A Websense® White Paper

Websense CloudMerge Ingestion Service

websense®
ESSENTIAL INFORMATION PROTECTION™

Table of Contents

Introduction

Preserve Chain of Custody While Building a Single, Online Email Archive

Since email dominates today's business environment, often containing vital pieces of information that may assist a company in tracking key events, employee behavior and information exchanges, it can often be valid legal documentation. In the event of an investigation or lawsuit, they must be securely stored, easily accessible and produced in a timely manner.

Given this possibility, mass amounts of "old email" can no longer be simply deleted and forgotten. Old messages are becoming increasingly important for corporations. The past few years, many companies have been required to search through backup tapes to recover old emails in response to requests from their legal team, human resources department or other divisions within the enterprise. These companies are left with the option of recovering email from backup tapes, which is possible in many cases, but extremely difficult for a number of reasons, including disruption to normal IT procedures; the need to create a recovery server for restorations; time spent searching for needed content; and time required of IT personnel for all of the related tasks. And time is money when you it comes to preparing your company's defense in legal matters.

Legacy Data

lLegacy data is a generally referred to as information stored in an old or obsolete format (or computer system) that is, therefore, difficult to access or process. In the email archiving world, legacy data refers to email data sitting on:

- Local archives (e.g., PST/NSF files) on desktops, laptops and servers
- Email storages (e.g., Exchange/Lotus) on servers
- Legacy email archives (e.g., Symantec Enterprise Vault)
- Backup tapes

By ingesting your legacy email into Websense's archiving solutions, you have a complete, living record of ALL your email history, both past and present for that ingested email. But, it is critical to preserve the chain of custody when you move your email data from one store to another.

Chain of Custody

If "chain of custody" is a new or a vague concept to you, you are not alone. In the simplest terms, a proper chain of custody establishes the integrity of a piece of evidence, showing that it wasn't tampered with or otherwise altered since it was first collected. Chain of custody has significant practical implications for IT managers and other professionals.

Chain of custody plays a critical function in litigation, especially when opposing counsel is challenging the authenticity of evidence – particularly where digital evidence and emails are involved. We've found that very few organizations adequately account for chain of custody or understand its importance until the authenticity is called into question.

This whitepaper walks you through the process of how we carefully ingest your legacy data and preserve chain of custody. Websense has performed thousands of legacy data uploads, so we've established some best practices that are reflected in the processes outlined below.

Websense CloudMerge Ingestion Service

1. Extraction
Client extracts data from existing repositories (e.g. , email store, PST/NSF files, legacy email archives)

2. Shipment
Client ships data on encrypted media with Chain of Custody Form

3. Verification
Data is verified against Chain of Custody Form

4. Data Import
Data is securely imported to Websense's cloud based archive

5. Chain of Custody Documentation
Websense completes and sends a copy of the Chain of Custody Form to the client

6. Data Returned or Destroyed
Data is either shipped back to the client or the data is destroyed per the client's wishes

The Websense Data Import Process

As the illustration above depicts, the Websense Data Import Service process involves six key steps.

1. Extraction: In the extraction phase, the client extracts legacy email data from existing repositories. This can include email stores sitting on an email server, PST/NSF files, legacy archives (from a previous email archiving solution) and/or backup tapes. This is a critical first step and one that provides the greatest opportunity to break the chain of custody. So, it's important to document the number of original files/PSTs/NSFs sitting in your existing stores in order to compare to the total number of files/PSTs/NSFs imported into the archive.

Websense can ingest a variety of email and archiving platforms, including Microsoft Exchange, Lotus Domino and on-premise/hosted email archiving systems. We also support the following formats:

PST
• Files need to be 2 GB or less in size
• PST files cannot be password protected
• Files can be zipped with a password
• Ability to preserve folder structure upon request
EML
• Does not preserve folder structure

MSG
• Longer conversion and ingestion time
• Does not preserve folder structure
NSF
• Ability to preserve folder structure upon request
• Longer conversion and ingestion time
• Additional cost

Physical Media: Based on our experience, Websense recommends the following guidelines when sending your legacy email data on physical media. If the volume of data is less than 10 GB, send the data on CD/DVD, Flash Drive or via S-FTP (secure FTP). If the data volume is greater than 10 GB, send the data via a USB hard drive (high speed USB 2.0 with external power recommended).

Encryption: Websense requires a basic level of encryption (i.e., password-protected zip file) for all email data transferred and strongly recommends using TrueCrypt Freeware (truecrypt.org) or PGP (pgp.com) to encrypt data before shipping it to us as a best practice. Zipping the data with a password is also acceptable, but less desirable, since it delays the import. We do not accept password-protected PSTs, since they are generally not secure and cannot be processed by our legacy upload system.

2. Shipment: During this phase of the project, Websense outlines the key steps for your project. This includes having you complete our Legacy Upload Order Form , which captures key chain of custody information, such as the date and time sent, number of files, number of emails, format of legacy data (e.g., PSTs), amount of data and contact information. Then you ship the data along with the Legacy Upload Order Form to us, addressed to the attention of the Websense Data Import Center.

3. Verification: Upon receipt of your data, we log, initial and date the Chain of Custody Form (part of the Legacy Upload Order Form) under received data. The case is then assigned to a one of our specialists, who reviews the details and adds your project to our schedule. The completion date is based on the current volume of legacy uploads in our queue. In the meantime, Websense places your data in its current format in our data vault for safe storage.

After your data is extracted and uploaded, your specialist verifies that the data volumes match the information you provided on the Chain of Custody Form and updates you of any deviations. If there is a material difference in the email received/extracted vs. the data on the form, we then explore the reasons for the missing data. Note: If the data does not match the Chain of Custody Form or the data is corrupt, you will be notified to inform you of our findings and determine the next steps.

4. Data Import: After your legacy data is prepped, your specialist imports it into your archive. Websense securely captures and indexes all messages (and attachments) imported into one of our archiving solutions – Discovery Archive orEmail Archive. Once indexed, administrators and users can quickly and easily search the entire contents of archived emails and attachments, using a variety of search criteria, including to, from, date, subject, message body, message attachments and other message properties. As a final check, your specialist verifies the data and signs off on the Chain of Custody Form.

5. Chain of Custody Documentation: Once your import is complete you will be notified with the results. ny deviations from the original Chain of Custody Form you completed prior to ingestion are noted and reviewed.

6. Returned Data/Destroyed Data: The final step in the process involves Websense either returning the original data to you or destroying it. In either case, we document the process and note the date the media was either returned or destroyed per your wishes.

- Returning Your Original Data: Websense returns legacy email data to you in the same format it was originally sent. If the data was encrypted, the data remains encrypted upon delivery. When the data is returned, we include a copy of the completed Chain of Custody Form documenting all of the steps in the process.
- Destroying the Data: Optionally, Websense can destroy the data after it has been ingested. Any optical media, such as CDs and DVDs, is physically destroyed by securely shredding the media while flash and USB drives are wiped clean. We record the destruction date and send a copy of the completed Chain of Custody Form back to you for reference. We also retain the final copy of the Chain of Custody Form in our locked data vault for our records.

Final Thoughts

The transition of evidence from paper to an electronic format imposes new requirements upon litigants to ensure that a proper chain of custody is maintained. Today there is increased emphasis on the collection of email and other electronically stored information (ESI) in the civil litigation arena. Therefore, it is important to properly maintain and document a chain of custody in the acquisition, processing and submission of ESI. From the initial collection of evidence through to its eventual introduction in the courtroom, a properly documented and maintained chain of custody greatly assists in the submission of electronic evidence.

At Websense, we take special care to ensure that your email data is protected every step of the way. If your email data is ever called into question or submitted as evidence, your case is bolstered by having the proper documentation and processes in place to demonstrate your data was protected and unaltered during the migration process.

Top Nine Things to Know About Websense Data Import Service

1. **Use of FTP:** Websense recommends using the FTP option when shipping data for only small data sets (i.e., less than 10 GB).
2. **Supported Formats:** Websense supports a variety of email formats, including PST, EML, MSG, NSF and MBOX. PST is the required format if you want to preserve your folder structure for the data that is imported.
3. **File Sizes:** Websense can process PSTs up to 20 GB, but we prefer files that are 2 GB or less. Smaller PSTs have less chance of corruption and process faster.
4. **Media Types:** Though we accept DVDs, Websense recommends physical hard drives or flash drives when you ship data to us.
5. **Encryption:** While Websense requires a basic level of encryption for email data you ship to us, we do not accept password-protected PSTs, since they are generally not secure and cannot be processed by our legacy upload system.
6. **Mapping Archiving Accounts:** When we import your PST files, we first scan them to determine how to map your users to individual archive accounts. If a sender's email address is found in more than 80 percent of the messages in the PST file, the emails are archived under the sender's address. This process, however, is error prone, so we recommend providing us with a CSV file for mapping purposes. If a CSV file is not provided, the system automatically determines the mapping.
7. **Supported Message Types:** We currently only archive post items. Non-post items, such as calendars, contacts, system messages and voicemails, are not currently supported.

8. **Legacy Unassigned Archive**: If we cannot match an email address or common name to an archived account, we place the emails in the Unassigned Legacy Archive. Emails in the Unassigned Legacy Archive must have a common name (unmatched) or an email address with a recognized domain in the archived account.
9. **Corrupted Data**: Before sending us data, we encourage our clients to run scanpst.exe, a utility included with Microsoft Outlook, to determine whether or not your PST file is corrupted. We cannot import corrupted PST files.

About Websense

Websense, Inc. (NASDAQ: WBSN) is the leading provider of unified content security. We are the global leader in unified Web, data, and email content security solutions, and provide the best security for modern threats at the lowest total cost of ownership to tens of thousands of enterprise, mid-market, and small organizations around the world. Distributed through a global network of channel partners and delivered as software, appliances, and Security-as-a-Service (SaaS), Websense content security solutions help organizations leverage new communication technologies and enable collaboration and the productive use of Web 2.0 business tools. We do this while protecting organizations from advanced persistent threats, preventing the loss of confidential information, and enforcing Internet use and security policies. Websense is headquartered in San Diego, Calif., and has offices around the world.

Questions?

Want to learn more about our legacy upload process?
Please contact us today at http://www.websense.com/content/phone-support.aspx. We can review your needs and address any questions.