

Increasing your Data Security with Referenced Backup technology

What is a Referenced Backup and what are its advantages compared to a traditional backup?

Content

1.	Problems Of Conventional Backup Procedures	2
1.1	Renaming Files	2
1.2	Moving Files Within A File System	2
1.3	Deleted Files	2
1.4	Interrupted Backups	2
1.5	Backing Up Large Volumes Of Data	2
1.6	Restoring Data Saved By A Conventional Backup	2
3.	The Solution - 'Index Referenced Backup'	3
4.	Application Examples	3
5.	Operational Considerations	4
6.	Synthetic Full Backup	4

1. Problems Of Conventional Backup Procedures

Normally incremental backups work purely by adding data that has changed since the last backup. The decision about which data to save is made by scanning a file system and comparing the creation and modification dates with those from the last backup. If the dates of the files are newer, the file will be backed up and an entry in the index will be added.

Most backup systems employ this additive method - despite the fact that it has some intrinsic problems. It is therefore no longer appropriate in modern systems for the following reasons.

1.1 Renaming Files

If a file is renamed, this has no effect on its modification time. An incremental backup employing the additive method will therefore not be aware of this change and will not save the new file during an incremental backup. The new name of the file remains effectively unknown to the backup.

1.2 Moving Files Within A File System

Files moved within a file system will not be detected by traditional backup software since, again, their modification date is not affected by the move. Therefore the files will still be found at their original location within the index and hence restored to their original location.

1.3 Deleted Files

Again, file deletions are not reflected in the backup index. Although they are no longer a part of the file system, they are still included in the index at the place where they once were and hence restored to their original location.

1.4 Interrupted Backups

A conventional backup procedure considers an incomplete backup to be null and void. Consequently, not all data saved up to the point of termination, although written to storage media, will be available to be restored. In any case, the next backup will have to save the same data again, since the backup index contains no information about the failed backup.

1.5 Backing Up Large Volumes Of Data

Experience shows that backup windows are getting smaller since production takes more time and produces exponentially larger amounts of data. The conventional backup procedure, which considers interrupted backups as incomplete and broken, is consequently hitting its limits. It is becoming increasingly difficult to save large volumes of data by means of a conventional backup within the decreasing backup window.

1.6 Restoring Data Saved By A Conventional Backup

The issues with the conventional backup as described above, have further

consequences on the restore procedure.

Renamed files can only be restored into their old locations and with their old names, since their new names and locations are not known to the backup system.

Furthermore, a user will not be able to search for and locate a file within the backup index, unless its previous name and/or location are known.

A restore of a moved file will only be performed into its old location, since the backup system is not aware of where it had been moved. When performing a full restore, files which have been deliberately removed, will be restored regardless, since there is no information available to the backup system about deleted files.

3. The Solution – 'Index Referenced Backup'

By implementing an 'Index Referenced Backup', PresSTORE 2 solves the problems and addresses the weaknesses of conventional backup technologies.

The incremental backup employs the available backup index which contains a detailed 'snapshot' of the state of the filesystem last time a backup was performed. This is used as a reference to determine the exact changes that have taken place in the file system since the last backup and therefore which files need to be included in this backup. This method additionally enables detecting renamed, moved and deleted files. When restoring files, these will be correctly named and located in the filesystem according to their original location. Deleted files are marked as such in the index and will not be restored, unless restoring from a date prior to their deletion.

By employing the more sophisticated index, PresSTORE 2 realises a true 'adaptive backup', which registers all the changes and adapts to the actual state of the file system over time.

4. Application Examples

The new 'Referenced Backup' in PresSTORE 2 maintains an accurate and complete footprint of the file system and is the basis of many new features and options within the product. For example, it is now possible to backup laptop computers although they are connected to the network for relatively short periods at a time. The moment the laptops appear on the network, the backup commences. The first time around, a full backup will be taken, since there are no records in the index at this point. Thanks to our Referenced Backup, the laptop owner does not have to wait for the backup to complete, but can at any time disconnect his laptop from the network.

Any conventional backup software, would consider the first full backup as 'broken' and reinitiate a new full backup the next time around when the

computer next appears on the network. Laptops, which are rarely attached to the network long enough for a full backup, would thus almost never be backed up in their entirety using tradition methods and therefore go without being backed up at all!

Thanks to its index, even when only partly filled, PresSTORE 2 can determine which files have changed and therefore need to be backed up anew. Any files not yet saved or changed since the last backup will be detected and backed up.

PresSTORE 2 implements a fully interruptible backup. For PresSTORE 2, no backup is broken. It will pick up where it left off and does not have to restart from the beginning.

This opens new possibilities where large data sets are involved. A backup, which would take longer than the allotted time window and interfere with the daily production can be interrupted and continued on the following evening. Backups with intentionally introduced pauses are possible without impacting the safety of the data. Furthermore, a restore of a snapshot of the file system can be done at any time with all the renaming, moving and delete operations that took place on the file system considered.

5. Operational Considerations

Practically speaking, it becomes sufficient to initially execute a full backup and then to continue regularly running incremental backups without impacting the safety of the data. It is, still recommended, nevertheless, to execute a full backup at regular intervals. A full backup creates a new clean and lean index, which releases disk space and boosts operational performance. Without a full backup, the index would keep growing and the restore of an entire file system would take an unnecessarily lengthy period. Additionally, a full backup will occupy new media, thus making the old media and all the media used to store increments available for recycling.

A full backup in PresSTORE 2 can be done with help of the index. This 'Synthetic Full Backup' enables a speedy backup without penalising server and network performance – see below.

6. Synthetic Full Backup

PresSTORE 2 includes a unique new feature, which enables the creation of a full backup at any point in time by only employing the current index and using the data available on the media. The 'synthetic backup' does not access the servers and their filesystems; but uses the data available on the media as the source for the backup. From this data and the backup index, PresSTORE creates a new full backup on a fresh set of media without imposing load on the network and impacting resources and the routine production on the server. The fully automatic media manager identifies, and exchanges the required media to read and write, provided the necessary media is available in the tape-library.

For companies where the data volume exceeds the available backup window, Synthetic Full Backup is an ideal and speedy solution to obtain a new full backup without producing extra load on the network and computers. Especially during time-critical periods of intensive production, a high degree of security for the data can be provided.

ARCHIWARE® PresSTORE®

PresSTORE® Synchronize • Backup • Archive

The Modular Data-Management Platform