



hp NonStop TMF
software

tape handling for
disaster recovery

a technical brief
from hp

configuration alternatives to optimize disaster readiness and recovery

This technical brief examines disaster planning and disaster recovery options and alternatives using HP NonStop™ Transaction Management Facility (NonStop TMF) software. Its purpose is to help companies using NonStop servers better understand and use the NonStop TMF catalog and tape handling subsystems to recover from a catastrophic system failure (for example, fire in the computer room, earthquake, or flooding). This brief is also intended to elicit comments about the value of possible enhancements to NonStop TMF recovery functions.

NonStop TMF components and operation

NonStop TMF software consists of several components. Those discussed in this brief are the master audit trail (MAT), auxiliary audit trails (aux trails), and the tape handling subsystem, including the catalog, online dumps, and audit dumps.

All transaction boundaries (begintransaction, endtransaction) are recorded in the MAT. Depending on the NonStop TMF configuration, before and after images of database changes from one or more disk volumes can be recorded in the MAT or in one or more aux trails. It is the information in these audit trails that allows for database roll forward and transaction backout (reversal). The use of multiple audit trails provides parallelism and enables selective dumping of files and audit information (database changes) for those files.

NonStop TMF software can read and write two kinds of tapes: *online dumps* and *audit dumps*. An online dump is similar to a fileset backup, except that it can be run safely while the application is running. Using wildcards, an online dump can contain from one to all of the files on your system and can fully use and span multiple tapes, regardless of their size.

Audit dumps are offline backups of the NonStop TMF audit trail disk files. That is, audit trails are backed up only after they are closed. The maximum audit trail size is fixed at 2 gigabytes.¹

To recover a database file or disk volume, NonStop TMF software performs a roll forward. It restores files from the latest online dump(s) and succeeding audit trails. It then applies the after images in the audit trails to the base files to arrive at a consistent

¹ Changing the size of the audit trail also requires changing the format of the audit trail records. This has a major impact on other system software as well as third-party products. Thus, increasing the size must be evaluated carefully and should be considered a long-term enhancement..

point in time. Using the catalog, NonStop TMF software performs these steps automatically.

fast and slow, fresh and stale

Central to a discussion on the use of the NonStop TMF tapes for recovery are the terms *recovery time objective* (RTO) and *recovery point objective* (RPO). RTO, also called the *recovery window*, defines the delta time from when a disaster is declared until the business process—or in this case, the computer application—must be available again. RPO, also referred to as the *freshness window*, describes how much work in progress can be lost—or in this case, the point in time to which data must be recovered. Stale (old or obsolete) information no longer reflects the state of the company. A business impact analysis (BIA) is used to develop requirements for how stale the data for a recovered system can be and how much data can reasonably be lost when it is made available again (that is, how close in time the RPO is to the disaster incident).

The RTO and RPO are not coupled. For example, a database could be recreated from backup tapes in two days (RTO), but the backup tapes might be one week old (RPO). Would it make a difference if the tapes were only three days old and could be reloaded in six hours? Applications can be positioned, according to a company's recovery needs, into the four quadrants of long and short RPO and long and short RTO. The table shows a matrix of how some example applications might fit into the quadrants.

RTO/RPO matrix

	<i>RPO</i>	
<i>RTO</i>	<i>short</i>	<i>long (stale)</i>
short (fast)	stock exchange	ATM
long (slow)	funds transfer	payroll

The stock exchange application example is the most demanding in terms of RTO and RPO, requiring both short RTO and short RPO.

The ATM application is less demanding in terms of RPO, because ATMs keep their own hard-copy audit trails and the missing data can be recovered from them when the machines are balanced at night. However, if they are down for a long period, customers may gravitate to the competition.

The payroll application is the least demanding example. Many companies outsource their payroll to a third party. In that case, if a company's system goes down, the third party can cut checks that are the same as those for the previous period, and adjustments can be made when the system and data are restored. If a company does not outsource, it can use last week's payroll report to generate new checks.

The funds transfer application requires a short RPO, but not a short RTO. If a funds transfer system loses a transaction, the bank can be liable for billions of dollars. But

alternatives such as faxed or verbal transactions are available if it is down for a period of time.

If your application falls into the upper left quadrant of the matrix, consider using NonStop Remote Database Facility (NonStop RDF) software to replicate critical data and ensure uninterrupted service in the face of natural disasters. If your application falls into the lower left quadrant, consider configuring NonStop TMF and the tape library for the least loss of data—for example, by using remote tape drives. For applications that fall in the upper right quadrant, configure NonStop TMF software for the fastest possible recovery.

planning for recovery

backing up the configuration subvolume

The shorter your RPO (the fresher your data needs to be), the faster online dump and audit trail tapes need to be protected from disaster. To afford this protection, you need one additional tape: the NonStop TMF configuration subvolume. This subvolume stores the NonStop TMF configuration and media catalog, which must be available for system recovery. If this information is not available, you must recreate it on your backup system—a time-consuming, if not virtually impossible, task. To capture the necessary information, back up the NonStop TMF configuration subvolume (`<config-vol>.ztmfconf.*`) with the open option after every online dump is completed. Note that as of Release Version Update D30.02 of the NonStop Kernel operating system, the user can change the configuration volume to one other than `$$SYSTEM`. The default configuration volume is `$$SYSTEM`.

physical safety and accessibility

A key issue in disaster recovery planning is physical safety and accessibility. Backup hardware should be located as far as possible from the computing hardware it is protecting. At a minimum, the hardware should be separated by fireproof walls and automatic fireproof doors.

NonStop TMF software creates tapes, which must be protected from physical damage; otherwise, the tapes cannot be used for recovery. Many companies use tape silos, some of which are fireproof. However, external events can affect the tapes' accessibility. For example, suppose there is a fire in your facility and the tapes are stored in a fireproof tape silo in the computer room. Your computers will be molten pools of metal, but the tapes will be safe because the time and/or temperature limit was not reached during the fire. You should be able to pull the tapes from the silo, move them to a backup site with a mirror of your primary system, and use them to recover.² But, if the fire was of a suspicious origin or the authorities don't think the room is safe, your tapes could become inaccessible. And they could remain in this state for hours, days, or even months.

With a short RPO, tapes need to go offsite quickly so that they cannot be destroyed or rendered inaccessible. Appending audit dumps requires that tapes stay onsite until they are full, which makes them more vulnerable to inaccessibility.

² Currently, NonStop TMF software can recover files only to a system and a disk volume that are identical to the original system and disk volume from which the files were dumped. NonStop Storage Management Foundation (SMF) software can be used to create virtual disk volumes with matching names on the backup system.

Depending on the results of the BIA, tapes should be removed from the computer room or tape backup site and sent offsite as soon as possible. Of course, the ultimate in minimum RPO involves a product like NonStop RDF software, which ensures that data is copied from the system as soon as it is written.

disaster recovery using NonStop TMF software

This section discusses how recovery of a damaged database with NonStop TMF software might work and how tape is used for audit dumps. See Support Note S96048B for detailed instructions on performing disaster recovery using NonStop TMF software. This note is available from Total Information Manager. Select Support (NonStop Kernel operating system), Date: General, Support Notes, For 2000. Or contact your HP representative.

the power of parallelism

In the event of an unplanned outage, NonStop TMF software uses high-performance, parallel operations to recover lost files or tables. Because this software runs operations in parallel, it can recover a disk or data volume quickly. For example, it can use all tape drives connected to a node, concurrently and in parallel, to restore the node's database.

The shorter the RTO, the more parallelism is required for recovering the database. This means that several proper subsets of database files should be dumped using multiple online dump threads to create parallel restore sets. In addition, more frequent online dumps need to be taken, so that fewer audit dumps are required once the online dump restores are complete. The shorter the RPO, the more often audit trails need to be rolled and dumped. Obviously, the risk of collision between RTO and RPO increases as more and more audit dumps need to be applied.

To achieve a short RTO during recovery, the operator issues multiple instances of the TMFCOM RECOVER FILES command and uses multiple tape drives to process the roll forward threads. Audit trails are needed to roll a database file forward from the last online dump. Accordingly, NonStop TMF software requests that the tape with the corresponding audit trail be mounted, whereupon the trail is copied to disk. Assuming sufficient disk space, audit trails are held for about 5 minutes from the time they are closed in case they are needed again.

appending audit dumps on tape

One concern expressed by some companies is that the NonStop TMF catalog does not allow for appending online dumps or audit dumps. That is, once a tape has been written, it cannot be updated with additional files. Online dump sets can be organized to take advantage of the largest tape sizes available. However, because the maximum audit trail is 2 gigabytes, the maximum data that can be stored on an audit dump tape is also 2 gigabytes, regardless of the size of the tape.

Some risks are associated with appending audit dumps on tape. For example, if an audit dump is written to a randomly chosen tape, one of the roll forward threads could be blocked during recovery because another thread is restoring a different audit trail on the same tape. If the RTO is short, audit trails should not be appended on tapes randomly because of the possibility of blocking. In the worst case, if a tape with

multiple audit dumps is lost or damaged, then recovery can become impossible—even with parallel dumps enabled. Also, a tape may not be released because the last dump on the tape is still required, even though the previous 25 dumps are not.

Appending audit dumps makes tape management much more complicated and still doesn't guarantee that roll forward threads won't block one another. More investigation and failure simulations are needed to establish confidence that appending audit dumps is safe for recovery from a disaster.

what are the alternatives?

Appending audit dumps can be a risky proposition. So, what are the alternatives? One alternative may be to use intelligent appending so that every (or every other) audit dump from the same master or aux trail is stored on the same tape. This would allow serial restoration of audit trails, using multiple drives for parallelism. However, building the heuristics into the catalog to support intelligent appending would be a complicated process.

You should audit your use of NonStop TMF software. Have you evaluated your application to assess selected dumping of aux trails? You may have database transactions that must be consistent but do not need to be recovered. If this is the case, you can segregate these files onto disks logging to an aux trail and not dump the aux trail.

Some companies dump audit trails to disk and then run independent backups to copy several audit trails to the same tape. Although this is suitable for recovering damaged files onto the same system, it may not be optimal for disaster recovery. If the system is destroyed, so are the audit trails needed for recovery. Additionally, operators are now responsible for keeping track of which trails need to be restored to disk and in what order, and for inserting audit trails back into the NonStop TMF catalog using the ADD DUMPS command or the SNOOP utility. Mistakes are easy to make in a disaster situation.

using NonStop TMF software with an offline third-party hot site

This method of disaster recovery is especially good if you have a hot site contract with a third-party vendor that has agreed to let you configure NonStop TMF software to dump to disk on one of their systems (which can have multiple other primary systems dumping to it as well). When you declare a disaster at your primary site, the NonStop TMF configuration subvolume from the primary site is copied to another remote system, which has been renamed the same as your primary system. Following is an overview of how this would work.

In this example, the original system, A, is called `\SOURCE`. The system used for remote dump to disk, B, is called `\DUMP`. And the system used for system recovery if system A is lost is system C. When failing over to system C, system C must first be renamed as `\SOURCE`. The NonStop TMF catalog from system A is then overlaid on system C. Because the dump entries in the catalog already point to system B (`\DUMP`), nothing needs to be done except to add any missing dumps to the catalog and issue a Recover Files command to NonStop TMF software.

where do we go from here?

HP product management seeks feedback on which features are valuable to you for disaster recovery of systems using NonStop TMF software.

- Does this paper change your views on the use of NonStop TMF software for disaster recovery?
- If one enhancement were to be made to this software's tape handling, what would you want it to be?
- Do you feel that the risks involved in appending audit dumps are acceptable and that this option should be added to NonStop TMF software?
- Should audit trail and audit dump tapes be placed into separate tape pools? This could allow smaller and perhaps geographically separated tapes to be used for audit dumps.
- Should the audit trail size be increased?

Send your comments and suggestions to Ron LaPedis, NonStop TMF software product manager, at ron.lapedis@hp.com.

For more information, go to www.hp.com/go/nonstop.

July 2002, first published 2001. All product names mentioned herein may be trademarks of their respective companies. HP shall not be liable for technical or editorial errors or omissions contained herein. The information is subject to change without notice. The warranties for HP products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

02-0430

©2002 Hewlett-Packard Company

