Enterprise Vault.cloud™ Legacy Import Guide

Enterprise Vault.cloud: Legacy Import Guide

Last updated: 2018-09-04.

Legal Notice

Contents

Legacy Import Guide

This chapter includes the following topics:

About legacy import

Every organization has email messages that are located in active user mailboxes, personal stores, document management systems, or other communication libraries. You can consolidate some or all of these legacy email message sources into Enterprise Vault.cloud. This guide outlines the legacy import process and explains how the legacy email messages for your organization can be imported into Enterprise Vault.cloud.

About legacy email messages

Legacy email messages refer to the email messages that are stored on local archives on desktops and laptops, on servers, or on backup tapes. These email messages may be on old or obsolete computer systems and are not used on a daily basis. Importing your legacy email messages to Enterprise Vault.cloud provides you with

a complete record of all the email messages for your organization, both past, and present.

Legacy email messages have context, unlike the active email messages that are journaled to Enterprise Vault.cloud. Journaled email messages are captured in transit and are saved in their original form. Legacy email messages may have been altered before they are transferred to Enterprise Vault.cloud. For example, a legacy email message may have been copied into a folder or forwarded to other recipients after a user received it. Legacy email messages may also have been copied into PST files resulting in altered metadata fields.

The Veritas data management team uses a specific import process to capture all the relevant information for legacy email messages. Using this import process, we can ensure all the contextual elements of your legacy email messages are preserved.

Legacy email message extraction

The legacy email messages must be collected at your site location and then sent to the Veritas data management team for import into Enterprise Vault.cloud. Enterprise Vault.cloud supports the import of legacy email messages from any Microsoft Exchange, IBM Domino, or on-premises/hosted email systems. Additionally, Enterprise Vault.cloud supports the import of legacy email messages from other email archiving services or stores.

The email message formats that can be imported into Enterprise Vault.cloud currently include the following:

- PST

- EML

- NSF (requires conversion and may take longer to process)

- MSG (requires conversion and may take longer to process)

If you use PST files for the import, note the following:

- We recommend that you divide your legacy data into PST files no larger than 20 GB. Small files are less prone to corruption and are faster to process.

- We recommend creating separate PST files for each user. Name the PST file with the user's email address, such as **user1@yourcompany.com.pst**.

- If a user requires more than one PST file, append an underscore and the number of the PST file to the user's email address, for example **user2@yourcompany.com_1.pst**

- PST file names should not contain commas, special characters, or double-byte characters.

- If you create separate PST file for each user, create a CSV mapping file to aid the mapping of the PST files to the appropriate mailboxes. The CSV mapping file should identify each PST file with its owner email address, in the following format:

  PSTFileName*,*OwnerEmailAddress

  For example:

  PSTFileName,OwnerEmailAddress

  user1@yourcompany.com.pst,user1@yourcompany.com

  user2@yourcompany.com_1.pst,user2@yourcompany.com

  user2@yourcompany.com_2.pst,user2@yourcompany.com

Legacy email message transfer

For information about legacy email message transfer, refer to the PST ingestion guidelines at the following URL:

https://www.veritas.com/support/en_US/article.100043544.html

In addition, you can send your data to Veritas in the following ways:

- Secure File Transfer (SFTP)
  There is no size limit for uploads using SFTP. For uploads greater than 1024 GB, you should coordinate with a Data Management team member.

- Physical media
  Veritas recommends that you ship your data on USB 3.0 or NAS devices that can be read by the Windows operating system. For NAS devices, you must use a standalone device that can be connected to the Veritas network via an ethernet cable. The NAS device should have a static IP address assigned to it in the range of 192.168.1.50 to 192.168.1.99. Additionally, Veritas will require the share name and the login credentials to access the share on the NAS.

Note: We recommend that you consult your legal counsel before your organization destroys or removes legacy email messages from their original source location.

To protect the privacy of your legacy email messages, all physical media should be either encrypted or password-protected before transfer. After the Veritas data management team receives the physical data, you can provide the team with the password for any password-protected data.

Note: Do not password-protect individual PST files. The Veritas data management team cannot access these files.

Legacy email message assignment

After the Veritas data management team receives your legacy email messages, we securely capture and index the messages before we import them to Enterprise Vault.cloud. The indexing process assigns the legacy email messages from your organization to individual Enterprise Vault.cloud accounts. The assignment of email messages to individual accounts is based on various parts of a message such as the message sender or message recipient. During the indexing process, a single email message may be matched and assigned to several different accounts. To ensure that email messages are assigned to the correct account, we use the unique hash key or the common name that is associated with them.

We associate a unique hash key with each legacy email message that we receive. The unique hash key is also known as an identification code. The unique hash key contains some numbers that represent the following information:

- Company ID

- Subject

- Date

- Sender

- Recipients (including Forward, CC, or BCC recipients)

We use the unique hash key to assign all legacy email messages for your organization and check them against all Enterprise Vault.cloud accounts for your organization. During the indexing process, the unique hash key is compared to all other unique hash keys in an account. If the unique hash key already exists, the email message is discarded as a duplicate. If the unique hash key does not already exist, the email message is assigned to the account.

Note: We use this deduplication process to discard multiple copies of the same email message. This process helps to speed up the search speeds for your Enterprise Vault.cloud instance.

Interoffice email messages refer to messages that are sent between two users within the same email domain. Microsoft Outlook and other email applications often label interoffice email messages using the common names of the message sender and recipient. For example, a message that is sent from **joe.smith@example.com** to **jane.doe@example.com** would be labeled from **Joe Smith** to **Jane Doe**. During the indexing process, the Veritas data management team can assign legacy email messages to individual Enterprise Vault.cloud accounts using these common names. However, we can only use common names if your organization provides us with a list of all common names in a CSV file. The CSV file must contain the first name,

last name, email address, and Enterprise Vault.cloud account user names for all the users in your organization.

Note: The common name CSV file is different from the CSV mapping file that your organization provides to the Veritas data management team during email message extraction.

During the indexing process, the common name legacy email messages are compared to the first name and last name for an account. However, the following limitations prevent deduplication of the email messages that are assigned to an account using common names:

- The unique hash key for a common name legacy email message differs from the hash key for the same message containing the full email address. The deduplication process does not discard either message as a duplicate because the unique hash keys for the two messages are different.

- When more than one user in your organization has the same common name, the indexing process cannot assign and discard duplicate email messages properly. If two users have the same common name, but one user no longer belongs to your organization, all legacy email messages for both users are assigned to the user that is still part of your organization. If two active users have the same common name, the legacy email messages for both users are assigned to the unassigned legacy account. The system cannot differentiate between the two users.

Note: Due to these limitations, we do not recommend using common names to assign legacy email messages to user accounts. If possible, all legacy email stores should be matched to user accounts using PST files.

Legacy folder structure

If your organization provides Veritas with PST files that contain Microsoft Exchange folder structure information, the folder structure can be replicated in Enterprise Vault.cloud. In addition to user-created email folders, the following folders can be replicated in Enterprise Vault.cloud:

- Inbox
- Drafts
- Sent items
- Deleted items

Other types of folders that are available in Microsoft Outlook such as Calendar items, Contacts, Tasks, and Notes cannot be replicated in Enterprise Vault.cloud.

Note that only sent email messages are imported or journaled to Enterprise Vault.cloud. Therefore draft Microsoft Outlook email messages do not appear in Enterprise Vault.cloud, and users may find that the Drafts folder in Enterprise Vault.cloud is empty.

Note: Enterprise Vault.cloud currently only supports replication of legacy folder structure from PST or NSF files. This functionality is currently not available for other file types.

Unassigned Legacy account

During the indexing process, some email messages may not match any Enterprise Vault.cloud account but still be imported into Enterprise Vault.cloud. The import process associates with your organization any email messages that contain your domain. For example, legacy stores and old PST files may contain email addresses from former employees, or may have corrupt email headers that contain your domain. In addition, the CSV file that is used to link Enterprise Vault.cloud accounts to PST files many contain misspellings or duplicate email addresses that contain your domain. In any of these cases, the import process may be unable to assign a legacy email message to a user account. Rather than discard these email messages, the import process assigns them to the Unassigned Legacy account.

The Unassigned Legacy account is a catch-all account for your organization. After the legacy email message import process is complete, the Enterprise Vault.cloud administrators for your organization can search and review the Unassigned Legacy account.

Note: We recommend that you do not disable the Unassigned Legacy account because it prevents email messages from being lost during the import process.

Chain of custody

Email dominates the business environment today and can be used as legal documentation since it contains vital information about your organization. In the event of an investigation or lawsuit, email messages must be stored securely and must be easily accessible, especially during transfer from one location to another.

To ensure the integrity of your legacy email messages, the entire legacy import process is fully documented using a chain of custody form. The chain of custody

form lists all data information and any special instructions. The Veritas data management team records each step of the import process including the date an action occurs and the individual completing the action. As a result, the chain of custody is never broken for your data. Your legacy email messages remain secure, and the original content remains intact.

Legacy Import Frequently Asked Questions

This chapter includes the following topics:

- Legacy email message extraction FAQ

- Legacy email message security FAQ

- Legacy email message assignment FAQ

- Unassigned Legacy account FAQ

- Appearance of legacy email messages FAQ

Legacy email message extraction FAQ

The following frequently asked questions provide more information about legacy email message extraction.

Question: What tools can I use to extract legacy email messages from my current email environment?

Answer: Veritas recommends Microsoft ExMerge for Microsoft Exchange email environments.

Question: How should our organization provide our legacy email messages to the Veritas data management team?

Answer: You can send your legacy email messages in the following forms:

- You can transfer your data using the Veritas S-FTP site. Restrict this method to collections that are no larger than 1024 GB.

For security reasons, Veritas limits the IP addresses that are allowed to connect to the S-FTP site. Therefore, your organization must provide Veritas with the IP addresses to be used for transferring your data.

■ For collections that are larger than 1024 GB Veritas recommends that you send the collection on external USB 3.0 or NAS drives.
For security reasons, Veritas requires that your organization encrypts the data using Bitlocker. The use of other encryption methods may cause delays, or your data may be returned without being ingested.

Question: What is the preferred media for sending the legacy email messages?

Answer: Veritas prefers that your organization sends your legacy email messages on an external USB 3.0 drive, to provide the fastest data extraction.

Question: What is the preferred method for encrypting the legacy email messages?

Answer: Veritas recommends that you use BitLocker encryption.

Question: What is the maximum email message size that can be imported into Enterprise Vault.cloud?

Answer: The maximum email message size that can be ingested is 200 MB for messages that are ingested from PST files, and 200 MB for EML messages.

Question: Are there any types of legacy email messages that cannot be imported into Enterprise Vault.cloud?

Answer: Veritas cannot import the following:

■ Emails that exceed the maximum message size for import.

■ Corrupted emails

■ Encrypted emails

■ Inaccessible emails

Legacy email message security FAQ

Question: How does the Veritas data management team ensure the security of our legacy email messages during the transfer process?

Answer: Veritas uses a chain of custody form to document the transfer of all legacy email messages. The chain of custody form records information about your organization, any return shipping information, and your import instructions. In addition, Veritas records on the form all the actions that are performed.

Question: What does the Veritas data management team do with the physical media if we do not need it returned?

Answer: If your organization does not need your media returned, Veritas physically destroys the media at its facility. Veritas destroys the media in a manner that prevents a third party from retrieving any data after the import process.

Legacy email message assignment FAQ

Question: Which email formats can be imported into Enterprise Vault.cloud?

Answer: Enterprise Vault.cloud supports PST and EML message formats. While NSF and MSG files are also supported, importing of these file types requires conversion, which increases the import time.

Question: What is the preferred email format?

Answer: Veritas prefers that your organization sends your legacy email messages in PST format. PST or NSF format is required for the replication of legacy folder structures.

Question: How does the Veritas data management team ensure that PST files are assigned to the correct Enterprise Vault.cloud account?

Answer: Veritas uses the CSV mapping file that your organization provides to map the PST files to the correct user accounts. The mapping process ensures that all the legacy email messages belonging to a user are assigned to that user account.

Question: How can our organization ensure that interoffice email messages are assigned to the correct Enterprise Vault.cloud account?

Answer: In addition to the CSV mapping file for PST files, Veritas requests that your organization provides us with the common names for all user accounts. With this information Veritas can ensure that all the legacy email messages belonging to a user are assigned to that user account.

Question: What if my legacy PST files are not linked to individual users?

Answer: If import is performed without a CSV mapping file, note the following consequences:

- Email messages rely on a sender or a recipient email address or a common name to map correctly to a Enterprise Vault.cloud account.

- Email messages where the sender or the recipient email address does not contain the domain name for your organization or a common name are not imported into Enterprise Vault.cloud.

- Email messages that cannot be mapped are assigned to the Unassigned Legacy account for the life of the account. Reassignment of messages to a user account cannot be performed.

Unassigned Legacy account FAQ

Question: What is the Unassigned Legacy account?

Answer: The Unassigned Legacy account is a catch-all account that is automatically created for your organization. An email message is assigned to the Unassigned Legacy account if the message contains the domain for your organization but does not belong to an active user account. The Unassigned Legacy account ensures that all email messages that are sent to your domain are archived in Enterprise Vault.cloud.

Question: Do we need to request an Unassigned Legacy account?

Answer: No, the Unassigned Legacy account is automatically created for all customers.

Question: How can I disable the Unassigned Legacy account?

Answer: Veritas recommends that you do not disable the Unassigned Legacy account. The account ensures that all email messages that are sent to your domain are archived in Enterprise Vault.cloud.

Appearance of legacy email messages FAQ

Question: What is the limit for the number of legacy email messages that can be imported?

Answer: The number of legacy email messages that can be imported into Enterprise Vault.cloud is not limited.

Question: How long does the legacy email message import process take?

Answer: The amount of time that is required to complete the import process depends on the following factors:

- The size of your legacy email message collection.

- The number of customers ahead of your organization in the import queue.

- The number of PST files in your collection that are corrupted.

Veritas recommends that you consult your implementation specialist or the Veritas data management team for an estimate.

Question: Why do legacy email messages from the Sent Items folder in Microsoft Outlook appear in the Inbox folder in Enterprise Vault.cloud?

Answer: During forced-mapped legacy data ingestion, any messages that are linked to a user are automatically placed in the Inbox folder.

Question: Why is the Drafts folder for users empty?

Answer: Only sent email messages can be imported into Enterprise Vault.cloud. Since draft email messages were not sent, they are not imported into Enterprise Vault.cloud. For this reason, many users may find that their Drafts folder does not contain any email messages.

Question: Can my legacy contact information be imported into Enterprise Vault.cloud?

Answer: No, Enterprise Vault.cloud currently only supports archiving of email messages. Other legacy data such as contact information, notes, tasks, and calendar items cannot be imported or archived to Enterprise Vault.cloud.

Question: Can my legacy folder structure be replicated in Enterprise Vault.cloud?

Answer: Yes, PST or NSF files that are imported into Enterprise Vault.cloud retain their folder structure. The legacy email messages that are imported appear in the same email folders in Enterprise Vault.cloud as they did in your active email inbox.

Note: Enterprise Vault.cloud currently only supports replication of legacy folder structure from PST or NSF files. This functionality is currently not available for other file types.

Question: Why do some users see duplicate legacy email messages in Enterprise Vault.cloud? Can these users delete duplicate copies of legacy email messages?

Answer: Users may see duplicate copies of legacy email messages if these messages were imported using common names. Duplicate copies of email messages cannot be deleted.

Question: Are legacy email messages imported into Enterprise Vault.cloud according to date?

Answer: Yes, dates are used when Veritas imports email messages into Enterprise Vault.cloud since dates determine the order in which messages appear. By default, Veritas uses the sent date of a message. If the sent date is unavailable, the received date is used. If the sent date and the received date are both unavailable, the import date is used.

Legacy Import Expected Behavior

This chapter includes the following topics:

Legacy email message extraction expected behavior

- Corrupted legacy data can delay the import process. Even though we can process PST files up to 50 GB, we recommend that your organization divides your legacy data into 20-GB sizes or smaller. Smaller files are less prone to corruption and are faster to process.

- We cannot import messages that are larger than 50 MB for messages ingested from PST files, or 75 MB for EML messages.

- Legacy data in EML or MSG format must contain fewer than 50,000 files per folder. Any folders that contain more than 50,000 files may delay the import process.

- Since we cannot import individual PST files that are password-protected, legacy data in PST format must not be password-protected. However, you can password protect hard drives or other media containing legacy data during the transfer process. After the Veritas data management team receives the physical data,

your organization can provide the team with the password for any password-protected data.

Legacy email message assignment expected behavior

- We recommend that your organization provide the Veritas data management team with a list of user email addresses and their respective PST files. The Veritas data management team can use this list to map the PST files to user accounts.

- Duplicate copies of interoffice email messages or messages that use common names may be retained in Enterprise Vault.cloud. The duplicate copies appear when the same message is imported and journaled to Enterprise Vault.cloud. Each message that is added to Enterprise Vault.cloud is given a unique hash key containing information about the Company ID, Subject, Date, Sender, and Recipient of the message. Since an imported interoffice message with only a common name has a different hash key than the same journaled message, the messages are not recognized as duplicates.

- Legacy email messages that are sent to a Global Address List (GAL) contact are assigned to a user account using a common name. Since the message header of an email message that is sent to a GAL contact contains only a common name, there is no SMTP address in the message header. As a result, duplicate copies of these messages may be assigned to multiple user accounts.

- If two users have the same common name, legacy email messages may be assigned to the wrong user account if assignment by common name is used. A forced-mapped PST file legacy import is required to address this issue.

Appearance of legacy email messages expected behavior

- By default, we organize legacy email messages in Enterprise Vault.cloud using the sent date. If the sent date is unavailable, we use the received date. If the sent date and the received date are both unavailable, we use the import date.

- Users cannot move imported legacy email messages between folders in Enterprise Vault.cloud. Any legacy email messages that have been imported into a legacy folder remains in that folder.

- A small percentage of legacy email messages may have corrupted header information, meaning the sender, recipient, or date information is corrupt. Legacy

email messages with corrupted header information are gathered during the extraction process before the data is sent to Veritas. Any legacy email messages with corrupted header information cannot be imported into Enterprise Vault.cloud.

Legacy folder structure expected behavior

- Enterprise Vault.cloud currently supports the replication of a legacy folder structure from PST or NSF files. This functionality is currently not available for other file types.

- Legacy folders are static. Users cannot add, delete, or move the email messages in these legacy folders.

- Each PST file with a legacy folder structure imported into Enterprise Vault.cloud appears as a unique folder tree. The folder structure from multiple PST files cannot be merged into a single folder tree. The Veritas data management team can import multiple PST files with the same file name into one folder tree in Personal.cloud. However, the files must be in different folders on the media that is provided to Veritas because the files must have the same file names.

- If a legacy email message cannot be placed in a legacy folder, it is saved to the Inbox folder in Enterprise Vault.cloud.

- Users can only perform searches on their entire archive and cannot search with a single legacy folder.

- Personal.cloud supports up to 6,000 folders. When users access Personal.cloud, they may experience a significant loading time if their mailbox has a large number of folders. For more than 4,000 folders, a user may experience a loading time of up to 5 minutes.