10 Point Plan to Eliminate PST Files

Executive Summary

When it comes to assuring a comprehensive corporate data retention and litigation readiness plan, no single data set seems to present more challenges than the PST file. Those pesky personal email storage folders seem to create more challenges for today's IT operations than just about any other database or application.
Microsoft Outlook PST (or personal-storage-table) files can be generated for a number of reasons. They may be the result of restrictions imposed by the organization itself. Others are a result of poor user awareness. Some are even created to ease the handover of data from one employee to another. Yet, regardless of the reason they have been created, PST files can be a major headache for IT teams and the company.

Many PST files reside on laptops and workstations that are not typically backed up, and they can present enormous backup, recovery, corruption and discovery challenges. They are not often included in an organization's data protection strategy, so the risk of data loss can be very high. Compounding the challenge, they are often prone to excessive size that can lead to data corruption. Presenting an even higher potential corporate risk is the challenge of including PST file data in an electronic discovery request.

With the release of Exchange 2010, Microsoft has begun to improve its practices when it comes to PST files. Exchange 2010 allows for larger mailboxes and secondary mailboxes where additional data can be archived. Yet, this secondary mailbox solution is only found with the higher-cost "Enterprise Client Access License" and still there is no current feature or service to aid in locating and ingesting existing PST files or managing the content within them. This is potentially putting volumes of critical corporate data at risk.

This white paper provides a 10-point action plan to eliminate PST files so that they can be efficiently integrated into an overall corporate retention policy for assured data protection, corporate compliance and litigation readiness. But, first things first. You need to understand why you want to get rid of PST files before you can set out to eliminate them.

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

Understand why you want PST files eliminated

This may seem very obvious. However, understanding why you want to get risk of PST files will be the key to how you execute and fund the entire process. Are you looking to reduce the cost of maintaining users with PST files? Are you more interested in the corporate risks of eDiscovery? Are you in need of finding data at any time to satisfy legal or HR? Do you need to adhere to corporate retention policies? Or, are you just looking to reduce the cost and complexity of having data residing off a central server? The answer may help you uncover the budget you need to solve the problem.

We have heard the complaints time and again:

- *"We have imposed mailbox quotas, but users have gotten around the quotas by creating PSTs, now our SAN is full of PST data!"*
- *"We have eDiscovery requests and know that there is data contained in a PST that is not visible to IT or legal and it requires manual intervention every time."*
- *"Our corporate policy on compliance says that we must reduce risks on data. Unknown data in PST files causes unknown risks we can't accept"*
- *"Corporate policy says we should delete emails greater than X years old. We can't do this with PST data because users are in control, not IT."*

The issues can go on and on. But, ultimately understanding your organization's largest pain points will enable you to build the best strategy for solving it. For most organizations that may be struggling with corrupted PST file data or high client storage consumption, PST file elimination is almost certainly the best solution. Once you know your motive for PST removal, you can then begin the 10-step plan below to achieve PST freedom. For a few organizations, where eDiscovery and adherence to corporate retention policies are the primary drivers, PST file elimination may not even be necessary.

1. Create a strategic plan for eliminating PSTs (Where, How and When)

Once you have committed to eliminating PST files, you need to create your strategic execution plan – including where, how and when.

Where: Evaluate where you want to store the recovered PST data – if indeed you want to keep it at all. Will you recover it to Exchange, into an archive, on a SAN or other secondary storage unit, or under a managed solution? What limits do you have in place today that might restrict the recovery of PST data to this location, and what cost implications are involved if you need to expand your existing resources?

If you are considering a move into a cloud-based solution, be sure you understand the entire upload and storage costs in advance. If you are migrating into Exchange 2010, you may not want to put old PST file data into the new server – you may find that it quickly compounds your storage requirements and can even result in three times more data than you started with, if you follow the Microsoft recommendations for DAGs. If it's Microsoft 365 you're considering, be sure you know just how much data it will allow you to ingest. You may be surprised that it can't accommodate as much as you might need it to.

How: How do you want to achieve the PST elimination process? If you have an on-premise email archive, determine if it can discover PST files. If it can't, you may want to consider better options. If it can, identify if it requires you to move PST files manually to a single location. There are alternative options – even ones that can work with an existing email archiving solution – that can automate the PST data consolidation process, without manual PST file movement.

Ideally, you will want a process that requires no user interaction and minimal work from IT. There are solutions that can automate the process without requiring users to make decisions or take actions. You may also want to ensure that you are recovering lost PSTs. You can even opt for a solution that will upload the data without moving PSTs to a central location, which can achieve your compliance and eDiscovery goals without a heavy impact on storage resources.

When: Given all of the factors above, identify the timescale of your PST elimination project. Outline the process for PST data discovery, as well as the time for uploading the data, archiving the data or ingesting the data to Exchange, if that is your planned path. Be sure to identify a reasonable timescale, without overloading system or staff resources. Weeks, months or even years in the case of a large installation is not unacceptable as long as the processes work in the background. After all, it may have taken 10-15 years to accumulate this data, so it will not disappear overnight.

2. Establish the location and details of all PSTs

With your "where, how, and when" strategy in place, you next need to seek out all of the PST files in your organization so that you can make informed and comprehensive decisions about the data they contain. PST files can hide in many locations, from file servers and client desktops to client laptops and company servers, where ex-employee data may be residing. Consider a solution that detects and seeks out remote PST files, regardless of whether or not they are loaded or associated in a user's Outlook profile. The location and state should be irrelevant as they can be automatically found. The result is a complete report that will identify the scale of the issue, how many PST files are in the enterprise and their size, location, message counts and ownership details. This is vital information to validate your “where, how and when” strategy to assure that you have the right storage capacity and location for your project.

3. Establish the owners of “lost” PSTs

PST files can often become de-coupled with their owner’s Outlook application. While this is most frequent with former employees, it also happens with active employees with multiple PST files. Finding lost PST files and pairing them back with their respective owners can be a time consuming, manual task. However, this too can be automated. This is important when employees are missing valued PST data. It can also be important during an eDiscovery event when specific employee data must be reviewed for potential legal hold and review. Without proper PST ownership information, PSTs with valuable data that may be pertinent for a legal discovery request might be overlooked.

4. Re-evaluate your strategic plan

Now that you have a comprehensive understanding of the scale and complexity of your PST elimination project, revisit your strategic-plan to see if it is feasible and the budget is available. Be sure to evaluate your system resources, complexity and timescales, based on the full assessment of your organization-wide PST file data. Refine the plan to meet your needs.

5. Check ‘Legal Hold’ requirements

Check with your organization’s legal and human resources organizations to understand what “legal hold” requirements need to be in place. Your organization will likely have pre-determined data sets (such as financial data or executive data) that must be retained. There may also be existing investigations or eDiscovery projects already underway that need to be taken into consideration. Data associated with these legal hold requirements should be flagged so that it is not removed or deleted during your PST elimination project.

6. Implement corporate retention policies

Your organization may already have corporate retention policies in place for the automatic removal of data of a specific age. It is common to see that these policies will not have been adhered to on an employee’s client machine PST. During the elimination process, once legal hold requirements have been identified, the removal of data that has expired beyond your retention policy may save volumes of storage. Assessing this will be an important component to update your strategic plan.

If your organization doesn't have a corporate retention policy, this may be a good time to establish one. Work with your corporate counsel to determine what timeframe is appropriate for your business and industry. Then, apply it consistently across your entire enterprise.

7. Archive data from PST files

Only after you have completed the above steps is it best to begin your PST archive process. Some solutions will require you to archive all PST data, or ingest it, before you search, assign legal holds and apply corporate retention policies. The problem with this is twofold: first, it consumes time, system and budget resources that you don't need to expend; second, it creates an archive "fingerprint" of the data – even if you wished to delete it. More specifically, it creates a searchable metadata record that the data existed in the archive which is a risk that, perhaps in some cases, your organization might not want to assume. Instead, it is more efficient to discover your PSTs first, eliminate old, unnecessary data, and then archive what is required. In fact, you may even determine that, after searching and discovering the data in all of your PST files before you archive, the data you need to move into your email archive could be a mere fraction of what you anticipated. This can save you thousands of dollars on storage costs and can reduce your actual migration time significantly.

8. Migrate the remaining PST content to the chosen location

Once the data that you need to retain has been archived, you can migrate the necessary PST content into a chosen location for user access: Exchange mailboxes or Office 365. Be sure that you don't overload your own enforced mailbox quotas by only migrating necessary data that's in active use; migrating all data risks storage bloat. Migration can be scheduled and automated with very little IT staff intervention. As a special note on Exchange 2010, Microsoft does enable PST files to be brought back directly into a secondary archive mailbox. However, the administrator must still discover PSTs and Microsoft does not provide an automated approach. Larger organizations may require more comprehensive email retention rules only available in third party email archiving solutions.

9. Compact PST files

Once the data in your corporate PST files has been either deleted, retained in an archive or migrated back into Exchange or Office 365, it's finally time to compact the PST files to free-up the white space. The PST file will now contain a lot of white space (gaps) where data has been deleted, archived or migrated out of it. This white space can be compacted, to reduce the size of the PST file and therefore the risk of corruption.

10. Remove PST files and prevent users from using PST files at all

Because the best solutions only need to be performed once, you now need to assure that you will remain free of PST files for good. To remove the dependence on PST files entirely, it is necessary to automatically remove the PST files from the user's Outlook profile, delete them and PST new file creation can be disabled for the user.

PST files don't have to put your valued corporate data at risk

Whether you decide to live with PSTs and proactively discover and manage the PST data where it resides or you choose to eliminate PST files altogether, you have options that can set you free. And, those options don't have to put an unnecessary burden on your IT staff, your system resources or your users. In 10 easy steps, you can eliminate PST files so that they can be efficiently integrated into an overall corporate retention policy for assured data protection, corporate compliance and litigation readiness.