Categories of Backup Software

Categories of Backup Software

There are hundreds of backup software applications out there for Windows and Linux. We try to keep track of a lot of them at R1Soft and even still we are surprised to discover new ones. It's nice to know that ALL Linux and Windows backup software fits into three major categories.

1. Legacy Backup Applications

Legacy Backup Application Performance	
Backup Method Used	Full / Incremental / Differential
Data Lost In Disaster	Days
Method for Computing Deltas	File Attributes and Check Sums
Backup Window Length	Hours or Days
Online Backup (Snapshots)	No (possible with third party application such as OTM)
Bare-Metal Restore	Not Possible
Backup Window Critical Path	= Time to Read All Data from Disks

2. Online Backup Applications

Online Backup Application Performance	
Backup Method Used	Full / Incremental
Data Lost In Disaster	Days
Method for Computing Deltas	Check Sums
Backup Window Length	Hours or Days
Online Backup (Snapshots)	Yes Must have to be Online
Bare-Metal Restore	Possible

Backup Window Critical Path	= Time to Read All Data from Disks to compute deltas

3. CDP Backup Applications

R1Soft CDP	
Backup Method Used	Virtual Full
Data Lost In Disaster	Minutes
Method for Computing Deltas	near-Continuous
Backup Window Length	Minutes when CDP is in Sync
Online Backup (Snapshots)	Yes Integral Part of the Product
Bare-Metal Restore	Yes
Backup Window Critical Path	= Only Time to Read Deltas

Backup Software Performance Attributes

Backup application performance can be characterized by the following criteria: backup method, Amount of data Lost in a Disaster, Method used for Computing Deltas, Backup Window Length, Online Backup, Bare-Metal Restore, backup window critical path, and file or block based backups.

Backup Method

- Backup Method - Full Backup
- Backup Method - Incremental Backup
- Backup Method - Differential
- Backup Method - Virtual Full Backup

Amount of data Lost in a Disaster

The amount of data potentially lost in a disaster is directly related to the frequency of the backup schedule and the backup window length. Backups that take a long time (backup window) and have a severely negative impact on server performance mean the backup can not be performed very frequently. Most applications can only be run on a daily basis. For some environments this

is actually weekly depending on performance sensitivity and the size of the data set. This is categorized by a backup that can only be scheduled one or more days apart and a backup that can be scheduled minutes apart.

The Method used for Computing Deltas

Deltas are simply defined as the data that has changed since the last backup run.

The three known methods for computing Deltas are:

1. Computing Deltas - File Attributes
2. Computing Deltas - Check Sums
3. Computing Deltas - near-Continuous (CDP)

Backup Window Length

The time needed to perform a backup operation in the typical case. Most backup applications were designed when a 100 GB drive was unheard of and the Internet was considered a fad. Most backup applications assume businesses shutdown every night and make servers unavailable to users so backups can be performed.

Online Backup (Snapshots, Open File Backups)

The ability for the backup application to take a point-in-time picture image of a live running disk volume on a server. This is also called Open File Backups. Without point-in-time snapshots the backup is completely inconsistent. Imagine a full server backup taking place at night while a web site administrator is uploading the latest version of a massive web site. The backup applications is likely to get a copy of some of his uploaded files and at the same time copy older versions of files that were overwritten while the backup was taking place. Worse yet there is usually nothing preventing the backup application from reading files while they are being written to and this can cause files to easily be corrupted.

In a large multi-user system which applies to most Internet applications this is a recipe for almost guaranteed corruption. Worse yet the backup application will pass verification as it is completely unaware of the application level changes to files and the inconsistent state of files will not be noticed until after they are restored which is usually too late.

See: Block Based Backup Technology, File Based Backup Technology

Bare-Metal Restore

Bare-Metal restore is the ability for the backup application to restore an entire server to a previous state directly from backup media. To be truly a Bare-Metal a backup application must restore a block level disk image directly from backup media without requiring that the O/S or

backup application be installed first. Any easy way to identify if a backup application supports a real bare-metal restore process is to determine if it reads data at the file level or at the file system or raw disk block level. Usually this can be determined by reading the feature list or documentation without ever even attempting a test bare-metal restore with real hardware and the vendor's application.

See Bare-Metal Restore for more details.

Backup Window Critical Path

The backup window critical path is the step in the backup process that takes the longest amount of time in the average case of a server backup. For example a backup may be sending data to a very slow tape driver or across an extremely slow network so that no matter what level of performance the backup application uses the slow network or slow tape drive may always be the biggest limiting factor. This is not the average case. The average case considered for this document is a server with a large number of files (200,000 or more) on a very fast network (e.g. fast Ethernet or Gigabit Ethernet) with disks that are high performance and a data set that is moderately large (e.g. 100 GB or more).
The backup critical path for most backup applications is almost always even in the best case the time needed to compute Deltas. This means the time it takes the backup application to figure out what changed since the last backup is almost always the single biggest limiting factor. And in most cases this is always equal to the time it takes the backup application to read all data in the files. The exception is R1Soft's near-Continuous method which does little more than read changed data.

File or Block Based Backups

Backup applications can read data from the disk in one of two ways using files or blocks.

- Block Based Backup Technology
- File Based Backup Technology