



PST Files and Archive One

What is a PST?

A PST is a 'Personal Storage file, for use with Microsoft Outlook which stores Email, Calendar and Contact data. It is usually used as an alternative to storing the data in Exchange mailboxes.

The Problem with PST files

PST files are often created to dodge mailbox quotas, configured storage limits applied to the Exchange servers' storage. Users who do not wish to delete email rely on these files to store some older data in order to keep within the designated quota. However, when a user stores data in a PST file this creates a number of issues for administrators:

- PST files are often stored on the users local hard drive and therefore is outside of the corporate backup policy
- Administrators have very little control over the content of the PST files (libelous statements, company confidential data etc)
- There are difficulties involved in search/discovery of the content in the face of legal search requests
- Little or no security is applied to the PST file contents, PST passwords can be hacked easily

Archive One Policy from C2C addresses these issues through discovery, analysis, archiving and migration of these PST files.

Processing PST files with Archive One

Archive One combines its Policy service component with a client component to process PST files. This client component, the PST Processor, is loaded into the memory of each PC client and then with the service component to allow the processing of the PST files in a controlled manner. There is no need to install the PST processor on each user.

The screenshot shows the 'Archive One Policy - [Archive One Policy] PSTs Discovered' window. The table below represents the data shown in the main pane of the application.

Mailbox	Title	PST Size	Mailbox Size	Mailbox Quota	Process	Migrati...	Machine	File Path	First Seen	Last Seen
Ken Hug...	Dossiers pers...	23825K	1775098K	Unlimited	No	N/A	SALES01	C:\Documents ...	01 February 200...	01 Februa...
Ken Hug...	BUG Test	144K	1775098K	Unlimited	Yes	Finished	SALES01	C:\Documents ...	01 February 200...	01 Februa...
Philip Bu...	Old_Support_...	57553K	34487K	Unlimited	Yes	N/A	PHILIPB	U:\Support_m...	10 May 2007, 10...	26 June 2...

Administration console

PST Discovery

The PST Processor will check every minute or so for a live Outlook session and, when it finds one, will determine the PST files that are loaded into the Outlook profile.

When it has determined the PST files it then communicates the discovered data back to the administrator for collation and display to the administrator. The discovered data sent to the service component includes machine name, mailbox name, PST file name, PST file size, PST file path and the date/time of discovery.

This staged PST processing can be run on its own as a method of understanding the size of the PST problem/data in an environment.

The PST Processor is also configurable in a number of ways by command lines switches. These allow the administrator to:

- Ensure that the network or Exchange servers are not 'flooded' with data from many workstations
- Hide the user interface so that the process is completely transparent to the end user
- Process data between specific times of the day (for example only at lunchtime)
- Introduce a delay before beginning the processing, allowing Outlook to complete its other processing first/as a priority (downloading of OST data etc)

PST Analysis

Although it is advantageous to archive and import PST content back to Exchange, this should only be performed with a full understanding of the impact to the online storage of Exchange servers.

Using any process or solution to migrate a large quantity of PST data directly back into Exchange is likely to create problems without process limits and controls. Understanding and applying process limits avoids:

- Excessive growth of the Exchange Information Stores
- Information Store dismounts due to drives becoming too full
- Excessive Exchange storage consumption causing backups failure
- Reaching mailbox quota limits
- User frustration and lost productivity

Archive One's administration console shows detailed information about each PST, its size, the associated mailbox, mailbox size and quotas etc. In fact, it provides all the information required to make an intelligent decision regarding whether the PST should be processed or not, without intruding on Exchange operations.

PST Archiving

To archive the content of PST files, administrators need to create a Policy which specifies PST files as one of the targets (a simple checkbox option) and check one or many of the discovered PST files as 'enabled for processing'.

The Policy defines which data should be archived out of the PST file using Archive One's granular email matching criteria (50+ items of metadata about the email such as size, delivery date, sender, attachments and many more).

Archive One also allows, in the policy definition, the administrator to specify:

- On which dates the PST Archiving policy should run
- A maximum amount of data to be processed in one run
- Time limits for the archiving process
- Which PST files are to be include for processing

When the policy has been defined the administrator initiates the policy run, the PST Processor finds PST data items which match the policy criteria and submits them to the service component via a web service. The original item is left intact until it has been confirmed as committed to the archive.

Submitted PST data item is stored in a 'staging area' until it is processed. This continues until the client component has completed looking at all data items in the PST file or until the user logs out of Outlook (or one of the policy limits is reached).

On the server side, the service runs the secondary stage of processing every hour. This secondary stage picks up the submitted PST data items from the staging area and then indexes, compresses and archives the data, replacing the item in the staging area with the final shortcut link. This continues until all the data items in the staging area have been committed to the archive.

The next time the PST Processor runs on the workstation it recognizes the submitted messages which have been committed to the archive and converts the original data items in the PST to archive link messages. These link messages provide the end user the ability to retrieve the original data item from the archive simply by double-clicking the link messages (as you would a normal message) within Outlook.

PST Migration

If organizations wish to remove the reliance on PST files, Archive One provides the ability to import the content of the PST back to the users' mailbox, leaving the original PST hierarchy intact. After importing the data, Archive One can compress the now empty PST and remove it from the Outlook profile to avoid further use.

In scenarios where a large number of PSTs are to have their content migrated back into Exchange, Archive One provides a number of checks, rate controls and throttling parameters to prevent the Exchange server overloading from excessive volume of data. In addition a warning is displayed when the volume of PST data to be imported would exceed mailbox quotas. This is all easily configured and administrated from the management console.

Summary

Archive One provides total PST management – other vendors only offer some level of PST processing. Archive One can find, archive, migrate, and report on PSTs irrespective if they are located on network storage or local drives. C2C offer two methods of removing the data and reclaiming the white space left behind by the PST:

- Copying the data directly to the central secure archive repositories and providing message links to this data back to the PST
- Importing the data from the PST pre, post or mid phase of archiving directly back to the mailboxes

Archive One provides a high level management console to determine the best actions to take with the PST data and to schedule how and when to work with the PST without causing system bottlenecks through processing too much at one time. This is all achieved through a simple 'right click' operation. Furthermore, the processing of the PST data is invisible to the users.

Contact C2C:

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

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