

Rule Based Backup Sets

Introduction

To determine what files to back up, Clients (meaning end computers) are assigned backup sets. A Client can have multiple backup sets, which allows you to assign several application-specific sets to a single user. Each Backup Set contains all the information necessary to decide when and what to backup as well as how long to keep copies of that particular data on the RBA. Specifically, we break a backup set into a Rule Set, a Schedule, and a Retention Policy.

Rule Sets determine which files are included in a given backup set and are a collection of individual rules. Each rule includes a root directory, an indication of whether or not to include subdirectories, and a filter that will cause the rule to match certain files. Each rule can either include or exclude files that match the rule. *Excludes rules that match on a file always supersede (trump) include rules.*

Schedules determine when a backup set will be run. This includes typical options like once, daily (or a frequency like every three days), weekly (on certain days of the week), or monthly (on a certain day of the month). If you want a more exotic scheduling policy, we offer an advanced “cron” scheduling ability or you can create another backup set using the same Rule Set and assign it a different schedule. Due to certain issues handling databases and other large, frequently changing files, we do not currently support CDP (Continuous Data Protection). However, our architecture is designed to fully support CDP and this feature will be added based on customer demand.

Retention Policies determine how long a particular version of a backup set is stored on the RBA. Versions can either be retained based on a certain number of versions or based on age. While each backup set can have only one retention policy, we have designed the time-based retention rules to allow for several overlapping rules. For example, you could keep one version a day over the last week (for restore purposes) plus one version a month for three years (for archiving purposes).

Each backup set is completely separate and exclude rules that have been added to one backup set do not impact any other backup sets assigned to that client. Each time a backup set runs, the appliance acquires a snapshot of all files that match that particular backup set. If a file is included in multiple backup sets, each set will create a new point-in-time snapshot for the file and those different snapshots will be kept based on the retention policy of the backup set that created the snapshot.

Defining Rules – Includes and Excludes

Please note that the “/” separator is used when describing “path matches” below. Either “/” or “\” may be used to mean the same thing on the system when creating filters.

3X Backup has four five types

- All Files
- By file type
- By file name
- Single file
- Directory name match (manager only)

Because directory name match is more like a regular expression, it will be discussed, in depth, in its own section.

All Files

To include/exclude all files, select the root directory in the right or left pane of the file navigator and decide whether or not to include subdirectories. Ensure that the “all files” rule type is selected. You will not need to enter any additional information in the rule match box and can select to either include or exclude this rule.

Including subdirectories will include all files in any subdirectory of any depth in the rule.

Given a directory structure:

```
C:\file0.xxx
C:\Directory1\file1.xxx
C:\Directory2\file2.xxx
C:\Directory1\Subdirectory1\file3.xxx
C:\Directory1\Subdirectory2\file4.xxx
C:\Directory1\Subdirectory1\SubSubDirectory1\file5.xxx
C:\Directory1\Subdirectory1\SubSubDirectory2\file6.xxx
```

If you include all files in C:\Directory1 and **do not** include subfolders (bold indicates reason for failure):

- × C:\file0.xxx
- ✓ C:\Directory1\file1.xxx
- ✓ C:\Directory1\file2.xxx
- × C:\Directory1**Subdirectory1**\file3.xxx
- × C:\Directory1**Subdirectory2**\file4.xxx
- × C:\Directory1**Subdirectory1**\SubSubDirectory1\file5.xxx
- × C:\Directory1**Subdirectory1**\SubSubDirectory2\file6.xxx
- × C:**Directory2**\file7.xxx

If you include all files in C:\Directory1 and include subfolders (bold indicates reason for failure):

- ✗ C:\file0.xxx
- ✓ C:\Directory1\file1.xxx
- ✓ C:\Directory1\file2.xxx
- ✓ C:\Directory1\Subdirectory1\file3.xxx
- ✓ C:\Directory1\Subdirectory2\file4.xxx
- ✓ C:\Directory1\Subdirectory1\SubSubDirectory1\file5.xxx
- ✓ C:\Directory1\Subdirectory1\SubSubDirectory2\file6.xxx
- ✗ C:**Directory2**\file7.xxx

Example in Practice:

Exclude the Windows directory from the backup

Rule Type	Exclude
Root Directory	C:\windows (and all subdirectories)
Rule	Exclude All Files

Include a Shared Directory located on a Server

Rule Type	Include
Root Directory	D:\shared (and all subdirectories)
Rule	Include All Files

File by File Type

To select files based on a particular file extension, use the “File Type(s)” option. Like the All Files rule, you must select the root directory in the right or left pane of the file navigator and decide whether or not to include subdirectories. Once you have ensured that “File Types(s)” is selected, enter one or more file types into the box. To match multiple file types place a space between each file type.

Note: It is not necessary to enter a dot or asterisk in the rule.

Example in Practice:

Include all Word and Excel files located in any directory

Rule Type	Include
Root Directory	C:\ (and all subdirectories)
Rule	File Type(s) “doc docx xls xlsx”

Exclude all MP3s from the backup set

Rule Type	Exclude
Root Directory	C:\ (and all subdirectories)
Rule	File Type(s) “mp3”

File Name

To select files based on their file name, use the “File Name” option. Like the All Files rule, you must select the root directory in the right or left pane of the file navigator and decide whether or not to include subdirectories. Once you have ensured that “File Name” is selected, you’ll be presented with several options for file name matches: equal, dont equal, contain, dont contain, start with, dont start with, end with, and dont end with. Once you have selected the correct match criteria, enter the desired file name (or part thereof) into the next box. This box **does not** accept wildcard characters.

Example in Practice:

Exclude Microsoft Office temporary files (files starting with the tilde, “~”)

Rule Type	Exclude
Root Directory	C:\ (and all subdirectories)
Rule	Exclude file name start with “~”

Single File

To include a single file, navigate to the directory containing the file and select it in the right hand window. Once the file is select, you need only indicate whether to include or exclude the file.

Note: If the file has been excluded by any other rule, including the file through an individual rule **will not** supersede the exclusion rule. To backup this file, create another backup set that includes the individual file (see the multiple backup sets section).

Example:

Backup Set 1	
Rule1	Exclude all mp3 files in c:\ and all subdirectories
Rule2	Include File c:\Important Files\File.mp3

The file **will not be included** in the backup set.

Directory Name Match

Directory name match is closer to a regular expression

While you can think of a directory match matching on all the files in a particular directory, directory name match actually matches on the path for each file. The rule automatically checks all subdirectories for files where the paths that match the rule. This means that a rule could match a particular path, but not its parents or children.

Matching a Path

? matches any single character in a directory name

* matches any number of characters in a single directory

So...

?	Matches any directory with a one character name
??	Matches any directory with a two character name
*	Matches any directory of any name
**	NOTE: This has a special meaning see the next section.
<text>	Matches any directory with the exact name <text>
*<text>	Matches any directory ending with <text>
<text>*	Matches any directory starting with <text>
<text>	Matches any directory that includes <text> in the name

From this point on, we'll assume that <match> implies a directory whose name matches with our single directory rule.

Matching a Directory Tree

A path <match> or * alone will require that a directory be present and can be used to match on a certain directory depth

** matches a subdirectories of any depth (including zero or “present in the root”)

Specifically,

<match>	Includes all files in any folder in the root directory where the folder name matches <match>
*/<match>	Includes all files in any subfolder of a root folder where the folder name matches <match>
**/*/<match>	Includes all files in a second depth subfolder of the root folder where the folder name matches <match>

**/<match> Includes all files in any folder (of any depth including depth zero or “located in the root directory”) where the folder name matches <match>

*/**/<match> Includes all files in any folder (of any depth one or more) where the folder name matches <match>

Since our rule ends in <match>, notice that a directory ending in Folder / file.xxx is *never* included because the filename path cannot match. To include these folders, we can place the wildcards after the <match>:

Recall that:

<match> Includes all files in any folder (where the folder is in the root directory) where the folder name matches <match>

Adding additional directories to the end will also match subdirectories of the <match> root:

<match>/* Includes any subfolder (only the next depth) of root/<match>

<match>/**/* Includes any subfolder (only the next depth) of root/<match>

<match>/**/* Includes any file in any subfolder (of any depth including depth zero) of root/<match>

When we compare these different rules to a sample directory structure shown on the left, we can see how these various rules impact our backup policy:

	<match>	*/<match>	*/**/<match>	**/<match>	*/**/<match>	<match>	<match>/*	<match>/**/*	<match>/**/*	
<match> / file.xxx	✓	x	x	x	✓	x	✓	x	x	✓
<match> / Folder / file.xxx	x	x	x	x	x	x	x	✓	x	✓
<match> / <match> / file.xxx	x	✓	x	x	✓	✓	x	✓	x	✓
<match> / <match> / Folder / file.xxx	x	x	x	x	x	x	x	x	✓	✓
<match> / <match> / <match> / file.xxx	x	x	✓	✓	✓	✓	x	x	✓	✓
Folder / file.xxx	x	x	x	x	x	x	x	x	x	x
Folder / <match> / file.xxx	x	✓	x	x	✓	✓	x	x	x	x
Folder / <match> / Folder / file.xxx	x	x	x	x	x	x	x	x	x	x
Folder / <match> / <match> / file.xxx	x	x	✓	✓	✓	✓	x	x	x	x

Directory Match Rules in Practice

Matching My Documents and all subfolders (in XP):

This data is typically included in a directory like c:\Documents and Settings\\My Documents and its subfolders.

Rule Type: include
Root Directory: typically the root directory on the hard drive like c:\
Rule: Documents and Settings/*/My Documents/**

Matching Desktop and all subfolders (in XP):

Rule Type: include
Root Directory: typically the root directory on the hard drive
Rule: Documents and Settings/*/Desktop/**
Note: If you fail to include ** at the end, the rule will not backup any directories that have been placed on the desktop.

Excluding My Pictures Folder and all subfolders (in XP)

Rule Type: exclude
Root Directory: typically the root directory on the hard drive like c:\
Rule: Documents and Settings/*/My Documents/My Pictures/**
Note: If you fail to include ** at the end of the rule, it will only exclude files in the root directory of My Pictures. Files in a subdirectory like “My Pictures/family” will not be excluded!

Multiple and Overlapping Backup Sets

Each backup set is defined and run as a completely independent unit, while our proprietary deduplication technology ensures that files included in multiple backup sets are not transmitted or stored multiple times.

Example 1: Applying Different Schedules or Retention Policies to Each Backup Set

Separating a backup policy into multiple backup sets can substantially simplify the administration of backup by separating backup policies into logical units with different schedules and retention policies. For example, end users often benefit from keeping daily or weekly backups of their My Documents folder for a month or more, but probably not a year. By contrast, a business critical database may need a few days of backup (to defend against damage), but might also benefit from or require a periodic archive for a year or more.

Example:

Backup Set	My Documents and Desktop
Schedule	Run Daily
Retention	Keep 1 per day for a week plus 1 weekly backup for a month

Backup Set	Email
Schedule	Run Daily
Retention	Keep 1 per day for a week plus 1 backup a month for 2 years

Example 2: Simplify Backup Set Contents

This approach also simplifies backup set administration by limiting each set to simple, cognitive units. Even if you use the same retention policy and schedule, you can separate “My Documents”, email, and application specific backups into their own backup sets. Not only does this avoid long, complex backup policies, but it facilitates easy reporting and analytics of the impact of different data types and sources.

Backup Set	My Documents (excluding a data directory)
Schedule	Run Daily
Retention	Keep 1 per day for a week plus 1 weekly backup for a month
Analytics	Retaining many versions does not significant increase storage space

Backup Set	My Documents Data Directory
Schedule	Run Daily
Retention	Keep 1 per day for a week plus 1 weekly backup for a month
Analytics	Retaining many versions substantially increases storage space

Backup Set	Local User Access Database
Schedule	Run Daily

Retention Keep 1 per day for a week plus 1 weekly backup for a month
Analytics Retaining many versions does not significantly increase storage space

Backup Set Email
Schedule Run Daily
Retention Keep 1 per day for a week plus 1 weekly backup for a month
Analytics Retaining many versions mildly increases storage space

Backup Set *need other kinds of data to analyze*
Schedule Run Daily
Retention Keep 1 per day for a week plus 1 weekly backup for a month

Based on this information, you might choose to alter the retention policies for different data types.

Example 3: Protecting Files Included in Global Excludes

Having multiple backup sets also allows an administrator to be more liberal with exclude policies without missing data.

Backup Set C:\ (and all subdirectories) excluding all MP3s
Schedule Run Daily
Retention Keep 1 per day for a week plus 1 weekly backup for a month

Backup Set My Documents\Critical Corporate MP3 Directory and all subdirectories
Schedule Run Daily
Retention Keep 1 per day for a week plus 1 weekly backup for a month

Excluding MP3s in the broad backup ensures that you are not protecting user's music files, while critical corporate audio files are included in a separate set. Even if other files are placed in the Corporate Audio directory, our deduplication technology ensures that you will not need to transmit or store this file a second time (unless it has changed between backup sets).