

# File System Feeder

The *File System Feeder* component periodically monitors a number of directories (and option sub directories), and looks for files that have changed since the last scan. A change means new additions, modifications, or deletes. Changed files are published to an Aspire pipeline manager. It monitors one or more directories, and periodically polls them to look for updated files (with an optional file name filter).

The feeder builds up a snapshot of the directory structure (optionally including the subdirectories), and compares this against the snapshot created the last time the feeder polled the directory. A list of new, updated, and deleted files is built. These files are published to an Aspire pipeline manager. When all of the changes from the scanned directory have been processed, the feeder processes the next directory. When no more directories exist, the feeder sleeps for a period of time before polling the directories again.

**Note:** This feeder is based on the [Simple Feeder](#)

File System Feeder	
Factory Name	com.searchtechnologies.aspire:aspire-filefeeder
subtype	fileFeeder
Inputs	The files in the monitored directories.
Outputs	An <a href="#">AspireObject</a> containing the path to the discovered file in the monitored directory in the <url> and <fetchUrl> tags and the action in the "action" attribute, published to the configured pipeline manager.

## Configuration

This feeder takes all parameters from the [Simple Feeder](#) plus the following:

Element	Type	Default	Description
feederLabel	string	FileFeeder	The feeder label submitted in the <feederLabel> of the published document.
scanLocations		None	The configuration of the folders to monitor. See below.

## Folder Configuration

The file system feeder monitors one or more directories, periodically polling them to look for changed files. The folder configuration is shown below.

Element	Type	Description
scanLocations /scanLocation	parent tag	Holds all of the information for a single directory. Each <scanLocation> tag holds the location of the directory plus all of the parameters (wildcard patterns, etc.) necessary for processing the files. Note that you can have multiple <scanLocation> tags in the same file feeder, as many as you'd like, to handle multiple folders from the same feeder.
scanLocations /scanLocation /@baseDirectory	string	The root of the directory tree to monitor. Files found in this directory (and optionally in it's sub directories) when the feeder polls will be published.
scanLocations /scanLocation /@match	String	A regular expression detailing the names of the files in the scanned directories that will be processed. If the file name is not matched by this expression, the file will be ignored. If this option is not specified, all files will be processed.
scanLocations /scanLocation /@recursive	boolean	If <b>true</b> , changed files in the <b>baseDirectory</b> and it's subdirectories will be published. If <b>false</b> , only files in the <b>base Directory</b> are considered.
snapshotLocation	string	If set, the files holding the status of the disk paths being fed ("snapshots") are located in the configured directory. Otherwise they are located in the directory given by the environment variable \$ASPIRE_HOME

## Metadata Mapper Configuration

The hot folder feeder maps some metadata fields to fields in the [AspireObject](#).

Field	Default Output Field	Description
fileName	fileName	The filename of the published file.
path	fileName	The path to the file.
fullFileName	fileName	The full filename (including the path) to the file.
fullPath	fullPath	The full path to the file (excluding the file name).

## Example Configurations

### Simple

```
<component name="FileFeeder" subType="fileFeeder" factoryName="aspire-filefeeder">
  <branches>
    <branch event="onPublish" pipelineManager="/system/StandardPipeManager"/>
  </branches>
  <snapshotLocation>testdata/com.searchtechnologies.aspire.feeders.filefeeder</snapshotLocation>
  <scanLocations>
    <scanLocation recursive="true" baseDirectory="c:\temp\fp1"/>
    <scanLocation recursive="false" match=".*\.doc" baseDirectory="c:\temp\fp2"/>
    <scanLocation match="[0-9a-z]*" baseDirectory="c:\temp\fp3"/>
  </scanLocations>
</component>
```

### Complex

```
<component name="FileFeeder" subType="fileFeeder" factoryName="aspire-filefeeder">
  <feederLabel>myFileFeeder</feederLabel>
  <metadataMap>
    <map from="fileName" to="fileName"/>
    <map from="fullPath" to="fullPath"/>
  </metadataMap>
  <autoStart>${autoFeedArc}</autoStart>
  <loopWait>43200000</loopWait>
  <feedWait>30000</feedWait>
  <branches>
    <branch event="onPublish" pipelineManager="/system/StandardPipeManager"/>
  </branches>
  <snapshotLocation>testdata/com.searchtechnologies.aspire.feeders.filefeeder</snapshotLocation>
  <scanLocations>
    <scanLocation recursive="true" baseDirectory="c:\temp\fp1"/>
    <scanLocation recursive="false" match=".*\.doc" baseDirectory="c:\temp\fp2"/>
    <scanLocation match="[0-9a-z]*" baseDirectory="c:\temp\fp3"/>
  </scanLocations>
</component>
```