Charlie's Tips

(Tips and Resources from a Veteran CFML developer)

by Charlie Arehart

Text File Processing

If you need to read a text file into CFML for processing, you have several choices to consider, each with its own strengths and weaknesses. The following will work in ColdFusion 5 and above or BlueDragon:

• CFFILE Action="read": This tag will read in a file, creating a single variable holding the entire file contents, which is fine for some applications. Learn more at http://livedocs.macromedia.com/coldfusion/7/htmldocs/00001622.htm. If you want to process the file line-by-line, or process the results like a query, it's tempting to use this tag to read the file and then manipulate the contents of that one variable using CFML functions, but the following alternatives may prove more appropriate.

• CFQUERY of an ODBC Text datasource: It may surprise you to learn that CFML has long been able to read in text files using CFQUERY. In CF5, you will find a text datasource type among the available datasource drivers. In CFMX or BlueDragon, you will need to use the Windows "DataSources" tool (located in Control Panel>Administrative Tools) to create an ODBC datasource. Select the "Microsoft Text Driver" option and then the directory location for the file(s) that you will be reading. You'll then need to define a DSN pointing to it in the CFMX or BD Admin .

In your CFQUERY statement, use "SELECT * FROM filename.txt", pointing to the file to be read, and use or dump the query results. Learn more about the ODBC setup at http://www.c-sharpcorner.com/Code/2002/Nov/AccessTextDb.asp, in the section, "Accessing a Text File". Those using CF5 should see another option at http://www.adobe.com/go/tn_17980.

• CFHTTP with NAME attribute: You can read a file into a CFML query result by using CFHTTP. If you use the NAME attribute with CFHTTP, it doesn't load the results into CFHTTP.FileContent. Instead, CFHTTP creates a query of the given NAME. An example could be: <cfhttp url="http://someurl/somefile.txt" name="get">

By default, ColdFusion expects that the page being read will be in CSV (comma-separated value) format, which is great for importing data that's been exported from other programs. But you can define how the input file is formatted using the DELIMITER and TEXTQUALIFIER attributes. And the file doesn't need to be a .txt file; it can be any file extension. If it's retrieving a CFML page, make sure to use CFCONTENT to set the TYPE to "text/plain". Finally, you can also control whether the file being read starts with a record that lists column names or not.

This CFHTTP approach may appeal if you're not using Windows, or where you can't or don't want to create a new datasource, or if you're reading a file from a remote source (though it's certainly not restricted to remote files only.) Learn more at: http://livedocs.macromedia.com/coldfusion/7/htmldocs/00001610.htm

• Java FileReader: While CFFILE reads in the entire file as one variable, what if you wanted to instead read the file in line-by-line? Since CFMX and BlueDragon are built atop Java (and Java can also be integrated into CF5), you can use a Java FileReader object to read a file on a line-by-line basis. See http://coldfusion.sys-con.com/read/86121.htm for some sample code. CF5 users can see this resource for another approach: http://www.adobe.com/devnet/server_archive/articles/leveraging_java_classes_cf.html

Charlie Arehart recently took the position of CTO at Garrison Enterprises in Charlotte, NC where he is helping open an additional Atlanta office.