

Optimal Designs for Managing Documents in SharePoint

The original version of this paper, which MacroView Founder and CEO Noel Williams wrote in July 2012, attracted widespread interest and positive feedback. In this update Noel looks at recent developments in MacroView DMF that enable even more ways to design SharePoint so that it is optimal for managing documents and emails.

Introduction

As more organizations explore the potential of using Microsoft SharePoint for document Management and Email Management a question that arises frequently is “What is the best Information Architecture for a SharePoint DM store?” In other words, what is the best way to arrange the various SharePoint 2010 ‘building blocks’ – site collections, sites, document libraries, document sets, folders and metadata columns – to come up with a design that is optimal in terms of volume handling, manageability and ease of use? That last objective – ease of use – is critical because if the users do not find it easy to interact with the document store they will likely store their documents in file shares or local drives and their emails in Outlook folders and the SharePoint DM Project will not be successful.

Deeply Nested folder Trees

There is now general recognition that deeply nested trees of SharePoint folders in one or two large document libraries is NOT a good design. When SharePoint 2007 was first released these designs appeared to offer an easy means of migrating existing File Shares into SharePoint – the tree of folders from the ‘P:’ drive were simply reproduced, one for one, with SharePoint folders. The fact that the full SharePoint folder names become part of the URL for a document stored in SharePoint rapidly led to URLs exceeding their maximum permitted length of 255 characters – especially if the folder names were long and meaningful.

These designs also led to very large numbers of documents being stored in a single document library, which in turn caused poor performance when viewing the lists of documents in those libraries. In an attempt to prevent performance degradation, with SharePoint 2010 Microsoft introduced List View Throttling. Put simply, if generating a view of a document library requires processing more than 5,000 items (items can be documents or folders or document sets) SharePoint will not display all the items (or depending on configuration will not display any items). This might avoid poor performance, but it is far from optimal from the user’s perspective.

Design Alternatives

So if deeply nested folders are bad, what alternative design is best? How about using document sets instead of folders? Isn’t it better to use metadata columns – particularly Managed Metadata columns - instead of either folders or document sets? Don’t you need to have lots of site collections if you want to store a large volume of documents? Maybe we should just use search and not attempt to have a hierarchy of storage containers at all?

At MacroView we specialize in SharePoint-based solutions for document and email management. We sell our MacroView document Management Framework (DMF) add-on to customers across the world. These customer organizations range from small to very large, so we encounter many different designs for SharePoint DM environments, including all those referred to above.

The first point I would make is that there is no single design that is right for every organization. To come up with an optimal SharePoint DM design you do need to take specific business requirements into account. A second point is that size does matter – lots of designs work well enough if you have only a small volume of documents or users, but good design makes a major difference when there are hundreds of terabytes of documents and / or thousands of users. A related observation is that document stores tend to grow very quickly. Even if you don’t have much volume stored in SharePoint today, you have only to look at how your file shares and Exchange stores have expanded in size over the past several years to realize that you should be designing your SharePoint DM store to cope with volume.

Viewing and Navigating the Structure of a SharePoint document Store

MacroView DMF is the best-available tool for viewing and navigating the structure of a Microsoft SharePoint document store.

MacroView DMF shows you the **complete tree structure** of your SharePoint document store – you simply register your SharePoint web applications and MacroView DMF then automatically (and efficiently) displays all areas of the store that contain document content for which you have access permission, or where you have permission to save new documents and emails.

Unlike the OOB SharePoint web browser UI, MacroView DMF does not just show the tree in a particular site. Instead the DMF display starts at the web application level and shows site collections, sites and sub-sites; for each site you can see the document libraries, and in turn the document sets and the tree of folders and sub-folders contained in each of those document libraries. The MacroView DMF tree-view display extends down to the hierarchies of those metadata columns that have been defined as being available for **Metadata Navigation**.

Unlike the OOB SharePoint web browser UI, the MacroView DMF tree display does not stop when it encounters a **break in permission inheritance**. Instead the DMF display extends all the way down to show any node that contains content for which you have access, or in which you have permission to save – even if there are multiple breaks in inheritance permission on the way down to those nodes.

Type	Name	Document ID	Version	Modified	Checked Out To
PDF	CaseStudyLargeScaleSharePointDM.pdf	MVLEGAL-236079001	1.0	7/9/2012 4:29 PM	
PDF	CaseStudyMSECDMF.pdf	MVLEGAL-236079002	1.0	7/9/2012 4:29 PM	
DOCX	CaseStudyWinstonStrawnDMFLegal.docx	MVLEGAL-236078996	1.0	7/9/2012 4:29 PM	
PDF	CaseStudyWinstonStrawnDMFLegal.pdf	MVLEGAL-236079003	1.0	7/9/2012 4:29 PM	
PDF	MacroView Email Handler - Release Notes.pdf	MVLEGAL-236079004	1.0	7/9/2012 4:29 PM	
DOCX	MacroViewDMF.docx	MVLEGAL-236078997	1.0	7/9/2012 4:29 PM	
PDF	MacroViewDMF.pdf	MVLEGAL-236079005	1.0	7/9/2012 4:29 PM	
DOCX	MacroViewDMFLegal.docx	MVLEGAL-236078998	1.0	7/9/2012 4:29 PM	
PDF	MacroViewDMFLegal.pdf	MVLEGAL-236079006	1.0	7/9/2012 4:29 PM	
PDF	MacroViewDMFv7 5NewFeatures.pdf	MVLEGAL-236079007	1.0	7/9/2012 4:29 PM	

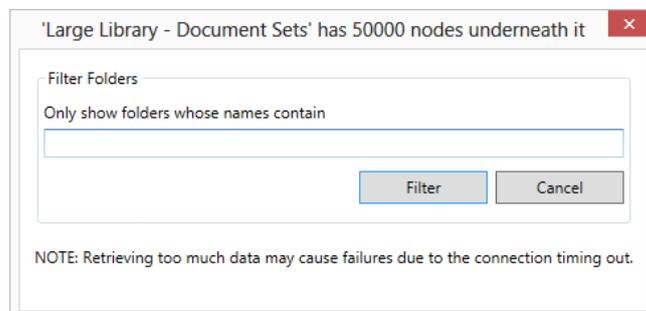
MacroView DMF displays the full SharePoint site and library tree, from web application level down to metadata navigation 'virtual folders'.

Efficient Display and Navigation of Large SharePoint Trees

We designed MacroView DMF to cope with **very large SharePoint document stores**. By 'large' we mean SharePoint stores with a large tree of site collections, sites, libraries, folders etc as well as large numbers of documents. A key aspect of the MacroView DMF tree display is that it never attempts to display more than a threshold number of sub-nodes when you click to expand any node in the tree.

Consider a site that contains hundreds of document libraries. MacroView DMF will prompt you to enter some characters contained in the titles of those document libraries and then display only those libraries whose titles do contain the nominated characters. Another example would be a document library that contains thousands of document sets, or thousands of folders. When you click to expand the document library node, MacroView DMF will prompt you to enter some characters contained in the names of the document sets or folders, and then display only those document sets or folders whose names do contain the nominated characters.

This filtering has excellent performance and minimal bandwidth consumption, thanks to purpose-built MacroView DMF components running at the SharePoint server. The overall result is efficient, not just in terms of machine resources but also for the 'real human' user – who is able to quickly drill down to the library, document set or folder that they want to work with. If the title for a Project library contains both the code number or ID for the Project and the name of the Project (e.g. 'AA1234567 – Sale of Paradise Palms Condominiums') then the user can quickly drill down to a wanted Project library using part of either the *Project ID* or the *Project Name*.



MacroView DMF prompting to filter a document library that contains 50,000 document sets.

As the screen shot above shows, this filtering works even when the number of document sets or folders present in the library exceeds the **List View Threshold**. It also works if you were to expand a folder that contained more than the List View Threshold number of sub-folders.

Familiar, Even Better User Experience

MacroView DMF makes using SharePoint feel familiar for users coming from Windows file shares and Outlook folders.

MacroView DMF makes the SharePoint document store look like a tree of containers, which is familiar to users based on their experience of viewing a tree of folders in Windows Explorer or Outlook. If anything, the experience is better than that in Windows Explorer or Outlook because MacroView DMF provides a number of ways of efficiently navigating around a large SharePoint tree to find the container that they want to work with – e.g. a specific library, folder or document set.

The filtering of sites, libraries and folders described above is one such way; another is the *Search Site Tree* feature of DMF, which enables direct navigation to a site or library by searching for the title (or part thereof) of that site or library. In our example from above, this approach would allow a user to locate the library for a particular Project if he / she can remember either the *Project ID* or part of the *Project Name*. There is no need to first navigate to the site that contains the library, and it does not matter where or how deeply the Project library is nested in the SharePoint tree. MacroView DMF uses the SharePoint or FAST search engine to perform an indexed search, so performance does not degrade as the SharePoint store grows.

Much Better Than the OOB SharePoint web Browser UI

Compared to using the Out-Of-the-Box SharePoint web browser UI, MacroView DMF provides a much better user experience for viewing and navigating the tree structure of a SharePoint document store. MacroView DMF respects the SharePoint security model and so it will not show you any content that you would not have been able to see by using the web browser UI (and many more keystrokes). However MacroView DMF has the following advantages over the web browser UI:

- Lets you view and navigate the SharePoint document store while you work in familiar applications such as Microsoft Outlook, Word, Adobe Reader, etc.
- Shows the complete SharePoint tree, but not 'furniture nodes' such as *Pages*, *Images*, *Site Collection Pages*, *Site Collection Images*, *Style Library*, etc.
- Copes with breaks in permission inheritance.
- Avoids issues with the List View Threshold.
- Is not subject to the 2,000 sub-site limitation.
- Facilitates efficient navigation around a large SharePoint tree.

MacroView DMF Enables Optimal Designs

The inherent limitations of the OOB SharePoint web browser UI tend to have a major and negative impact on the way organizations design their SharePoint document stores. As an example, with the OOB web browser UI you can easily view and click around the tree of libraries and folders and document sets that is contained within an individual site, but navigating to another site is more difficult – unless it's a sub-site you typically need to key in a URL. This shortcoming of the OOB web browser UI leads folks to designs that use lots of folders.

By making it easy for the end user to view the full tree of sites and libraries for which he / she has permission, and efficient to navigate across those sites and libraries, MacroView DMF enables designs that use more sites and libraries.

These designs avoid volume handling issues such as the List View Threshold, minimise issues with excessively long URLs and eliminate the need to delete and re-create storage containers when key metadata such as names of Clients and Projects change during the business lifecycle.

MacroView DMF enables optimal designs that are not really usable if all you have is the OOB SharePoint web browser UI.

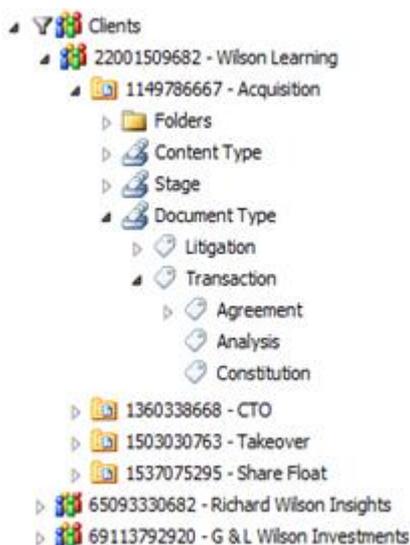
MacroView DMF enables a number of designs for a SharePoint document store that can be optimal in terms of volume handling, manageability and ease of use, but which would not really be usable if all you have to work with is the OOB SharePoint web browser UI. Let's take a look at the following designs in some more detail.

- Tree of sites, sub-sites and libraries
- Libraries with large number of document sets
- Large number of site collections

Tree of Sites, Sub-Sites and Libraries

This design uses a tree of sites and sub-sites; the site tree reflects the fundamental structure of an organization – generally the various lines of business or other activity. Document libraries occur freely across those sites. This design is relevant to a professional services organization – there is a separate site for each Client and a document library or maybe even a sub-site corresponding to each Project.

The document libraries contain whatever metadata columns are relevant to the organization. Because they are centrally managed, user extensible and can be hierarchical, Managed Metadata columns are a good choice for those metadata attributes that are common across multiple libraries. These Managed Metadata columns should be also available for Metadata Navigation, so that the user can 'slice and dice' the documents in the library on the basis of their metadata values.



MacroView DMF displaying sites for selected Clients expanded to show four Project Libraries for the first Client.

Wherever possible these metadata columns should be automatic – i.e. their values are recorded accurately to facilitate subsequent searching for documents, without needing to prompt the user as he / she saves / uploads to that library. There are various techniques for implementing these automatic metadata columns, all of them are facilitated because libraries are granular – i.e. each library can contain automatic metadata values that are specific to a particular Project.

An important aspect of this design is that the sites and libraries should each have a Title that can be used to navigate to the site or library using the filtering and searching capabilities of MacroView DMF. Because sites and libraries have a Title that is separate to their URL the Titles can be long and meaningful – such as the concatenation of *Project ID* and *Project Name* in the example above. There is no pressure to keep the Title short because it does not form part of the URL (we are thinking here about the 255 character limit on length of URLs).

The site and library Titles can be adjusted without having to change the URL of the site or library. This is very useful in business transaction lifecycles, where the name of a Client or Project may become more exact – e.g. as a Project passes from *Prospect* to *Sold* status.

This design has a number of strengths:

- Minimizes risk of exceeding the 255 character limit on URL length.
- Good volume handling – spreads the total document volume across a large number of document libraries, with low risk of exceeding the List View Threshold in any one library.
- Easy to View and Navigate – a MacroView DMF user can filter or search to locate to a specific Client site using part of either the *Client Code* or *Client Name*, or to a Project library using part of either *Project ID* or *Project Name*.
- Low Profiling Fatigue – thanks to the automatic metadata defined at library level, the user does not have to respond to as many (any?) prompts when saving a document.
- Good Governance – new sites and libraries must be created by suitably authorized staff (i.e. not Contributor-level end users).

Libraries with Large Number of Document Sets

Many organizations are attracted to designs wherein document libraries contain a large number of folders. One of the attractions of such designs is that, unlike sites and document libraries, folders can be created by end-users – i.e. any user with Contributor-level permissions.

“Now that we have implemented MacroView DMF, our users are no longer contacting the Help Desk saying they don’t know where their documents are.”

Himanshu Pandya
Manager, PMO & Governance
AEGIS Insurance Services

MacroView DMF is able to display very large folder trees including ones wherein there are multiple breaks in permission inheritance (e.g. the user has permission to certain sub-folders, but not to upper-level folders or to the library itself). As you expand the document library or a folder, MacroView DMF prompts you to filter if there are more than a threshold number (e.g. 50) of folders or sub-folders. MacroView DMF performs this filtering using server-side logic which is not subject to the List View Threshold. As a result, DMF enables designs where there are many thousands of folders at the root of a library, or in any folder within the library.

However as we have already noted, the risk with deeply nested folder tree designs is that the URLs for documents may exceed the permitted maximum length of 255 characters – particularly if folder names are long and meaningful.

Designs that use document sets rather than folders can often be more optimal. The document set is broadly similar to a folder, but from a design perspective has several key advantages over a folder:

- The document set can have its own metadata values. Selected values can be shared to the documents within the document set – i.e. as they are stored documents automatically inherit those shared metadata values.
- Document sets cannot be nested. This addresses one of the key weaknesses associated with design that use libraries with folders, which is that deeply nested folders can lead to document URLs that are longer than 255 characters.
- The document set can be treated a single unit for the purpose of workflow processing.

As you expand a document library that contains a large number of document sets MacroView DMF uses the same prompt-to-filter and server-side filtering logic as described above for libraries with large folder trees. This makes it easy to work with document libraries that contain many thousands of document sets (see the screen shot above from a library containing 50,000 document sets). Provided you are careful to define an appropriate default View for the document set, the document library overall can contain way more than the List View Threshold number of documents. MacroView DMF also has nice support for creating a new document set, which includes capturing its metadata.

This design shares many of the same strengths as the sites, sub-sites and libraries design described above. A key difference is that end-users are able to create new document sets, which may be an advantage in some situations.

Large Number of Site Collections

This design can be relevant to an organization that needs to store a very large volume of documents, because each site collection can potentially be mapped to a separate Content Database. In other words having a large number of site collections maximizes flexibility in terms of managing SQL Server storage. This design is also relevant to organizations that have deployed Personal Sites (also known as MySites) because each Personal Site is implemented as a site collection. Best practice is for these Personal Site Collections to be stored in a separate web application.

As you click to expand a web application that contains a large number of site collections, MacroView DMF uses the same prompt-to-filter and server-side filtering logic as described above for sites, document libraries, folders and document sets. This filtering can be made even more efficient by using the server-side caching feature of MacroView DMF. You will not see a site collection if it does not contain content for which you have access permission, or areas wherein you have permission to save.

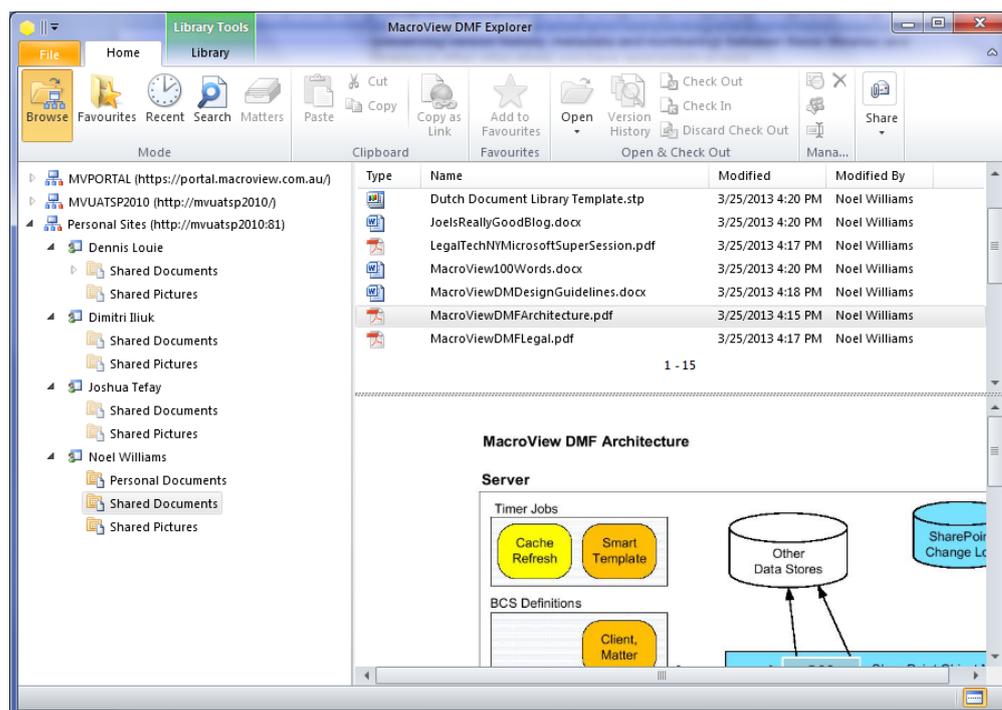
“Thanks to MacroView DMF, user productivity was restored, even improved in several situations. That in turn has helped to make our new SharePoint-based document management a success.”

Mark Buttice
CIO, Mountain States
Employers Council

The *Filter Site Collections by Favorites* command is another way in which MacroView DMF makes it quick and easy to navigate to the particular site collection that you want to work with. This command shows only those site collections that contain one or more favorite nodes (where a favorite can be a site, document library, folder or document set).

MacroView DMF automatically filters the display of site collections that relate to Personal Sites so that as a user you see only your own Personal Site Collection and those for your Colleagues. This can be a real advantage in a large organization where there are thousands of Personal Sites. MacroView DMF respects SharePoint security, so that you will see only documents that are stored in libraries within your own Personal Site and in the Public libraries of your Colleagues' Personal Sites. DMF lets you drag and drop to move documents (preserving version history, metadata and numbering) between these libraries and libraries in other sites where you have appropriate access.

Designs that use large numbers of site collections occur for good technical reasons; MacroView DMF makes such designs user-friendly because the site collections appear in a familiar and intuitive way as top-level nodes in the MacroView DMF display of the SharePoint tree structure.



Web Application with Personal Site Collections displayed in MacroView DMF Explorer. Note Colleagues' sites shown as Read-only and libraries such as *Style Library* not displayed.

Overcoming Objections

Let's take a look at some alternative design concepts for a SharePoint document or email store and consider how they compare to a MacroView-style design. BTW, all these alternative design concepts are drawn from actual implementation Projects.

I need to use document sets because their Metadata helps me navigate...

Some architects protest that they need to use document sets rather than libraries (e.g. a document set for each Project), because the document sets themselves can have multiple metadata attributes that can be used to filter and select the storage area for a particular Project.

It is certainly the case that you can use metadata to filter a View that shows document sets and thereby select relevant document set(s) that you want to access. This filtering is well supported in the File Lists that are displayed by MacroView DMF. However if the sites, sub-sites and libraries style of design is otherwise preferable an alternative approach is to include that key identifying metadata in the Title of the Project library (or site), while also having automatic metadata columns for those attributes defined in the library (or site) as it is created. This will allow the use of the *Filter Sites and Libraries* and *Search Site Tree* features of MacroView DMF to quickly navigate to a wanted site or library.

Why not use a managed metadata hierarchy to handle Client/Project structure?

Other architects argue that support for Managed Metadata by SharePoint 2010 is a good thing, so much so that they seek to use a Managed Metadata column based on a Hierarchical Term Set to capture the complete Client / Project structure for their organization. In an organization that has thousands of clients this design is a case of a good thing taken too far. The resulting Hierarchical Term Set has thousands of entries at the top level (to represent Clients), which causes issues with the display and navigation of the taxonomy. It can also lead to List View Threshold issues as more and more documents are stored in a single large document library.

The optimal alternatives include having a site or library or document set for each Client or Project. These approaches all work efficiently in MacroView DMF and have the further advantage of needing nodes in the tree corresponding to only active values (where a metadata navigation tree displays all possible values).

I want to use Search rather than a tree of SharePoint containers...

Architects who are migrating to SharePoint DM from a traditional DM system such as iManage Worksite or OpenText eDocs / Hummingbird DM, sometimes feel that a search-based approach is preferable to bothering with setting up and maintaining any container hierarchy in SharePoint. Often they also opt to use a very small number of libraries – because this is standard practice in a traditional DM system.

These designs might sound appealing on the basis that they are familiar to users moving to SharePoint from a traditional DM system. Indeed MacroView DMF enables very convenient and intuitive searching for documents across a SharePoint store.

However these designs are decidedly sub-optimal, essentially because they are at odds with the inherent structure of SharePoint, which is a tree of storage containers. When you work within one of these containers (e.g. a document library or document set) you see an accurate list of documents and can filter those documents based on their metadata. Searching in SharePoint 2010 is not so precise, because the SharePoint Search index is not dynamic – it is not updated immediately as documents are added, deleted or moved.

Furthermore, designs that rely on a small number of libraries will almost certainly run into volume handling issues (including issues related to List View Threshold).

Instead of attempting to hide the container tree, it is much better to design it so that it is intuitive to the user and efficient in operation. Hopefully the observations in this article will alert you to the techniques for the design of that tree.

“Email integration is what first attracted me to MacroView, but I have come to realize that the unique value of MacroView is the way it enables users to manage documents and emails across the whole SharePoint farm. Providing users with the same ability to quickly find, filter and use their sites (especially while working in Outlook) with all security pre-applied would otherwise require massive customization. The volume handling and ease-of-use that MacroView provides is a key success factor for large scale SharePoint DM Projects.”

Kyle Connell
ECM Specialist
RKO Business Solutions

For more information visit:

www.macroviewusa.com
www.macroview.com.au