Step 1: Create an archive file using tar.

% tar cf tarfile file-1 ... file-N dir-1 ... dir-N

Step 2: Store the archive file created above.

% rcp tarfile \${msas}:dirname

Step 3: Remove the local copy.

% rm tarfile

Using rcp to retrieve files and/or directories (without a tar file):

% rcp \${msas}:dirname/file-1 ...

% \${msas}:dirname/file-N .

Using rcp to retrieve directories (with a tar file):

Step 1: Retrieve archive file tarfile.

```
% rcp ${msas}:dirname/tarfile .
```

Step 2: Extract from the archive file tarfile.

```
% tar xf tarfile
```

Step 3: Remove the tarfile.

```
% rm tarfile
```

3.5.3 ftp

ftp commands cannot be incorporated into shell scripts due to security restrictions. Only Kerberized ftp (kftp) can be used to transfer files to and from the archival server.

Using ftp for *storing* files and/or directories (without a tar file):

Step 1: Open an ftp session to the archival server.

```
% ftp $ARCHIVE_HOST
```

Step 2: Press the return key to accept the default user name.

```
username:
```

```
Name (msasX:username):<CR>
```

Step 3: Change to the destination directory dirname on the archival server.

```
% cd dirname
```

Step 4: Put the file(s) to the archival server.

```
% put file-1 or mput file-1 ... file-N
```

Step 5: End the ftp session:

```
% quit
```

Using ftp for *storing* files and/or directories (with a tar file):

Step 1: Create an archive file using tar.

```
% tar cf tarfile file-1 ... file-N dir-1 ... dir-N
```

Step 2: Open an ftp session to the archival server.

```
% ftp $ARCHIVE_HOST
```

Step 3: Press the return key to accept the default user name.

```
username:
```

```
Name (msasX:username):<CR>
```

Step 4: Change to the destination directory dirname on the archival server.

```
% cd dirname
```

Step 5: Store the archive file tarfile on the archival server.

% put tarfile

Step 6: End the ftp session.

% quit

Step 7: Delete the local copy of the archive file tarfile.

% rm tarfile

Using ftp for *retrieving* stored files and/or directories:

Step 1: Open an ftp session to the archival server.

% ftp \$ARCHIVE_HOST

Step 2: Press the return key to accept the default user name.

username:

Name (msasX:username):<CR>

Step 3: Change to the source directory dirname on the archival server.

% cd dirname

Step 4: Retrieve the file(s) from the archival server.

% get file-1 or mget file-1 ... file-N

Step 5: End the ftp session.

% quit

Step 6: Extract from the archive file tarfile and remove it.

% tar xf tarfile

Step 7: Remove tarfile.

% rm tarfile

3.5.4 “archive” command

The archive command is a locally installed utility that provides a uniform interface to the archival servers at DoD DSRCs. The archive command can be used from within batch scripts.

Using the archive command to *store* file(s) (without a tarfile):

archive put -C dirname file-1 ... file-N

Using the archive command to *store* directories or projects (using a tarfile):

Step 1: Create an archive file using tar.

% tar cf tarfile file-1 ... file-N dir-1 ... dir-N

Step 2: Store archive file.

% archive put -C dirname tarfile

Step 3: Delete local copy.

% rm tarfile

The archive command can also create archive files using tar implicitly. See the man pages for more details.

Using archive to *retrieve* files and/or directories (without a tarfile):

% archive get -C dirname file-1 ... file-N

Using archive to *retrieve* files and/or directories (with a tarfile):

Step 1: Retrieve archive file tarfile.

% archive get -C dirname tarfile

Step 2: Extract from the archive file tarfile.

% tar xf tarfile

Step 3: Remove the archived tarfile.

% rm tarfile

The archive command can also extract from archive files using tar implicitly. See the man pages for more details.

3.5.5 safeftp

The safeftp command: the safeftp command is a locally installed utility that provides an interface similar to rcp and scp, but using ftp, which allows for some error checking and restart capabilities. Since it uses ftp, it cannot be used in batch scripts. Since the form is similar to rcp and scp, it is not described.

3.6 Other Commands

This section provides a brief description of other commands that may be useful for users on the archive system.

3.6.1 Standard Long Listing (sls)

The sls command is an extended version of the Unix ls command. It can be used to help determine the status of files on the archive server.

The following commands can be used for a single line listing:

```
% ls -l $ARCHIVE_HOME/test.dat (NFS mounted)  
or  
% sls -l $ARCHIVE_HOME/test.dat
```

produces:

```
-rw----- 1 user user 1000 Sep 10 12:50 /msas031/user/test.dat
```

The following commands can be used for two line listing:

```
% sls -2 $ARCHIVE_HOME/test.dat
```

produces

```
-rw----- 1 user user 1000 Sep 10 12:50 /msas031/user/test.dat
```

```
O-a----- --- sg
```

In addition to the output of the standard long listing, the “O” in the second line indicates that the file is off-line and must be staged before being accessed.

The following commands can be used for detailed listing:

```
% sls -D $ARCHIVE_HOME/test.dat
```

produces:

```
/msas031/user/test.dat mode: -rw----- links: 1 owner: user group: user
```

```
length: 1000 inode: 608
```

```
offline; archdone;
```

```
copy 1: ---- Sep 10 12:51 27eb4.1 sg 000305
```

```
access: Sep 10 12:50 modification: Sep 10 12:50
```

```
changed: Sep 10 12:50 attributes: Sep 10 13:52
```

```
creation: Sep 10 12:50 residence: Sep 10 13:52
```

refer to the on-line man pages for more information regarding sls.

As an alternative, a wrapper for sls exists on the AFRL DSRC Application Servers. This wrapper starts a remote shell on the archival server and passes any arguments entered to the indicated command. For example: sls -l is equivalent to rsh \$ARCHIVE_HOST 'sls -l'.

NOTE: when using these wrapper utilities, arguments that contain wildcards (e.g., *, ?, [,]) need to be quoted to avoid being expanded by the shell on the local machine.

3.6.2 Non-Kerberized rsh on the AFRL DSRC Application Servers

These commands are available via rsh or ssh to the Sun servers. The command line used for these commands should be in the form:

```
rsh|ssh $ARCHIVE_HOST 'command options filelist'
```

Only non-Kerberized rsh can be used in batch scripts. Only Kerberized rsh or ssh can be used to connect to the archival server from machines outside the AFRL DSRC.

System	rsh Location
Hawk	/usr/bin/rsh
Falcon	/usr/bin/rsh

4.0 Customer Service

4.1 Consolidated Customer Assistance Center (CCAC)

For customer assistance, call the Consolidated Customer Assistance Center (CCAC) at:
Toll Free: 1-877-CCAC-039 (1-877-222-2039)
Local: (937) 255-0679
DSN: 785-0679
e-mail: help@ccac.hpc.mil.

If you have any questions regarding the AFRL DSRC, contact CCAC first. If your problem or question is beyond the scope of the CCAC, they will refer you to the appropriate resource.

4.2 AFRL DSRC Support

In-depth technical inquiries and problems are forwarded to the AFRL DSRC Customer Assistance and Technology Center (CATC), which pursues such inquiries and problems through resolution as rapidly as possible. The AFRL DSRC CATC will attempt to determine the nature of the problem, then identify and coordinate whatever resources are needed to resolve the problem.

4.3 AFRL DSRC Website

The AFRL DSRC website (<http://www.afrl.hpc.mil>) is the best source for current AFRL DSRC information. Some of the topics found on the website include:

- **APPLICATIONS** (<http://www.afrl.hpc.mil/software/>)
Short and long descriptions of current AFRL DSRC applications
- **SYSTEMS** (<http://www.afrl.hpc.mil/hardware/>)
Information on AFRL DSRC servers and Archival Storage
- **CONSOLIDATED CUSTOMER ASSISTANCE CENTER (CCAC)** (<http://www.ccac.hpc.mil>)
Available CCAC Services
- **ONLINE DOCUMENTATION** (<http://www.afrl.hpc.mil/customer/userdocs/>)
Listings of the AFRL DSRC User Guides are available for viewing.
Instructions are given on obtaining postscript versions.
- **TRAINING** (<https://okc.erdc.hpc.mil/index.jsp>)
Current course offerings and schedule
- **FREQUENTLY ASKED QUESTIONS**
Read about various topics (such as "Customizing Your Environment")
- **POLICIES AND PROCEDURES** (http://www.afrl.hpc.mil/overall/policy_procedure/)
The latest policies regarding usage of the AFRL DSRC resources

