**Requirement:**

* The scope of the requirement is to move all the files which are more than 30 days old into documentum server.
* The batch job needs to run every 15 days for moving the files to documentum. (so the file will be in success folder for a maximum of 44-45 days because there might be files which are 29 days old and for those files the retention will happen in the next run I.e. after 15 days).
* Only those files which have been successfully processed by sanity check needs to be moved to documentum. (from success folder only)

**Design:**

* Create a batch application which will automate the process of moving success files to documentum.
* The job will first check all the files in the success directory and pick only those files whose creation date is more than 30 days.
* Each such file which is created more than 30 days back will be move to documentum server
* Once successful transfer of the file to documentum is complete, we will delete the file from the success folder (will require permissions to delete those files)
* If there are any errors while transferring the file to documentum, the technology team will be notified and the file will stay in the success folder as is.
* We will schedule a nightly job which will run at 12:00 am once every 15 days (we can discuss on the time and days).

**Program Architecture**

**Step 3** (on fail only)

If retention fails

Notify technology team

Notify

**Step 2**

Move files to

Documentum

**Step 1**

Read files more

than 30 days’ old

PROGRAM

SUCCESS

DIRECTORY

Documentum

Server

Retention Application

**Scenarios:**

**Scenario 1: success flow**

* Let’s says below are the files in success folder:

1. File1.xlsx (3 days old)
2. File2.xlsx (20 days old)
3. File3.xlsx (35 days old)
4. In the above case in step 1 we will get only 1 file (File3.xlsx).
5. In step 2 we will transfer the file to documentum and finish the retention process

**Scenario 2: No old files**

* Let’s says below are the files in success folder:

1. File1.xlsx (3 days old)
2. File2.xlsx (20 days old)
3. File3.xlsx (15 days old)
4. In the above case in step 1 we will not get any files older than 30 days.
5. So nothing will be moved to documentum

**Scenario 3: Partial Success flow**

* Let’s says below are the files in success folder:

1. File1.xlsx (33 days old)
2. File2.xlsx (40 days old)
3. File3.xlsx (25 days old)
4. In the above case in step 1 we will get 2 files (File1.xlsx & File 2.xlsx).
5. In Step 2 lets says the first file File1.xlsx is successfully moved and there was an error while moving File2.xlsx.
6. Then we will only delete the File1.xlsx and keep the File2.xlsx as is in the success folder. And we will also notify the actual error to the technology team on why File2.xlsx could not be transferred.

**Scenario 4: All Failures**

* Let’s says below are the files in success folder:

1. File1.xlsx (33 days old)
2. File2.xlsx (40 days old)
3. File3.xlsx (25 days old)
4. In the above case in step 1 we will get 2 files (File1.xlsx & File 2.xlsx).
5. In Step 2 dues to the documentum server being down or any other reason if all the files fail while moving to documentum.
6. Then all the files will remain as is in the success folder and the error details will be notified to the technology team.

**Scenario 5: Moving a file which already exists in Documentum server**

* Not sure if there could be such cases. But let’s say we are trying to move a file with name abc.xlsx to documentum in step 2, and in case in documentum there is a file in the directory with same name, I am assuming documentum will throw an error saying duplicate file and we will not move that file due to failure. The file will stay in the success folder
* The problem with the above such files is that until and unless we manually change the filename or remove the file, every 15 days the process will try to move the file and fail with a notification.
* We need to figure out a way to handle such files which have same name in success folder and documentum server:

**Proposed solutions:**

we can check the error details and see if we can verify that the error happened due to duplicate file name**.** If that is the case, then we have 3 options:

1. Delete the file from success folder. (may not be appropriate if the data in documentum file and our file vary)
2. Rename the file by appending some text (like “\_<currentdatetime>”) and try moving to documentum again.
3. Handle such files manually on a case to case basis (which may not be an ideal solution but every file is different and the case could be different).

**Questions:**

1. What happens to the files sitting in fail folder? While sanity check we will notify failures but how long will they stay in the fail folder? Do we need any strategy to delete failed documents after a point?
2. Need the documentum server details for dev testing
3. Email specification (like subject, body template, sender, host details etc.) will be needed
4. In case of failure a consolidate 1 email will be sent for all the failures I.e. even if there are multiple failures which transferring the files to documentum, we will consolidate all the errors and only send 1 email. Please confirm.