INFORMATION MANAGEMENT BRIEF

Mimosa Delivers Much Needed R&R (Retention and Recovery) for SharePoint

Date: March, 2009

Author: Brian Babineau, Senior Analyst, Lauren Whitehouse, Analyst

Abstract: Microsoft SharePoint is quickly becoming the next 'killer' enterprise application. Although mostly used for file sharing, SharePoint has plenty of features that will most certainly alter how information is managed in the future. The lingering question for many companies is how to manage SharePoint in the face of rapid data growth and strict business requirements such as regulatory compliance and electronic discovery. Mimosa Systems has added a module to its NearPoint platform to facilitate more efficient retention and recovery in SharePoint environments.

Introduction

Recent ESG research of over 1000 worldwide organizations indicates that nearly one out of two respondent companies has deployed or is planning to deploy Microsoft Office SharePoint Server (MOSS).[1] MOSS's success—it's a $1B-plus business for Microsoft—can be attributed to its versatility, which enables customers to solve a range of problems from replacing corporate file shares to automating business process workflows. Regardless of how companies are using MOSS, employees are using it to store important information, contributing to the estimated 25% annual data growth of MOSS implementations.

One the biggest myths in the marketplace is that MOSS is just another application that can be added to IT's management, security, and data protection schema. While it does have to be managed by IT—no easy task given its multi-tier, multi-part architecture including the underling SQL Server database—it is not just 'another application.' Of those organizations that have implemented SharePoint, over 52% indicated that it was one of their top *five* initiatives over the past twenty four months and an additional 30% said it was one of their top *ten* initiatives over the same period.

Already overburdened IT staffs must find ways to manage this new, or in some cases, just implemented critical application. MOSS data growth could slow application performance, which in turn, could impact productivity. The increase in data stored within MOSS is also likely to elongate backup and recovery times. Further, a subset of the information contained within MOSS, such as an employment contract, may be subject to record retention regulations or be requested as part of an electronic discovery matter, forcing IT organizations to retain it for a specified period of time.

To address these challenges, Mimosa Systems has expanded its NearPoint archiving and data protection solution to support MOSS environments. The newest addition to the NearPoint family enables customers to easily retain the right information while providing quick recovery for any subsets of MOSS application data in the event of data loss and corruption.

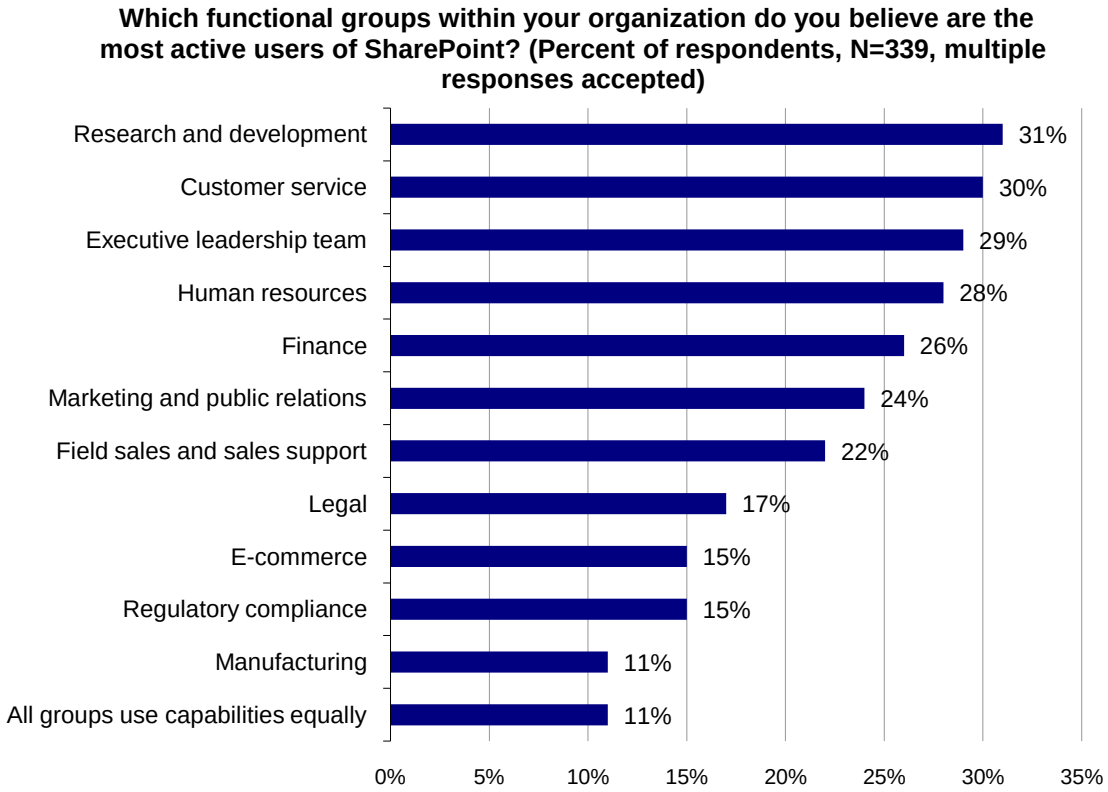
MOSS for the Masses

The primary reason organizations are implementing MOSS, according to ESG research, is to enhance internal project collaboration and improve information sharing between employees located in geographically dispersed offices. MOSS's web interface, document versioning capabilities, and customizable workflows make it an ideal solution for almost any department to centralize data as part of a project or simply to facilitate more efficient information sharing (see Figure 1). Some workgroups may leverage MOSS's advanced capabilities, such as the

[1] All statistics within this Brief are sourced from: ESG Research Report, *The Impact of Microsoft SharePoint on IT Infrastructure and Information Management*, March 2009, unless otherwise noted.

Business Data Catalog which enables employees to connect to and analyze information from external applications via a MOSS web part.

FIGURE 1. MOST ACTIVE MICROSOFT OFFICE SHAREPOINT USERS, BY DEPARTMENT



Source: ESG Research Report, The Impact of Microsoft SharePoint on IT Infrastructure and Information Management, March 2009

As MOSS becomes more popular, its drawbacks are starting to manifest—some of the issues arise due to its ease of use. For example, MOSS sites are fairly simple to set up, enabling project managers to quickly establish a location where data can be posted for all team members to access. Individual departments may have several MOSS sites as different projects include sensitive data that must be segregated for security purposes. The ease of setup can lead to MOSS site sprawl—a situation that is starting to emerge as 44% of current MOSS users are running the application in at least six different locations. While MOSS is designed to centralize information, its simplicity can spur the separation of content by department or office location.

Such sprawl, however, is an example of a manageable implementation issue. Unfortunately, customers do not have any control over MOSS's architectural limitations—some of which can make the application very difficult and costly to manage. As an example, MOSS does a great job at document versioning as employees can always see the latest copy of a file and whether or not someone is currently working on that version. Although only the most recent version of a document is displayed, MOSS saves all prior versions within the backend content database. If a particular site has several thousand or millions of files, MOSS stores all of the initial copies, as well as all of the versions. This can slow MOSS performance as well as other application services such as search—both of which negatively impact information access and employee productivity. Further, IT has to buy additional storage capacity to keep up with data growth as a result of all the files and versions of files being saved.

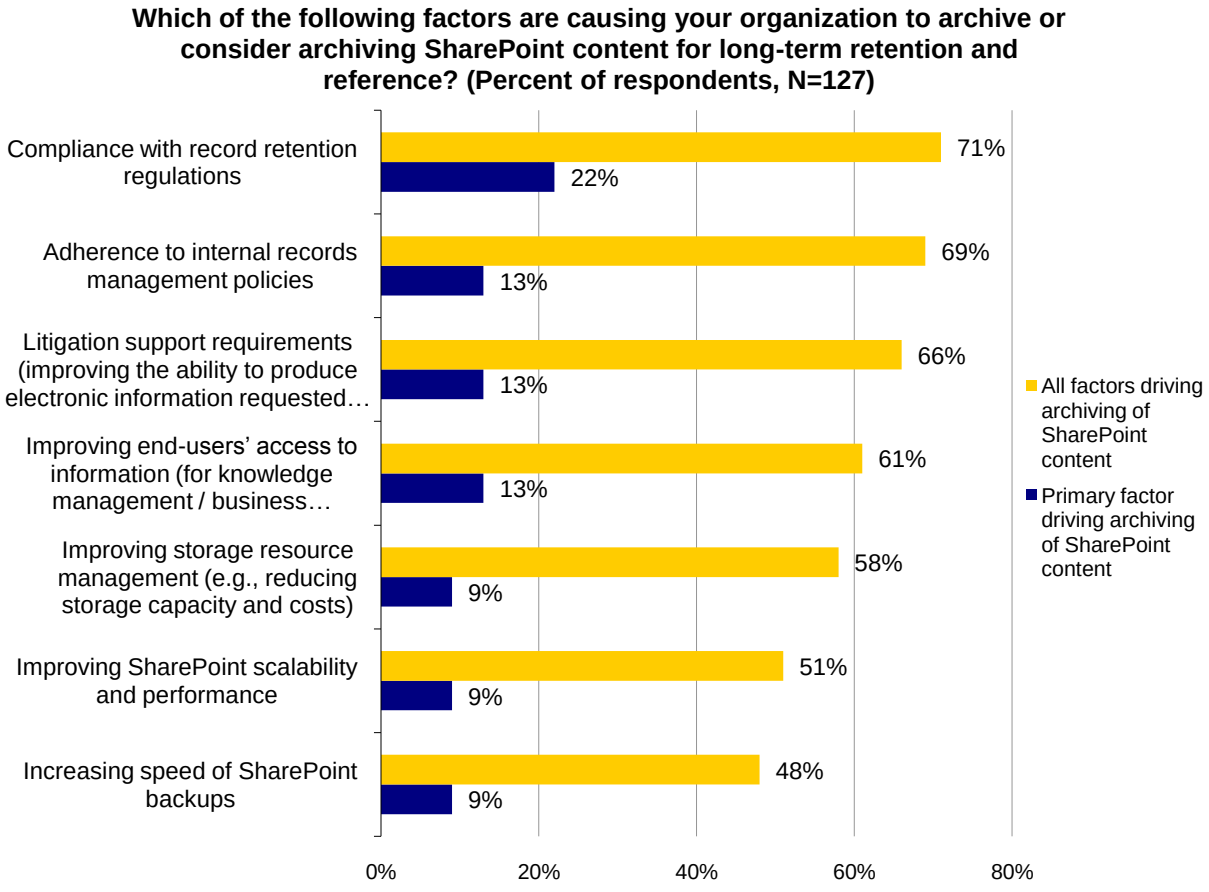
MOSS's architecture also presents entirely unique problems—it is made up of several components, including several databases, web servers, search files, amongst other things. In an estimated 75% of current MOSS implementations, customers implement these components on more than one server, creating a MOSS 'Farm.' All components are interrelated and there is a hierarchy for how they work together. IT must find a way to manage the environment as well as the information in context of the environment. Failure to do so can result in application downtime and data loss or corruption.

The Not so Perfect Storm

IT departments have enough to worry about when it comes to keeping MOSS implementations up and running. Trying to ensure the application is optimally performing, has adequate storage capacity to save all the information, and then figure out a way to create a consistent point-in-time copy of all the components adds layers of complexity and cost that most organizations are often unprepared for. Also overlooked is the fact that MOSS is often used as a repository for critical business documents and records. This content is discoverable at any time as it falls under the amended Federal Rules of Civil Procedure, which govern civil litigation processes—including the discovery and management of electronically stored information—in the United States. Corporate counsel may call on IT to assist in the search for specific information within MOSS and then put relevant content on legal hold.

Data growth, distributed implementations, and MOSS's multi-tier architecture make it a difficult application to back up. Sixty-eight percent of organizations said that MOSS has significantly or moderately increased the amount of data that they have to back up. Electronic discovery, along with other external factors such as record retention requirements, force nearly half of current MOSS users to archive data (see Figure 2).

FIGURE 2. FACTORS CAUSING ORGANIZATIONS TO ARCHIVE AND RETAIN MOSS CONTENT



Source: ESG Research Report, The Impact of Microsoft SharePoint on IT Infrastructure and Information Management, March 2009

A Unique Solution for a Unique Application

Existing MOSS implementations are going to grow and become more complex. Over a third of organizations that plan to adopt MOSS are in the mid-late stage of product evaluation. These trends mean more and more IT departments are going to quickly become familiar with opportunities and ongoing operational challenges that arise with MOSS deployments. New and existing MOSS customers can try to treat MOSS as just another application and integrate it into security and data protection operations—a decision that is unlikely to address most of the aforementioned issues. Another alternative is to evaluate how MOSS-aware retention and recovery management solutions such as Mimosa NearPoint can efficiently protect and optimize MOSS environments.

NearPoint archives data from a primary MOSS environment, moving site content as well as the context (metadata including the environment hierarchy, security permissions, site templates, etc.) into its repository. Once the information is centralized within NearPoint, customers can:

- **Apply and enforce retention policies.** These can be based on who created the content, what site it came from, what type of file it is, and several other criteria.
- **Save data on a less costly storage infrastructure.** The NearPoint repository is a logical entity that can be stored on several storage systems, including devices that save data in WORM format. As such, companies can deploy lower cost devices or purpose-built systems (in the case of WORM) to control costs and meet specific regulatory and legal requirements.
- **Facilitate seamless, permission-based access.** Employees can access the archived content via a stub (or link) left in the primary MOSS environment, while corporate counsel and other authorized individuals can utilize the NearPoint eDiscovery application to search across the repository to find data relevant to a specific legal or regulatory inquiry.

By archiving MOSS information, customers can reduce the size of the primary application environment, which boosts performance and lowers storage costs—the latter is possible because the archived information is single instanced (removal of duplicate content) before it is stored and the data can be saved on less expensive storage systems (when compared to those that support the primary application).

In addition to MOSS content, Mimosa NearPoint customers can also archive e-mails and file system data. All of the information can be stored within the same repository, enabling consistent enforcement of retention policies for records management and legal hold processes. In addition, by building an integrated content archive with NearPoint, companies create a system of record where compliance officers and corporate counsels can go to quickly locate information. This is a drastic improvement over the alternative methods for meeting compliance and legal requirements, which entailed searching and random sampling individual IT applications and disparate application repositories.

Mimosa Has Recovery Covered

NearPoint customers configure the MOSS sites and associated content they want to archive, as well as triggers for the execution of archive policies. For example, an organization may choose to archive a Corporate Finance-specific web part, which can be a component of a broader Accounting site. Whenever a new spreadsheet is added to the web part, an archive capture event is triggered. When modifications are made to the spreadsheet and versioning is tracked, an archive policy can be executed—the older versions can be stubbed and moved permanently to NearPoint.

As Mimosa captures specific content for archiving, it gathers the content, as well the context of that content, within a MOSS environment. In continuing with the example above, when Mimosa archives versions of a Corporate Finance spreadsheet, it also captures the site template where it was created and stored, the document list that the file was a part of, who has permission to access the file, what content database the file was originally stored, and the hierarchical relationship between the web part and the broader MOSS environment. While capturing all of the contextual information along with the content helps maintain the integrity of the information (a process needed if and when the data is requested during an electronic discovery event) and secures it in the archive (content permissions follow the data into the archive), it also enables customers to quickly recover MOSS information in the event of corruption or deletion.

To archive information, customers must set up NearPoint to capture changes within the Event or Change Logs or both—depending on the archive policies. As the changes are captured, they are processed (indexed, single instanced, etc.) and then saved for the appropriate retention period. By tracking all of these modifications along with their context, NearPoint can facilitate 'Fine-grained' data recovery of specific MOSS items. For example, if a MOSS document list is deleted, IT can simply go into MOSS, locate a copy of the list, and then restore it to its original location—or to a new location. All of the permissions associated with that document list remain intact. In order to make such a task feasible, the company would merely need to make sure it is archiving document lists.

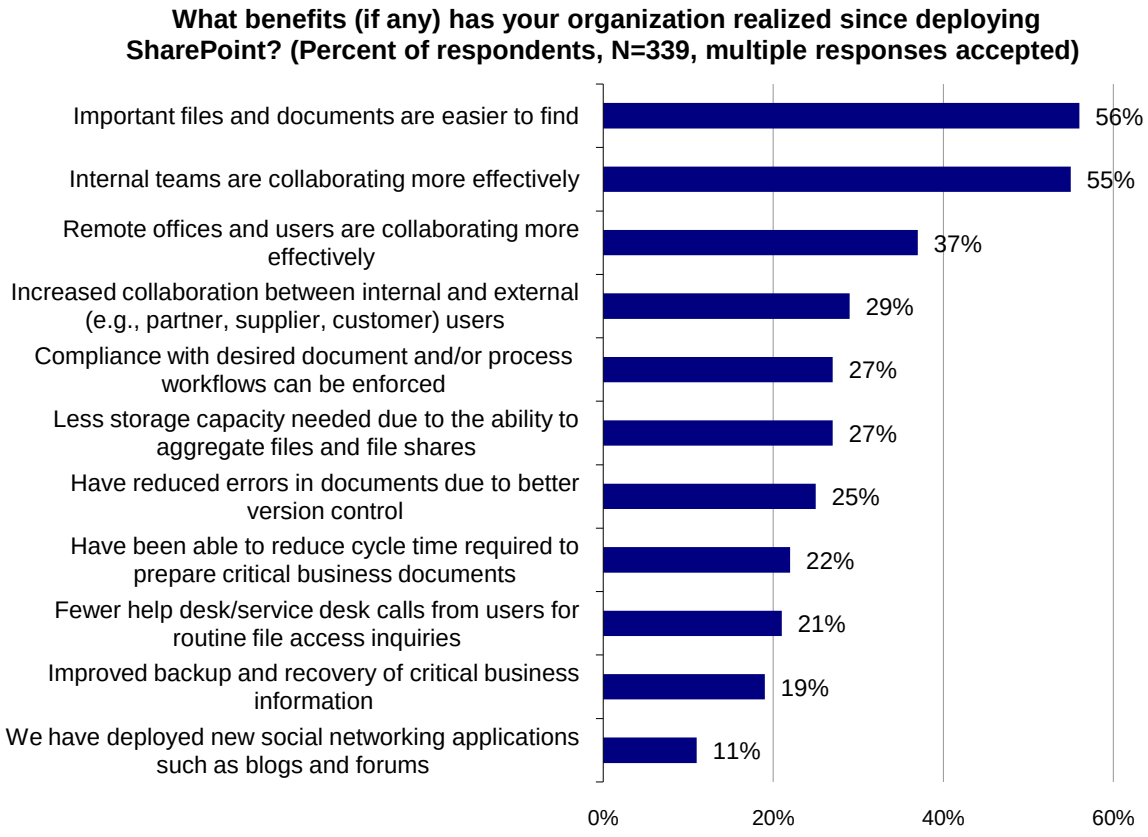
Fine-grained recoveries prevent against data loss of individual MOSS items and provide customers with the flexibility to restore those items to any MOSS Farm while maintaining the context (permissions, relationships,

hierarchy, etc.) of the information. Customers can also configure additional archive capture processes to enhance data recovery options. When customers initially set up NearPoint's capture parameters, they can establish a full copy of the MOSS environment within the archive. This is easily done by using Microsoft's Volume Shadow Services (VSS), which can take a snapshot of the entire environment and store that copy—which contains all of the necessary contextual information from a MOSS Farm—within NearPoint. For those that do not want to use VSS, NearPoint can 'crawl' a MOSS Farm, which also allows the appropriate information to be gathered in the archive. These full copies are kept constant by applying the aforementioned Event and Change Log changes that are also captured by NearPoint. If any of MOSS's underlying databases—including the Content Database or the search index files—are corrupted or deleted, NearPoint customers can complete a 'Coarse-grained' recovery. The latest full copy of the MOSS environment is restored and then all of the incremental changes that have been made since that copy was made are applied, allowing IT to recover the entire MOSS Farm up to the most recent transaction or change. To minimize the time it takes to complete a 'Coarse-grained' recovery, customers should execute a VSS copy or a crawl on a regular basis as this minimizes the amount of changes that will have to be applied during a restoration operation.

The Bottom Line

Current MOSS users are seeing significant improvement in collaboration and important corporate content—once scattered across employee PCs, file shares, and other systems—is now much easier to find. These productivity improvements, along with the other benefits made possible by MOSS (see Figure 3), cannot be ignored and companies will start to roll out MOSS feverishly. However, customers must realize that MOSS is not just another enterprise application and it must be managed differently.

FIGURE 3. MOSS BENEFITS EXPERIENCED BY CURRENT USERS



Source: ESG Research Report, The Impact of Microsoft SharePoint on IT Infrastructure and Information Management, March 2009

A unique application mandates a new approach to management and Mimosa delivers just that for MOSS. NearPoint's ability to improve application performance, lower storage costs, and centralize information for easy access via its archiving capabilities can enhance any MOSS implementation. When customers then realize that

they can significantly mitigate the risk of data loss and reduce the time it takes to recover data with the same NearPoint platform, they will realize that Mimosa addresses many of MOSS's current operational challenges.

Even though data protection and archiving should be separate processes (as they solve different problems), over three fourths of organizations surveyed by ESG stated that they are managed by the same IT group.[2] There is no reason why, in MOSS implementations, they cannot automated by the same software solution.

[2] Source: ESG Research Report, *Data Protection Market Trends*, January 2008.