

HPC DMF Basics

What is DMF?

Data Migration Facility (DMF) is SGI's hierarchal storage manager solution that migrates files between online, near line and offline storage based on defined criteria. At USQ we have a Storage Array attached to the Fawkes HPC for the online storage and a tape library for the near line storage. Currently there is no offline archive facility. DMF manages the disk and tape storage components to create file systems with very large capacities.

DMF provides data protection but is not a backup system.

DMF provides high performance and scalable access to very large volumes of data in a HPC environment. Data is protected by writing multiple copies of each file to independent storage volumes. DMF provides data protection but does not use all the mechanisms of a traditional backup system because it has different design goals.

Traditional backup systems support recovery of files from an arbitrary point in time using backup media which has been allocated an appropriate protection interval. For example a backup system can recover a file from an annual backup from two years ago providing the backup media is available. DMF does not support this type of data recovery.

How does DMF work?

DMF provides very large storage capacities by moving data between disk and tape storage tiers as required. The disk tier operates as a cache of the newest and most recently used files while the tape tier provides bulk storage. When a new file is written into a DMF file system it initially exists only on disk. Typically within 24 hours two copies of the new file will be written to tape. While sufficient disk capacity is available copies of files on disk are retained. Once disk space falls below a defined threshold files that have not been recently accessed will be deleted from the disk tier until the required amount of free storage is recovered. It is important to understand that files that have been removed from disk **have not been deleted** from the file system. When a file that does not have a copy on disk is required it will be retrieved from tape automatically without user intervention.

Which Fawkes file systems are managed by DMF?

There are three file systems on our SGI HPC system that are DMF enabled. These are:

- /research
- /home
- /usr/local/opt

How do I use DMF?

A user does not need to know that their data is on a DMF file system to use it successfully. DMF can be used like any other file system.

To make the best use of the DMF file systems and the Fawkes HPC environment it is necessary to understand how DMF can affect file operations.

What happens when I delete a file?

When a file is deleted from the file system, it is marked in the DMF database with a 'soft-delete' timestamp. The duration between marking 'soft-delete' and removing it completely from the database is controlled by the DUMP_RETENTION parameter which is currently set to 32 weeks. For 32 weeks after deletion the record of a deleted file is held so it is possible to recover the file. After 32 weeks the chances of recovery of a deleted file are low.

The behaviour of files deleted from the DMF is why it is described as providing data protection and not data backup.

How do I restore a deleted file?

To restore a deleted file the metadata of the file must be recovered from the DMF database. The restored metadata is then used to retrieve a copy of the file from tape. The retrieval task can only be performed by a DMF administrator and is likely to be successful if no more than 32 weeks have elapsed since the file was deleted.

How do I restore a previous version of a file?

To restore a previous version of a file the metadata of the file must be recovered from the DMF database. The restored metadata is then used to retrieve the required version of the file from tape. The retrieval task can only be performed by a DMF administrator and is likely to be successful if no more than 32 weeks have elapsed since the required version of the file was modified.

How do I check the status of my files?

You can issue the command "dmls -l [filename]" to see the status of your files. A sample output is below:

```
[hpc-dmf-prd-t1]$ dmls -l almutils.tar  
-rw-r--r-- 1 root root 62699520 2015-08-05 12:47 (DUL) almutils.tar  
[hpc-dmf-prd-t1 ]$
```

In this sample the file is dual-state as indicated by (DUL). DMF files are typically in one of three states. These are:

DUL - indicates a dual-state file which means that copies of the file exist on the disk file system and on tape.

REG - indicates a regular file which means that a copy of the file exists only on the disk file system.

OFL - indicates an offline file which means that a copy of the file exists only on tape.

How do I access my files if it is only on tape?

If your file has an offline status (OFL) then you can use the "dmget" command to retrieve it from tape onto the disk file system. The syntax is "dmget [filename]". The retrieval process may take some time depending on the size of the file. After the retrieval is complete the file's status will be dual-state. The alternative to explicitly retrieving an offline file with "dmget" is to access the file as you would any regular or dual-state file. The program accessing the file will simply wait for the file to be retrieved from tape before continuing.

How do I manually push a file(s) to tape so it doesn't use up my home drive space?

The following command will remove the copy from disk by performing a release operation and leaves only a file stub on disk but a full copy still on tape. When you issue a "dmls -l [filename]" command you will then see the words (OFL) in front of the filename and this indicates that the file is offline (i.e. not on disk file system).

```
[hpc-dmf-prd-t1]$ dmls -l almutils.tar
-rw-r--r-- 1 root root 62699520 2015-08-05 12:47 (DUL) almutils.tar
[hpc-dmf-prd-t1]$ dmput -r almutils.tar
-rw-r--r-- 1 root root 62699520 2015-08-05 12:47 (OFL) almutils.tar
[hpc-dmf-prd-t1 ]$
```

Where can I get more information on DMF

There is general information available on SGI's webpage about DMF. <https://www.sgi.com/products/storage/idm/dmf.html>

If you have specific questions on DMF the HPC support team will help you with your enquiry, we also have access to SGI support staff to assist with DMF queries.

How is the HPC environment backed up

The SGI HPC consists of admin node, login node, DMF (Storage) node and compute nodes. The admin node has a nightly file system dump that is written to the DMF storage. This backup is copied off to a backup gateway where it is stored for 3 months and written to tape for 6 months. The login and compute nodes are built from the same image and the image is stored on the admin node. These image files are included in the backup process of the admin node. The DMF server operating system is backed up using xfsdump commands and is on a 6 months retention on tape.