

Archive & Backup Revealed

Revised 4 March 1985

What is Archive?

There seems to be some basic confusion on the nature and purpose of a file archive system. On the DEC-20, DEC has implemented a file archive system that is the envy of many Unix and VAX/VMS users. Very simply, an archive system is a means for providing off-line storage of data which is not referenced frequently enough to justify the costs of keeping it on-line. There are costs associated even with the off-line archival storage, as you might expect.

What is Backup?

Much of the confusion seems to stem from failure to understand the basic process of data backup as practiced on every computer everywhere. *We do not rely on or even utilize the Archive system as any part of the data backup procedure.* There is no need to archive files merely to guard against the eventuality of the computer crashing and destroying disk. Furthermore, there is no need to archive a file to guard against your own error. The window over which archives are done (as much as two weeks after requests) does not provide protection from accidentally destroying transient data. **Merely keep one or more extra versions of a file for this kind of safeguard.**

Daily Dumps

We do a daily dump of all files that have been changed since the previous day. These dumps are saved for a **one month** period.

Full dumps

We do a weekly full dump of all files on the system, regardless of how recently they were created. These dumps are saved for a **three month** period. Together with the daily dumps, these dumps provide the best protection against user error in accidentally destroying data.

Off-site dumps

See the NIC Catastrophe Plan for more information on this. Basically, alternate weeks of full dumps are kept in storage at Lake Tahoe. This provides protection against data loss in the event of the destruction of our Menlo Park Facility.

Perm dumps

Every three months, one full dump is set aside, relabeled, and stored in the vault in perpetuity. This is an automatic form of archival storage. It also protects against damaged data which goes undetected over a very long period of time. This system has been in effect since June of 1982, long before the current computer facility staff joined the group. For information, the list of such perm dumps appears below.

NIC Permanent dump record

Machine	# Tapes	Date
NIC (F4)	21, 1 SAVASS	6/21/82
NIC (F4)	20, 1 SAVASS	9/20/82
NIC (F4)	23, 1 SAVASS	1/3/83
NIC (F4)	20, 1 SAVASS	3/4/83
NIC (F4)	21, 1 SAVASS	4/8/83
NIC (F4)	24, 1 SAVASS	8/24/83
NIC (F4)	13, 1 SAVASS	12/21/83
NIC (F4)	10, 1 SAVASS	4/18/84
NIC (F4)	10, 1 SAVASS	7/12/84
NIC (F4)	13, 1 SAVASS	10/30/84
NIC (2060)	3 SS:	8/25/83
NIC (2060)	4 PS:	8/30/83
NIC (2060)	1 SRC:, 4 PS:	9/26/83
NIC (2060)	1 SRC:, 4 PS:	12/23/83
NIC (2060)	1 SRC:, 4 PS:, 2 SS:	4/16/84
NIC (2060)	1 SRC:, 4 PS:, 2 SS:	7/11/84
NIC (2060)	1 SRC:, 4 PS:, 4 SS:	10/20/84

User Archive

Users can identify data to which they no longer need instant access. They can then store this data off-line through use of the archive system. **This is not a replacement for the delete command.** Some data is transient in nature and should be deleted rather than archived. For example, the multiple copies of a file created for protection against a user's own mistakes would not normally be archived, but would rather be deleted. By leaving such files online overnight, a user can know that they will be on a daily dump tape scheduled for one month retention, and so there is no need to archive the data just to provide emergency access over such a short period.

The best information on a user's use of the archive system is in the TOPS-20 User's Guide. See the Index for a pointer.

In addition, refer to the following message for a succinct summary of the archive commands, and some motivation for using them.

Date: Tue 19 Feb 85 02:35:39-PST
 From: David Roode <ROODE@SRI-NIC.ARPA>
 Subject: DISK SPACE LOW
 To: nic-staff@SRI-NIC.ARPA
 cc: Operations@SRI-NIC.ARPA
 Location: EJ286 Phone: (415) 859-2774

We currently do not have room on the system disk for a dump and subsequent normal operation following a crash. This is a bad situation and we need to free some disk space. Please go through your directories, as well as those of any you are responsible for, and delete unneeded files.

We have some plans underway to assign responsibilities for directories, and to identify files which are going unused, but your assistance in the meantime is very helpful.

Also, please note, I am not asking you to delete anything you really need, and especially people who use less than 500 pages need not take this message to heart (but every little bit does help).

The archive system is also available to store files which you do not need at present but believe you will need in the future. For information on this facility, do HELP ARCHIVE. Briefly,

mark a file for archive with

QARCHIVE filename <== not ARCHIVE FILE filename

Check names of archived files with

QVDIRECTORY ,
 Qarchive
 Q

Retrieve a previously archived file with

QRETRIEVE filename

Check status of retrieves requested with

QINFO RETRIEVE

Send mail to Action to hurry retrieves along.

Once archived and retrieved, a file is not writable, but you can edit it and write out a new version number with changed data.

Thanks for paying attention to housekeeping...

Migration

To aid users who are too busy to go through their files and identify those currently unreferenced, a facility is provided known as Migration. After a period of non-reference (usually around 30 days), files migrate to off-line storage. This is non-destructive because the user can get them back. In addition, studies have shown that few files so migrated are ever sought again. This service is an AID to the user and is also critical to a system maintaining sufficient disk space for continued operation.

The commands for accessing migrated files are the same as those for accessing user-archived files. The only difference is that migrated files still appear in directory listings (with the ;OFFLINE attribute), and their names are accessible by programs on the system. A user may change that at any time through use of the SET FILE INVISIBLE command, after which the files are in a state identical to the one they would be in if the user had issued the archive request instead of the system. Until this is done, when someone attempts to look at the file, he will see a message stating that the files are in off-line storage, and he may then request their retrieval. This is an improvement over the BSYS archive system in that archived files simply disappeared and the user had to know to look for them specially.

What is Archive

There seems to be some basic confusion on the nature and purpose of a file archive system. On the DEC-20, DEC has implemented a file archive system that is the envy of many Unix and VAX/VMS users. Very simply, an archive system is a means for providing off-line storage of data which is not referenced frequently enough to justify the costs of keeping it on-line. There are costs associated even with the off-line archival storage, as you might expect.

What is Backup

Much of the confusion seems to stem from failure to understand the basic process of data backup as practiced on every computer everywhere. *We do not rely on or even utilize the Archive system as any part of the data backup procedure.* There is no need to archive files merely to guard against the eventuality of the computer crashing and destroying disk. Furthermore, there is no need to archive a file to guard against your own error. The window over which archives are done (as much as two weeks after requests) does not provide protection from accidentally destroying transient data. **Merely keep one or more extra versions of a file for this kind of safeguard.**

Daily Dumps

We do a daily dump of all files that have been changed since the previous day. These dumps are saved for a **one month** period.

Full dumps

We do a weekly full dump of all files on the system, regardless of how recently they were created. These dumps are saved for a **three month** period. Together with the daily dumps, these dumps provide the best protection against user error in accidentally destroying data.

Off-site dumps

See the NIC Catastrophe Plan for more information on this. Basically, alternate weeks of full dumps are kept in storage at Lake Tahoe. This provides protection against data loss in the event of the destruction of our Menlo Park Facility.

Perm dumps

Every three months, one full dump is set aside, relabeled, and stored in the vault **in perpetuity**. This is an automatic form of archival storage. It also protects against damaged data which goes undetected over a very long period of time. This system has been in effect since June of 1982, long before the current computer facility staff joined the group. For information, the list of such perm dumps appears below.

NIC Permanent dump record

Machine	# Tapes	Date

NIC (F4)	21, 1 SAVASS	6/21/82
NIC (F4)	20, 1 SAVASS	9/20/82
NIC (F4)	23, 1 SAVASS	1/3/83
NIC (F4)	20, 1 SAVASS	3/4/83
NIC (F4)	21, 1 SAVASS	4/8/83
NIC (F4)	24, 1 SAVASS	8/24/83
NIC (F4)	13, 1 SAVASS	12/21/83
NIC (F4)	10, 1 SAVASS	4/18/84
NIC (F4)	10, 1 SAVASS	7/12/84
NIC (F4)	13, 1 SAVASS	10/30/84
NIC (2060)	3 SS:	8/25/83
NIC (2060)	4 PS:	8/30/83
NIC (2060)	1 SRC:, 4 PS:	9/26/83
NIC (2060)	1 SRC:, 4 PS:	12/23/83
NIC (2060)	1 SRC:, 4 PS:, 2 SS:	4/16/84
NIC (2060)	1 SRC:, 4 PS:, 2 SS:	7/11/84
NIC (2060)	1 SRC:, 4 PS:, 4 SS:	10/20/84

User Archive

Users can identify data to which they no longer need instant access. They can then store this data off-line through use of the archive system. **This is not a replacement for the delete command.** Some data is transient in nature and should be deleted rather than archived. For example, the multiple copies of a file created for protection against a user's own mistakes would not normally be archived, but would rather be deleted. By leaving such files online overnight, a user can know that they will be on a daily dump tape scheduled for one month retention, and so there is no need to archive the data just to provide emergency access over such a short period.

The best information on a user's use of the archive system is in the TOPS-20 User's Guide. See the Index for a pointer.

In addition, refer to the following message for a succinct summary of the archive commands, and some motivation for using them.

Date: Tue 19 Feb 85 02:35:39-PST
 From: David Roode <ROODE@SRI-NIC.ARPA>
 Subject: DISK SPACE LOW
 To: nic-staff@SRI-NIC.ARPA
 cc: Operations@SRI-NIC.ARPA
 Location: EJ286 Phone: (415) 859-2774

We currently do not have room on the system disk for a dump and subsequent normal operation following a crash. This is a bad situation and we need to

free some disk space. Please go through your directories, as well as those of any you are responsible for, and delete unneeded files.

We have some plans underway to assign responsibilities for directories, and to identify files which are going unused, but your assistance in the meantime is very helpful.

Also, please note, I am not asking you to delete anything you really need, and especially people who use less than 500 pages need not take this message to heart (but every little bit does help).

The archive system is also available to store files which you do not need at present but believe you will need in the future. For information on this facility, do HELP ARCHIVE. Briefly,

mark a file for archive with
 @ARCHIVE FILE filename

Check names of archived files with
 @VDIRECTORY ,
 @archive
 @

Retrieve a previously archived file with
 @RETRIEVE filename

Check status of retrieves requested with
 @INFO RETRIEVE

Send mail to Action to hurry retrieves along.

Once archived and retrieved, a file is not writable, but you can edit it and write out a new version number with changed data.

Thanks for paying attention to housekeeping...

Migration

To aid users who are too busy to go through their files and identify those currently unreferenced, a facility is provided known as Migration. After a period of non-reference (usually around 30 days), files migrate to

off-line storage. This is non-destructive because the user can get them back. In addition, studies have shown that few files so migrated are ever sought again. This service is an AID to the user and is also critical to a system maintaining sufficient disk space for continued operation.

The commands for accessing migrated files are the same as those for accessing user-archived files. The only difference is that migrated files still appear in directory listings (with the ;OFFLINE attribute), and their names are accessible by programs on the system. A user may change that at any time through use of the SET FILE INVISIBLE command, after which the files are in a state identical to the one they would be in if the user had issued the archive request instead of the system. Until this is done, when someone attempts to look at the file, he will see a message stating that the files are in off-line storage, and he may then request their retrieval. This is an improvement over the BSYS archive system in that archived files simply disappeared and the user had to know to look for them specially.