SharePoint Migration
Cleanup and Pre-Migration Checklist for Success
SHAREPOINT MIGRATION CLEANUP AND PRE-MIGRATION CHECKLIST FOR SUCCESS

CONTENTS

INTRODUCTION .................................................................3

PRE-MIGRATION OF USERS AND GROUPS .........................3
Identify Influential Users .................................................3
Remove Orphan Users ....................................................4
Cleanup Unused SharePoint Groups .................................5
Find the Inactive Users ...................................................6

SCRUBBING YOUR SHAREPOINT CONTENT ....................6
Cleanup Unused Sites .....................................................6
Breakup Large Site Collections .......................................8
Manage Content Database Sizes .....................................8
Find Large Sites (and promote to Site Collections) .............9
Find Small Site Collections (and demote to subsites) ........9
Identify and Manage Large Lists ....................................9
Cleanup Content ..........................................................10
Remove Duplicate Files ..................................................10
Archive the Audit Log .....................................................10

SHAREPOINT DESIGN ELEMENTS .............................11
Web Part Decisions .......................................................11
Where are the Workflows? .............................................11
Are the Content Types being used .................................12
Insight into the SharePoint Hierarchy ..............................12

ABOUT METALOGIX ..................................................13
INTRODUCTION

No migration to SharePoint 2013 should be taken lightly. SharePoint content has grown at a remarkable rate of 75% annually. With that growth, your company’s users demand 100 percent up time and expect few or zero lapses in their ability to access business-critical sites. The IT department, meanwhile, must ensure that only the right people can access certain content while overseeing a smooth, error-free transition that will minimally impact their workload.

After helping hundreds of clients migrate to SharePoint 2013, we offer a simple message: “Cleaning up your current SharePoint environment today and planning out your future migration to SharePoint 2013 will save you time, money and headaches.” The complexity lies, however, in determining what to cleanup and how to plan for a successful migration.

Cleanup offers two major gains for SharePoint 2013 migration. First, it will increase the performance of your current environment by removing resource-intensive tasks. Second, it will expedite the overall time required to move from your current SharePoint environment to SharePoint 2013.

Cleaning up SharePoint is ultimately a two-part process: analysis and action. After extensive work with clients, we’ve created two powerful products that address each part: ControlPoint and Content Matrix. In this guide, we’ll demonstrate how to use ControlPoint to gain a deeper understanding of a SharePoint environment by analyzing what users have created, identifying sites that need optimization and providing a comprehensive overview into what users should be accessing. Once that analysis is complete, we’ll show how Content Matrix can be used to make changes to the environment, such as breaking up site collections and migrating the content.

PRE-MIGRATION OF USERS AND GROUPS

No SharePoint environment would be complete without users. During the pre-migration analysis, identifying inactive and active users and groups not only helps you understand their post-migration expectations, it also gives administrators an opportunity to ask why these users are not using SharePoint for their collaboration needs. The result is less overhead and improved communication with your SharePoint users as you migrate to SharePoint 2013. The following sections explain how to identify these users.

IDENTIFY INFLUENTIAL USERS

Including your influential users in the planning and implementation stages of your migration can prove beneficial to the success of the migration. Users can provide both “on-the-ground” feedback during the migration and communicate their SharePoint 2013 requirements and expectations.
Start by identifying the most active SharePoint users and sites in your environment using the Most/Least Activity report in ControlPoint. Communicate with these users throughout the process and help get early “buy-in” so that they will in turn communicate any changes or new features to their teams.

Alternatively, ControlPoint can also be used to create a select SharePoint Migration group of trusted users who have Full Control. That group can then be used as a resource during the migration – for example, as a way to communicate with business owners about the status of their migration.

REMOVE ORPHAN USERS

One often overlooked area of the cleanup process is removing orphan users. These are employees who are no longer with the company and were removed from Active Directory but still have a SharePoint user account and associated permissions for sites or content.

Orphan users may also have active SharePoint alerts or My Sites. While orphan users can no longer access the content, if they’ve created alerts, My Sites, and other mechanisms in your environment, those can impact SharePoint’s hardware resources. Cleaning out these users before your migration will help teams gain back resources and better predict SharePoint user growth.

Use the Orphan Domain User report in ControlPoint to run reports across your entire farm or single sites or lists. This will ensure quick identification of users that have been granted permissions in SharePoint but are no longer in Active Directory. If your SharePoint policy is to “disable” users in Active Directory, the ControlPoint report can ignore those users and find only the ones that your company considers orphans.
Once identified, orphan users can be deleted automatically or manually removed on an individual basis. When you delete a user from a site, list or site collection, you can also choose to delete a user’s SharePoint alerts, their “My Site” or completely remove them from the entire Site Collection, i.e. delete them from the site collection’s All People and Groups List.

**CLEANUP UNUSED SHAREPOINT GROUPS**

When new site collections are created, often the default SharePoint Groups (owners, contributors, etc.) are also created (intentionally or not). Frequently, only a few of those groups are used. ControlPoint’s *SharePoint Group* analysis will highlight all unused SharePoint groups. Migrating these unused groups can be costly and delay the migration. It’s recommended that any migration plan include a task to review group activity.
**FIND THE INACTIVE USERS**

ControlPoint’s *Inactive Users* report can be used to identify all users in the environment that have no activity during a specified period of time. This information is valuable as you plan for your migration and can help reduce the number of Client Access Licenses (CALs) needed and reduce your user footprint. It’s also an opportunity to reach out to inactive users who are still with the organization to ask why they’re not using SharePoint.

**SCRUBBING YOUR SHAREPOINT CONTENT**

When planning a migration, SharePoint teams make decisions about what content to move and what to leave behind, for example, decisions need to be made about how to handle low activity sites or site collections. Duplicate or deprecated content as well as large site collections will also need to be cleaned up in order to reduce database size. Using ControlPoint’s *Site Storage Analysis*, administrator’s benefit from a clearer view of their SharePoint content, and can quickly identify where to find savings and where to improve the user experience.

**CLEANUP UNUSED SITES**

Unused sites that contain almost no content or have very limited activity are excellent candidates for archiving as you prepare the migration process. ControlPoint makes it easy to isolate sites that demonstrate either of these characteristics. For example, you can find under-utilized sites within any scope (Site Collection, Web Application or even the whole farm) using the *Most/Least Storage* reports. You can even filter the report to display certain site definitions.
The *Most/Least Storage* report will return the sites or site collections in size-rank-order and will also include information about when the content was last updated. You may find that a site doesn’t have a lot of content (yet) but that it recently was updated. This information may lead you to further analysis – what sort of activity has this site exhibited? Is it a new site that is just starting to be used/configured?

<table>
<thead>
<tr>
<th>Select</th>
<th>Rank</th>
<th>Name</th>
<th>URL</th>
<th>Created By</th>
<th>Size (MB)</th>
<th>Page Views</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>1</td>
<td>foxrot</td>
<td>(<a href="http://sarabh/sites/beta/foxrot">http://sarabh/sites/beta/foxrot</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>2</td>
<td>bravo</td>
<td>(<a href="http://sarabh/sites/beta/bravo">http://sarabh/sites/beta/bravo</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>3</td>
<td>charlie</td>
<td>(<a href="http://sarabh/sites/beta/charlie">http://sarabh/sites/beta/charlie</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>4</td>
<td>delta</td>
<td>(<a href="http://sarabh/sites/beta/delta">http://sarabh/sites/beta/delta</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>5</td>
<td>echo</td>
<td>(<a href="http://sarabh/sites/beta/echo">http://sarabh/sites/beta/echo</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>6</td>
<td>golf</td>
<td>(<a href="http://sarabh/sites/beta/golf">http://sarabh/sites/beta/golf</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
</tbody>
</table>

Use the *Most/Least Activity* report to find under-utilized sites. This report will identify sites with the least activity over a certain period of time and include information about site activity, size and when it was last updated. It might be prudent to do some further investigation into these sites – just to be sure before you decide they are no longer needed.

In the following demo environment report you can see that both the “Accounting” and “Alpha Snack Foods” sites haven’t had any page views, however, both have been recently updated and contain some content. In instances like this, users may be archiving sites where content needs to be stored for compliance reasons only.

<table>
<thead>
<tr>
<th>Select</th>
<th>Rank</th>
<th>Name</th>
<th>URL</th>
<th>Created By</th>
<th>Size (MB)</th>
<th>Page Views</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>3</td>
<td>alpha</td>
<td>(<a href="http://sarabh/sites/beta/alpha">http://sarabh/sites/beta/alpha</a>)</td>
<td>SHAREPOINT System</td>
<td>0.26</td>
<td>0</td>
<td>10/11/2012</td>
</tr>
<tr>
<td>Select</td>
<td>4</td>
<td>bravo</td>
<td>(<a href="http://sarabh/sites/beta/bravo">http://sarabh/sites/beta/bravo</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>5</td>
<td>charlie</td>
<td>(<a href="http://sarabh/sites/beta/charlie">http://sarabh/sites/beta/charlie</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>6</td>
<td>delta</td>
<td>(<a href="http://sarabh/sites/beta/delta">http://sarabh/sites/beta/delta</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
<tr>
<td>Select</td>
<td>7</td>
<td>echo</td>
<td>(<a href="http://sarabh/sites/beta/echo">http://sarabh/sites/beta/echo</a>)</td>
<td>SHAREPOINT System</td>
<td>0.22</td>
<td>0</td>
<td>7/16/2011</td>
</tr>
</tbody>
</table>
Site Collections (or sites) often outgrow their original size. This suggests that users are actively engaged in SharePoint. It also points to a potential resource issue. Very large site collections can impact performance which in turn will likely hinder your migration plans.

Typically, it makes sense to split the Site Collection during the migration process. Or you may want to split the content so that some of your users can migrate to SharePoint 2013 and start using it, while other groups use SharePoint 2010.

Using the Most/Least Storage analysis in ControlPoint, you can quickly ascertain which Site Collections are the largest.

ControlPoint’s Content Database Storage report gives insight into both the size of the Content Databases and details about the content contained in each. For example, the above report details the size of the audit log, the size of the log file and each site collection in the content database.

If the administrator finds content that needs to be moved elsewhere, ControlPoint’s Move to another Content Database action lets you re-balance the Content Database before migration. Alternatively, the information can be used to identify opportunities for moving site collections into different content databases on the destination environment.
SHAREPOINT MIGRATION CLEANUP AND PRE-MIGRATION CHECKLIST FOR SUCCESS

FIND LARGE SITES (AND PROMOTE TO SITE COLLECTIONS)

As part of a migration plan, you might decide to reorganize your sites and promote some of them to site collections based on their size. ControlPoint Most/Least Storage report lets you identify the sites with the largest amount of content. The report can be configured to only show sites that are above a certain size based on your company’s storage policy. For example, all sites above 100 GB might be candidates for promotion to Site Collection.

Microsoft has documented best practices for the size of Site Collections that should be followed. http://technet.microsoft.com/en-us/library/cc262787(v=office.15).aspx#SiteCollection

FIND SMALL SITE COLLECTIONS (AND DEMOTE TO SUBSITES)

Another opportunity for cleanup is to identify small site collections and demote them to subsites. Often user requests for sites are satisfied by creating site collections. Over time, it might become apparent that those site collections could or should have been created as subsites of a different site collection. Using the ControlPoint Most/Least Storage report, you can easily rank the sites collections by size and determine the candidates for demotion from site collections to sub-site.

Another outcome of this analysis might be that you choose to just delete those site collections since they have very little use.

IDENTIFY AND MANAGE LARGE LISTS

Microsoft’s best practices for numbers of items in a list or folder are often exceeded due to the unanticipated growth of SharePoint. Using the Most/Least Storage reports in ControlPoint, you can identify lists with excessive number of items. During the migration you are then in a position to split those lists.

You can also use the List Property report to analyze the columns to determine whether they are approaching the Microsoft recommended list view thresholds (2,000 or more items with 20 or more columns).
CLEANUP CONTENT

ControlPoint also provides insight into how current the content is. For example, the List and Library Storage report includes last accessed and/or last modified metadata. With that knowledge, you can choose to archive some of the content as part of the migration effort and/or the migration can be configured to skip older content that is no longer needed.

Another effective way to maintain control over content growth is to implement policies around versioning. Using ControlPoint, you can gain insight into the storage impact of the document versions. The more versions an item has, the slower the migration of the item will be. With ControlPoint, you can use Set List Properties to configure versioning across any scope in the farm. ControlPoint will configure SharePoint to automatically prune the versions from the content.

REMOVE DUPLICATE FILES

It’s not uncommon for SharePoint to be the repository for the same document in multiple places. To improve the migration performance and reduce user confusion, use the Duplicate Files report in ControlPoint to identify and remove duplicate files.

ARCHIVE THE AUDIT LOG

Using the Audit Log archiving capability in ControlPoint, audit data can be removed from source content databases and stored in a standalone SQL database. This provides access to audit information even after you’ve moved to the new SharePoint version. Once the data is in the archive, you can use ControlPoint’s powerful audit log reports. SQL Server Reporting Services or other tools can also be used to create audit reports after migration. This ensures a continuous audit trail even after the farm is migrated.
SHAREPOINT DESIGN ELEMENTS

WEB PART DECISIONS

Web parts are used everywhere – your SharePoint environment probably includes Microsoft web parts, custom web parts, and third party web parts.

As you plan your migration, it’s important to get an accurate assessment of the impact of any custom web parts used in your environment. This step will help you determine whether to update a custom web part so it works in 2013 or deprecate the component as its functionality is no longer needed or has been replaced by a new capability in SharePoint 2013.

Web part usage analysis will also assist in determining if support/maintenance on third party web parts should be continued or not.

Using ControlPoint’s Web Part analysis, you can gain global insight into which web parts are being used and which sites they are being used on. The report can even filter for the Closed or Hidden web parts.

The Web Part report can be run by site administrators or distributed to them so that they can understand which components will be migrated and which will not. This will set end user expectations and prevent miscommunication during migrations.

WHERE ARE THE WORKFLOWS?

The Workflow analysis in ControlPoint can be filtered by a number of parameters including the state, date, duration, etc. The report includes details about workflow definitions, instances and their state. This information is particularly important for migration since it displays the details for all the in-progress workflows that cannot be migrated.
ARE THE CONTENT TYPES BEING USED

Content types are a powerful tool in SharePoint that enable organizations to organize, manage, and handle content in a consistent way across site collections. Gaining insight into the custom content types and where they are being used will help determine how the data will be managed once you migrate. Investigating the use of Content types will also help to highlight opportunities for creating enterprise content types that will provide broader scope.

The Content Type report in ControlPoint can be filtered to show all content types or ones that are defined but not used. This information can be used to clean up the content types prior to the migration.

INSIGHT INTO THE SHAREPOINT HIERARCHY

The SharePoint Hierarchy report provides a simple hierarchical view of all the SharePoint objects and highlights all uniquely secured content with in the scope selected. This information can drive an assessment of the permissions model being used in the existing SharePoint environment and whether it should be preserved or modified when you migrate.

Note:
The reports mentioned in this document can be filtered by any scope in your SharePoint environment even across farms and by many different criteria including size ranges, users, timeframes, site definitions, URLs, wildcards, etc.
ABOUT METALOGIX

Metalogix is the leader in management tools for Microsoft content and collaboration platforms, including SharePoint, Exchange, and Office 365. Metalogix tools provide a better way to monitor, migrate, store, synchronize, archive, govern and backup content.

14,000 customers on 7 continents rely on Metalogix’ best of breed suite of tools and 24x7 LIVE support to improve availability, reduce risks, and lower costs of their mission-critical Microsoft platforms.

Metalogix is a Microsoft Gold Partner, Microsoft Public Sector managed partner and certified GSA provider. Headquartered in Washington DC, Metalogix is backed by Insight Venture Partners and Bessemer Venture Partners.