Solution Brief
Managing Corrupted PST Files

PST files are widely used for storing copies of messages, calendar events, and other similar items. They are particularly unreliable and prone to corruption, which sometimes causes the loss of valuable data. Additionally, the time taken in resolving issues with PST files can be a significant drain on resources for IT administrators.

Many organizations are now using solutions such as Barracuda PST Enterprise to locate these troublesome files wherever they exist, and to migrate the data elsewhere so that use of them can be completely eliminated.

This solution brief provides a background to PST files, and explores the reasons why they get corrupted so easily. It then discusses the strategies that should be employed by IT administrators to address the problem of corrupted files. Finally, it outlines the role that PST Enterprise can play in helping to resolve issues with corrupted files during a PST migration and elimination project.

Overview of PST Files

PST files are created by the Microsoft Outlook desktop client. They are used to store local copies of email messages, calendar events, contacts, task lists and other similar items.

Each PST file is a container that comprises a folder structure and a number of individual items. As emails and attachments are often quite large, and some PST files will contain many items, these files can sometimes grow to a size of several GB each. The maximum size allowed for a single PST file is 50GB.

PST is an acronym for Personal Storage Table, but they are more commonly referred to as an Outlook archive or as a personal archive file.

How they are used

PST files are created by the "Auto Archive" function in Outlook, which automatically removes older messages and data from Exchange and archives these locally in a PST file. This feature was enabled by default in Outlook 2003 and Outlook 2007.

PST files are also often used when email needs to be moved from one place to another, such as when migrating between mail servers, or when an organization needs to retain a separate copy of data (such as when an employee leaves an organization). They are also quite widely used simply as a convenient way for end users to store and organize their older email data.

Why they are a problem

The PST file is not a robust file format, and was not designed for reliable long-term storage of large amounts of data. These files often grow to an excessive size, which makes them particularly liable to corruption — especially if they are accessed over a network. The time taken

in resolving issues with PST files can be a significant drain on resources for IT administrators.

With the release of Exchange 2010, Microsoft recognized the problems these files caused and introduced enhancements to eliminate the need for users to have separate local archive storage. However, it is still possible for users to create and work with these files, even in the latest versions of Outlook.

Administrators can disable the ability to use them, but many organizations have a large number of PST files still in existence, and they often still allow them to be used for a number of different reasons.

How PST Files Get Corrupted

The PST file is a complex binary file format and is not designed to be fault tolerant. Any interruption in processing while a file is being created or updated may result in a partial update, which causes a corruption. The risk of this increases the larger a file gets.

A PST file may be corrupted if it is open in Outlook when the application terminates abnormally. This can happen if either Outlook or Windows crashes, or if the machine is powered down without quitting Outlook and Windows normally.

Some organizations have PST files stored on a central file server with users accessing them over the network, even though this configuration is not supported by Microsoft. In this case, any interruption in network connectivity may result in a corruption, and simultaneous access by more than one user is also likely to cause problems.

PST files can also get corrupted if the total size approaches the limit for this file type. This was more often a problem with files created in older versions of Outlook where the maximum file size was 2GB.

Types of corruption

In some cases, the structure of a PST file will be corrupted. These files are easy to identify as it is not possible to open the file at all within Outlook — an error message will be displayed to the user.

In other cases, just a few individual folders or messages within the file will be corrupted. These files are more difficult to identify as they can be opened correctly within Outlook, and only cause a problem when the user tries to access the specific folders or messages that are corrupt. Everything else about the file appears to be normal; therefore, these corruptions can remain undetected for long periods.

Dealing with Corrupted PST Files

Corruption of PST files is such a prevalent problem that a wide range of both free and paid for tools are available to repair them. Often, it may be feasible to repair a corrupted PST file; however, this will not always be possible and at least some (if not all) of the data contained within a particular file may be unrecoverable.

It is important to approach the problem by reviewing all corrupted files with the relevant users. This helps you to understand the importance of the data contained in them, and to discuss the risks involved in recovering it. This will also help you identify the best way to resolve the problem for each corrupted file.

1. Repair the PST file

    Microsoft provides a free "Inbox Repair Tool" (called ScanPST) that will examine the PST file headers and structure, and attempt to repair any errors it finds. This will often work, but sometimes the nature of the corruption is such that it is unable to be resolved. The process of resolving one corruption may also expose additional corruptions in the same file that will then need to be addressed, so it may be necessary to run the tool more than once.

    Many of the third party PST repair tools that are available claim to be better than the free Microsoft repair tool. They include additional functionality that can help administrators in both repairing corrupted files and recovering data from these files, so they are also worth considering.

    It is important to bear in mind that whichever tool you use, there is a probability that the repair process may not be able to retain all items from the file and some data may be lost. It is also a good idea to take a backup of the file before you start the repair process in case it fails and renders the file unusable.

2. Skip the errors

    The easiest way to deal with individual corrupted items in a PST file is to simply discard them, which may be appropriate if it is not critical to retain every item from the file, or if you know there are items within the file that cannot be repaired or recovered.

    If you import the contents of the PST back into a mailbox in Exchange, you can set the 'Bad Item Limit' to ignore up to a certain

number of corrupted items. It will skip over these items during the import process, and you can then export the remaining items back out to a new PST file.

Alternatively, if it is possible to open the file in Outlook, then either the IT administrator or the file owner can use Outlook to copy all data that is accessible from the corrupted file into a new PST file.

3. Restore to a backup

If the PST file has been backed up, then you can of course restore back to the last good version of the file. One problem with this is that corruptions within a PST file can remain undetected for an extended period of time, so it may be difficult to identify which version to restore back to. Also, you may not even have retained a backup that goes back far enough. Another problem is that PST files are often stored locally on the user's machine where they are typically not backed up at all.

4. Delete the whole file

Before you start repairing a PST file or recovering the data from it, make sure that the user definitely still needs it. A large number of PST files are kept "just in case" and may no longer have any value to the business.

Managing Corrupted Files with PST Enterprise

The Barracuda PST Enterprise client will scan every message in every PST file that it finds. This means that it will discover all corrupted files, and all corrupted items in each of those files. It will even discover corrupted items within files that appear to be healthy because they are open in Outlook and are being accessed successfully by an end user.

When PST Enterprise discovers a corrupted file, the Health indicator in the PST Enterprise Admin Console for that file will be set to "Corrupt", and it will move on to the next file. No further processing will be attempted for the corrupted file at this point. PST Enterprise does not attempt to initiate the repair of any corrupted files itself, as this could potentially result in valuable data being lost.

As discussed above, the IT administrator should review each corrupted file individually with the end user and decide the most appropriate action to be taken for that file. They should also bear in mind the importance of the data within it and the risk of losing at least some of it. Other actions that could be taken instead of trying to repair a file or skipping the corrupted items will include restoring to a backup or just deleting the file.

Repairing PST files

If the decision is to attempt a repair on a corrupted file and retain as much data as possible, Barracuda recommends the IT administrator run Microsoft's free Inbox Repair Tool (ScanPST) against that file.

If this resolves the corruption, you can then mark the file as "Healthy" in the PST Enterprise Admin Console, which will trigger the PST Enterprise Client to scan the file again. If it discovers additional corruptions in that same file, it will again be flagged as "Corrupt", and you will need to follow the same process again.

Skipping corrupted items

PST Enterprise can be configured to skip up to a set number of individual corrupted items in each file before that particular file is considered corrupt. However, this option should be used with caution as it will result in data from from the corrupted items being lost.

Minimizing the problem

The PST Enterprise client includes two features that IT administrators can use to help minimize future problems with corrupted PST files. First, it can be configured to modify each user's Outlook profile so that they cannot create any new PST files or update existing ones. Second, it can compact any existing PST files it discovers that are currently open in Outlook. This will eliminate any unused space embedded within the file and reduce the overall file size, both of which will lessen the risk of future corruptions.

Conclusion

There are several ways of dealing with the problem of corrupted PST files, and tools such as Microsoft's free "ScanPST" Inbox Repair Tool are available to help. While these can go a long way in addressing the problem, it's not feasible to automate the whole process. The IT administrator should play a key role in identifying the most appropriate resolution for each file to minimize the loss of valuable business data.

If an organization is using Barracuda PST Enterprise for their PST migration and elimination project, this will discover all corrupted PST files wherever they exist. It can then assist administrators with managing the process of identifying and resolving the problem for each file.