Sophisticated Regex Replication Enhancements

Replication provides BASIS DBMS application users the ability to replicate all or part of their database to one or more locations for the purpose of backup and/or data warehousing. A replication job essentially consists of a list of files and directories to include in the job and a list of files and directories to exclude specifically. While this generally covers the needs of most users, it is not uncommon for a user to need to include or exclude items that match particular filter criteria. In version 14.0, previewed in BBj® 13.10, replication users have the ability to add regular expression and glob filtering to the replication inclusion and exclusion lists.

By Jeff Ash
Software Engineer

Adding a Glob or Regular Expression Filter

Glob filters are the easiest to use and probably most familiar to users, e.g., wildcard filters like *.tmp, prefix*.*, etc. They provide a simple way to identify entire groups of files based on the name of the file. **Figure 1** shows a wildcard exclude filter that will instruct the replication job to ignore all files with the 'tmp' extension.
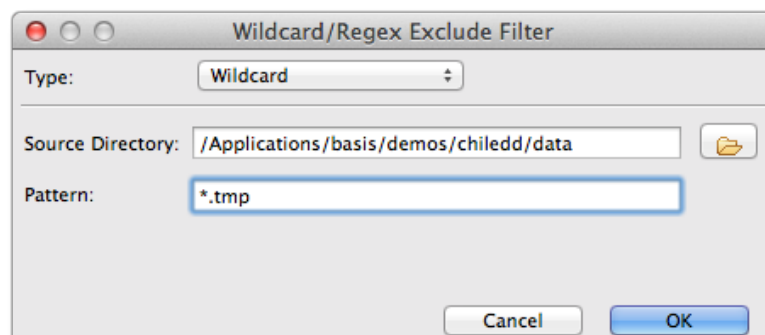

Figure 1. Specifying a wildcard exclusion

Regular expressions are also available as a filter type and while they can be more complex, they are also far more powerful and provide sophisticated searching and pattern matching that goes well beyond simple wildcards. Regular expressions allow for matching character classes and boundaries, alternation and grouping, repetition, and more. For more information on regular expressions, check out the Java Regular Expression Tutorial.

To add a filter, follow these simple steps:

1. Click the wildcard or regular expression button on the Replication Job Wizard on the appropriate page (included or excluded as appropriate) or while editing an existing replication job, on the appropriate included/excluded tab.

2. Select the directory where the filter applies.

3. Specify the glob or regular expression filter as appropriate.

Now, replication will validate file names with the filter before including or excluding them. An example of this appears in **Figure 2** that specifies a wildcard pattern of 'C*' for the types of files to include in the replication job.

Figure 3 shows how this file filter appears after creation and how it is available in the list of directories and files to include in a new replication job.

Summary

If you have ever spent any amount of time building a list of files to exclude from a replication job, such as temporary files generated by reporting or monthly closing processes, you have probably wondered if there was a better way to accomplish the task. With the introduction of powerful filtering support built into the replication interface, it is now possible to easily include or exclude entire groups of files based on naming conventions, even those that may not exist at job creation time. Whether you choose to use a simple glob filter, or pull out the big guns and use regular expressions for complete control, it is now a cinch to indicate categories of files to ignore or include. ■
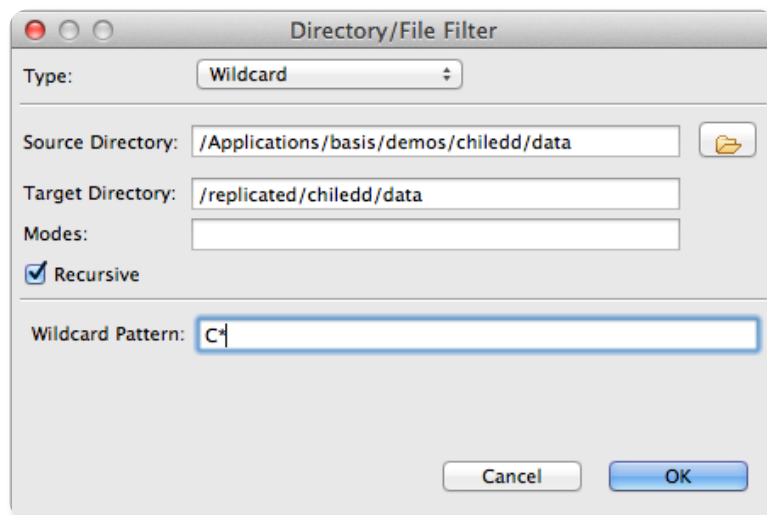

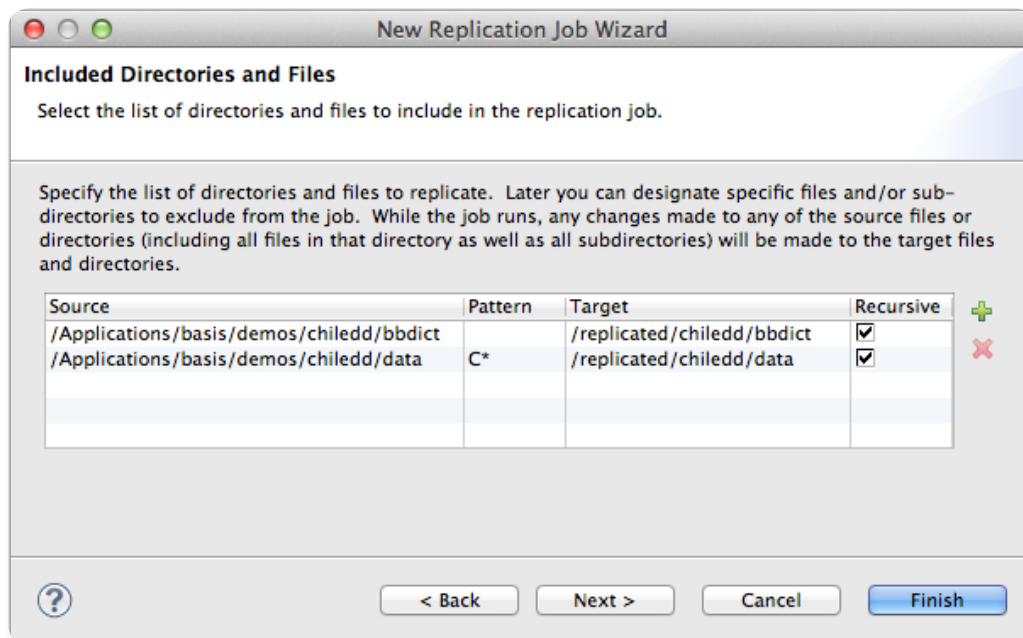**Figure 2.** Specifying a wildcard inclusion

Figure 3. Creating a new replication job with the newly defined directory/file filter

For more information, read
- *Replication Introduction* in the online help at links.basis.com/replication
- *Anatomy of a Replication Job* at links.basis.com/11replication