The Challenge of Managing PST Files in 2012

The Problems in supporting PST files and an examination of the solutions.

2012 Update

This paper was first published in 2007 as we saw a trend towards closer management of PST files. In 2011, we expected most organizations to aggressively seek to eliminate PST files from within the environment and, with the advent of a usable repository in Exchange 2010, this seemed feasible. The reality is, this has not been the case.

Many organizations have Exchange 2010 in plan but not yet deployed and, rightly, would not consider importing PST data back into the existing email system. An existing email archiving system offers a plausible alternative, but the organization still needs to decide how much of the data residing in these PST files really needs to be kept and for how long. And who is responsible for determining which records are kept - the user, the legal group, the administrator or some programmatic process.

There remains ambition to get better control over PST files, but elimination is difficult – users are accustomed to having local, immediate access to their email so they can connect to email records when they need to – IT & Legal see uncontrolled email records scattered across the organization with little retention or discovery capability.

A number of solutions have evolved to accommodate better co-existence or elimination of these files, some of which we offer. These may address simply backing up PST data from the local workstation to a central location or providing a means to eliminate PST files completely. But this involves an acquisition cost and project effort to properly identify and resolve the problem - it can be more than many organizations are prepared to justify.

The PST Problem remains and grows. The question becomes is the appetite big enough to focus on this above other activities.

The Challenge of Managing PST Files

The ease-of-use of today's email systems means end-users are using the email system as their primary document management system. Users save items in folders relating to projects or subjects so they can find that email again in seconds. As a result, email is kept in a mailbox longer than perhaps it should – as a reference document – leading to significant load on the email infrastructure; wasted space, slower servers etc.

In a Microsoft Exchange environment, PST files have evolved into one of the most popular personal storage areas – they are easy for the user to define and very easy to manage.

However, over time PST files become a significant issue to the IT Support group. They can be saved anywhere on the network, they may contain valuable or confidential information that needs to be available to other users, and most likely they will not be regularly backed up, resulting in the risk of data loss or corruption.

This white paper is designed to help you understand the issues surrounding the use of PST files and offers some choices on how to alleviate these problems. This paper will provide invaluable planning information together with Best Practice Advice, built upon significant experience gained by ReSoft International in assisting its clients with implementation of Email Retention & Management solutions.

What is a PST file and how is it used?
A PST file is a Personal Storage File. PST files were designed around older versions of the Microsoft Exchange email server, at a time when performance was slower and internet bandwidth was narrow. It allows a user to move messages through Outlook from their server-based Exchange mailbox to a PST file on the desktop, either though drag-and-drop or the Outlook AutoArchive feature. Email residing in a PST file on the desktop can then be read without being connected to the network, which is beneficial for users when traveling.

The PST architecture was also adapted to support the SMTP and POP internet protocols which allow any user with a Windows PC to save email into an Outlook format from their personal ISP account.

When users exceed the server mailbox quota, set by IT in an effort to prevent mailbox sizes from getting too large, they often use PSTs as a means of removing email from their Exchange server mailbox.

Where does a PST typically reside?
PST files were designed to be resident on the user's desktop. However, in many cases an organization with many users and many PST files will place these files on a network drive accessible to each user, even though Microsoft officially does not support this configuration. This approach makes it easier should the IT Group wish to back up and restore PSTs. (It should also be noted that many IT groups do not backup files on a user's desktop PC.)

A PST operates as a file-based store rather than as a database. This means that read/write operations to PSTs are slower and place a higher demand on the server than reading and writing directly to a database such as the Exchange mailbox. A large number of users who are continuously updating network based PST files can place a significant load on the network, leading to slower response times.

Additionally, (depending on the backup software being used) if a user happens to leave Outlook open and connected to a PST overnight, the PST file may either not be backed up at all or may be invalid. If so, it would be unusable if the need arose to restore it.

How many PSTs can I have?
As many as you wish – each PST, however, needs to have a pointer set within the Outlook profile so that Outlook knows where that PST is. But this is not a difficult task, and most users quickly learn how to create PST files.

How easy is it to find information in PST files?
Somewhat limited! Outlook has a search capability allowing you to scan through the content of PST files by using keywords. But there is no facility to scan across multiple PST files in a single search and the search (prior to Outlook 2007) does not look for keywords residing in an email attachment

How are messages stored in a PST?
A PST file is a single file containing many messages. This gives the user a central point to get access to messages but leads to a lot of duplication and repetition when managing, backing up and restoring PST data.

What happens if my PST file gets corrupted?
As PST files get larger they are prone to corruption. Microsoft does provide a utility called the Inbox Repair Tool to recover data from a corrupted PST file, but this has limitations. If the repair fails you most likely will lose all messages contained within that PST file.

How large can a PST file get?
Generally, the larger a PST file gets, the slower access will be to the emails it contains. Older versions of PST files (Outlook 2002 and below) will not allow a PST file to get larger than 2 gigabytes before corrupting, and will prevent the user from retrieving any messages from that PST. Later versions (Outlook 2003 and on) allow substantially larger sized PSTs (up to 20 Gigabytes) so size-related data corruption should be less of an issue. In the event of a corruption however, the Inbox Repair Tool may not fix the problem. Instead, you will need to use a utility called the Oversized PST and OST Crop Utility, which also has some serious limitations. Again, a failed repair will most likely mean loss of all messages.

A further consideration is that you probably don't really want to manage a single 'database' of email of up to 20 gig in size for each of your users.

So why are PSTs a problem now?
The PST file's primary benefit is portability - allowing a user to locate it pretty much anywhere on the network. But in the current era of litigation, compliance and because so much mission-critical data is now stored in personal email, it is no longer safe to rely on each individual user to store data in individual PST files scattered across the network.

Should I plan to move PSTs into Exchange 2010 or into a separate Email Archive solution?
Microsoft has built Exchange 2010 to increase the ability of mailboxes to store large volumes of data. They claim that mailboxes can be larger by factor of 4, 5 or more, that data can be stored on less expensive disk and that backup/restoration problems inherent with older Exchange implementations will be eased by the use of Database Availability Groups (DAG). Added to this is a specific PST import capability where a user can elect to import their PST emails into the Exchange 2010 Personal Archive mailbox (albeit one at a time).

One of the primary factors behind the demand for a larger mailbox is the need to accommodate larger file sizes, such as unified messaging data. Large attachments such as voice messages and video will generate huge burdens on mailbox sizes, and very quickly soak up any increased efficiencies and increased quota space. Adding PST files back into the mailbox will consume prime mission-critical storage for storing old and very infrequently accessed email information. And as Exchange 2010 does not support the single instancing of email messages there will be a large volume of duplication in these lower value emails and across the various DAG replicas. So, the reality is that pressure on mailboxes size is not going to go away under Exchange 2010. This leads us to believe there is still a very good case for implementing a separate Email Archiving Solution.

What are my Choices with PST files?

1. First, take an inventory of where your PSTs live to size the problem
You need to understand the scope of the problem before you can decide an action. PST files are often scattered across the network so it is important to build an inventory of where those PSTs are and how actively used they are. Software tools like **PST Station Corporate** or **PST Enterprise** will reach out across the network to identify where PSTs reside, who owns them, how large and how old they are.

2. Leave PSTs as they are and live with the consequences
As long as you understand the consequences of living with PSTs, continued use can be achieved and there are tactical third party tools to help in this process. Organizations taking this approach may move or copy PSTs off the desktop onto a central network server to be able to access for e-discovery and backup. Tools such as **PST Station Corporate** automate the copying and monitoring of PST files to prevent them from reaching large sizes by automatically creating a new PST when the original reaches a pre-determined size.

For old format PSTs, put PST files onto a dedicated NTFS partition and use Windows 2000/XP's disk quota management feature to generate a warning when the user is getting near the 2 GB limit. Alternatively, Outlook 2003 allows you to use Group Policy to control the size of both the old and any new PST files.

Most likely, and not to be underestimated, there will still be a requirement for an E-Discovery tool, such as the **Discovery Attender** solution, to search through the content of emails in multiple PST files simultaneously to meet regulatory, legal or acceptable-use auditing.

3. Eliminate PSTs completely
By importing PST data back into Exchange 2010 or introducing an Email Archiving Solution, users can get access to their data without needing PST files. This involves a fairly sizable migration of the data, building ownership information for each PST file into the process and setting policies for access. Some email archiving systems do have a disconnected mode to allow users to maintain a cache of their archived email for disconnected access. To prevent users from creating new PST files you can set the DisablePST registry settings to prevent a PST file from being created, and DisablePSTGrow to make existing PSTs read-only during the migration process.

4. Allow PST Co-existence
This option allows a copy of the email to be kept in a central repository while allowing the user to maintain their own copy of the data on the desktop. By implementing a scheduled synchronization process between the two, the user has disconnected access while IT is able to manage a copy of the data, meeting their data retention/management standards.

PST Station Corporate allows you to introduce an automatic synchronization process between a user PST on the desktop and a central archived copy, thereby allowing IT to protect the copy in the event of a local PST failure.

SUMMARY OF PST PROS & CONS

• Reduces size of Exchange mailboxes	• Prone to loss of all data in a PST through corruption
• Allows disconnected access to email	• Difficult to perform E-Discovery
• Easy to create	• PSTs contain no ownership information
• Easy to use	

What should I do now?
ReSoft International is here to help you. We are an independent consultant, integrator and reseller of best-of-breed software technologies that address Email & Instant-Messaging Security, Regulatory Compliance and Storage Relief.

1. **Get Educated** – ReSoft has put together a comprehensive White Paper Library addressing many of the issues you are probably facing. You can download any of these white papers free of charge.

RESOFT WHITE PAPER LIBRARY SAMPLE LIST: http://www.re-soft.com/wp.htm

All the ReSoft solutions are at http://www.re-soft.com/pst

Other useful links:
Personal folder files are unsupported over a LAN or over a WAN link
http://support.microsoft.com/kb/297019

Network Stored PST files … don't do it!
http://blogs.technet.com/b/askperf/archive/2007/01/21/network-stored-pst-files-don-t-do-it.aspx

Description of sample tools for managing PST files in your Outlook Profile
http://support.microsoft.com/kb/2000021

Understanding Personal Archives in Exchange2010
http://technet.microsoft.com/en-us/library/dd979795.aspx

> *"Frost & Sullivan believes that Re-Soft's main competitive advantage lies in its ability to offer a complete solution for the content security space that more specialist vendors are not able to match. This strength is likely to have a positive effect on the company's growth in the market."*
> *Frost & Sullivan*

2. **Contact ReSoft for further discussions on how we may be able to help you.** ReSoft has, for over **17 years**, acted as a trusted advisor to hundreds of organizations, applying the relevant technologies and techniques from its broad set of Email & IM Security tools to reduce liability risks and save time. More details on our tools can be found at:
   http://www.re-soft.com/pst

Conclusion

This paper has set out to illustrate ReSoft's significant coal-face experience in providing professional services and solutions that have helped many organizations to define & implement policy-based Archiving & Discovery systems across industries and organizations of all sizes. We have defined what we believe are the key questions to be asked when planning what to do with PST files. These questions are part of an Archive Planning Workshop we offer to assist organizations in determining whether they need to deploy email archiving, what components would be required, and how they would be justified.