Things to consider when moving to Exchange 2013

January 2013

Prepared by: Andrew Abbate, Principal Consultant and author of Exchange 2013: Unleashed

Convergent Computing – *http://www.cco.com*

1450 Maria Lane, Walnut Creek, CA, 94596

Tel: 925.933.4800 – Fax: 925.933.4801

Exchange 2013 has been released and the necessary service pack for Exchange 2010 to support a migration to it is due out soon, thus it's an excellent time to consider what you should be doing when migrating to your new Exchange 2013 environment in order to take advantage of the new features and to offer the best experience to your end users. While you're at it, you might as well also think about how to make this environment easier for IT to manage and secure.

Clean up your Active Directory: Exchange has been dependent on Active Directory since Exchange 2000 and this doesn't change in Exchange 2013. Active Directory is the foundation on which Exchange 2013 will operate. If the foundation isn't stable, nothing built on top of it can be stable. Take the migration to Exchange 2013 as an opportunity to do a thorough health check and address any issues that exist in your directory. Common problems include inaccurate site/subnet mappings, legacy Exchange objects for servers that are no longer online and replication problems. By addressing these issues, making sure information is accurate and cleaning up unnecessary objects before moving to Exchange 2013, the odds of a successful and uneventful migration are greatly increased.

Reverse proxy: TMG has been retired, but performing pre-authentication and not publishing your CAS servers to the Internet is still a really good idea, especially if it can be used to benefit other critical applications like Lync or SharePoint. A reverse proxy, if you aren't using one already, allows for secured publishing of web based services. In the Exchange 2013 world, this would mean OWA, Activesync, Exchange Web Services and Outlook Anywhere. By using a reverse proxy that supports pre-authentication, one is able to validate that a user is authorized to access resources before their packets reach a protected system. Reverse proxies are traditionally not members of the Active Directory domain and are essentially built to be firewalls. A good reverse proxy will also support 2-factor authentication to further protect the user experience. If you aren't currently using a reverse proxy, strongly consider one. It can provide the same levels of protection to other web published applications like Sharepoint or Lync, making the investment even more worthwhile.

A relatively new use of reverse proxies is to take advantage of advanced authentication mechanisms to identify the user and the device and use that information to make intelligent decisions around granting access. For example, it's been a classic complaint in Exchange that all users are allowed to access Activesync by default and in order to control it, it must be enabled or disabled on a per user basis. Clever administrators have utilized Active Directory groups and custom PowerShell scripts to control this setting through group membership but it's not the easiest way to control access, nor does it give one the ability to say "Bob can access Activesync from a Tablet, but not from a Smartphone". By taking advantage of intelligent Reverse Proxy systems, one can configure rules in a very granular manner to decide which IIS subsites of Exchange a user can access, from what device, and whether or not to require a health check from the device in question. Having this type of functionality available can greatly improve the process of controlling access through simple group memberships.

Securing OWA: As much as OWA is "secure" and over SSL, that doesn't help when your users drop into a web café or use a kiosk that is running a key logger… Similarly, is it a good idea to just have port 443 open internally w/o any inspection? With mobile devices that are constantly exposed to insecure networks, is this really any less dangerous than having port 443 available externally? SSL won't protect you from a key logger, but 2-factor authentication will. By requiring a second factor to your authentication, the classic "something you know and something you have", you can protect against key loggers and network sniffers. 2-factor authentication can include methods like a One Time Password generating token, a digital certificate that's installed on an approved device, or systems that will "text" an OTP to a mobile device, to ensure it's the correct device. For environments that need an extremely high level of security around OWA, consider making all users VPN into the network in order to access OWA and don't publish it to the internet directly at all. While this may limit the devices that can access OWA, that can often be the tradeoff for increasing security. Also be very aware that OWA in Exchange 2013 has an offline mode that allows a user to interact with OWA even when disconnected from the network. This means that mobile devices like tablets may be holding e-mail content in its browser. This risk needs to be mitigated as well through tactics like encrypting the device's local storage or through clearing the browser's cache.

Mobile device security: With the proliferation of new phones and tablets, mobile device access to Exchange 2013 will only continue to grow. Employees' need to be constantly in contact with e-mail means that the risk to IT is greater than ever, as very often personal devices that are not managed by IT are required to be able to access e-mail either via Activesync or through a browser to access OWA. The biggest risk here is exposing internal systems to devices that are potentially unprotected or even compromised in some way. The only way to really protect the systems is to inspect the traffic before it reaches Exchange and to provide access only to systems that have passed some level of a health check. This allows one to perform a layer of intrusion detection to determine if the device connecting is doing anything suspicious, outside of a normal OWA or EAS communication. Similarly, one can create device access rules to enforce things like "Only let a device connect if the device has an approved anti-malware solution with a signature version x.z.y or higher and is running an approved version of operating system." This gives IT a powerful layer of security to prevent unprotected or unsupported devices from connecting and potentially compromising a system.

Load balancing: Exchange 2013 CAS functions change the way in which load balancing is used, but just because load balancing can be moved from Layer 7 to Layer 4 doesn't mean that the need for robust and stable load balancing has gone away. While it's true that any Exchange 2013 CAS can proxy communications to the Exchange 2013 CAS closest to a mailbox, that doesn't mean it's always a good idea to do so. By placing network layer logic that will connect a user to the most appropriate Exchange 2013 CAS, one can avoid unnecessary WAN traffic from clogging up expensive MPLS links. Similarly, if one opts to remove the Layer 7 logic from their load balancing strategy, one opens the door for poor utilization of systems. For Exchange 2013, the Microsoft Exchange team implemented hidden web

pages in each Exchange service sub directory (OA, OWA, EAS, EWS, AutoDiscover, etc) this hidden page was implemented to enable service level (layer 7) monitoring of each service. TCP or L4 monitoring can only monitor a TCP level connection to an IP address on a NIC, whereas Layer 7 can decide service by service if it's up and prevent a single service hiccup from dropping the entire CAS from the LB group, thus maximizing resource availability. While it's been suggested that one could replace load balancing in Exchange 2013 with DNS Round Robin, this is a recipe for trouble in the event that a CAS goes down, as "1/n" (n being the number of CAS systems) of client systems would still receive references to a CAS that's down.

Mailbox management policies: Upgrades and migrations are a great opportunity to revisit mailbox retention and Data Loss Prevention policies. Exchange 2013 offers native DLP and continues to support archiving and mailbox cleanup policies in both on-prem and off-prem. By implementing retention policies prior to migrations to Exchange 2013, one can greatly speed up the process of moving by not having to move potentially Terabytes of Deleted Items. There's nothing that depresses an Exchange administrator more than learning that 30% of their storage is holding Deleted Items that "e-mail hoarders" refuse to clear out. Having rules that regularly flush Deleted and Sent items is an excellent way to control mailbox growth.

Exchange 2013 also offers some very impressive rules for Data Loss Prevention that map directly to standards like PCI, SOX or HIPPA. By placing advanced Hub Transport rules, that utilize context in addition to pattern matching, and by layering this with workflow logic, Exchange 2013 is able to detect and control the flow of protected information. For example, one could set up a rule to prevent Social Security Numbers from being sent outside the company. If the person composing the e-mail containing SSNs were on Outlook 2013 and Exchange 2013, they would receive a Mail Tip popup that would tell them "you appear to be sending protected materials outside the company". Based on the policy set by the administrator, the e-mail would either be prevented from being sent or the user could click a link indicating they wish to send the message anyway. This could either trigger a workflow forward to a Compliance Officer who could approve the mail going out or it could be set to allow the message to send and simply notify the Compliance Officer that it happened. There is also the ability to place a link in the Mail Tip to allow the user to declare a "false positive" and have the Exchange administrator review the compliance rule to see if it's picking up false positives.

Layering the retention policies with the native archiving allows Exchange administrators to effectively extend the size of a user's mailbox without incurring the same expenses associated with giving the user a larger traditional mailbox. What this means is that by creating an archive mailbox to pair with the traditional mailbox, Administrators have the ability to create "cheaper" mailboxes, by placing them on less expensive storage. Similarly, one might choose to apply a less strict SLA on the archive mailboxes allowing them to be implemented at a lower cost. A common configuration is to place primary mailboxes on SAS or SAN storage and maintain 2 copies in the primary datacenter and 1 copy in a disaster recovery datacenter, then place the archive mailboxes on large SATA disks and maintain only 1 copy in the primary datacenter and potentially 1 copy in the DR Datacenter. This results in a much lower

cost per GB of mailbox for the archive environment. Taking this primary/archive approach is also very helpful for clients running Outlook 2007 or 2010 because only the primary mailbox can be placed into cached mode. Users today with very large primary mailboxes often complain about local performance because their OST file is very large and typically their laptop hard drive is relatively slow. By maintaining say a year of data in the primary mailbox and the remainder in the archive mailbox, the OST is kept fairly small resulting in excellent performance for the laptop user. This concept was extended in Outlook 2013 to allow users to control how much of their primary mailbox is cached, which is an excellent option for environments that don't implement archiving.

Access Auditing: Exchange 2013 (and 2010) offers the ability to audit administrative or delegate level access to mailboxes. This means that if enabled, when a "non-primary mailbox owner" access a mailbox, it is possible to create an audit trail of who accessed the mailbox and what they did. It can track things like moving a message, deleting a message or changes to its read/unread status. While this is very useful to determine if a mailbox is being accessed, it's also useful to understand what devices are accessing mailboxes in order to maintain compliance with industry security standards like SOX, HIPPA or PCI DSS. So while it's one thing to know that a delegate account was used to read messages and set them back to "unread", its entirely another to know that it was done from a device that doesn't belong to the delegate. By layering the ability to track device level access on top of account level access, one has the ability to create a very accurate and comprehensive view of what materials are being accessed, by whom and from where. This type of information is critical in order to remain in compliance with some industries' regulatory requirements.

WAN optimization: With improved mailbox density offered in Exchange 2013, more and more companies are consolidating their Exchange environments into fewer datacenters. While this reduces support costs, it places an increased load on the WAN. Between the increase in users accessing data across the WAN and the added traffic of mailbox database replication, WAN optimization in the areas of caching, compression and SSL offloading are more important than ever.

WAN optimizers can do some pretty amazing stuff with Exchange and Outlook. The two primary benefits of WAN optimization in Exchange are in the areas of DAG replication and User traffic. With Exchange 2013 (and 2010) most environments employ DAGs and almost all have at least 1 copy of databases in a WAN connected datacenter. The common reaction of the network team is "you want to replicate how much traffic?" It's not unusual to see 5-20 Mbps of log shipping generated during business hours, and this is traffic that needs to replicate to another site. While Exchange has mechanisms to allow queues to build up and complete when they are able, and this typically fits into peaks and valleys of traffic in Exchange and on the WAN, the concern is that if a primary site fails and there were very large copy queues, there might be more missing in the DR site than the Transport Dumpster can cover. In these cases, messages could be lost, so most environments would prefer to be able to keep up with replication in real time. This is where WAN optimization can be especially helpful

as an unencrypted and uncompressed DAG configuration can be compressed by as much as 75% via 3rd party WAN optimizers. Turning a 20Mbps requirement during peak hours down to a 5Mpbs requirement is much easier to accommodate and greatly reduces the risk of the queues falling behind the protection level of the dumpster.

The other area where WAN optimizers really shine are in situations where offices don't have local Exchange servers. Since a large percentage of e-mails are between users in the same physical location, one of the concerns with centralizing Exchange services is that in situations where a user in a remote office is sending an attachment to another user in that office, the content has to be sent to a WAN connected site and then retrieved over that WAN connection. In the case of WAN optimization that's been configured to support Exchange 2013 and Outlook 2013, the message is sent normally, with some compression and optimization from the WAN accelerator and then when the recipient goes to pull the message, the WAN optimizer looks at the incoming content and thinks "wait, I've already seen most of this message, I'll just grab the changes to the envelope and I'll send my locally cached bits rather than pulling them over the WAN" which results in a 90% or more reduction in the retrieval of the message. This can be very significant in terms of overall performance. WAN optimizers (also called WAN accelerators) are typically less expensive to implement than a comparable increase in bandwidth, resulting in an excellent ROI on these devices.

Namespace Consolidation: One of the big complaints in Exchange 2010 was the need for many environments to maintain multiple namespaces for various services. Seeing things like "NA_OWA.domain.com" and "EMEA_OWA.domain.com" were fairly common occurrences as it was necessary to resolve users to their correct entry point for OWA or other web related services. Exchange Server 2013 helps IT departments move toward a single namespace design. This is a more simple architecture in some ways, but admins have been using namespace as a way to segregate users, so with this going away in 2013, customers need a new way to support certain requirements such as sending certain groups of users to CAS arrays in certain locations versus others. While it's true that the CAS architecture in 2013 allows each CAS server to operate like a stateless proxy for connections to the mailbox servers, the fact is that in geographically dispersed deployments, it could force more traffic than necessary across WAN links. To respect these WAN constraints as well as other more straight forward requirement such as sending half of users to West coast and half to East coast to reflect 2 DAG separation (and co-locating CAS near mailbox servers), a network solution can be used to identify and route users based on those rules in a single namespace environment. Being able to consolidate into a single namespace makes support easier for both users and the helpdesk as there is no longer a need to figure out where someone is located before determining where they should connect. By moving the logic to the network layer and to some degree to Exchange, it's a few less things for IT to worry about. It can also save a few bucks on Subject Alternate Name certificates.

Planning the migration itself: With all the focus on surrounding technologies and strategies for security the new environment, it's easy to lose track of the migration event itself. One of the most important things an administrator should do is to establish a good pilot group. These pilot users should be aware of the implications of being moved to a new system and that they might have to suffer the occasional reboot as systems and processes are tested and tuned. These pilot users should also represent an accurate cross section of the environment and they should be users who aren't afraid to tell IT if there's something they don't like or if some part of the process is impacting them. The other critical thing to gather during this pilot is how long it takes to move content. As much as mailbox moves are a background event between Exchange 2010 and Exchange 2013, it's still important to understand how long it's going to take to complete the migration in order to set realistic expectations. Moving mailboxes across a WAN connection may result in very different throughput than moving across a LAN. Being able to predict how long each location might take (especially if consolidating into fewer datacenters) is an important piece of the overall project.

Exchange 2013 has further improved its native Move Mailbox tools to help manage the process. One of the really nice improvements is the concept of Batch Mailbox Moves. In Exchange 2013, all mailbox move jobs get a "batch name" and have the ability to send notifications during the move with reporting. The updated tool also adds functionality when moving a user with an archive, as administrators can pick different targets for the primary and archive mailboxes. Mailbox moves can be prioritized and Exchange 2013 now supports incremental syncs to destination mailboxes to "pre-move" the bulk of the content so that mailbox moves can happen very quickly. In the past, the ability to pre-seed and incrementally sync mailboxes was exclusive to third party migration tools.