EFFECTIVE BACKUP STRATEGIES FOR THE HP3000

Bud Beamguard
Syntex Corporation
3401 Hillview Avenue
Palo Alto, CA 94304

Let us start out with a thought:

"Risk is a theoretical concept until you have been burned."

A backup is the transferal of data from a more volatile medium to a less volatile medium, in order to provide a temporary second copy of the data offline. On the HP3000, this usually means storing off the contents of discs to magnetic tape, using SYSDUMP, STORE, DESTORE, etc. With SYSDUMP a complete copy of the system can be made.

We do backups of various kinds in order to preserve the integrity of the system:

1. The System Directory
2. System I/O configuration, system tables and parameters
3. System software: FOS and subsystems from HP
4. Third-party packages
5. In-house software
6. User data

Several additional reasons why backups are done:

1. Because HP tells you to
2. Legal liability
3. Government regulations
4. Loss of financial records
5. Non-reconstructable data, e.g., data from real-time instruments
6. Corporate/governmental auditing requirements

There are plenty of things that can happen to a system, both at the hardware and software levels:

1. Disc head crashes and other disc problems
2. Corrupt system directory
3. Volume table destroyed, must perform RELOAD
4. Inappropriate use of PM capability resulting in carnage
5. Not enough free space on LDEV 1 during load

6. Sabotage and security breaches
7. Fire, water, acid, etc. in computer room
8. Earthquakes, tornadoes, floods, meteor strikes, lightning, war, the End of Civilization As We Know It, etc.

Any of these can develop into a recovery situation, i.e., situations where data (or even the system itself) must be recovered from previously prepared backup tapes.

By their nature, emergencies are unforeseen. For purposes of developing a backup strategy, it is not important what causes emergencies: software, hardware, or human error. Rather we attempt to be prepared as best we can WHEN and IF they happen. It cannot be emphasized enough that no universal solution exists, particularly where extensive Turbo-IMAGE databases are in use.

Our goal is to be able to:

1. Bring MPE back up with as little delay as possible
   (Obvious? How are you going to get anything done with a dead system? How long will it take to recreate MPE from you last MPE upgrade tapes?)
2. Bring back the accounting structure/system directory/UDC structure/passwords
   (How long will it take to reconstruct all of this by hand? How complete is the documentation?)
3. Recover the system I/O configuration, system tables, system operating environment
4. Recovery software (MPE, subsystems, user-developed)
5. Recover user data
   User data gets the most attention in discussions of backup strategies, but it can be useless without a coherent system to run it on. Applications are often highly dependent upon a particular system environment.

Although backup tapes are frequently useful for recovering files which have been accidentally purged or otherwise ruined during day-to-day operations, this function is secondary. Backups should not be approached as archival operations, either.

Ideally, we should be able to restore the system to the exact state it was at the time of the outage. Even more ideally, we should be able to do this on a completely different hardware setup, should the old machine be physically destroyed. Usually we have to settle for a system as it was during the most recent backup.

A backup strategy goes hand-in-hand with a disaster recovery plan. A disaster recovery plan will clarify your needs and goals based on YOUR specific environment, and make you aware of the often cold realities of recovery operations. With a good set of backup tapes and a reasonable level of technical competence, you can recover from an emergency and come out looking like a genius. The exact opposite is also possible if you are not prepared! Develop a backup strategy that will provide you with the means to approach an emergency with confidence. Draw up a comprehensive recovery plan. A routine reload of the system, done to repack disc files, is an excellent opportunity to test such a recovery plan and to eliminate the glitches which occur during reloads. If you feel up to it, get out your DUS tape and practice loading and using SADUTIL. SADUTIL can be a life-saver in that it is able, in some cases, to rescue files which would otherwise be lost; sometimes there are vital files for which a timely backup copy does not exist.

A few general considerations about backups:

1. Backups are a form of insurance.
   Unless you intend to go naked, you are going to have to get some of this insurance.

2. Backups are an attempt to keep risk within acceptable limits. They DO NOT eliminate risk. Even tapes have their problems and limitations.

3. Like insurance, backups cost money. The cost is directly proportional to the coverage obtained.

4. Remember that system downtime is a major cost factor.

5. Only YOU know the acceptable level of risk in your situation.

6. The acceptable level of risk is going to be different for every set of files or databases on the system.

7. You are going to have to balance cost versus coverage when deciding on a backup strategy.

8. Backups must be a design requirement in any application system.

9. Be very careful about promising more than you can deliver. Never underestimate the potential complexities of a recovery, and never overestimate your technical abilities.

Backups require downtime, or at the minimum intervals during which applications and data are not available for use. It is this non-availability which is the major problem for most system managers.

Several suggestions to deal with the downtime dilemma:

1. Make sure that upper management understands what backups are and why they are necessary. It is your job to help them appreciate this necessity.

2. Make certain that backups occur on an iron-rigid schedule. Users can get used to anything as long as you are punctual and consistent. Discipline on your part helps them to utilize downtime for other purposes, to best advantage.

3. See to it that backups are included during the requirements phase of system/application planning. Beware of situations where backups are unfeasible due to uptime demands, because this is a no-win situation for you.

4. Try to present backups in a positive manner, not as a waste of time and money.

5. What goes up (MPE) must come down: nobody ever promised either a fail-safe operating system or a perpetual-motion hardware setup. Make sure that everybody realizes this. Sometimes people expect reliability levels out of a computer which they would never ask from a car or airplane. Hasten to correct these delusions.

SYSDUMP

Every system manager is familiar with SYSDUMP. An understanding of its operation can provide ideas for an effective backup strategy.

SYSDUMP performs the following steps (in order):

1. Asks ANY CHANGES?

2. Asks for dump date.
   Will backup all files created or modified ON or SINCE this date, on the basis of their file labels.

3. Asks for dump file set.
   Identical with a STORE command parameter string.

4. Asks whether to do a dump list.

5. Dumps the System Directory, the system SL, the I/O configuration, the FOS files from PUB.SYS, etc. onto the tape.

6. Activates STORE as a son process, which carries out the remaining three steps:

7. Searches the System Directory and disc file labels for files matching the date and file-set parameters specified.
   Prepares a list of these in a temp file called GOOD.

8. Stores the files listed in GOOD to tape.

9. If specified, prepares a printed list of the files, using GOOD.

Here is an example of a job stream to perform a full backup:

```
!JOB SYSDUMP,MANAGER.SYS,OPERATOR;HIPRI;outclass=lp,4,1
!COMMENT This jobstream does a full backup of the system.
!FILE T;DEV=TAPE
!FILE DUMPLIST;DEV=LP,4,1
!SYSDUMP *T,*DUMPLIST
NO                          << any changes >>
0                           << dump date >>
@.PUB.SYS,&                 << dump file set >>
@.@.SYS-@.PUB.SYS,&         << dump file set >>
@.@.@-@.@.SYS;&             << dump file set >>
FILES=24000;&               << dump file set >>
PROGRESS=1                  << progress message every 1 minutes >>
YES                         << list files dumped >>
!EOJ
```

The dump file set specification in this job stream will cause the contents of PUB.SYS to be dumped first, then the SYS account minus PUB.SYS, then everything else minus the SYS account. All files are dumped only once. The tape set thus created allows you to restore "first things first" during a recovery or a reload. For example, COMMAND.PUB.SYS is extremely important because it contains the UDC structure of the system, and can be restored quickly since it is toward the beginning of the first tape reel. Similarly, the PUB.SYS group can be restored and a measure of order restored to the system environment before the user accounts are recovered. It may be possible to release certain applications even while the recovery is still in progress.

Several frills to add to your backup jobstreams:

1. STARTCACHE commands

2. Do a run of BULDACCT.PBU.TELESUP. This program will create three jobstreams, JOBACCT, JOBACCTB, and JOBCUDC. These streams can be used to recreate the System Directory. They contain ALL your passwords!

3. LIMIT and JOBFENCE commands to keep stray users off the system.

4. Tape validations using VALIDATE or FCOPY. (These take time, but may be worth it.)

An enhanced version of a full backup:

```
!JOB SYSDUMP,MANAGER.SYS,OPERATOR;HIPRI;outclass=lp,4,1
!COMMENT*****************************
!COMMENT This jobstream does a full backup of the system.
!COMMENT*****************************
!LIMIT 0,0
!JOBFENCE 14
!COMMENT*****************************
!COMMENT Make sure disc caching is turned on.
!CONTINUE
!STARTCACHE 1
!CONTINUE
!STARTCACHE 2
!CONTINUE
!STARTCACHE 3
!CONTINUE
!STARTCACHE 4
!COMMENT*****************************
!COMMENT Execute BULDACCT
!PURGE JOBACCT.OPERATOR.SYS
!PURGE JOBACCTB.OPERATOR.SYS
!PURGE JOBCUDC.OPERATOR.SYS
!RUN BULDACCT.PUB.TELESUP
!COMMENT*****************************
!SHOWTIME
!COMMENT*****************************
!FILE T;DEV=TAPE
!FILE DUMPLIST;DEV=LP,4,1
!SYSDUMP *T,*DUMPLIST
NO                      << any changes >>
0                       << dump date >>
@.PUB.SYS,&             << dump file set >>
@.@.SYS-@.PUB.SYS,&     << dump file set >>
@.@.@-@.@.SYS;&         << dump file set >>
FILES=24000;&           << dump file set >>
PROGRESS=1              << progress message every 1 minutes >>
YES                     << list files dumped >>
!COMMENT*****************************
```

```
!PURGE JOBACCT.OPERATOR.SYS
!PURGE JOBACCTB.OPERATOR.SYS
!PURGE JOBCUDC.OPERATOR.SYS
!LIMIT 5,45
!JOBFENCE 2
!COMMENT*****************************
!SHOWTIME
!COMMENT*****************************
!COMMENT validates the tape(s) using VALIDATE.PUB.TELESUP
!RUN VALIDATE.PUB.TELESUP
N     << PRINT THE TAPE DIRECTORY ? >>
N     << PRINT THE FILE LABEL INFO ? >>
Y     << VALIDATE THE ENTIRE TAPE ? >>
N     << PRINT LIST ON LINE PRINTER (LP) ? >>
!COMMENT*****************************
!SHOWTIME
!COMMENT*****************************
!EOJ
```

A widely used backup strategy is to use SYSDUMP to perform a full system backup once a week (often on weekends), with a partial backup on each working day thereafter, using the date of the full backup as the dump date of the partial; all files modified on or after that date are then dumped. This strategy is so popular that the commands FULLBACKUP and PARTBACKUP were recently added to MPE to facilitate operations. This is an excellent approach to backups: the tape sets thus created make recoveries quick and straight-forward. If downtime and operator time are not a problem, this is certainly a recommended backup strategy.

Nevertheless, this strategy contains plenty of redundancy (i.e., overkill):

1. MPE, the System Directory, I/O configuration, etc. are repeatedly backed up, perhaps needlessly should the system be static.

2. The "@.@.@" fileset specification will cause the machine to examine EVERY FILE LABEL ON THE SYSTEM, which takes plenty of time; again perhaps needlessly.

3. A great many possibly unimportant files will get backed up, usually several times; this takes more downtime.

4. A large number of backup tapes are generated, with resulting storage/security problems.

5. FULLBACKUP and PARTBACKUP use an unqualified "@.@.@" as the fileset, creating tapes which may be clumsy to use in a recovery. It is usually better to use SYSDUMP and fix up your own sequence.

If your site is under pressure to keep downtime to a minimum, there are alternatives which may or may not be appropriate for your situation.

1. Perform partials based on the date of the previous partial (or full, whichever is more recent). This will tend to reduce redundancy and thus shorten backup time. The downside of this approach is that during a recovery situation a RESTORE must be run on each of the partials, in correct order by date; an error-prone and time-consuming process.

2. Extend the time between full backups, e.g., to the first of each month. This may be a good plan in situations where the same set of user files gets modified and hence backed up day after day, while the system environment is stable.

3. In situations where the MPE environment is stable, consider the following strategy:

   a. Prepare a SYSDUMP tape containing @.PUB.SYS, based on a dumpdate of 0. This tape can be used to recover MPE and its environment.

   b. Use STORE to perform full and partial backups.

   The redundant backup of the MPE environment is eliminated.

4. Establish an Account structure in which active or critical files are kept in special groups by themselves. Examples would be major database, or a source code library. During the backup use a dump file subset which contains only these selected groups. The amount of time consumed by the directory search will be greatly reduced. This is at some risk to files NOT in these groups (which may be tolerable).

5. Sometimes a TurboIMAGE database contains datasets which are extremely volatile, and other datasets which are static. Consider breaking off the volatile datasets into a second separate database. In this manner the static sets will not be backed up merely because of activity in the volatile datasets.

6. If there are several large, unrelated applications all on the same system, consider dividing the applications among two or more networked (new) machines, thus reducing the overall time that the applications are

unavailable. This approach is especially good if one of them "just has to be up", since the others will not be exposed by its uptime demands.

7. Remember that an old, slow, antique tape drive costs lots of money in downtime and in operator time. Use this fact to cost-justify a new drive. Your HP Sales Rep will be only too glad to help.

8. Use one of the third-party data-compression packages that are available. This can greatly shorten your backups. But before you buy, be sure to test your ability to recover your system with the resulting tapes, preferably by doing an actual full RELOAD. Do not buy unless you are happy, since it is not worth the worry. Find out what HP has to say about the package.

A few considerations about backup tapes:

1. Store your tapes offsite or in another building. Obviously do not keep them in the computer room, where they could go up in smoke along with the machine.

2. The best place to keep tapes is in a fire-proof vault with a lock, even though vaults are very costly and there never seems to be enough room in them.

3. Remember that SYSDUMP tapes contain the entire System Directory, and hence ALL your passwords. If you use BULDACCT, the passwords are even easier to find since they are contained in the JOBACCT, JOBACCTB, and JOBCUDC jobs. Keep the tapes in a secure place (like a vault).

4. Tapes do wear out. Replace them every three years or so.

5. Data stored on tapes can fade away magnetically after a period of years. Use TDTCOPY or TAPECOPY (in TELESUP) to create new copies.

6. The current "live" set of backup tapes (full and partials) should be kept in a definite, easily found place where they can be retrieved quickly. Do not mix them with other tapes. You may have to tell people where to find them -- over the telephone.

7. Use good-quality tapes for your backups.

8. Have a retention schedule for backup tapes.

Discs which contain lots of dead wood are obviously going to take plenty of backup time, especially over the long haul:

1. Do not unwittingly use your discs as an archive. Insist that inactive data be removed to TAPES. Get rid of seldom-referenced accounts: TELESUP and the CSL accounts are often in this class. If certain programs are used from these accounts, copy them to special groups in SYS.

2. Make sure that the application people are not creating huge TurboIMAGE databases which remain mostly empty. Sometimes dataset sizes are specified on the basis of what might happen years down the road. During the meantime YOU are going to have to backup all that empty space.

3. Do not assume that your users have an understanding of what disc space means. Watch out for people who create monster files without realizing what they are doing. Consider putting limits on sector usage.

4. REPORT will tell you who is eating up your disc space. If you do not know why, find out.

A few suggestions to keep you prepared for a recovery:

1. Retain the System Coldload Tape and the Diagnostic Utility System tape created during the most recent MPE upgrade, as instructed by Hewlett-Packard.

2. Using SYSDUMP, create a separate coldload tape with a dumpdate of 0 and a fileset of @.PUB.SYS.
   If it turns out that your backup tapes cannot revive MPE, this tape can be used for a coldload, and a great deal of time saved (not to mention nerves).
   It can also be used to restart the system after a system failure.
   Update this tape every time a configuration change is made, and keep the tape in the computer room.

3. Keep up-to-date copies of the system I/O configuration in your files and taped to the computer room wall. SYSINFO will provide a comprehensive listing of all system parameters, both for you and for your CE.

4. Keep a list of the Disc Volume Table in your files and taped to the computer console. Ditto for the LOAD, START and DUMP parameter settings for the system, even if default settings are used.

5. Occasionally practice loading your DUS tape, and keep up to speed on the use of SADUTIL. SADUTIL can be of considerable help during a recovery -- but that is not the time to figure it out from the manual.

6. Prepare a written recovery procedure based on RELOAD. Validate this procedure with a live test! Work the bugs out and know what you are going to do BEFORE the moment of truth arrives. Involve your operators, and remember that plenty of recoveries have had to be done over the phone.

7. Maintain control over your backup tape library, and know exactly where everything is.

8. Do not put too much stock in the war stories that people tell about recoveries. You only hear about the successes (just like the stock market). Do not be duped into believing that recoveries are easy.

It is my heartfelt wish that you will never have to recover from a flames-and-ashes cataclysm (or even from a "routine" system failure). Hopefully, these suggestions will help you to be ready for a quick and confident recovery, should it come. In the meantime, you will sleep better knowing that you have addressed some of the more dreadful possibilities in the life of the system manager.

wp/3410d