

FS2DS

File Server to DocuShare Content Migration Program
for use with Xerox DocuShare
from
Criteria First

<http://www.CriteriaFirstWare.com>

FS2DS

File Server to DocuShare Content Migration Program

Introduction

FS2DS is a program that lists all the directories and files in the specified directory and processes them in the specified chunk sizes. It's advised that once the migration process begins that the content being uploaded is not moved around until everything is uploaded.

Requirements

Xerox DocuShare 6.0 or later

All Windows platforms

Temporary disk space to accommodate documents being processed as a batch based on your preferred chunk size (number of files being processed per batch).

Downloaded program setup

- Download the trial to a common directory to view in Windows Explorer.
- Open the compressed folder or unzip the compress file.
- Extract the files to a folder that won't be part of the uploaded files.
- Open a command window to that folder.
- Enter FS2DS to review the commands and determine what syntax to use.

Registering FS2DS

When starting **FS2DS** for the first time you will be presented with a dialog window to register **FS2DS** for ongoing use. You may select Cancel if you wish to evaluate **FS2DS** for up to 15 days. If you've already purchased or choose to do so, you will be asked to submit the software code in the top field along with your purchase so that we can send you the liberation key to unlock it for perpetual use. If purchased online, email your **FS2DS** software code to Criteria First support at support@criteriafirst.com for your liberation key.



Configuring FS2DS

FS2DS can be used in either Windows (GUI) or command line interface (CLI) modes. The Windows mode will be preferred by casual users and ad-hoc usage by technical personnel. The command line mode is useful for repetitive operations and routines.

Note: FS2DS relies on another freeware software package called 7-Zip to create the content intake package that DocuShare ingests, so you'll need to download that separately from <http://www.7-zip.org/>. You may install 7-Zip before or after installing FS2DS, but both must be installed before FS2DS will work for you.

We will first review the Windows interface. Starting FS2DS will display a dialog window for you to make selections for the content you want to upload, where you want it uploaded to in DocuShare, and some controls for how uploads are processed.

Source directory: type to enter or click the triple-dot button to select the top level folder on your PC or

FS2DS

File Server to DocuShare Content Migration Program

server you'd like to upload. Both drive letters and UNC paths are both acceptable.

File extensions to exclude: By default, everything will upload with the exception of hidden system files, thumbnail markers, shortcuts, and other such files that have no use or purpose in DocuShare. This includes nested folders full of folders and their files. If, for example, you want to upload a folder containing both PDF and TIFF image files and want to upload only your PDFs, but not your TIFFs, enter "*.tif, *.tiff" without the quotes to tell FS2DS to exclude these files.

DocuShare Content Intake "input" directory: This is the shared folder on your Windows PC or server where DocuShare watches for your files to upload. Both drive letters and UNC paths are both acceptable. You'll need to set up a content intake profile in your DocuShare Admin Home, so if you're not the DocuShare site or content administrator you will have to ask that person for assistance to complete your setup. In general, we recommend you create several subcollections in this watched folder that the content intake module will use and let your DocuShare administrator know of these to complete the setup of the profile. They also will need to know the collection handle where you want your files to go, so let them know where your uploaded files' destination. Set up a main watched folder with some useful name, such as Accounts, with the following subfolders.

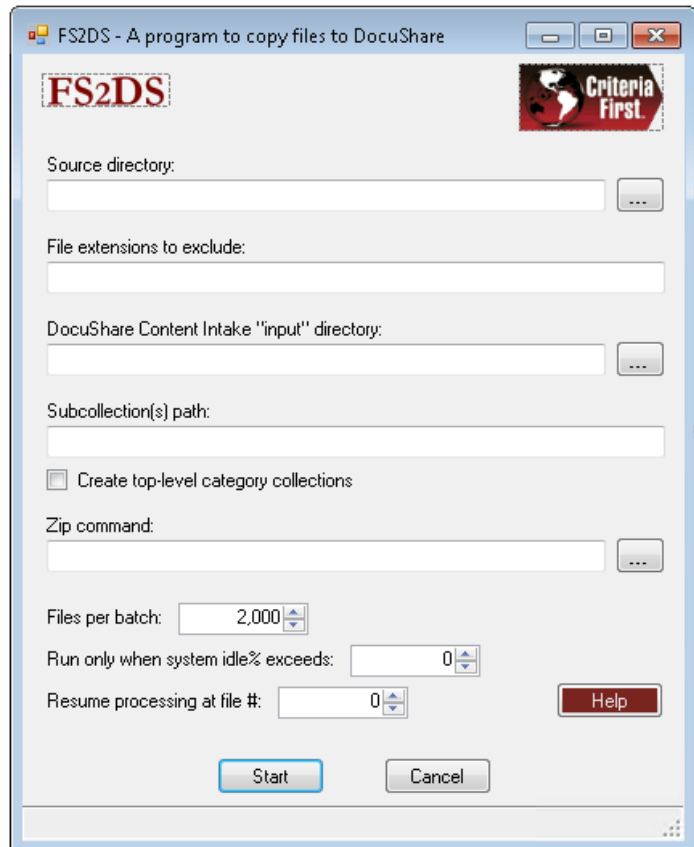
```
\\someservername\transfers\Accounts\  
├──Errors  
├──Input  
├──Logs  
└──Temp
```

You can use the main Accounts folder for input if you like, but the content intake module requires folders for errors, logs, and temporary processing.

Subcollection(s) path: This allows you to tell FS2DS to drill down one or more collections to begin loading your files. This keeps you from having to create a unique content intake profile for anything but your top level collections. So, if you want to upload your folders and files to a collection called Application Study Materials under the Customer Support collection for which you have a content intake profile, you'd enter Application Study Materials in this field (case sensitive). Use a slash to indicate multiple levels of collections as being intermediate, as in Application Study Materials/Workflow.

Create top-level category folders is a great feature to assist you in filing your content under an alphanumeric character by creating top level folders based on the first character of the file or directory. For example, "ACME, Inc." would be placed in folder "A". If no "A" folder exists, then one will be created for it. This option is incompatible with the "Intermediate Folders" option.

Zip command allows you to tell FS2DS where your 7-Zip program resides. FS2DS uses this to create the zip file packages containing your files and an xml file which tells DocuShare where they go and what to call them.



FS2DS

File Server to DocuShare Content Migration Program

Files per batch: can generally be left at 2000 documents, but you can raise or lower this number if you observe you get better performance from your combination of FS2DS and DocuShare's content intake module.

Run only when system idle % exceeds: n% If you're running this on a server that is often busy performing other functions, you can set this number rather high, say at 90%, so that it only processes upload batches when activity is relatively low.

Resume processing at file #: is useful if there's an interruption to the batch upload process due to your decision to stop FS2DS or in the event some other factors causes processing to stop. You should be able to see the "chunk" numbers in an open command window indicating how many chunks were processed before it stopped. It is best to not change the chunk size while uploading content to ensure that this count will be accurate in determining which files need to be processed from that point forward.

FS2DS command line syntax

FS2DS is also a command line driven program that requires no configuration other than to copy it to a directory convenient to the user.

```
C:\Temp >FS2DS
```

```
Usage: FS2DS /SOURCEDIR:sourceDirPath /DESTDIR:destinationContentIntakeDirPath  
[/RESUMEAT:fileNumber] [/FOLDER:rootFolderOnDocuShareContentIntake]  
[/IDLEMINIMUM:minimumIdle%] [/EXCLUDE:fileExtensionsToExclude]  
[/CHUNKSIZE:chunkFileCount]
```

The only required options are /SOURCEDIR and /DESTDIR. The commands are not case sensitive. If you must refer to a path with spaces in a directory name, quote the entire phrase as such

```
"/sourcedir:M:\Customer Accounts"
```

The DESTDIR refers to a directory that is watched by DocuShare for Content Intake. Refer to your DocuShare Admin Guide for instructions on setting up a watched folder under Admin Home > Services and Components > Content Intake > Configuration.

/CHUNKSIZE should be used when uploading large numbers of files, generally understood to be more than 2000. Chunks refer to the number of files processed to be uploaded as delivered in an individual package consisting of a zip file with an xml file containing information about all the documents to be uploaded and the documents themselves. The larger the chunk, the longer it takes to create each package and the longer it takes for DocuShare to ingest and present them as content. As a benchmark, a chunk of 2500 documents has been observed to take 6 minutes to create and 6 minutes to ingest for a total of about 12 minutes on a 3 GHz Xeon dual-core processor with two 7200 rpm drives in RAID 1 configuration.

/IDLEMINIMUM is used to detect the level of activity running on the current PC or server to avoid resource conflicts. It's recommended that you run DAVUpDoc and DocuShare on separate hardware to avoid competing for resources. If you run them on the same platform, be sure to enter a reasonable value to have the script wait until the machine isn't busy with other activities such as DocuShare's content intake background ingestion process. Even when on a separate machine, this option could be useful in avoiding resource conflicts with virus scans, backups, and other activities. A reasonable value would be something like 80 thereby indicating that the machine's System Idle Process indicates it is mostly at rest and available for you to process your files.

/RESUMEAT provides a way to restart long-running content transfers that, for whatever reason, terminate. As FS2DS runs, it echoes information about the files currently being transferred. If FS2DS terminates abnormally, after the underlying problem is corrected you can restart the transfer by using the same command line as before, with the added /RESUMEAT parameter indicating (by count) with which file FS2DS should restart.

/FOLDER provides a way to insert an arbitrary path fragment on the destination file location. For

FS2DS

File Server to DocuShare Content Migration Program

example, the files uploaded should be placed in a special branch of the destination. For example, if the content intake configuration says to place the files in the Accounting/Inbox collection path, but you want the files to appear in the Accounting/Inbox/Uploaded collection, then you would use the /FOLDER:Uploaded option.

/EXCLUDE provides a way to exclude files by file extension. For example, to exclude MS-Access files you would use the /EXCLUDE:mdb,ldb option.

To process your accounting files on your PC or server, enter a command like the following:

```
FS2DS "/SOURCEDIR:C:\Accounting\Transactions" "/DESTDIR:K:\WatchDir"  
/CHUNKSIZE:2000
```

Other functions within FS2DS include splitting the chunks where it has found that two files of the same name exist within the boundaries. DocuShare's content intake will reject the file if they have the same name even if they exist in different folders. This has been escalated to Xerox as a bug to be fixed.

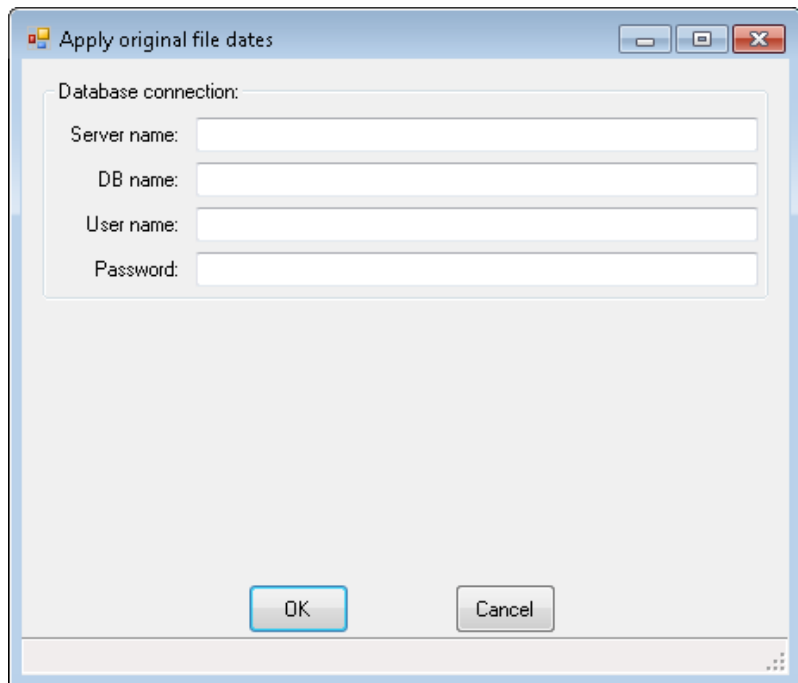
Setting up DocuShare for content intake operations is beyond the scope of this document, so please refer to Xerox documentation on this matter.

Correcting dates after uploading

The DocuShare Content Intake mechanism creates files using the current date and time for both the file creation date and file last modified date. FS2DS records the original file's creation and modification dates in a special format in the "keywords" property of the uploaded files. Once the Content Intake module has imported all of the files, run the included FS2DS_Dater program. This program sets the DocuShare files' dates to the values specially recorded in the keywords properties.

To use the FS2DS_Dater program, it will require read/write access to the SQL Server database used by DocuShare. The FS2DS_Dater program will update the files' dates while DocuShare is up and running, although the users should be cautioned against modifying any of the uploaded files until after FS2DS_Dater program has finished running.

Once FS2DS_Dater has finished, it is necessary to rebuild the content index for those uploaded files, since the indexer's values for the file creation and modification dates are "today", while FS2DS_Dater has changed them in the database to be the correct values. The documents will display correctly, but you won't be able to search on them until you re-index. The content index can be rebuilt for the entire DocuShare site (if the total number of documents is not large), or for only specific collections and their contents.



Rebuilding the entire content index

Use the command:

```
dsindex index_all
```

FS2DS

File Server to DocuShare Content Migration Program

Rebuilding the content index for a specific collection and its contents

Use the command:

```
dsindex -c Collection-### index  
(where "Collection-###" is the collection's handle).
```

Error messages

If you see messages that occur like the following:

```
C:\Documents and Settings\Doug\Local Settings\Application  
Data\Temp\ZQXTQITXFP.vbs(158, 2) (null): The system cannot find the file  
specified.
```

... this is probably due to your not having yet installed 7-Zip for use with FS2DS. The name for the user (Doug in this case) will change and the reference to the program (ZQXTQITXFP.vbs in this case) will change every time.

Operational Considerations

If your documents don't appear using Content Intake Mode, check your DocuShare log files associated with the zip files prepared for you to see if there is an explanation of why they may have failed to load. If you have problems, please contact Criteria First Customer Support for assistance.

If DocuShare appears to stall after a while and the CPU appears to be in a resting state, try changing the Polling Interval value a little and clicking the Submit button to see if that kicks off the content ingester again.

Content Intake Batch mode requires that the DocuShare administrator set up Content Intake in the Admin Home before implementation. Please read the DocuShare Administrator Guide for information on setting up content intake. Essentially, you will need to log in as a site or content administrator, go to the Admin Home, navigate to Services and Components > Content Intake > Configuration and set up a new profile using:

Function	Option or Entry
Category	Imaging Subsystem
Type	DocuShare Image Manager
Name	[whatever you want to call this profile that relates to your application or target collection]
Enabled	checked
Default User Handle	User-2 [or another handle with content admin group membership]
Target Root Collection Handle	Collection-123 [the top level destination collection handle]
Watch Directory	[the folder to be watched for content to upload]
Log Directory	[where you want the DocuShare content intake logs to write]
Confirm Document Upload	checked
Error Directory	[where rejected files are placed]
Temporary Directory	[where DocuShare places files being processed]

FS2DS

File Server to DocuShare Content Migration Program

Polling Interval	[set this low for testing and then sufficiently high to avoid cycling too often]
------------------	--

These fields are all required and the others are used only for other specialized purposes.

Click the Create button and watch for confirmation. If you need to make changes use the Edit option below Configuration.

Customer Support

Customer Support is available Monday through Friday from 8:00 AM to 6:00 PM CST via email to support@criteriafirst.com or by calling 972-492-4428.