Metalogix

metalogix.com

WHITE PAPER

Addressing Item Level Restore Scenarios for Microsoft Office SharePoint Server 2007 and Windows SharePoint Services 3.0 with Metalogix Selective Restore Manager

By Michael Noel

Published: April 2009

CONTENTS

IMPROVING SHAREPOINT RECOVERY WITH METALOGIX SELECTIVE RESTORE MANAGER

Microsoft SharePoint Products and Technologies have become an integral part of many organizions document management and collaboation environment in recent years. Indeed in many cases, SharePoint has become as I as email or phone functionality for organizations heavily invested in the technology. It is subsequently surprising to find that enterprise backup and restore capabilities are not always included in SharePoint projects. Some of this is due to the expectation that the integrated tools in SharePoint would provide satisfactory enterprise restore capabilities. Unfortunately, however, many of these organizations are discovering that the built-in tools are, in fact, woefully inadequate for many restore scenarios, and are specifically lacking for common usage scenarios such as item-level restore requests.

There are several different methods of backing up and restoring SharePoint content out of the box, all of which differ in their approach in execution. In fact, much of the confusion stemming from SharePoint Backup and Restore is a direct result from the fact that there are so many ways to backup a SharePoint environment and that different backup techniques have implications for how the data can be restored.

Metalogix Selective Restore Manager is a tool that extends SharePoint native backup and restore capabilities to include functionality not available out of the box. It allows for item level restore functionality using a fully supported process that works without the need for a recovery farm and also allows for the ability to quickly compare any backup set with real-time production data. Subsequently, many organizations are finding it to be a useful tool that greatly improves SharePoint restore capabilities.

This whitepaper delves into the various approaches to backing up and restoring SharePoint content using the built-in tools, and how those approaches can be supplemented and improved with the use of Metalogix Selective Restore Manager. It is meant to be used by SharePoint architects and administrators who want to get a better understanding of how backup and restore processes work with SharePoint and how they can be improved to fit the needs of an enterprise organization.

WHY CHOOSE METALOGIX SELECTIVE RESTORE MANAGER ?

A large number of organizations using SharePoint have been making due with simple SQL Server database backups and/or SharePoint farm backups and may be wondering why it would be necessary to consider implementation of new software. Unfortunately, the answer often comes at a time when SharePoint content needs to be restored and administrators are faced with some of the more glaring limitations within the native backup tools.

What makes Metalogix Selective Restore Manager a good value proposition is that it is not a replacement for existing backup strategies, but instead offers significant restore capabilities not possible with the SharePoint restore tools built into the software. It allows administrators to browse through older backup files, select individual items for restore, compare them against current data, and import them easily via a fully supported SharePoint API process. Indeed, it can even extend the restore capabilities of environments that already use advanced backup software such as Microsoft's System Center Data Protection Manager (DPM) 2007.

Before understanding how Metalogix Selective Restore Manager can improve restore capabilities, however, it is first important to understand how the out of the box SharePoint Backup and Restore methods work and what limitations each approach has.

OUT OF BOX BACKUP APPROACHES

SharePoint 2007 has no less than four distinct backup mechanisms available out of the box, and each backup approach uses a unique process for backup and restore. Further complicating the matter is the fact that other SharePoint components on web front-ends have distinct backup needs that aren't met by the integrated tools. Drilling down into the specific backup needs of SharePoint is subsequently critical to understanding how to restore a SharePoint farm or specific SharePoint content.

SHAREPOINT COMPONENTS REQUIRING BACKUP

The following SharePoint components should be included in a comprehensive farm backup plan:

1. SharePoint Content Databases

The most critical components required for backup in a SharePoint environment are the content databases, stored on a local or remote Microsoft SQL Server instance. These content databases contain 99% of all content in SharePoint, including document libraries, documents, lists, web parts, and site structure. They are subsequently the most critical piece in any backup and restore strategy. These databases are often set up to be backed up using SQL Server Maintenance Jobs that allow for Full and Incremental backups to be scheduled.

2. Additional SharePoint Databases

In addition to the Content Databases, SharePoint uses a wide range of miscellaneous databases. Some databases, such as the Config Database, cannot be used to restore a ShrePoint farm, and are subsequently not typically included in backup plans. Other databases, such as the Search databases and the Shared Services Provider database(s) are needed for a full farm restore of Search, Index, Profile, and other farm settings.

3. SharePoint Indexes

SharePoint stores a full text index of all content that is crawled by the SharePoint indexer and made available for searches by its search engine. These indexes are stored as flat files on servers that run the SharePoint Index and Query roles. Backing up the Index files through a supported SharePoint farm backup process such as that provided with the SharePoint Central Admin tool is needed to be able to restore them in the event of failure. If the indexes are lost, they can be regenerated, however, by re-indexing the content, though this can take a while for large environments.

4. IIS Configuration on Web Front-ends

Any custom configuration changes made to Internet Information Services (IIS) on SharePoint Web Front-ends should ideally be backed up, particularly if a large amount of customization has taken place. The IISBACK. VBS script included in Windows Server allows for scripted backup of the IIS Configuration, and can be used to restore the configuration if required.

5. The '12 Hive' on Web Front -ends

The '12 Hive' on web front-ends, which is the sum of files located in the '\Program Files\Common Files\Microsoft Shared\Web server extensions\12' directory and sub-directories contains any custom assemblies, web parts, image files, and other custom components installed directly on web front-ends. Ideally, this folder and its contents would be backed up on a regular basis to prevent content databases from not rendering properly after restoration.

NOTE: Both SharePoint 2007 products, Windows SharePoint Services (WSS) 3.0 and Microsoft Office SharePoint Server (MOSS) 2007 have similar architectural components and similar backup needs. They are subsequently grouped together in this document generically as 'SharePoint 2007.'

INTEGRATED SHAREPOINT CENTRAL ADMIN FARM BACKUP

OVERVIEW OF THE SHAREPOINT CENTRAL ADMIN FARM BACKUP PROCESS

The one backup option directly available from the SharePoint Central Admin Tool is the Farm Backup Option. This option allows for a manual (not scheduled) backup of the entire farm or specific components of a farm.

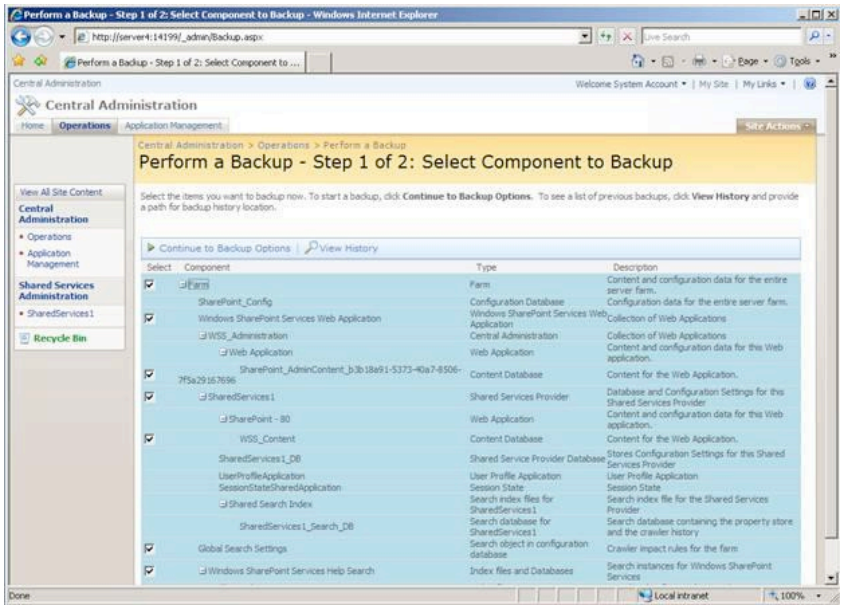

Figure 1 – MOSS 2007 Central Admin Backup Dialog

This process has the following key characteristics:

- A Farm Backup allows for both SharePoint Content Databases, Search, SSP, and Indexes to be backed up.
- While the GUI does not have a mechanism for scheduling a backup, the same process can be initiated using the STSADM tool discussed in more detail later in this document.
- The XML file stored in the root of the backup file location must be maintained or the backup files are effectively unusable, do not lose or delete this file.

This type of backup is primarily used for full farm, disaster recovery or 'point in time' backups of a SharePoint environment.

Under the surface, SharePoint Central Admin backup uses native SQL backup mechanisms for the content databases, which makes them accessible to tools that understand how to restore from SQL backup files.

To initiate a farm backup using the SharePoint Central Admin Tool Farm Backup option, perform the following steps:

1. From the SharePoint Server, open the Central Administration tool (Start – All Programs – Microsoft Office Server – SharePoint 3.0 Central Administration)

2. Select the Operations Tab

3. In the Backup and Restore section, click on Perform a backup

4. From the Select Component to Backup page, check the box next to Farm. If only specific components are needed to backup, narrow down the selection by unchecking boxes.

5. After making the selections needed for backup, click the link labeled Continue to Backup Options

6. In the Select Backup Options page, select a Full backup (Differentials are provided for delta backups of changes made since the last full backup). Enter a backup location and click OK.

7. After starting the backup, SharePoint will display the Backup and Restore Status screen. It may take several minutes for the backup process to appear on the page. The progress can be monitored. Wait for the Progress field to show 'Complete' for all items. Backup History can be viewed by clicking on the View History link.

To initiate a farm restore using the Farm Backup option in SharePoint Central Admin, perform the following high-level steps:

1. From the SharePoint Server, open the Central Administration tool (Start – All Programs – Microsoft Office Server – SharePoint 3.0 Central Administration.)

2. Select the Operations Tab.

3. In the Backup and Restore section, click on Restore from backup.

4. On the backup location page, enter the folder location where the backup manifest file is located.

Click OK when entered.

5. Select the specific backup that will be restored from the list and click Continue Restore Process.

6. Select the particular components that will be restored, keeping in mind that the Configuration database and Central Admin Content database cannot be restored with this utility. When finished selecting restore options, click on Continue Restore Process.

7. In Step 4 of the restore process you can select to either restore the content onto the same configuration (overwrite the existing data) or restore onto a new configuration (which writes it to a new Web Application, allowing both the restored data and the current data to coexist simultaneously. Once the option is chosen, click OK to start the restore process.

8. Follow the progress of the Restore from the Backup and Restore Status page. Just as with the backup, you may need to wait several minutes before the status appears. The restore is complete after the progress indicator shows 'complete'.

CHALLENGES WITH THE SHAREPOINT CENTRAL ADMIN FARM BACKUP PROCESS

There are some significant limitations to the integrated SharePoint Central Admin Farm Backup process, namely the following:

- Issues with any one component within a farm can cause the backup to fail.
- The lack of a scheduling mechanism within the GUI limits the usefulness of the process.
- Loss of the XML file in the root of the backup folder can make the backup worthless.
- Changes to farm topology can make backups obsolete.
- Performing a Farm backup using these tools does not remove the need for file-level backups of the 12 hive or the IIS configuration of front-ends.
- There is no ability to restore granular items, such as individual documents or sites. Central
- Central Admin Farm backup is solely good for disaster recovery scenarios.

SQL DATABASE BACKUP

OVERVIEW OF THE SQL DATABASE BACKUP PROCESS

By in large, the most common method for backup up a SharePoint environment involves backup of the physical SQL database files themselves. This typically includes at least the SharePoint content databases, but often also includes the non-content databases as well.

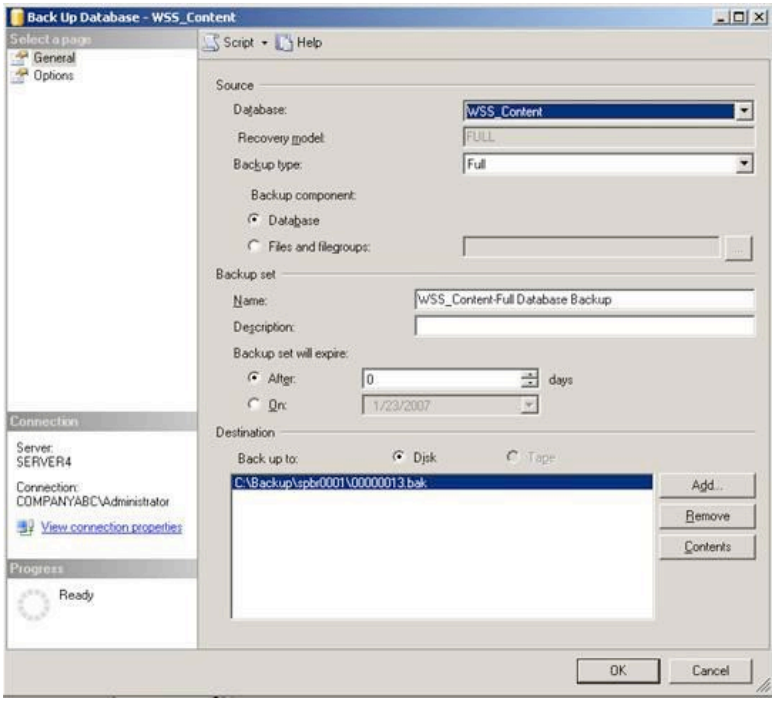

Figure 2 – SQL Server Management Studio Backup Dialog

Backup of SQL content databases is relatively straightforward, and typically involves the creation of a SQL Maintenance Plan that performs the backups. Some organizations also use SQL specific backup utilities that backup the individual SQL databases to a central backup server. Manual backups of SharePoint content databases can also be performed through the following procedure:

1. Choose Start, All Programs, Microsoft SQL Server 2005, SQL Server Management Studio (SQL 2000 or 2008 can also be used, but the process is slightly different.)

2. When prompted, connect to the SQL database where the SharePoint database files are housed and click Connect.

3. In Object Explorer, first expand the desired server and then expand the database folder.

4. Select the desired SharePoint database or databases to backup.

5. Right click the database, select Tasks, and then select Backup.

6. On the General page in the Back Up Database window, review the name of the database(s) being backed and validate that the Backup type option is set to Full.

7. Type the desired name and description for the backup. Select the option Database.

8. In the Destination section, choose the Disk option. Accept the default backup location or remove the existing path and click Add to select a new destination path for the backup.

9. In Select Backup Destination, type in the path on the hard disk where the database backup will be created including the backup file name. Click OK. Alternatively, a Database Administrator can also choose a Backup Device instead of storing the backup on hard disk.

10. It is possible to now initialize the backup or enter advanced backup options by clicking the Options in the Select a page pane.

11. In the Options page and Overwrite media section, maintain the default settings, Back up to the existing media set and Append to the existing backup set.

12. In the Reliability section, choose the option to Verify backup when finished, Perform checksum before writing media and Continue on error. Click Ok to execute the backup.

13. Review the success or failure error message and click Ok to finalize.

CHALLENGES WITH THE SQL DATABASE BACKUP

While providing for a minimum level of backup protection for SharePoint content, there are several limitations to the SQL Database backup approach:

- Restoring SharePoint content requires restoration of the entire SharePoint content database, even if only a single file needs to be recovered.
- SQL Server Maintenance Plans are not enterprise level or centralized, and include only limited notification capabilities.

- If restoring to the same location, the entire Content Database file must be replaced.
- There is no mechanism for determining changes made to a production content database since a
- SQL Backup has taken place.
- Similar to the Central Admin Farm backup, there is no ability to perform granular restore operations of individual sites, libraries, or documents.

STSADM SITE COLLECTION BACKUP

OVERVIEW OF THE STSADM BACKUP PROCESS

Microsoft includes a multi-use command line tool with SharePoint called STSADM. This tool can be useful for multiple SharePoint administration tasks, including SharePoint backup and restore operations. The tool is installed by default into the '\Program Files\Common Files\Microsoft Shared\Web server extensions\12\bin\' directory on a SharePoint server, and is invoked from the Windows command prompt.

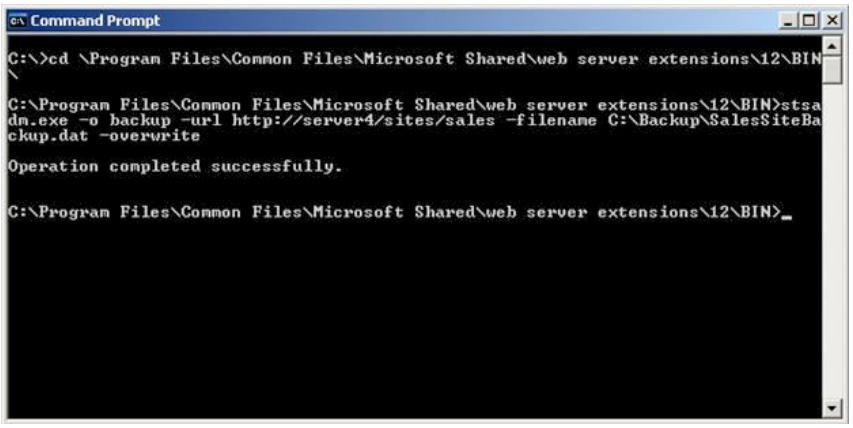

Figure 3 –STSADM Command Line Backup

There are three completely different backup and restore processes that can be invoked using STSADM. The first backup process, called a 'catastrophic backup', is effectively the same type of backup process as that which is invoked using the SharePoint Central Admin Tool. Backup of a farm using the catastrophic backup approach uses syntax similar to the following:

- Stsadm.exe –o backup –directory \\fileserver\Farmbackup –backupmethod FULL

A separate, completely unique backup process is possible with the use of STSADM. This type of backup process converts Site Collections to flat file backups, which can be ported to different site collections or even different farms. These types of backups are full fidelity, meaning they preserve all aspects of the source site collection, and can be invoked using syntax similar to the following:

- Stsadm.exe –o backup –url http://server1/sites/testsite -filename z:\testsitebackup.bak

To further complicate things, STSADM has a third type of backup process that also uses a separate mechanism than the other two approaches. This third backup method allows for backup of individual sub-web elements, such as individual document library backups or backups of individual sites. It uses a mechanism identical to that which is used by the SharePoint Designer 2007 tool, which allows for partial

fidelity backups of individual elements directly from running sites. This mechanism was partly derived from the original Content Deployment API.

Because this third type of backup is partial fidelity, not all site elements are preserved, and some aspects of a site such as individual document level security is not maintained. As a technology, the Content Deployment API style backups never fully matured, and moving or restoring data using this approach is rife with issues that have been encountered over the years. It is subsequently only useful for very specific circumstances where specific elements are backed up. The following sample syntax is used for backup using this STSADM process:

- Stsadm.exe –o export –url http://server1/sites/testsite/doclib1 -filename z:\doclibbackup.bak

CHALLENGES WITH THE STSADM BACKUP PROCESS

The challenges to catastrophic backup using STSADM are the same as those outlined with the SharePoint Central Admin tool process. Additionally, the following challenges exist for Site Collection backups using the STSADM tool:

- Backups cannot be made to be any more granular than the Site Collection Level.
- The process does not scale well toward very large Site Collections.
- Site Collection Backups can be time-consuming.
- There is not automated way of backing up Site Collections with STSADM, it must be scripted via batch files or some other automated approach.

Finally, the third STSADM sub-web backup process has the following distinct disadvantages:

- Partial fidelity backups are not useful in most cases as not all site structure and permissions are retained.
- The same automation and scalability concerns exist.
- This type of backup mechanism can place heavy load on the SharePoint server for large sites, and often must be scheduled during maintenance windows.

MICROSOFT SYSTEM CENTER DATA PROTECTION MANAGER (DPM) BACKUP

OVERVIEW OF THE DPM BACKUP PROCESS

Microsoft has released a SharePoint aware backup product called System Center Data Protection Manager (DPM) 2007 in recent years. This product has some unique characteristics which make it popular among clients invested heavily in SharePoint, namely the ability to ‘snapshot’ SharePoint farm content on a regular basis, as often as four times an hour. It does this by integrating with the Windows Volume Shadow Copy Service (VSS), which keeps track of the block level changes made on a server and synchronizes the delta of those changes only, allowing for a much more efficient backup process.

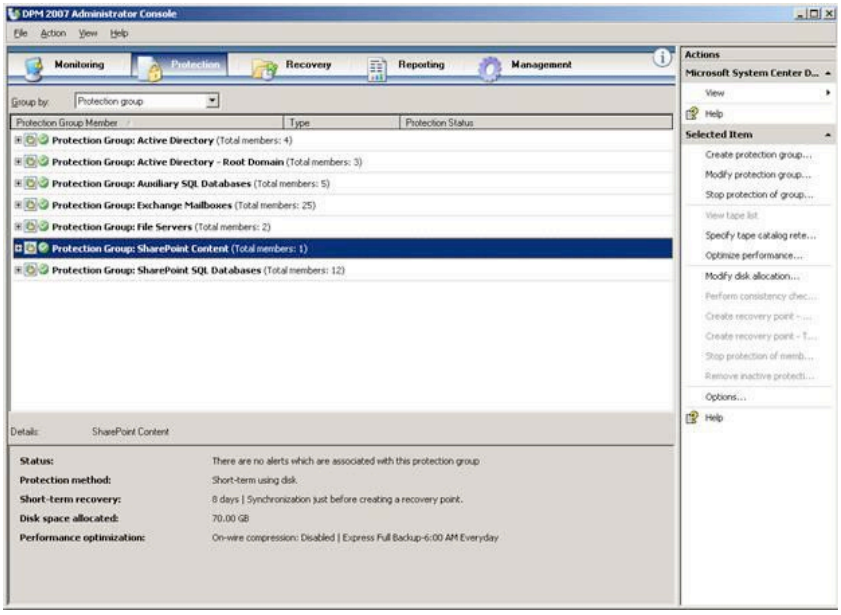

Figure 4 - Microsoft Data Protection Manager Admin Console

DPM allows for both short-term and long-term storage options for the SharePoint content. Typically, the short-term option is to backup SharePoint farm content to disk, where it is maintained for a specified period of time. After this time period, the content is then automatically transferred to long- term tape. DPM has the flexibility to restore SharePoint content databases at any snapshot point in time, allowing for a great deal of restore flexibility.

Challenges with the DPM Backup Process

The one significant challenge to relying solely on DPM to provide for SharePoint restore functionality is that DPM requires the use of what is called a 'recovery farm' to restore individual items into SharePoint. The recovery farm is a SharePoint farm that is specifically setup so that content databases can be temporarily restored to the farm. The entire content database must be restored to the recovery farm before an individual item can be extracted out of it and imported back into the production farm. This process, while made to be relatively seamless, does require additional hardware for a farm and can take a while to complete depending on the size of the content database. In addition, purchase of DPM software and agent licenses for each server being backed up is required when deploying DPM, which can add additional expense and overhead.

Table 1: Backup Method Comparison

	Full Farm Recovery (Including Index)	Full Fidelity Content Restore	Backup Automation Built-in	Item-Level Recovery	Ability to Restore from multiple backup types
Central Admin Backup	▶	▶			
SQL Maintenance Plan Backup		▶	▶		
STSADM Catastrophic Backup	▶				
STSADM Site Collection Backup		▶			
STSADM Content Deployment Export					
Data Protection Manager	▶	▶	▶	(Via Recovery Farm)	
Metalogix Selective Restore Manager + Out of box tools	▶	▶	▶	▶	▶

BENEFITS OF SHAREPOINT BACKUP WITH METALOGIX RESTORE MANAGER

Integration with Existing SQL, DPM, or SharePoint Farm Backup Files

One of the biggest advantages obtained by using Metalogix Selective Restore Manager lies in its ability to be able to restore directly from backup files created by SQL, DPM, STSADM, or SharePoint Central Admin. Metalogix Selective Restore Manager has the capability to integrate directly with any or all of the backup approaches previously discussed in this document, which provides for a great deal of deployment flexibility and allows it to integrate more easily in existing environments.

So, a SharePoint administrator can effectively continue to use an existing backup process, and as long as they have access to the original backup files. Environments that have significant investment in DPM or even third-party SQL backup software can take heart in the fact that their tools can simply be used in conjunction, rather than in replacement, to Selective Restore Manager. In addition, the fact that the tool also does not create or use any proprietary backup formats means that standard, supported tools can be utilized if needed to extract data.

Fully Supported SharePoint API Recovery Without the Need for a Recovery Farm

Another significant advantage to Metalogix Selective Recovery Manager is the fact that all restore operations performed with the tool strictly adhere to fully supported SharePoint API calls, and data is never directly injected into the SharePoint databases through unsupported methods. This ensures a high level of support from Microsoft, as they have indicated that they do not support any type of method of accessing or moving data that does not use the SharePoint API.

Metalogix Selective Restore Manager also does not require the use of a recovery farm for restore operations. This is in direct contrast to item-level restores performed from DPM, which require the use of the extra recovery farm as a temporary landing location for content databases. The fact that a recovery farm is not required, simplifies restore operations significantly and reduces overall hardware costs.

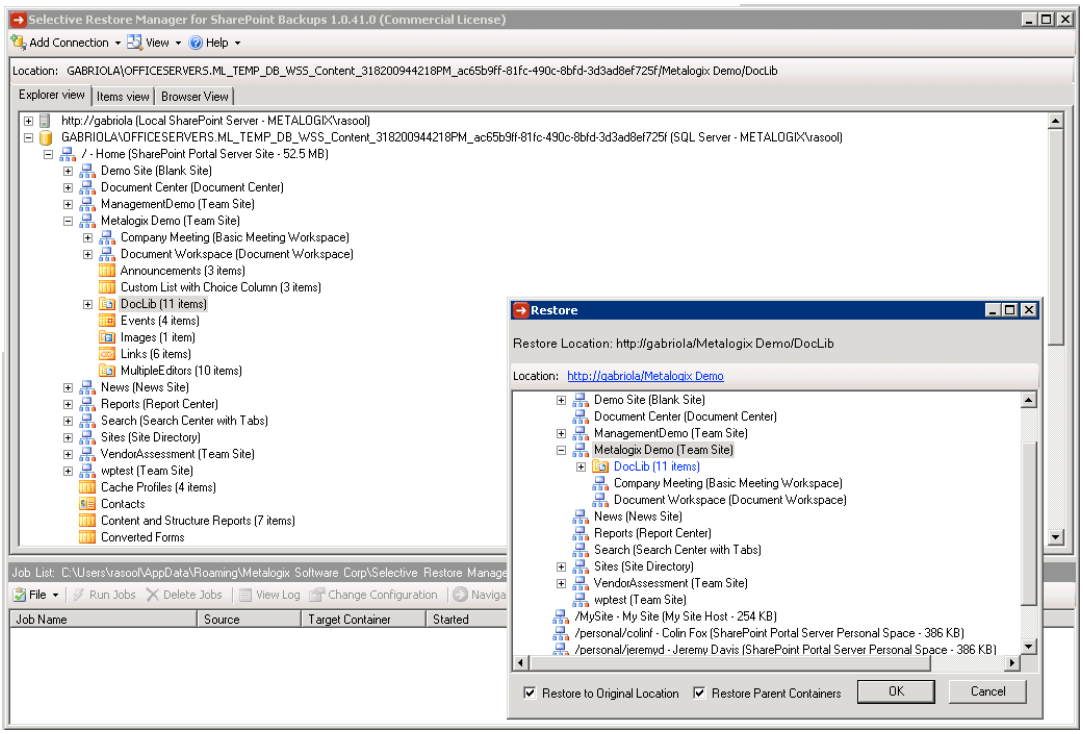

Figure 5 – Metalogix Selective Restore Manager (Restore Dialog)

Restore from Any Content Database to Any Farm

Metalogix Selective Restore Manager is highly versatile and flexible in that it allows content pulled from any content database to be restored to any farm, even those not related to the source of the backup files. The original source site does not even need to be running, the only requirement is that the database backup files be mounted to any running SQL server instance.

Flexible Recovery Options to Restore Entire Site Structures or Individual Items

One major advantage to using Metalogix Selective Restore Manager is the fact that you can use the tool to do full fidelity restores of specific sites, document libraries, or simply individual pieces of content. The tool allows for the flexibility to restore those subsets of content to the same location or to a new location in a different site or site collection. This type of flexibility allows for a restructuring of content or even a reorganization of sites within individual site collections.

Browse Through SQL Backup Files and Select Items Directly from the Metalogix Interface

As long as the backup files created by SQL maintenance plans, STSADM exports, or SharePoint Central Admin backups are available, they can be opened by the Metalogix Selective Restore Manager tool and their contents browsed as easily as navigating through file folders. No special tools are required and no catalogs need to be generated, which makes selecting files for restore that much easier.

Run Differential Reports to Quickly Determine What has Changed Since the Last Backup

Another huge feature enhancement with Metalogix Selective Restore Manager is the ability to run differential reports on the backup files compared to production data. This allows administrators to quickly determine which files have changed since the backup file was created, and lets them make intelligent choices on what to recover and what shouldn't be overwritten.

Simple and Easy Deployment, No Dedicated Server Required

Administrators will greatly appreciate the lack of a requirement for a dedicated console for the Metalogix Selective Restore tool. Indeed, all the tool requires is a workstation with a running Dotnet framework implementation installed. No special server agents, web consoles, or scripts are required, which greatly reduces the footprint on the environment and lessens potential conflicts on the servers. Windows OS versions supported to run the client include Windows Server 2003, 2008 (x86 and x64), Windows XP, and Windows Vista. Selective Restore Manager also supports remote deployment and recovery operations with the installation of a Metalogix web service on target SharePoint web front ends.

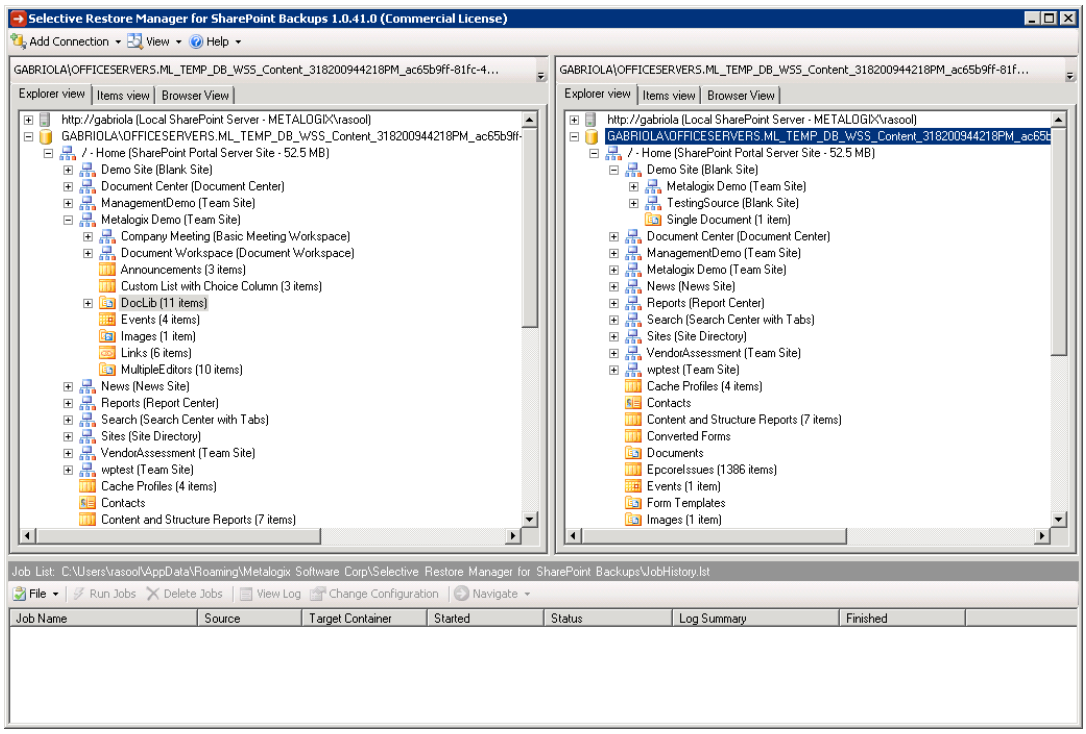

Figure 5 – Metalogix Selective Restore Manager

CONCLUSION

There are a myriad different ways of backing up SharePoint content, each with specific disadvantages that limit their ability to be an effective part of an enterprise backup plan. Fortunately, however, the vast majority of these disadvantages can be taken out of the picture with the integration of Metalogix Selective Restore Manager, which can take existing backup files and use them to perform restores of SharePoint content.

Using existing backup files allows the Metalogix tool to be highly flexible and does not require an environment to retool their backup approaches. Other key benefits of using the tool such as console-less installation, differential report creation, flexible recovery options, SharePoint API use, and full support of multiple backup formats positions it as the gold standard SharePoint recovery solution available on the market today.

ABOUT THE AUTHOR

Michael Noel – Partner/Principal Consultant – Convergent Computing

Michael Noel (MCITP) is an internationally recognized technology expert, bestselling author, and well known public speaker on a broad range of IT topics. He has authored multiple major industry books that have been translated into over a dozen languages worldwide. Significant titles include SharePoint 2010 Unleashed, Exchange Server 2010 Unleashed, Windows Server 2008 R2 Unleashed, ISA Server 2006 Unleashed, and many more. Currently a partner at Convergent Computing (www.cco.com) in the San Francisco Bay Area, Michael's writings and extensive public speaking experience across six continents leverage his real-world expertise helping organizations realize business value from Information Technology infrastructure.

ABOUT METALOGIX

Metalogix is the trusted provider of innovative content lifecycle management solutions for Microsoft SharePoint, Exchange and Cloud platforms. We deliver high-performance solutions to scale and cost-effectively manage, migrate, store, archive and protect enterprise content. Metalogix provides global support to thousands of customers and strategic partners and is a Microsoft Gold Partner, a managed partner in Microsoft's High Potential ISV Group and GSA provider. Metalogix is a privately held company backed by Insight Venture Partners and Bessemer Venture Partners.

METALOGIX
5335 Wisconsin Ave NW, Suite 510, Washington DC 20015
sales@metalogix.com | www.metalogix.com | 1.877.450.8667