✓Symantec.

Symantec Enterprise Vault and Exchange Server 2010

November 2011

Introduction

Exchange Server 2010 is the latest release of the highly successful messaging and collaboration platform from Microsoft. This new version of Exchange Server is very important to Microsoft and some of the new features show that they are expanding the purpose of the platform to include functionality historically offered by 3rd party vendors such as Symantec.

Exchange Server 2010 is also very important to Symantec Enterprise Vault (EV) as one of the leading vendors in the Active Email Archiving market.

The new features of Exchange Server 2010 offer some overlap with those traditionally offered by Enterprise Vault. This paper aims to discuss and dissect requirements around email archiving and E-Discovery to show that Enterprise Vault can still offer substantial benefit to Exchange Server 2010 customers.

Enterprise Vault Support and Compatibility for Exchange Server 2010

Enterprise Vault 9.0 (or later) offers full support for Exchange Server 2010 SP1[1]. Full support in this instance includes:

- Mailbox archiving, including support for Database Availability Groups (DAG) and all the same end user features and functionality offered for previous versions of Exchange;
- Journal archiving, including support for the new Exchange journal report decryption feature;
- Public Folder archiving;
- PST migrations;
- Outlook Web App (OWA) integration including support for the new streamlined user interface

---

[1] EV 9.0 does not support Exchange Server 2010 Original Release. Support begins with Exchange Server 2010 SP1 and later.

So What Is New In Exchange Server 2010?

Feature	Explanation
Personal Archive	Administrators can create an on-line personal archive for each active mailbox to which users can manually move or copy mail items or set up rules to automatically move data from their mailbox.
Database IOPS Reduction	The IOPS requirement for each Exchange DB store has been reduced by 70% (over the previous release Exchange 2007, and 90% over Exchange 2003). This means that customers can utilise much cheaper, slower disk to host Exchange stores.
Max Mailbox and Database Increased	The maximum mailbox size has been increased, and the maximum database size has been increased to 2TB.
Max Number of Items Per Folder Increased	The maximum number of items that can be practically stored in a single folder in a user's mailbox has been increased to approximately 100,000.
Outlook Web App Interface Streamlined	The OWA interface has been streamlined to make it more user friendly and more like the Outlook experience.
Compliance and E-Discovery	Keyword based searching across the Exchange organisation with a multi-mailbox search application is now available. This allows searching of data across multiple mailboxes and their associated personal archives.
Legal Hold	User mailboxes can be placed on legal hold if necessary in order to protect the data contained within from modification or destruction.
PST Migration	Command line based PST import capability to allow for the movement of PST data into users mailboxes or personal archives.

Do I Still Need Enterprise Vault?

The answer to this really comes down to requirements. What are the basic needs and therefore requirements behind the Exchange Server messaging infrastructure? Requirements such as the need to retain access to all data online or offline or to keep all email for a period of 10 years using the smallest possible storage footprint while providing the ability to discover this data alongside other pre-collected or active content within the same discovery platform will point towards needs that cannot be satisfied with Exchange Server alone. It's these requirements that need to be extracted, analysed and then satisfied. Let's take a look at some of the key requirements:

Platform Requirements

For many customers archiving is about more than just Exchange Server. It can mean many things to different customers but the bottom line is that solving storage management headaches, application by application, is an out-dated and short sighted way to resolve the larger problem. The speed and volume of information growth within organisations of all sizes is huge. Solutions which aim to help manage this growth need to offer more than just point solutions for point applications as data growth is rarely constrained to a single application – especially those which encourage users to collaborate. A classic example of this is the humble PST file. Many years ago this became the Microsoft solution to mailbox quotas. Many years later this solution has led to headaches not only for the Exchange team but the desktop, file server, backup and even legal teams.

The solution to data growth in Exchange Server 2010 is now the introduction of a personal archive mailbox to which users can move data from their primary mailbox. While a step forward from PST files, the value that PST files provided (i.e. removing data from Exchange) was undeniable and a very real and major reason why moving the data to another mailbox within the Exchange environment can be seen to be undesirable, especially for customers with long term data retention plans.

Enterprise Vault's unique value proposition means that it is positioned as a platform from which data from many different applications and environments[2] can be managed (e.g. file systems, SharePoint and Instant Messaging). While requirements for data management may stem from data within Exchange Server they later may encompass data from other sources. Platform based solutions are the only way to cater for this expansion of requirements.

> **Key Point –** Customer requirements may dictate data management from many data sources – Enterprise Vault's platform approach meets these requirements where Exchange 2010 does not.

---

[2] Such as SharePoint Server, Instant Messaging and File System environments.

Storage Requirements

Exchange Server 2010 has had many improvements to its disk IO requirements, resulting in databases that can be deployed and hosted on very cheap disk. This is message behind Microsoft's message that larger mailboxes are now the way to cater for future information growth. This combined with an increase in the maximum number of items that can be practically stored in a single folder[3] can appear to be a pretty compelling message.

While storage has become more inexpensive over the last few years and will undoubtedly become even less expensive this fact does not cater for the value of data and the need to manage the ever increasing volumes. Throwing storage at a problem will only work for so long. What then? The value of the data is only as good as the application used to access that data. Therefore any long term repository of data must provide both logic to allow for data to be classified according to what it is and the end user interface to allow users to access it easily. We'll cover both of these topics shortly but the storage element is worth looking into.

Migrating to Exchange Server 2010 from previous Exchange versions will incur a hydration cost as previously single instanced messages are hydrated back into separate entities. Microsoft's own guidelines (Source: Technet) quote up to a 20% increase in DB size as a guideline, Gartner's investigations show up to a 30% increase in size and various independent consultants have reported up to 100% database size increases during a migration. Whatever the right answer is one thing is for sure, migrating to Exchange 2010 will result in bigger databases.

---

[3] Although in reality this increase in maximum items per folder only applies to Outlook in online mode I,e, non-cached. If cached mode is enabled (which for most users it is) then this improvement is largely irrelevant.

With these bigger databases in mind, the solution to information management and long term retention within Exchange Server 2010 is to simply throw more storage at the solution and then to duplicate this storage many times over to provide a level of high availability. The greatly reduced disk IO requirements of Exchange Server 2010 and therefore the ability to use low cost disk comes at a cost – that cost is total loss of any single instance storage. Any duplicate item stored within the Exchange stores is simply stored again and again as the data is replicated. While the high availability features of Exchange Server 2010 in the form of DAG's is undoubtedly a useful and capable solution for availability, the level of duplication it creates can be extreme – particularly if you are considering a PST migration of any type. This duplication is not limited to just the information stored in the databases but also includes the searchable indexes of the data.

The solution to information management and long term retention when using Enterprise Vault is to first of all remove the data from the primary application (freeing it up to focus on what it does best) and then to provide compression and single instancing[4] of all the data it stores. This capability is entirely storage agnostic meaning it does not rely on any specific type of storage device. This also means that data can be stored on any combination of supported storage devices with very little restriction. Data can be stored across standard NTFS disk remotely or locally, WORM devices, Content Addressable Storage devices, tape, DVD and more. This open storage layer also allows for this data to be tiered and moved over time to ensure that as data ages it gets stored on the most suitable storage medium. These features ensure that archived data has the smallest possible storage footprint and is stored on the most suitable storage medium.

> **Key Point –** Customer requirements may dictate storage savings and/or use of storage types not compatible with Exchange Server. Enterprise Vault offers an open flexible approach to the use of storage.

---

[4] Down to the attachment level and across all Exchange Servers.

PST Migration Requirements

PST migration is about more than just the movement of data from PST files. A project to eliminate PST files is normally wide ranging and complicated by nature and therefore the tools to support such an effort must match this in terms of functionality and expectation. Expectations for such tools include the ability to locate PST files, identify their owners, collect the PST data to a central point, migrate the data into an archive, apply single instance storage and compression[5], post process PST files (e.g. delete them, provide end user notification of what is happening and provide a suitable level of error handling).

Exchange Server 2010 provides minimal functionality in this area in the form of a command shell cmdlet to allow for a single PST to be imported into a user's mailbox or personal archive mailbox. There are no GUI's or front ends to the tool, and it exists to solely move the data from a specific PST file to a specific users mailbox/personal mailbox archive. Even though this functionality exists one must question the feasibility of whether moving PST data back into Exchange Server is really a good idea. Any customer who is currently making use of PST files is well aware that they proliferate just about everywhere and are substantial in volume. Moving this data into Exchange Server will solve the PST headache but bring a bigger headache in the guise of managing a huge Exchange infrastructure. Lack of single instance storage and compression capability means that this data will be the same size and volume as it was when in PST form. Many customers realize that this cannot and should not be the final resting place for this large PST data.

Enterprise Vault provides a series of tools which gives the ability to conduct and control a PST elimination project from design to finalization while importantly not requiring that the PST data is ingested into Exchange Server. The inbuilt tools allow customers to find and migrate PST data from file servers and end users desktops while giving options around parallel processing of PST files, end user notification, post processing and error handling for corrupt or password protection.

> **Key Point –** PST Elimination is more than just migrating data. Workflow and flexible tools are necessary to make this a reality and Enterprise Vault provides both.

---

[5] The nature and use of PST files means that their contents are often highly duplicative.

Content Management Requirements

Exchange Server utilises many different content types. From standard email messages to calendar items, tasks, posts, journal reports and more, being able to store and manage these many different content types is important. Exchange Server 2010 archiving is restricted to the movement of just email messages from the user primary mailbox to their personal archive mailbox. Items such as calendar appointments are not moved as part of the standard archiving process and must instead be moved manually into the personal mailbox archive.

Public Folders tend to represent a significant volume of data and having no direct solution for controlling this growth in Exchange Server 2010 means that this data is open to unrestricted growth as with previous versions of Exchange Server. Finally if email journaling is required for regulatory or discovery purposes, then there is no solution in Exchange Server 2010 (as with previous Exchange Server versions) for storing this volume of information[6]. The suggested solution is to store the resulting journal stream into a series of mailboxes which over time (and most customers apply a reasonably long term retention to journaled items) cannot scale very easily, due in some part to lack of compression and single instance, and in others to the maximum Exchange Server 2010 database size of 2TB. This makes this solution not only difficult to implement but highly unattractive.

Enterprise Vault provides a solution to take this journal stream directly from Exchange and store it into an archive which is compressed and single instanced and optimized for very large data sets. Enterprise Vault also provides a solution for archiving <u>any</u> content from mailboxes and Public Folders. Enterprise Vault also makes use of index conversion filters during the archiving process which can convert and index over 400 different content types. Exchange Server 2010 does also offer indexing of content in users mailboxes and personal mailbox archives but can only convert and index approximately 50 different content types. This could be a major issue during EDiscovery and may not be legally defensible.

> **Key Point –** Providing a solution for all data types is one of the critical paths for archiving from Exchange Server. Enterprise Vault is optimised to be the long term respository for all archived data.

---

[6] Some of the larger Enterprise Vault customers capture 6+ million journal messages/day equating to around 400GB per day

E-Discovery Requirements

E-Discovery and compliance requirements can vary greatly from industry to industry and country to country. The Electronic Data Reference model[7] (EDRM) attempts to establish guidelines for conducting discovery requests against electronic data. The ability to search, preserve and review electronic data is key to the model.

Exchange Server 2010 offers basic keyword searching across primary and personal archive mailboxes. Searches can only be run manually, and there is no native capability to schedule searches to run automatically. The results of search can then be exported and duplicated to discovery mailboxes where compliance officers can conduct a review of the data. The problem with this model is that when searches bring back hits numbering more than a hundred or so items[8], reviewing the items quickly becomes practically impossible. The marking of items becomes largely a process of trawling through large lists of items inside a user's mailbox. This makes sharing the review work among more than a single reviewer difficult to achieve. Finally the application of a preservation order on data (legal hold) is possible but must be levied against entire users mailboxes (i.e. a single relevant item found in a user's mailbox means that the entire mailbox is placed on hold). This hold request is applied using the Exchange System Attendant, which runs on a schedule and hence a legal hold order may not actually be applied until several hours after it was placed.

By comparison, Enterprise Vault customers can utilise the Discovery Accelerator workflow[9] or the Clearwell EDiscovery platform to provide full control over searching of all content (including un-archived content) and then presenting the search results for easy review by one or more reviewers. Searches can be run manually or scheduled to run on a regular basis – as is very common to ensure the results set is as up to date as possible. The review workflow also provides an easy mechanism to mark and tag items as relevant or not. Legal hold requests are applied immediately, as a high priority, and can be applied at the individual object level.

---

[7] Edrm.net

[8] Which just about any E-Discovery search will in a company of any size. Some search requests with broad search terms can result in hundreds of thousands of hits.

[9] Including full Chain of Custody auditing and reporting

> **Key Point –** E-Discovery requirements vary from customer to customer. Symantec's EDiscovery toolset can be mapped directly onto the EDRM best practice model.

User Interface Requirements

Finally the all-important user interface. When it comes to end users changing or breaking their daily routines/workflows is a cardinal sin. User productivity is paramount to all customers, so whatever process is put in place to help deal with ever increasing data volumes (mailbox quotas and PST files was the first attempt by Microsoft to deal with this), it must provide a seamless experience to help them to work more efficiently.

The Exchange Server 2010 personal archive mailbox is an online entity only. This means that unless the user is somehow connected to Exchange, they will not have access to the personal archive or anything in it. This can present substantial difficulties for users who are regularly disconnected from the network and yet forget to move important archived data back to their mailbox so it is available off line. The personal archive is only available for OWA, Outlook 2010 and Outlook 2007[10] users. Outlook 2003 or Mac based Entourage/Outlook users have no access to the archive.

Outlook cached mode provides users with access to their email data when offline, but even this capability is not without its limits. This limitation is one of the most compelling reasons not to create very large mailboxes in Exchange Server 2010 and is clearly outlined on the Microsoft TechNet website[11] when talking about Exchange Server 2010 client deployment scenarios:

> *"When to use Outlook online (non cached) mode: When very large mailboxes on which performance considerations become an issue in Cached Exchange Mode."*

OST file scalability is an issue for all versions of Outlook. Due to this users cannot have very large mailboxes or their personal mailbox archives available offline.

---

[10] Requires an SP2 hotfix to be applied in order to make the archive visible however there are restrictions. For example searching needs to be done independently in the archive and in the users mailbox.

[11] http://technet.microsoft.com/en-us/library/cc179067.aspx#BKMK_Choosing

Enterprise Vault provides full access to archived items both on and off line. The user interface is provided via Virtual Vault which presents items to end users via an additional container in their Outlook client which represents their archive. Users can search, browse, view, archive, restore and delete items in their archive dynamically, using drag and drop, and importantly using the same workflow as they are used to when dealing with their primary mailbox and even PST files. This user experience does not change as the users connection state to Exchange changes. This applies to Outlook 2003, 2007 and 2010.

Access to archived items from a Mac based Entourage or Outlook clients is also fully supported; however, it is restricted to shortcuts access and EV search.

Key Point – User interface is one of the most important factors when end users workflow is impacted by archiving. Key factors such as offline access to data can be the critical difference between acceptance and rejection of a solution.

Summary

Enterprise Vault offers a broad archiving platform across many different content sources and data types, and its place as an archiving platform is very different from Exchange Server's version of a storage management solution. As a result choosing to deploy Enterprise Vault becomes a case of analysing customer requirements and determining whether Enterprise Vault archiving is necessary. In many cases customers will need a solution that comprises elements of both Exchange Server and Enterprise Vault.

About Symantec

Symantec is a global leader in providing security, storage and systems management solutions to help businesses and consumers secure and manage their information. Headquartered in Cupertino, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation
World Headquarters
20330 Stevens Creek Boulevard
Cupertino, CA 95014 USA
+1 (408) 517 8000
1 (800) 721 3934
www.symantec.com