



# Managing .PST Files from Discovery to the Archive

## ***Index***

Assessing the Facts .....	2
Who Needs Quotas? .....	2
Message Size Reduction .....	3
From Personal Stores to Policy Archiving .....	3
Discovering .PST files .....	4
Assessing Content in .PST files .....	4
Archiving from .PST files .....	4
Recovering Redundant Space .....	5
Conclusion .....	5
Action List .....	5
The C2C Approach .....	7

## ***Overview***

If you're reading this paper, you probably realize already that the .PST files scattered around your network contain messages and attachments which should really be part of your central data store. As the need to retain corporate data has increased to meet regulatory requirements — and best practice may dictate email meeting certain internal policies — the very existence of .PST files may put you at risk of failing to meet these legal requirements which could have significant implications'. So what can you do about .PST files?

This white paper outlines some ideas, methods and applications for living with legacy .PST files and avoiding the creation of new ones. It will help you:

- Assess the true facts about .PST files
- Manage existing .PST files
- Avoid the need for .PST files in the future

## Assessing the Facts

Users create .PST files as a secondary data store to save those emails they wish to retain and refer to later. But why is this necessary when they have access to mailboxes especially designed for this purpose?

The usual answer is "quotas." If mailboxes are restricted by size quotas, and users receive a notification that they're reaching the limit, they have a decision to make. Choosing to ignore the alert just provides a short-term breather. The user will soon be unable to send emails, and finally unable to receive emails. Typical email user reaction to hitting mailbox quotas is alarming. The following list from a survey by Osterman Research shows how users usually cope:

1. Delete email from inbox and/or folders
2. Create one or more personal archives
3. Complain to IT
4. Expand mailbox
5. Delete all "sent" email with attachments

These findings are contrary to e-Policy, best practice, regulation and productive use of time. If users have to delete email, but they need to retain information to do their job, they will spend longer and longer trying to find emails that can be deleted without impacting their work life. This time costs money. If, for example, this takes just 30 minutes per week, it amounts to thousands of dollars of lost productivity per employee per year. What's more, the user is likely to continue to come up against quotas and have to go through this process regularly.

This individual and selective approach also carries the risk that they will delete email critical to their work, or contrary to data retention policy. A common short-term solution, often advised by Exchange administrators, is to create a .PST file. End of story? Unfortunately, the practical limitations of .PST files are now well known and should not be ignored.

Making them a less-than-strategic fix, .PST files were not intended as a long-term, continuous-use method of storing messages in an enterprise environment. .PST files:

- Are not supported by Microsoft over a LAN or WAN link
- May not be part of the normal security backup and maintenance processes
- Are hard for the administrator to locate
- Cannot be monitored easily for content according to corporate policy
- Are inefficient in their use of disk space particularly in comparison to Exchange mailboxes
- Can only be accessed by a single user at a time, i.e. not available for teamwork
- Have known tendencies to corrupt when they near their maximum size
- Take time to repair when damaged
- Fill up local drives and server space
- Remove the ability for PC independence

## Who Needs Quotas?

So, if the driving force for the creation of .PST files is mailbox quotas set by administrators, but .PST files are a bad idea, why are quotas set? In simple terms, mailboxes have quotas set to ensure there is sufficient space on the Exchange server to hold all the user mailboxes. Although the cost per GB of disk space has reduced significantly over recent years, both the number of emails received has risen (by around 15-20%) and the average size of email has risen (by around 10-125% pa.) thus negating the savings and increasing the cost of managing the data. Back-up and restore times within SLAs and processing requirements of servers are also major factors.



So to reduce the need for mailbox sizes quotas, we need to minimize the total amount of data needed to be held in the Exchange server mailboxes and manage the information for longer term storage and easy retrieval.

This combined process is referred to as mailbox size management, encompassing automated email compression to reduce message size, combined with rules-based archiving to move data to more appropriate storage.

## Message Size Reduction

An automated zipping system which follows corporate e-Policy and which zips or unzips email at all the appropriate enterprise touch-points (client, OWA, Internet gateway and Exchange server) is recommended to maximize the benefits and minimize any impact on normal processes.

The immediate benefit of data compression is to reduce the physical space needed to store a particular message and attachment. The improvements are typically dramatic, reducing space requirements by around 50%

However, selecting the appropriate system to compress email messages and attachments is important, as anything which would impact the end users by involving extra actions — users needing to zip and unzip files or requiring additional training — should be avoided.

## From Personal Stores to Policy Archiving

Having made the first step towards reducing the need for .PST files by reducing message size, the next step is to archive according to e-Policy. This reduces the dependence on existing .PST files without compromising the user's access to previously stored items.

To achieve this part of .PST management you need to follow the strategy below:

- Discover all .PST files known to Outlook Mailboxes
- Manage content of .PST files either on a scheduled, exception-driven or one-time basis
- Archive messages and attachments from .PST files according to centrally controlled criteria
- Place small message links in .PST files for easy user discovery and retrieval of archived messages and attachments, which reduces the reliance on assistance from the IT team
- Compact .PST files after messages are archived, thereby recovering the unused disk space
- Index the archived .PST data for fast search and retrieval

This will:

- Bring data contained in dispersed .PST files under central administration for easier backup and restoration
- Allow user retrieval of archived email from choice of .PST or mailbox
- Reduce .PST file sizes or allow their entire removal
- Reduce risk of loss of data due to drive failure

To work through this list can be a long and complex manual procedure, but policy-driven archiving tools can help you to automate these processes. However, e-Policies are different in every organization. Beware oversimplified solutions that don't allow for multiple criteria to run simultaneously; even a complex rule-set should be easy for the administrator to set up, use and update.

Here's a comparison:

- Simple rules: Archive Bob's .PST email to storage
- Multi-criteria rules: Archive Bob's .PST email over 3 years old to low cost storage, and .PST email under 3 years old to high availability storage



## Discovering .PST files

Assuming the end users are not going to manage their .PST files themselves, the first challenge is to locate the .PST files. Where do you start? The most simple and assured way is to run a process that hooks into the user's Outlook session and automatically locates any current .PST files.

This process should also provide the identification of which file is associated with which user for later use (when messages are moved out to the archive). To ensure that the end user does not tamper with this process this application is best launched from a login script and without local configuration.

This will also ensure email integrity and security by identifying the ownership of .PST files correctly, by checking for a .PST file coupled to the host mailbox as defined within the Outlook profile.

.PST files are more liable to corruption the nearer they get to their maximum size so identifying the largest files is an essential step. A simple discovery recording the .PST file name, path and size, for later review will show the files to prioritize.

## Assessing Content in .PST files

Having determined the order in which to action the .PST files, the next step is to identify what information is contained in the .PST files. .PST files can contain a variety of different types of information, anything from personal data, items which might contravene your e-Policy, messages which include viruses and various other attachments. C2C recommends checking content before any further actions are taken.

Searching the files should cover known issues, again according to multiple criteria set by the administrator. Actions will be needed to ensure that only appropriate email is moved to the archive. A variety of actions can be taken including move, delete, copy or alert the administrator. A good e-Policy search tool should automate all of this for you. Once this 'housekeeping' step has been completed, the next step is to apply your archiving criteria to the remaining information in the .PST.

## Archiving from .PST files

Selections on date created, subject matter, attachment size, content, sent to and other simultaneous, multiple criteria should form the basis for moving the messages and attachments into a secure, long-term repository. These items are then stored according to the same criteria from the user's mailbox.

Your choices of storage for the email archives should reflect your normal data retention strategy. A good archiving vendor will be neutral to the storage you choose to use and integrate with a variety of storage management solutions to simplify processes and optimize the use of any investments already made. Keep in mind that end users still need to access the items moved over from the .PST files. The process should include the creation of message links (stubs) within the reduced .PST file. This will ensure that the user is directed to the new location of the message for easy viewing and retrieval.

## Controlled archiving or migration

Administrators need to be able to archive or migrate the content in a controlled process that works in 'harmony' to the Exchange server rather than in an uncontrolled and in many cases harmful manner. Administrators need to be able to apply actions to the content of the .PST files, specifically to allow or deny if the content can be archived or migrated back to the host mailbox. These actions can be applied from one .PST to all.

For example, if the Administrator wishes to migrate the content of the .PST files back to the host mailboxes, it is essential to know the impact on the Exchange server and user mailboxes prior to initiating the event. Without understanding and controlling the impact, you are likely to import so much content that mailbox quotas are exceeded and worse the Exchange server balloons to excessive size with the possibility of store dismount.

## Recovering Redundant Space

Now it is time to recover unused space. .PST files need assistance to reduce the space they occupy when data is removed. Outlook 2003 user's PST data is compacted automatically but for older legacy systems this may need to be automated by a third-party tool.

At the end of the Compaction process, you should have a .PST file which uses less disk space, still providing the access to old messages and attachments that the user created it for and free of any items that could be a company security risk

As a small .PST file, the risk of potential corruption is reduced, and the company information is incorporated into a secure, visible store with information from the user's mailbox.

This process is a complex task. For situations where thousands of .PST files are involved, this should be treated as an ongoing process. Start with the priority files and work through the entire structure. The ability to run these tasks on a scheduled, exception-driven or one-time basis becomes a major benefit to ensure the advantages are maximized and maintained.

## Removal of .PST files

To remove the dependence on .PST files entirely, it is necessary to import the remaining data back into the users' mailbox. A process should be undertaken (ideally automated) to migrate the residual user data. To avoid confusion to the location of the data, the imported data should maintain the original folder hierarchy as found within the original .PST file.

Now you can compact the empty .PST file and remove it from the user's Outlook profile.

## Conclusion

If .PST files are part of your environment, then bringing them in line with the requirements of your organization's e-policy, best practice and regulation without alienating your end users will certainly make life easier for email administrators and reduce the time spent on email 'housekeeping' tasks.

## Reference:

Enterprise Email Archiving: Market Problems, Needs and Trends  
An Osterman Research Multi-client Study, 2003 and survey 2005  
[www.ostermanresearch.com](http://www.ostermanresearch.com)

Administration Guide for Microsoft Exchange Server 2003, Microsoft Corporation  
<http://www.microsoft.com/downloads/details.aspx?familyid=98E45481-1458-4809-97D6-50D8AEED8A1&displaylang=en>

Personal folder files are unsupported over a LAN or over a WAN  
Microsoft Corporation.  
<http://support.microsoft.com/kb/297019/en-us>

## Action List

To reduce the need to create new .PST files:

1. Reduce the server space needed for mailboxes with email archiving
2. Less space required for mailboxes = Users less likely to hit mailbox quota limits
3. Less mailbox quota issues = Removes the need for .PST files

This practical, step-by-step list will help you to manage existing .PST files:

1. Locate the .PST files
2. Select the .PST files for priority action
3. Relate .PST files to users
4. Search and discover .PST contents
5. Select data for archiving as per your email retention policy
6. Move information to centralized storage
7. Archive messages and attachments
8. Insert message links in .PST file
9. Compact .PST File
10. Import residual data and de-couple .PST File

# The C2C Approach

We hope that this paper has helped you to plan your way forward. C2C's award-winning email archiving software Archive One is a modular solution that includes:

## Archive One Policy Manager

Policy Manager Analyses email and enforces rules-driven email control, retention and retrieval, for mailboxes, local and central .PST files and Public Folders. Policy Manager identifies .PST size and automatically applies archiving and compaction to the file — reducing potential damage, loss of data and disk space usage. There are a large number of customer selectable, centrally set criteria of what to archive.

Key benefits include:

- Improves management of storage
- Reducing mailbox quota frustrations
- Enforces email policies for mailboxes and .PST files
- Improves litigation support capabilities
- Improves email system performance
- Provides a secure method for long term retention of email
- Transparent user retrieval of archived messages
- Large range of archiving criteria for selective archiving
- Flexible repository based archiving
- Intuitive, comprehensive rules engine for flexible multi-policy archiving
- Storage management and hardware support for low TCO and integration with your storage and backup strategy
- Archives in standard compressed format for efficient storage
- Recovers messages to original status
- Helps to meet and improve SLAs

For more about Policy Manager and a free 30-day evaluation, visit <http://www.c2c.com/site/products/archiveone/policy/default.asp>

## Archive One Discovery Manager

Discovery Manager's forensic discovery engine can be used to search un-archived (un-indexed) historical email. The .PST file discovery and housekeeping features help enforce e-Policy and reduce the legal risk associated with email.

Key benefits include:

- Searches Exchange folders, mailboxes, local and central .PST files and attachments
- Multi-criteria, content-driven search and discovery of emails (search on folder type or name, content or property of message, item and more)
- Reduced time and costs with definable, automated, scheduled and repeatable auditing
- Searches single or multiple information stores
- Reduces risk of litigation associated with email
- Identifies who is sending or storing specific email
- Alerts user or administrator by email of the results of a search match
- Easy to use, multi-criteria search engine built for Exchange
- Removes or identifies non work-related and other unwanted email
- Searches and deletes attachments
- Helps to meet the latest e-disclosure legislation
- Reduces time required to respond to disclosure requests
- Finds information that is mislaid or even deleted

For more about Discovery Manager and a free 30-day evaluation, visit <http://www.c2c.com/site/products/archiveone/discovery/default.asp>



## MaX Compression

C2C also offers a rapid-response tool, MaX Compression for immediate Exchange system performance gains by rules-driven compression of email attachments, even within .PST files. Automated compressing and uncompressing of attachments is invisible to the user and shows measurable benefits to the organization. Running on over 2.5 million desktops world-wide, in 2000+ user organizations, MaX Compression is the benchmark for auto-compression of Exchange and has an impressive track record in all industry sectors.

Key benefits include:

- Invisibly and automatically zips and unzips email attachment, even in .PST files
- Reduces the Exchange Information Store content by up to 50%
- Reduces bandwidth requirements and transmission times for messages by up to 89%
- Reduces network bottlenecks between Exchange servers
- Benefits remote user productivity
- Invisible to user
- No user training required
- Centralized administration
- Maximizes return on investment (ROI) and cost savings

For more about MaX Compression and a free 30-day evaluation, visit <http://www.c2c.com/site/products/activemailtools/max/default.asp>

## Disclaimer of Liability

While every precaution has been taken in the preparation of this document, C2C Systems assumes no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

## Trademarks

All trademarks, trade names, service marks and logos referenced herein belong to their respective companies. Copyright C2C Systems 2007.

## About C2C

C2C offers email archiving and management solutions, which reduce risk, optimizes performance and minimizes compliance issues for over three million users at more than 2,000 organizations world-wide. Based on their in depth understanding of message management, C2C developed its award-winning Archive One suite to help organizations comply with industry regulations and minimizes mailbox size. C2C also offers rapid-response tools for email performance, security and crisis control.

The Company, a Microsoft Gold Certified Partner, supports organizations in the government, manufacturing, finance, education and healthcare industries, including Fortune 1000 companies. Established in 1992, C2C is a privately held company with US offices in Springfield and Westborough, Mass; and Reading, Berkshire in the UK.

## Where can I find out more?

For more information and free evaluation software, visit [www.c2c.com](http://www.c2c.com) or email [info@c2c.com](mailto:info@c2c.com).

1 Federal Street  
Bldg. 101-R 3W  
Springfield, MA 01105-1199  
T: +1 413-739-8575  
F: +1 413-739-4980

6 Richfield Place, Richfield Ave  
Reading, Berkshire  
UK RG1 8EQ  
T: +44 (0) 118 951 1211  
F: +44 (0) 118 951 1111

