



Managing PST Files from Discovery to the Archive

October 2004 Copyright C2C Systems

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.

Contents:

Managing PST Files from Discovery to the Archive	1
Contents:	2
Overview	3
Assessing the Facts.....	3
Who Needs Quotas?	4
Message Size Reduction – A Continuous Process.....	5
From Personal Stores to Policy Archiving.....	5
Discovery	6
Assessing Content in PSTs	7
Archiving from the PST	7
Recovering Redundant Space	7
Conclusion	8
<i>Where can I find out more?</i>	9
<i>Action List</i>	10
To reduce the need to create new PST files	10
To manage the PST files you already have	10
Appendix 1: MaX Compression Enterprise	11
Appendix 2: Archive One Policy.....	12
Appendix 3: Active Folders Content Manager.....	13

Overview

If you're reading this paper, you probably already realize that the Personal Storage Files (PSTs) scattered around your network contain messages and attachments which should really be part of your central information 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 PSTs is putting email more and more 'at risk.' So what can you do about PSTs?

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:

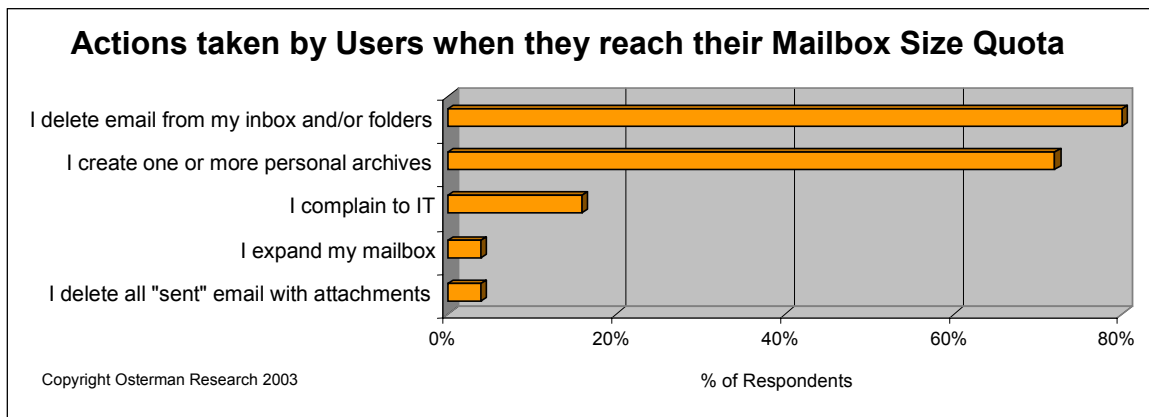
- 1) Assess the true facts about PSTs.
- 2) Help you if you have PSTs.
- 3) Help you to avoid the need for PSTs in the future.

Assessing the Facts

PST files are created by users as a secondary data store to save emails that 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.

The user reaction to hitting mailbox quotas can be alarming. This chart, courtesy of Osterman Research, shows how they cope.



The graph shows that decisions can be 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 this takes, for example, 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.

The common short-term solution, often advised by Exchange administrators, is to create a PST file. End of story? Unfortunately, the practical limitations of PSTs are now well known and cannot be ignored, making them a less-than-strategic fix.

PSTs are not reliable places to store company information because they:

- Are hard for the administrator to locate.
- Have known tendencies to corrupt when they near their maximum size (2GB for Outlook XP and earlier editions).
- Are not part of the normal security backup and maintenance processes.
- Cannot be easily monitored for content according to 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.
- Take time to repair when damaged.
- Fill up local drives and server space.
- Removes 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 PSTs 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. Other factors such as keeping backup and restore times within SLAs and processing requirements of servers also come into play, but disk space is the major bottleneck.

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 – A Continuous Process

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 50% or more.

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.

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.

(Max Compression from C2C, Appendix 1).

From Personal Stores to Policy Archiving

Having made the first step towards reducing the need for PSTs 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 PSTs according to centrally controlled criteria.
- Place small **message links** in PSTs for easy user discovery and retrieval of archived messages and attachments, which also reduces the space requirements of PSTs.
- **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:

- 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.

"PSTs created with Outlook 2002 and lower versions have a limitation of 2GB. If you do reach 2GB, it renders the PST useless and you have to do some tricky stuff to just be able to look at it," says Dan Gamez, LAN Administrator at Gerald Metals, Inc. in Stamford, CT. "It's important to discover and archive PSTs before they are irretrievable."

Discovery

Assuming the end users are not going to manage their PST files themselves, the first challenge is to **locate the PSTs**. 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 PST files they are currently using. 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). This application is typically launched from a login script and needs no local configuration.

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

Assessing Content in PSTs

Having determined the order in which to action the PST maintenance, 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. Best practice suggests checking content before any further actions are taken.

Searching the files should cover known issues, again according to multiple criteria set by the administrator:

- Are there instances for known email subjects that may contain viruses?
- Are there personal emails which should be moved to another location?
- Are there messages or attachments that need checking to ensure they meet the company policies?

Actions will be needed on these types of entries 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 the PST

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.

When moving items from the PST, keep in mind that the end user still needs to access this information. 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.

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. This can be achieved by an Outlook process called 'Compaction.'

At the end of the Compaction process, you should have a PST file which is:

- much smaller with less disk space used
- 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(s).

This process is a complex task. For situations where thousands of PSTs are involved, treat this 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.

Conclusion

PST files are part of your environment, but bringing them in line with the modern requirements of the 21st century without alienating your end users will certainly make life easier for email administrators and reduce the risks to the company e-Policy systems.

The following pages provide an action list and information about C2C's messaging management solutions to help you with PST management.

Where can I find out more?

For more information visit <http://www.c2c.com/>

C2C Systems Inc.
1 Federal Street
Bldg. 101-R 3W
Springfield, MA 01105-1199
T: 413-739-8575
F: 413-739-4980
info@c2c.com

C2C Systems Ltd.
6 Richfield Place, Richfield Ave
Reading, Berkshire
UK RG1 8EQ
T: +44 (0) 118 951 1211
F: +44 (0) 118 951 1111
info@c2c.com

All trademarks, trade names, service marks and logos referenced herein belong to their respective companies. Copyright C2C Systems 2004

References:

Enterprise Email Archiving: Market Problems, Needs and Trends.
An Osterman Research Multiclient Study. 2003. www.ostermanresearch.com

3 Reasons to Archive Email. C2C Systems 2004. <http://www.archiveone.com/3r.asp>

MaX Compression Enterprise: Gaining optimal benefits from compression
www.c2c.com/industry/whitepapers.htm

Exchange Server 2003 Administration Guide
Microsoft Corporation 2003
<http://www.microsoft.com/downloads/details.aspx?familyid=98e45481-1458-4809-97d6-50d8aeebd8a1&displaylang=en>

Action List

This practical, step-by-step list will help you reduce user's need to create new PST files, and manage the PST files you already have.

To reduce the need to create new PST files

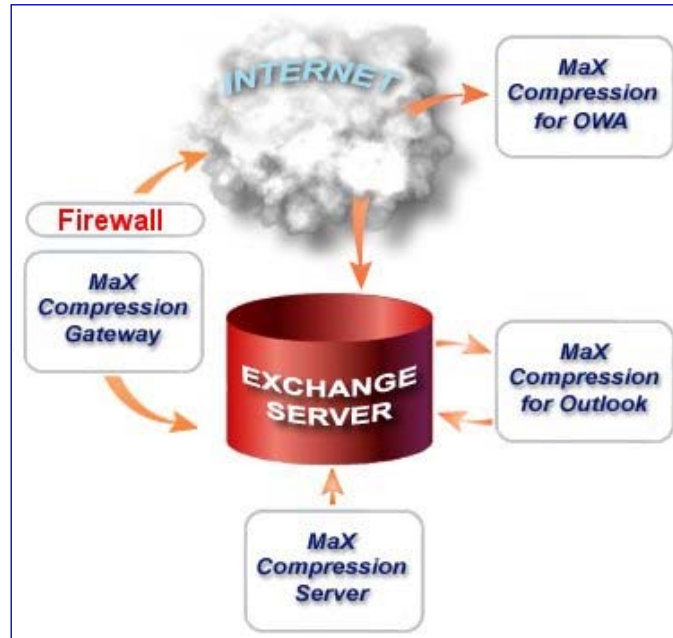
- 1) Reduce the mailbox sizes with Compression**
 - a. Recommended – *Max Compression Enterprise (see Appendix 1)*
- 2) Reduce the mailbox sizes with Email archiving**
 - a. Recommended – *Archive One Capacity or Policy (see Appendix 2)*
- 3) Smaller mailboxes = less need for mailbox quotas**
- 4) Less need for quotas = less need for PST files**

To manage the PST files you already have

- 1) Locate the PST files**
 - a. Recommended – *Active Folders (see Appendix 3) or PST option in Archive One Policy*
- 2) Select the files for priority action**
 - a. Recommended – *Active Folders or PST option in Archive One Policy*
- 3) Relate PST file to its user**
 - a. Recommended – *Active Folders or PST option in Archive One Policy*
- 4) Search and discover PST contents**
 - a. Recommended – *Active Folders or PST option in Archive One Policy*
- 5) Delete unwanted information**
 - a. Recommended – *Active Folders or PST option in Archive One Policy*
- 6) Move information to other files**
 - a. Recommended – *Active Folders or PST option in Archive One Policy*
- 7) Archive messages and attachments**
 - a. Recommended – *PST option in Archive One Policy*
- 8) Insert message links in PST file**
 - a. Recommended – *PST option in Archive One Policy*
- 9) Compact PST File**
 - a. Recommended – *PST option in Archive One Policy*

Appendix 1: MaX Compression Enterprise

MaX Compression Enterprise gains immediate Exchange system performance by rules-driven compression of email attachments. Automated zipping and unzipping 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.



1. Reduce the Exchange Information Store size by up to 50%.
2. Reduces bandwidth requirements and transmission times for messages by up to 89%.
3. Reduces network bottlenecks between Exchange servers.
4. Benefits remote user productivity.
5. Invisible to user.
6. No user training required.
7. Centralized administration.
8. Complete suite of products for a total benefits solution.
9. Maximize return on investment (ROI) and cost savings.

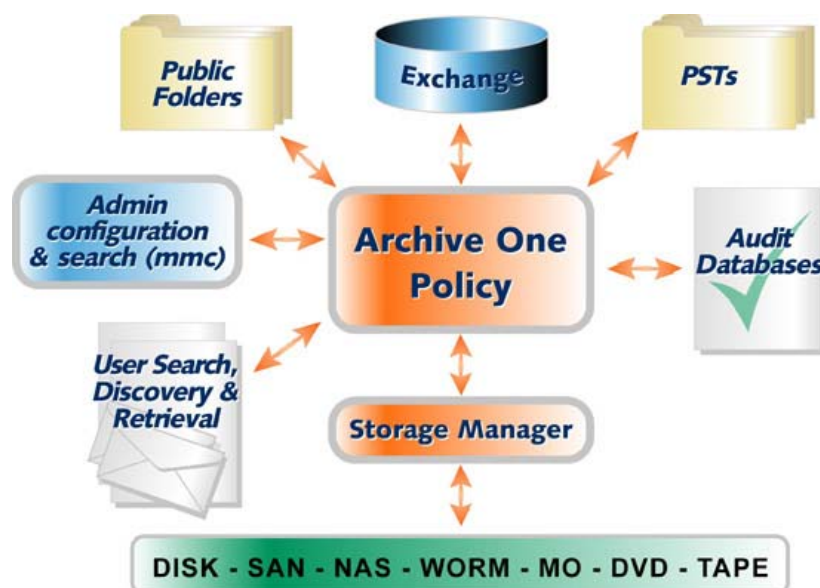
For more information and a free 30-day evaluation, visit <http://www.c2c.com/>

Appendix 2: Archive One Policy

Archive One Policy, a multi-criteria archiving solution, enforces e-Policy by managing and retaining e-mail in chosen storage media for secure access and retrieval.

Archive One Policy 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. This includes selections based on date, size, attachments, subject, etc., which can be varied according to the requirements of a particular user or group of users. You can also automate the selection of priority files.

Archive One Policy

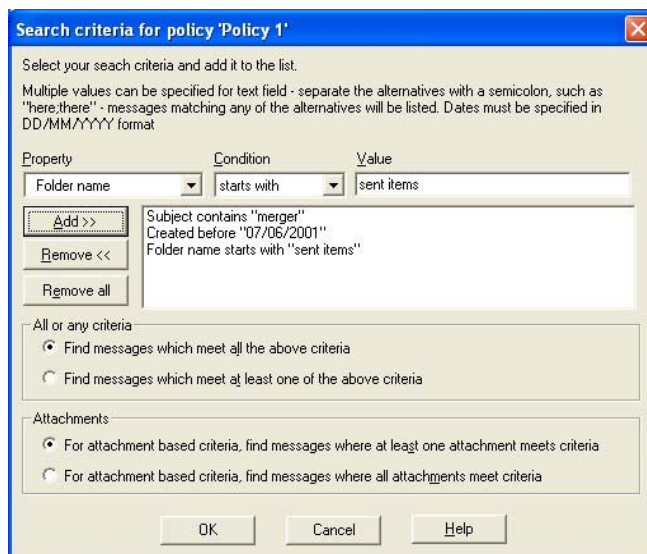


1. Enforces e-mail policies for mailboxes and PST files.
2. Long term, secure retention of e-mail.
3. Fast and easy, search and retrieve messages through user driven web interface.
4. Easy to install, use and manage.
5. Intuitive, comprehensive rules engine for flexible multi-policy archiving.
6. Large range of criteria for selective archiving.
7. Improves user productivity.
8. Reduces the need for mailbox quotas.
9. Writes archive and index to choice of SAN, NAS, Optical, Tape, WORM, to name a few, as selected by the user organization.
10. Archives in standard compressed format for efficient storage and recovers messages to original status.

For more information and a free 30-day evaluation, visit <http://www.c2c.com/>

Appendix 3: Active Folders Content Manager

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



1. Multi-criteria, content-driven search and discovery of emails.
2. Save administration time and money with definable, automated, scheduled and repeatable auditing.
3. Reduce risk of litigation associated with email.
4. Searches Exchange folders, mailboxes, local and central PSTs and attachments.
5. Search single or multiple information stores.
6. Identify who is sending or storing specific email.
7. Alert user or administrator by email of the results of a search match.
8. Free-up redundant storage.
9. Remove viruses when AV fails.
10. Easy to use, multi-criteria search engine built for Exchange.
11. Fast: can search millions of messages an hour.
12. Remove or identify non work-related and other unwanted email.
13. Search and delete attachments.
14. Search on folder type or name, content or property of message or item.

*"I'm *stunned* that Active Folders will open in-use .PST files! Fantastic product!" Senior Systems Administrator, Oak Ridge, Tennessee, USA.*

For more information and a free 30-day evaluation, visit <http://www.c2c.com/>