

File Specification Filter Exclude / Include Files & Folders

1. Overview

By default, During Distribution and Replication processes, R-1 will synchronize the entire directory structure (Recursively from the source Path) and all of its contents.

A given distribution process may require the exclusion of certain files or directories, or specifying a specific file types to transfer. The most common scenario includes temporary and log files that are system specific and should not be distributed to the target nodes. R-1 enables the filtering of files and/or directories from the distribution process, by defining exclusion of files and/or directories, as well as specifying which files/folders from the source tree should be transferred.

The user can specify both files to exclude as well as file specification – the combination of this filter specification will result in the list of files that will be transferred in the distribution process.

To ensure that the filter specification defined will perform as expected, it is advised to employ the **Manually confirm comparative snapshot** feature for the first distribution.

NOTE: In a Distribution Job, the filter is applied for all Edges. Files can be excluded from distribution to all Edge(s); they cannot be excluded from distribution to some of the Edge(s) within a distribution job. The same applies for File Specification (include).

Target Directories

The content of the Target directory after the replication process is over depends on more than just the filter specification. It is affected by the replication logic, its content before the replication, and other job properties.

In Distribution jobs, the different target directories on the different Edges may look different after the distribution job is complete, depending on their tree before the replication. But, the excluded files are excluded from the complete job, and these files will not be transferred to any of the target Edge(s).

As this document's focus is on the filter, it ignores other factors while describing the target directory's content, and concentrates more on which files and directories are transferred during the replication process, or excluded from it.

Table of Content

1. Overview.....	1
2. Specifications.....	3
2.1. Definitions.....	4
2.2. Examples.....	7
3. GUI.....	10
4. CLI.....	11
5. Conversions.....	13
6. Automatic Excludes.....	15
Appendix A Advanced Examples.....	16

2. Specifications

There are two types of filters: Include and Exclude.

You can specify multiple Include and Exclude filters within the same distribution job.

Each filter that you define consists of two parts – the **non-recursive** part, and the **recursive** part.

- **Non-recursive** - The “non-recursive” part can be thought of as a “base” specification. This is used to define a starting point for that particular include or exclude filter.
- **Recursive** - The recursive part can be thought of as a qualifier to be applied to that “base”, further controlling what is included or excluded within that base path.

For each file specification, specify whether the specification applies to files, directories, or both files and directories:

- **Files** – Replicate only files matching the specification
 - **Non-recursive** – Only files on the level specified will be replicated
 - **Recursive** – All files in the tree matching the specification will be replicated. Only sub-directories containing those files will be created.
- **Directories** – Replicate only folders matching the specification
 - **Non-recursive** – Only folders on the level specified will be replicated
 - **Recursive** – All folders in the tree matching the specification will be replicated. Only sub-directories containing those directories will be created.
- **All** – Replicate all files and folders matching the specification
 - **Non-recursive** – Only files and folders on the level specified will be replicated
 - **Recursive** – All files and folders in the tree matching the specification will be replicated. Only sub-directories containing those files and directories will be created.

NOTE: As all file specifications are relative to the source directory, when specifying folder names, do not put a \ (for Windows) or a / (for UNIX) before the folder name.

2.1. Definitions

Important notes about recursive and non-recursive include and exclude, files and directories specifications:

- **Wildcards** - The filter supports two wildcards:
 - * – Sequence of one or more characters, except of the dir separator,
 - ? – One character, except of dir separator.
- **Recursive** – Exclude / Include according to the Directory or File Specification from anywhere in the Source replication tree. Any given filter can be marked for recursive exclusion / inclusion.
- **Recursive part of specification** – Specifications can be recursive, non-recursive or start with a non-recursive part and continue with a recursive part.

Each file/directory specification for either exclude or include has two parts: non-recursive and recursive. The non-recursive part is always the first, followed by the recursive part.

- To specify a recursive specification, specify only the recursive part.
- To specify a non-recursive specification, specify only the non-recursive part.
- To specify a complex specification, fill in both the non-recursive and recursive parts.

- **Directory Specification** -
 - Folder names are always relative to the Job’s Source directory.
 - **Folder names** specified in the Exclude / Include should always end with a Directory Separator (“/” in UNIX, “\” in Windows) to indicate it is a folder:
 - foldername\sub-folder\
foldername/sub-folder/
 - Using the **GUI**, mark the specification as **Directories**.
 - Directory Specification may include wildcards and multiple sub-directories:

Non-recursive part	Recursive part
foldername\ foldername\subfolder1\subfolder2\ folder?\ foldername*\ foldername*\	sub*\ subfolder\ subfolder\

- To exclude / include first level folders from the source just enter in the field:

Non-recursive part	Recursive part
foldername\	

- To exclude / include sub-directories put the entire path relative to the source in the field:

Non-recursive part	Recursive part
foldername\subfolder\	

- Excluding a directory means, “Exclude the directory and all files and folders within that directory”.
- Including a directory means, “Include the directory and all files and folders within that directory, and not other files and directories”.
- **Recursive and Non-recursive include directory specification** – will include all directories that match the filter criteria.
- **Non-recursive exclude directory specification** – files and directories in all levels under the directory specification are excluded.

- **File Specification -**

NOTE: Using * as the exclude specification will result in all files in the source directory being excluded from the replication.

- File Specification may include wildcards and multiple sub-directories:

Non-recursive part	Recursive part
filename.txt	
	filename.txt
foldername\sub_folder\file.*	
folder*/sub*/	f?.*
folder/	sub/*

- Using the **GUI**, mark the specification as **Files**.
- **Exclude file specification** – Using File Specification no directories are excluded. The whole directory tree starting from the specified path is created on the target.

All files in the path matching the File Specification are excluded.

- **Recursive exclude file specification** – All files matching the File Specification in the Path specified and in all sub-directories **under** it, are excluded. No directories are excluded; the whole path is created on target.
- **Include file specification** – if the file specification is in a sub-directory, the complete path to the files in the File Specification is created on the target (i.e. all the sub-directories in the path are also included).

All files in the path matching the File Specification are included.

NOTE: Folders not in the path are not created on the target.

- **Recursive include file specification** – All files matching the File Specification in the Path specified and in all sub-directories **under** it, are included, except empty folders.

2.2. Examples

The following samples illustrate various exclusion scenarios.

For more advanced and detailed examples please refer to [Appendix A](#).

Source Directory Tree for all samples:

- o C:\Master_data\
 - o Server.log
 - o Server.tmp
 - o Data\
 - a.tmp
 - b.tmp
 - a.doc
 - summary\
 - logs\
 - data\
 - o 2002\
 - 2002.tmp
 - logs\
 - a.tmp
 - b.tmp
 - a.doc
 - summary\
 - o logs
 - o data
 - january\
 - o logs\
 - o data\
 - o 2003\
 - ...

NOTE: For simplicity, we will assume the target directory is empty and the Logic is Mirror.

2.2.1. Basic Files and Directories Inclusion

Would like to include:

- All files under data\ directory.

The job properties will exclude:

Non-recursive part	Recursive part	Includes
data\	*	Files

Files transferred:

- ALL files and directories beneath data\ directory, except empty folders.

To include also empty folders:

Non-recursive part	Recursive part	Includes
data\	*	All

2.2.2. Simple Files and Directories Exclusion

Would like to exclude:

- C:\Master_data\server.log
- C:\Master_data\2002\logs directory and everything under it

The job properties will exclude:

Non-recursive part	Recursive part	Includes
server.log		Files
2002\logs\		All

2.2.3. Recursive Files and Directories Exclusion

Would like to exclude:

- All files with a tmp extension throughout the source directory tree
- All logs directories within C:\Master_data and everything inside them

The job properties will exclude:

Non-recursive part	Recursive part	Includes
	*.tmp	Files
	logs\	Directories

2.2.4. Recursive Multi-Level Directories Inclusion

Would like to include:

- All *.doc files residing in all data\summary directories within C:\Master_data\

The job properties will include:

Non-recursive part	Recursive part	Includes
	data\summary*.doc	Files

Files transferred:

- C:\Master_data\data\summary*.doc
- C:\Master_data\data\summary\data*.doc
- C:\Master_data\2002\data\summary*.doc

- C:\Master_data\2002\data\summary\data*.doc

2.2.5. Recursive & Non-Recursive Multi-Level Directories Inclusion

Would like to include:

- All *.doc files residing in all data\summary directories within C:\Master_data\
- Server.log file from the source directory

The job properties will include:

Non-recursive part	Recursive part	Includes
2002\	data\summary*.doc	Files
Server.log		Files

Files transferred:

- C:\Master_data\2002\data\summary*.doc
- C:\Master_data\2002\data\summary\data*.doc

2.2.6. Combined (Include / Exclude) Filter

Would like to:

- Publish an update for a Web Site, update all files and subdirectories in \html and \images sub directories, but exclude all *.tmp files in these directories and all _vti_* files and directories.

The job properties will include:

Non-recursive part	Recursive part	Includes
html\		All
images\		All

The job properties will exclude:

Non-recursive part	Recursive part	Includes
	*.tmp	Files
	vti*	All

3. GUI

The new Filter implementation required some changes in the GUI. The Include / Exclude specification will now be at the same Job Properties Tab called **Files**.

The **Files** tab will contain two sub-tabs that will look identical, for the Include and Exclude, where the user will enter the file specifications for each.

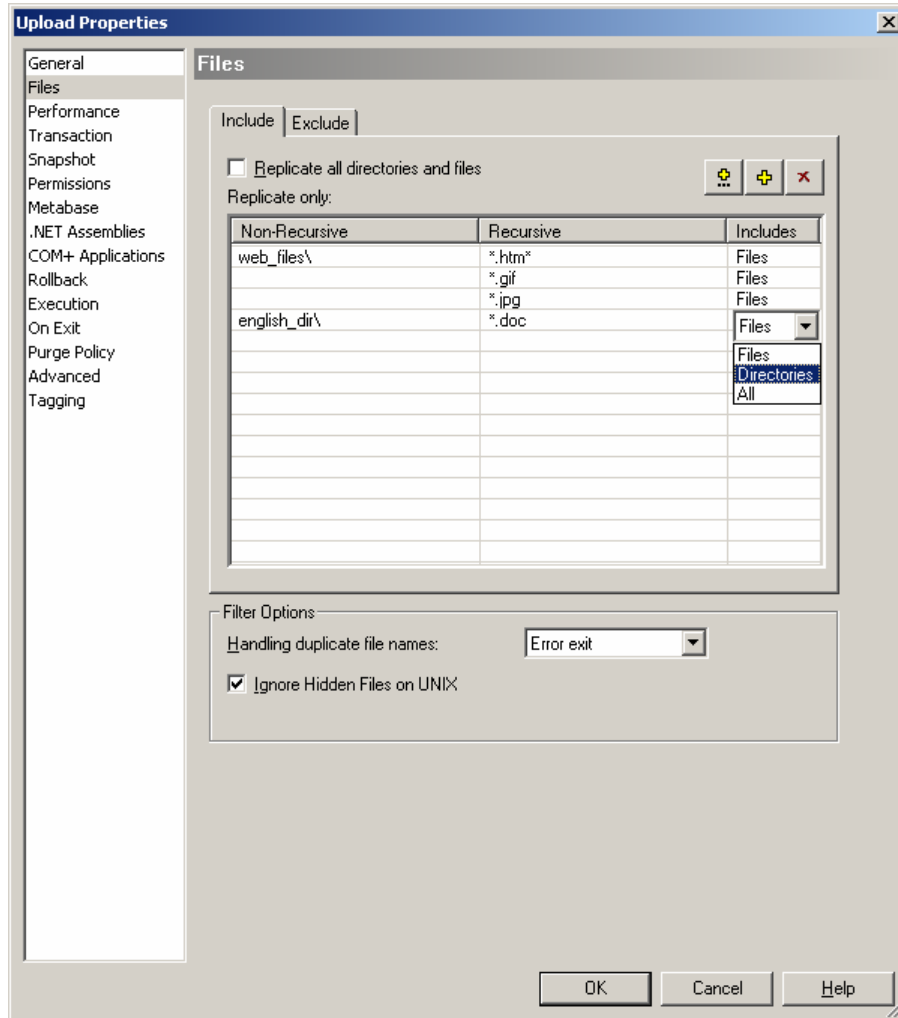


Figure 1 – New File Specification Screen

4. CLI

The command line is also changed to reflect the change of the Include / Exclude behaviour.

In previous R-1 versions, the specification to exclude files was in the following qualifiers: `-exclude` & `-recursive_exclude`; each received a list of file specifications separated by commas.

OBSOLETE:

`-file_specs=`**file_type**

Synchronize only the following types of files specified in **file_type**. These can include individual files, wildcards, or entire subdirectories. Must be used with `-noall_files`.

`-tree_recurse` | `-notree_recurse`

A switch, specifying whether replication should recurse through the entire source tree structure, or should only replicate the root directory specified without performing recursion.

`-exclude=`**file_type**

File types that match the criteria in **file_type** should be completely ignored during distribution.

`-recursive_exclude=`**recursive_file_type**

The following file types should be recursively ignored during distribution. This means any files or directories matching the **recursive_file_type** criteria will not be deleted or transferred.

As opposed to previous R-1 versions, where a File Specification could have been either recursive or non-recursive, and could contain wildcards only on the last part of the specification, every File Specification can now contain a non-recursive and a recursive part, and as many wildcards as needed, in every part of the File Specification.

The new CLI reflects the change, and the filter can now be specified as a list of items, where each item is a pair of the non-recursive and recursive parts.

NOTE: File Specifications may include a non-recursive and recursive part, separated by a comma.

-include_files=\ "[non-recursive part, recursive part], [non-recursive part, recursive part], ...\
-exclude_files=\ "[non-recursive part, recursive part], [non-recursive part, recursive part],...\
"

The curly brackets indicate an item in the list. Each item contains two parts: non-recursive and recursive. The comma must reside even if one of the parts is not defined.

Include Examples



```
>r1 submit ... -noall_files -include_files=\ "[*.doc,], [english\, *.htm],  
[,index.htm]\"
```

Include "Files" Example

The following command will include all Files under english folder.



```
>r1 submit ... -noall_files -include_files=\ "[english\,]\"
```

Include "Directories" Example

The following command will include all Directories under english folder.



```
>r1 submit ... -noall_files -include_files=\ "[english\\,]\"
```

Include "All" Example

The following command will include all Files & Directories under english folder.



```
>r1 submit ... -noall_files -include_files=\ "[english\,*]\"
```

Exclude Examples



```
>r1 submit ... -exclude_files=\ "[*.tmp,], [\*vti*\,*]\"
```

Exclude "Files" Example

The following command will exclude all Files under english folder.



```
>r1 submit ... -exclude_files=\ "[english\,]\"
```

Exclude “Directories” Example

The following command will exclude all Directories under `english` folder.



```
>r1 submit ... -exclude_files=\"[english\\,]\"
```

Exclude “All” Example

The following command will exclude all Files & Directories under `english` folder.



```
>r1 submit ... -exclude_files=\"[english\\,*]\"
```

5. Conversions

Using Templates and File Specification from previous R-1 versions, some conversions rules apply, as the new format changes the way some file specification work.

Note that templates from previous R-1 versions will behave exactly the same in R-1 3.4. However, filters specified in R-1 3.4 may be too complex for previous R-1 versions to handle. In these cases, manual intervention must be applied.

It is recommended to upgrade to the latest R-1 version as the filtering functionality has been greatly improved.

All R-1 components should run the latest version to utilize the new Filter to its full.

Some File Specifications that cannot be automatically converted:

1. **Only the final part of the file specification may contain wildcards:**

<u>Non-Recursive part</u>	<u>Recursive part</u>
<code>a*</code>	<code>*.doc</code>
<code>*english*\index.html</code>	

This specification is not supported in previous versions.

2. **Mixed recursive and non-recursive** File Specifications in the list (when there are multiple File Specifications):

<u>Non-Recursive part</u>	<u>Recursive part</u>
<code>a\</code>	<code>*.doc</code>
<code>index.txt</code>	

If all File Specifications in the list are either recursive or non-recursive than automatic conversion can be made.

6. Automatic Excludes...

R-1 may exclude files automatically during the Distribution process, based on user specification:

1. Files that could not be accessed during the snapshot generation phase. These files will be listed in the **Exclude** report.
2. During the Snapshot Generation phase, files are accessed for retrieving details only. If the user has access to retrieve details of these files, it will pass this phase successfully. However, if that user does not have access to read or delete these files, the job will fail at the Transfer stage.

To overcome access problems at the Transfer stage, uncheck the **Abort on error transferring files**.

3. **UNIX-hidden files** - Exclude files beginning with a period.
4. **Ignore Windows invalid file names** – If the source is UNIX and the target is Windows, UNIX files whose names include invalid characters in Windows will not be transferred. (If this option is not selected, and files of that kind exist, the job will fail.)

Appendix A Advanced Examples

Non-Recursive Examples

#	Non-Recursive part	Recursive part (N/A)	Included Directories	Excluded Directories	Include files	Exclude files
1	/a/b/*.c		/a and /a/b	No directories are excluded	All files under /a/b/ that match *.c	Same as include
2	/a/b*/*		/a/ and directories that match /a/b*	No directories are excluded.	All files under directories that match /a/b*	Same as include
3	/a/b*/*.c		/a and directories under /a that start with b	No directories are excluded	All files, under directories that match /a/b*, that end with '.c'.	Same as include
4	/a*/b/file1		Directories that match /a*/ and /a*/b	No directories are excluded	Files named file1 under /a*/b	Same as include.
5	/a/b/*c/		/a, /a/b and /a/b/*c	Directories under /a/b/ that match *c and all the directories under them (all levels).	Files immediately under the directories that are included.	Files under the directories that are excluded (all levels).
6	/a/b*/*/		/a/ and directories that match /a/b* and directories one level under them.	Directories under directories that match /a/b*/.	All files under directories /a/b*/ */ only one level (such as /a/bv/c/f1)	Files under directories /a/b*/*/ (at all levels)/(a/bc/d/f1 and /a/bb/d/e/f2 but not /a/bb/f3).
7	/a/b*/*.c/		/a, /a/b*/, directories under these that match *.c	Directories that match at all three levels and all directories under them. (No directories no first and second level)	All files, under directories that match /a/b*/*c/ (one level)	Files at all levels under directories matching all three levels /a/b*/*.c/
8	*		No directories are included.	Same as include.	All files directly under the source directory.	Same as include.

9	*/		All directories directly under the source directory.	All directories in all levels under the source directory (All directories directly under the source directory and all directories under them).	No files are included.	Files under excluded directories.
---	----	--	--	--	------------------------	-----------------------------------

Partial and Full Recursive Examples

#	Non-Recursive part	Recursive part	Include Directories	Exclude Directories	Include files	Exclude files
10	/a/b/	*.c	/a and /a/b and all directories under /a/b/ that contain *.c files.	No directories are excluded	All files at any level under /a/b/ that match *.c	Same as include
11	/a/b*/	*	/a/ , directories that match /a/b* and all directories under /a/b*/	No directories are excluded.	All files, at any level, under directories that match /a/b*	Same as include
12	/a/	b*/*.c	/a, directories that match /a/b* and contain *.c files.	No directories are excluded	All files, at any level, under directories that match /a/b*, that end with '.c'.	Same as include
13		/a*/b/file1	Directories that match /a*/, /a*/b and all directories under /a*/b/ that contain file1.	No directories are excluded	Files named file1, at all levels, under /a*/b	Same as include.
14	/a/b/	*c/	/a, /a/b and all directories under /a/b/ that match *c/	Directories at all levels under /a/b/ that match *c	Files at all levels under directories that match *c and located anywhere under /a/b/	Files under the directories that are excluded (all levels).
15	/a/	b*/*/	/a/, directories that match /a/b*/ and all directories one level under them.	Directories one level under /a/b*/	Actually all files under directories /a/b*/*/ (all levels)	Files under the directories that are excluded (all levels). (a/bd/r/f1 a/bd/f/f/f2 but not a/bd/f3)
16		/a/b*/*.c/	/a, directories that	Directories at all	Files at all levels	Files under

			match /a/b*/ and all directories under them.	levels under /a/b*/ that match *c and all the directories under them (all levels).	under directories that match *.c and located anywhere under /a/b*/	the directories that are excluded (all levels).
17		*	No directories are included.	Same as include.	All files in all levels under the source directory.	Same as include.
18		*/	All directories in all levels under the source directory.	Same as include.	No files are included.	All files under excluded directories.
19	*/	*	No directories are included	Same as include.	All files at all levels under directories that are directly under the source directory	Same as include
20	*/	*/	All directories at all levels under directories that are directly under the source directory	Same as include.	No files are included.	All files under excluded directories.
21	*	*	Illegal spec – unexpected behaviour – the non-recursive part of the spec must end with a dir separator.			
22	*	*/	Illegal spec – unexpected behaviour – the non-recursive part of the spec must end with a dir separator.			
23	/a/b	*c/	Illegal spec – unexpected behaviour – the non-recursive part of the spec must end with a dir separator.			

###

For any additional information, please contact us at support.repliweb.com