

SharePoint document migration challenges when migrating files and folders

Migration tool for Microsoft SharePoint – White Paper

There are several challenges when migrating documents to Microsoft SharePoint. While these challenges can be overcome, they are a real pain if the migration source, content and file systems are not SharePoint friendly. We will discuss below some of the common ones here. Folders and files exist in several sources - file shares, web-based sources, network / backup drives, personal drives etc.

1. Dealing with Special characters and Lengths in Folder and File names

SharePoint does not accept certain special characters (tilde, number sign, percent, ampersand, asterisk, braces, backslash, colon, angle brackets, question mark, slash, pipe, quotation mark - ~, #, %, &, *, {, \, ;, <, ?, /, |, "). Hm... that's a lot of special characters and certainly, it is not that uncommon to find some of these in filenames in file systems. Also, SharePoint does not allow folder and file name lengths to be longer than 128 characters in WSS 3.0. These aspects alone can be such a pain during migration of folders & files that contain special characters and long names. Windows folder / file names with special characters have to be replaced with SharePoint acceptable characters to avoid manual work in renaming folders and files. For files that contain special characters based on certain logic or a set of rules, this can easily be dealt with by using scripts or some tools. However, if the files contain special characters in a random manner with no orderliness about them, it can be a laborious task to rename the folders and files before migrating them to SharePoint. Similarly, long folder and file names have to be truncated to the prescribed length before moving them to SharePoint. A few nasty folders / files in random can put a spoke in a well planned, large and orderly migration. Here are two useful links to know more about SharePoint special characters, limits on URL lengths and long filenames.

<http://blogs.msdn.com/joelo/archive/2007/06/27/file-name-length-size-and-invalid-character-restrictions-and-recommendations.aspx>

<http://support.microsoft.com/default.aspx?scid=kb;en-us;905231>

2. Maintaining the same folder / file structure when migrating to SharePoint

Most companies will want to retain the same structures for files and folders to maintain operational consistency as well as business continuity. For example, an organization currently using a traditional Windows based file server platform for document collaboration could be deploying SharePoint as the collaboration platform for the users. In such situation, it will be easier to maintain the same folder and file structure in the newly setup SharePoint library without changing the user experience in handling folders and files. It makes the navigation intuitive, eases migration process, minimizes user training and improves operational efficiency.

3. Migrating a select set of document types / formats such as doc, xls, ppt, jpeg, dwg, pdf etc.

Sometimes, only documents of certain types (or formats) need to be moved to SharePoint libraries depending on the document templates or content types or file types allowed to enhance the document management framework. Everything else needs to be filtered. To selectively move files based on their types will require some programming, especially so if there are large folder trees (nested folders).

4. Migrating a large number of unstructured and poorly managed files.

From the perspective of file contents, business relevance and usage patterns, migrating a large number of unstructured and poorly managed files (remember some of those legacy file systems?) to a structured, organized and searchable framework within SharePoint is not easy. A lot of planning and meticulous reorg of files and folders is a must. This essentially means that there will be frequent rework of mapping and remapping the source folders and the destination libraries. The migration has to be broken down to several different tasks (batch processes), with several different rules to process the files and folders.

5. Using a mass document migration application to work off your desktop instead of running right on the SharePoint servers directly.

If you are using third-party tools for migration, a solution that can run either on the desktop and or on the server will be ideal. Take the case of incremental migrations. The SharePoint server could already be in production mode while several different libraries in SharePoint could still be under migration. Server performance will be compromised if the migrating application is going to run on the server. It is better if the application runs on a desktop performing actions like data cleansing, processing etc. and just do a final bulk upload into SharePoint.

6. Migrating and tagging the tens of thousands of documents from your legacy file folders into your new SharePoint repository while retaining the existing taxonomy or migrate to a new taxonomy.

Most file system users will not be familiar with the concept of document metadata. They are more familiar with the term file properties. The concept of document metadata originates from Document Management Systems (DMS), where documents are stored and made searchable on a wider set of keywords and phrases. Document libraries in SharePoint are akin to Document Management Systems in the way they store documents and properties for search and retrieval. However, these properties need to be presented to SharePoint during migration in a certain manner that makes the search and retrieval more powerful and elegant. You may migrate / propagate metadata fields from the document properties stored within the file or using an external metadata file / database. Mapping and tagging existing file properties and adding additional properties to documents in SharePoint, especially for a well structured taxonomy can be a laborious process.

7. Retaining (carry forward) the Created Date and Last Modified file attributes from the file system to maintain business continuity for users and minimize user training when collaborating in the new SharePoint environment.

This is another challenge that is constantly faced by SharePoint users. There are plenty of business reasons to retain the same Date field values once the documents have been migrated to SharePoint. Unfortunately, SharePoint falls short in this aspect and you need third-party tools or some in-house programming to carry forward the original date fields for documents.

8. Automating the migration process to reduce the time necessary and labor involved to move large file repositories to SharePoint.

Almost all mid-sized to large-sized migrations require automation, especially when incremental/batch migrations from different sources take place. Automated batch jobs help you to take complete control of the migration process by handling errors and triggering events that can be managed efficiently. Otherwise, you have to spend hours on ad-hoc problems and tracking down repetitive errors thrown by non-automated migrations.

The DocKIT solution for SharePoint File Migration

In the previous section we discussed in detail many of the Sharepoint document/file migration challenges. Here we will summarize how Vyapin's DocKIT solution addresses these issues.

Features

- Import folders & files and metadata
- Import documents from a batch descriptor file
- Apply metadata to documents from metadata file
- Import metadata of documents
- Import multiple file versions
- Automate Tasks using Tasks Manager
- Recover from file migration errors using task re-import
- Import to multiple destinations

Partial list of supported metadata from the source content

- Custom properties defined in the external metadata file
- Summary file system properties - Title, Subject, Author, Category, Keywords and Comments
- File properties such as Manager, Company, DateLastPrinted, DateLastSaved, RevisionNumber, Version, WordCount etc. in the case of MS-Office documents
- Original Created Date & Last Modified Date of source documents
- Author (Created By) and Editor (Modified By) of source documents
- Content Approval Status & Approval Comments

Please click the following link to know more about Vyapin's DocKIT:

http://www.vyapin.com/products/sharepoint_2007_document_migration_dockit.htm



Vyapin Software Systems Private Limited

Website: <http://www.vyapin.com/>

Sales Contact: sales@vyapin.com

Technical Support: support@vyapin.com



Copyright © 2009 Vyapin Software Systems. All rights reserved.

Document Import Kit for SharePoint 2007 is a registered trademark of Vyapin Software Systems. All other brand or product names are trademarks or registered trademarks of their respective companies.