Virtualizing Exchange 2016: the Right Way!

Michael Van Horenbeeck
Microsoft Certified Solutions Master
Introduction

Both Microsoft Exchange Server and the concept of virtualization have been around for a long time. Yet, the topic of virtualizing Exchange is still highly debated. Every now and then, someone will make a statement about why virtualization for Exchange is better than deploying on physical hardware — sparking another debate about the sense or nonsense of it all. While discussions often lead to good things, the intention of this white paper is not to make any bold statements, start a new controversy or leave you wondering about whether or not you should virtualize Exchange. After all, there are plenty of valid reasons to virtualize Exchange — some of which we will discuss in this white paper.

Microsoft does a decent job of documenting its recommendations for virtualizing Exchange. Rather than stating the same facts again, this white paper aims to provide a better understanding behind the rationale of those recommendations and requirements. We will also cover why some virtualization features make more sense than others, specifically because some of Microsoft’s recommendations can seem opposed to certain features and recommendations from hypervisor vendors.

Virtualization is often deeply rooted into an organization’s IT strategy. As such, it is unrealistic to assume that you should only deploy Exchange on physical hardware. On the other hand, you must keep in mind that Exchange is designed to run on physical hardware. That is why some features are designed the way they are and why some of those features might not always make a lot of sense from a virtualization point of view.

We will a look at virtualizing Exchange from the Exchange administrator point of view, providing administrators with the information required to help virtualize Exchange in the right way. Throughout this document, we will explore topics such as high Availability (HA), storage, sizing of the Exchange environment and many more. While some of the views expressed in this white paper might conflict with the views and recommendations of particular hypervisor vendors, it is impossible to take into account unique capabilities from each hypervisor platform. If you believe that one of your virtual platform’s features can help you virtualize Exchange in a better way, we encourage you to explore them.

Regardless of how you decide to deploy Exchange, you should always have the supportability of your solution in mind. While some hypervisor features might be unsupported by Microsoft, using these features could be supported by the platform manufacturer. Needless to say, you want to make sure there are no alternatives before considering going down the unsupported path. Running an unsupported configuration does not mean that Microsoft will not help you at all. However, if you ever need to escalate an issue to Microsoft, there might be additional work involved for you. It is not unreasonable of Microsoft Support to ask you to reproduce the issue on a fully supported configuration in order to rule out any impact potentially caused by the unsupported configuration.

Before you continue reading

- Throughout the white paper, various Exchange Server concepts and features are discussed. While most of these concepts are explained to a certain degree, the reader is assumed to have some experience in designing, configuring and managing an Exchange 2013 or 2016 architecture.

- This white paper is based on the information and recommendations available at the time of writing (March 2016). While we assure that the information in this white paper is correct at its release, we cannot ensure the validity in the future as a result of changing recommendations or new product capabilities.

- This document contains references to external resources. Veeam® does not control these external resources and therefore cannot guarantee the availability thereof, nor can Veeam guarantee the content on those external resources will not change in the future.

- Veeam conducted a webinar with members of its Technical Product Marketing Evangelist team and the author of this white paper, to discuss the various aspects of virtualizing Exchange. The webinar is a great companion to this white paper, because it elaborates on the information and concepts that are addressed in the following pages.
A quick introduction to Exchange 2016 and the Preferred Architecture

When Microsoft released Exchange 2016 to the public in October 2015, not many organizations immediately made the jump. However, as Cumulative Updates for Exchange 2016 and other highly recommended components (such as the Office Online Server) become available, organizations will start planning to move to Exchange 2016.

Although there are some technical and architectural changes too, most improvements revolve around an enhanced end-user experience which is now more closely aligned with Office 365:

- **Improved Outlook Web experience.** Outlook Web App is formally renamed to Outlook on the web. This change reflects the many improvements to the client as it attempts to more closely match feature capabilities in the full Outlook client. Although there is still a significant discrepancy, Outlook on the web is now — more than ever — a viable alternative for a lot of information workers thanks to improved collaboration features such as the side-by-side viewing of attachments, new compose features, inline images, additional themes and more.

- An on-premises SharePoint 2016 server unlocks **new collaboration features**, like saving attachments to OneDrive or linking OneDrive files instead of adding attachments to the message itself.

- Exchange 2016 continues to focus on **new and enhanced eDiscovery scenarios**. For instance, Public Folders now support In-Place eDiscovery and In-Place Hold. These existing and new eDiscovery scenarios are supported by improved Compliance Searches. Unlike Office 365, these compliance searches are only available through PowerShell using the various -ComplianceSearch cmdlets.

![Image 1: Outlook on the web's new side-by-side attachment preview](image1.jpg)
Along with the improvements for a better end-user experience, Microsoft has shifted to using MAPI/HTTP as the default connectivity protocol. Although RPC/HTTP is still available, Microsoft has started de-emphasizing its use in favor of MAPI/HTTP. MAPI/HTTP is the successor to RPC/HTTP (also known as Outlook Anywhere) and was built from the ground up to be more efficient and robust in today's interconnected world. For instance, MAPI/HTTP recovers more quickly from so-called micro-outages in the network — something that is not uncommon when using flaky Wi-Fi connections.

There is a lot more that has changed in Exchange 2016. Most of the improvements, however, have little to no impact on how to best virtualize Exchange. As such, these changes are beyond the scope of this white paper and not discussed here.

## Exchange 2016 Preferred Architecture

The architecture of Exchange 2016 is not dramatically different from its predecessor. However, there is one important change which potentially impacts how some organizations deploy Exchange. In accordance with prior guidance, Microsoft collapsed all server roles into a single unified role. The Exchange 2016 Mailbox Server role is now the only one left, if you don't count the Edge Transport Server role. There is no longer a separate Client Access Server. This should not come as a surprise as Microsoft's guidance to deploy multi-role servers has been around since Exchange 2010 Service Pack 1. Typically, only very specific situations called for a deviation from those guidelines.

In order to further emphasize the design guidelines for Exchange, Microsoft updated the *Preferred Architecture* which was first published for Exchange 2013 and details what the optimal Exchange deployment architecture looks like. As the name implies, the *Preferred Architecture* reflects the ideal way to deploy Exchange in order to reduce the cost of the messaging environment, maximize the features built into the product and the increase the overall Availability of the deployment.

Unlike support statements, which define exactly what you can and cannot do, the *Preferred Architecture* does not dictate anything. Instead, one should always attempt to adhere to the guidelines as closely as possible. It’s perfectly acceptable to deviate from the guidelines if they are not compatible with your organization’s technical or functional requirements, or if there is some constraint to deploying Exchange as depicted by the Preferred Architecture.

---


© 2016 Veeam Software
Some key points from the Preferred Architecture for Exchange 2016 are:

- Built-in data protection features that enable you to take advantage of low-cost, commodity hardware to deploy Exchange on the physical hardware instead of virtualizing. Server resources should not exceed 96 GB of memory and have a maximum of 24 CPU cores. If you use commodity hardware and disks, a battery-backed write cache controller is recommended.

- Use at least two, preferably three data centers. The third data center allows for automatic site failovers to occur. If you only use two data centers, manual intervention might be required to activate the second data center in case of a data center failure. Having multiple data centers dramatically improves the solution's resiliency and optimizes the use of certain HA features, like SafetyNet.

- Deploy a single Exchange namespace per client connectivity protocol and per data center pair. A data center pair is considered to be two data centers in which mailbox servers from a single Database Availability Group reside. If you have Exchange deployed in multiple regions, you might have multiple data center pairs. This could, for instance, be the case if you have locations in Europe and North America.

- Deploy a Database Availability Group with at least four members — each hosting a copy of a database for a total of four database copies: One active, two passive and a lagged database copy.

- Deploy Office Online Server for the full end-user experience. For HA reasons, you should deploy at least a pair of Office Online Servers in each data center that hosts Exchange 2016 servers.

Note: The items in the list above do not represent all recommendations. Microsoft’s guidelines go into much more detail, including the use of other data protection features.

The Preferred Architecture and Exchange virtualization

Many people find Microsoft’s recommendation to deploy Exchange on physical hardware somewhat bizarre. After all, virtualization has proven to be a solid technology with many benefits.

According to Microsoft, some of the reasons for recommending deploying Exchange on physical hardware are:

- Depending on the size of the hosts, and in order to handle failures appropriately, resource utilization on hosts should kept at a level that allows Exchange virtual machines (VMs) to move to those hosts within the virtual server farm in case of a failure. Because this potentially affects the overall guest density on those hosts. This is not a desirable outcome for most virtualized environments trying to maximize host resource utilization.

- Virtualization may add an additional layer of complexity and management overhead.

- There are little to no benefits from additional recovery and HA features, available through virtualization because these features generally only protect against a limited set of failures. Application-level HA on the other hand, offers both protection against hardware-level failures and failures caused at the operating system or application layer.

Virtualizing Exchange 2016: the Right Way!

While these arguments absolutely make sense from an Exchange Server perspective, they might need to be nuanced a bit in light of the topic of virtualization.

First of all, virtualization allows you to better use resources on a server. The hardware that is available today is often too powerful for a single application. Depending on the application, deploying it on a dedicated, physical server would not be a very efficient use of that server’s resources. As a result, the host server would be idling for most of the time. Instead, virtualization allows you to share those excess resources with as many applications as the host has resources for. As such, resource utilization goes up and you now use the available hardware more efficiently. Because you can host multiple applications on a single server, you have to buy less hardware and thus drive down cost.

Secondly, most hypervisor platforms have built-in HA features which allow you to migrate VMs from one host to another, often with limited to no downtime incurred. An important note here is that there is a difference between planned and unplanned outages — at least for Exchange. As described in more detail later, planned migrations can occur on the fly with little impact. Unplanned outages, and the resulting VM moves on the other hand, must result in a reboot of the VM, therefore causing a longer outage than one that is handled at the application layer (e.g., by activating a passive database copy). Regardless of the use case or how the move happens, having such mobility for your VMs is quite beneficial.

In order to take advantage of these features, you often have to deploy a cluster of servers, shared storage, a dedicated network, etc. which potentially leads to more complexity and, in turn, to overhead and more room for error. Complexity should always be avoided as much as possible. A less complex solution will always be easier to maintain and thus result in higher Availability.

An Exchange 2016 server should not be configured with more than 96 GB of memory and 24 cores. Anything more can adversely affect the server’s performance. While a machine with such specifications is already considered to be a high-performance server, it is not uncommon to see virtualization hosts that have access to even more memory and processor cores. Especially for the latter category where a host might have access to 40+ cores and 512 GB of memory, a single guest assigned with 96 GB of memory and 20 CPU cores using 80 percent of its allocated resources is not necessarily a problem. The story is obviously different if you have a host with 24 CPU cores and “only” 128 GB of memory.

There is no denying that virtualization adds a certain amount of overhead in terms of management, complexity and resource utilization, but it is very hard to quantify exactly how much. There’s a number of elements that influence how much overhead and complexity there is to maintain the virtualization infrastructure, but this topic is out of scope of the white paper. Besides, in any professional, well-run virtualized environment these complexities should already be dealt with as Exchange is probably not the only workload where certain rules must be adhered to in order to make virtualization deliver the best results.

If your organization has not already adopted virtualization, it would not make a lot of sense to do so solely for Exchange. If you previously invested in a virtualization infrastructure, it also does not make sense to let those investments go to waste in order to deploy Exchange on physical hardware no matter what. Whether or not it makes sense to virtualize Exchange is a case-by-case decision and depends on numerous things such as prior investments in a virtual infrastructure, the resources you have available and whether or not you are able to deploy additional physical machines at all.

Exchange 2016 contains a lot of features to help safeguard data and maximize Availability of the Exchange infrastructure. Many virtualization platforms also offer a variety of interesting HA features, and although the paradigm is similar for both, the main difference is that Exchange’s HA features operate with full knowledge of the application’s state and logic whereas HA features from the virtualization platform are application agnostic and tend to be only useful when hardware fails. Additionally, the time to recover from such failures is generally higher than when relying on the built-in Exchange capabilities.

None of the HA features offered through the hypervisor protects against failures within the Exchange guest or from logical corruption of the Exchange databases. Truth be told, there’s few occurrences of the latter these days, but that’s largely because of some the features built into Exchange to prevent it from happening in the first place.

**Best Practices for Virtualizing Microsoft Exchange Server**

Depending on who you ask, you might get a different answer to the question of how to best virtualize Exchange. An Exchange administrator is likely to follow the best practices from the application’s point of view whereas a virtualization administrator will focus on the hypervisor and platform best practices instead.

Often, vendors will publish their own views and recommendations on how to best deploy Exchange on their solutions. While vendor-specific white papers form a great basis, you must always keep one important thing in mind: The final solution should be supported by both Microsoft and your virtualization vendor. If a recommendation from the vendor does not align with Microsoft’s requirements or recommendations, you will have to carefully consider whether or not you want to follow that specific recommendation.

Hypervisor vendors can participate in Microsoft’s Server Virtualization Validation Program (SVVP) which allows them to validate their solution(s) with Microsoft. In return, Microsoft supports running its products on those validated configurations — provided that the implementation on the alternate hypervisor does not generate a conflict with any of the requirements and constraints as depicted by Microsoft.

For instance, Microsoft fully supports running Exchange on VMware vSphere 5. However, if the underlying storage technology is NFS-based, Microsoft will not support that part of the deployment — regardless of whether VMware supports it or not. The better choice would be to use a supported storage technology, like presenting block-based storage directly to the guest running Exchange and ensuring NFS is not used in any layer of the storage solution.
Designing Exchange for a virtualized environment

There are many design decisions involved with the development of a new Exchange architecture. Some of these decisions are influenced by the expected workload like the amount of users, while others depend on technical, functional or perhaps even legal requirements.

The simple truth is there is almost no difference in how you design a virtualized Exchange environment versus when deployed on physical hardware. There is nothing that should stop you from pursuing the Preferred Architecture at all times.

Sizing

Properly estimating resources for a new Exchange infrastructure is a very important task. Without the right amount of resources, an Exchange server will underperform and ultimately affect the end-user experience. In extreme cases, it can even compromise the stability of your deployment.

Sizing for an Exchange server isn't easy. The process is lengthy, lots of variables are at play and it is easy to make mistakes because of the complexity. Proof is the lengthy article from Microsoft outlining the entire process.

Exchange Server Role Requirements Calculator

To reduce the complexity of the process, Microsoft developed the Exchange Server Role Requirements Calculator early on which uses a wide range of input parameters to return information such as the amount of storage required, storage I/O, processor utilization and memory requirements. As with most processes, bad input leads to bad results. It is important that you provide accurate information so the tool can do its job correctly. If you don't, you could end up with a server that is dramatically under- or over-sized.

Processor requirements

As mentioned earlier, it is recommended to configure Exchange 2016 with no more than 24 CPU cores. Although there is nothing that stops you from assigning more cores to the guest running Exchange, it could lead to decreased performance.

One of the perceived benefits in virtualization is that you can assign VMs with more virtual cores than physically present or available for use in the host system. If you do so, you are over-subscribing resources — more specifically CPU cores — in this case.

When multiple guests exist on a single host, the hypervisor ensures that all VMs are granted access to the resources they have been assigned to. The logic behind over-subscribing resources is that applications rarely use 100% of their assigned resources and, if they do, it's highly unlikely they would do so at the same time. Nonetheless, over-allocating resources is still a risky business. What happens when VMs are actually consuming more resources than what's available, even when it only happens infrequently and during peak usage times?

---

Over-subscribing CPU cores can lead to so-called micro-delays in the CPU cycles assigned to the VM and have disastrous effects on the Exchange Server’s performance. When a CPU is busy, new tasks are queued for processing. When this happens, an Exchange Server VM must wait for a host’s CPU’s to finish executing current tasks so that CPU resources become available to the guest for execution. Queueing happens all the time, and because of the speed at which modern CPUs process instructions, often incurs no performance penalty. However, as queues grow larger, so do the resulting delays causing more noticeable and potentially problematic delays.

The ratio of total assigned virtual CPUs-to-physical cores in a system hosting Exchange VMs should never exceed 2:1. Even if you have a ratio lower than that, you must closely monitor the Exchange Server’s performance, and if necessary, adjust overall resource allocation on that system until you reach the point that performance is satisfactory. Obviously, not over-subscribing at all is even better and Microsoft recommends a 1:1 ratio for Exchange VMs. When considering CPU ratios, you must size as if Hyperthreading is not enabled. For example, if a host system had 16 physical CPU cores but had a total of 64 vCPUs assigned across all VMs hosted on the system, the ratio would be at 4:1 (not supported by Microsoft for hosting Exchange VMs). Alternatively, if a host system had 16 physical CPU cores but had a total of 32 vCPUs assigned across all VMs hosted on the system, the ratio would be at 2:1 (supported by Microsoft for hosting Exchange VMs).

**Hyperthreading (HT)**

The idea that a processor, and by extension a processor core, is capable of multi-tasking is nothing more than an illusion. Although it appears a CPU is handling multiple operations at a time, it executes them sequentially. HT is a mechanism that allows the processor to more efficiently schedule incoming tasks and thus process operations more quickly.

Whether or not an application benefits from HT depends on a range of elements. In case of Exchange, Microsoft determined that there isn’t a clear advantage to using HT because of the little performance gain of HT itself and the negative impact of increased memory allocation because of a higher number of logical CPUs.

Although Exchange does not benefit from HT and Microsoft recommends disabling it for physical deployments, you can use it in a virtualized deployment. The only requirement is that during the sizing process you consider the actual physical cores of the server — and not the logical processor cores yielded by HT. For instance, when a CPU has four physical cores, it might have eight logical cores when HT is enabled. For sizing calculations, you must use the four physical cores. As previously stated in the Processor requirements section, when determining CPU ratios, you must act as if HT is not enabled.

**Memory**

Exchange has always been considered to be a ‘memory-hungry’ application. This is mainly because of how the Exchange Information Store process works. In Exchange 2016, the amount of memory that is made available to a database is calculated during the Information Store service startup. This is for example one of the reasons why you must restart the service after adding additional database copies to a server.
Hypervisors typically add another memory management layer on top of in-guest memory management in order to efficiently share memory across several VMs. Not only does this allow hypervisors to more efficiently deal with memory allocations, but also to dynamically manipulate the amount of memory assigned to the guest. How this is done, varies from one virtualization solution to the other. Some platforms revoke memory without exposing it to the guest whereas other solutions will visibly lower the amount of memory. Taking away memory from an Exchange Server is never a good idea as Exchange does not handle it well.

Dynamic Memory, which is Hyper-V’s way of dynamically adjusting memory allocation, over-committing memory or any other way to manipulate memory allocation, is not supported because of the disastrous effects it has on Exchange. Therefore, Exchange VMs hosted on Hyper-V must use “Static Memory,” or in other terms, have Dynamic Memory disabled.

Over-subscribing memory is principally the same as over-subscribing CPU cores where you assign more memory to VMs running on a host than physically available to the host. When VMs on that host effectively utilize all their allocated resources at the same time, the over-subscription turns into over-commitment.

VMware uses a technique called ballooning, which runs in the background. Ballooning is different in that it does not expose memory being removed to the guest. Even though Exchange isn’t aware of the memory being revoked and used elsewhere, the performance impact is significant.

The following image illustrates the effect of dynamic memory on an Exchange server. The Available Bytes clearly depicts when Dynamic Memory kicked in. The RPC Averaged Latency represents the end-user experience. Note the correlation between the drop in the Available Bytes and the resulting spike in the RPC Averaged Latency.

![Image 2: Dynamic Memory at work. Source: Microsoft](image-url)
Another way to mitigate potential memory management problems is to work with memory reservations — if applicable to your platform. As with most features, there are benefits and downsides from doing so. Generally speaking, configuring a memory reservation is a safer choice because it signals the hypervisor that it must guarantee the assigned amount of memory to that VM — potentially at cost of other guests running on that host. The downside is that when you have a memory reservation set for a VM, it can only be moved to another host that has at least the amount of memory available as what is assigned to the VM. If you have large Exchange Servers, for example with 96 GB of memory, other hypervisors in the cluster must have at least 96 GB of available memory if they are to be considered as potential targets for moving those VMs to. VMware recommends not overcommitting memory on systems hosting Exchange VMs. If this cannot be avoided, setting a memory reservation to the configured size of the Exchange Server VM is recommended.

**Non-Uniform Memory Access (NUMA)**

Older server platforms often used a single system bus where one, or more, processors access the computer systems physical memory. As applications continue to grow more resource hungry and the amount of data that has to be processed becomes larger, congestion occurs on that system bus. This is especially apparent within multi-processor systems. Non-uniform memory access (NUMA) tries to solve that issue by dividing the system’s resources into different NUMA nodes. Each NUMA node is comprised of a processor and memory to which the processor has prioritized access. All the NUMA nodes in a system are linked together through an advanced controller which allows one processor to access memory from another processor if required. The interaction with such non-local memory is much slower than when a processor accesses its own locally allocated memory. When an application is NUMA-aware, it takes advantage of the architecture and schedules its operations within the boundaries of the single NUMA node. This avoids remote memory access scenarios and potentially speeds up processing.

Although Exchange is not NUMA aware, it can take advantage of the host system’s support for NUMA — if available. If you want to have the optimal performance, it is recommended not to exceed the resources available to a single NUMA node. For instance, if a NUMA node has access to 16 cores and 64 GB of memory, the Exchange VM should ideally not exceed those specifications. Depending on the size of your deployment, this means that you might have to artificially limit the amount of resources assigned to the VM. In turn, you might have to compensate for that by scaling out and adding more Exchange servers to handle the entire workload.

If you can’t limit the resources to a single NUMA node, it’s recommended to expose the NUMA topology to the guest. Although it cannot guarantee that all operations are confined to a single NUMA node, the efforts of the Windows operating to try and intelligently schedule resources within the boundaries of a single NUMA node can be helpful and improve performance.
Storage and Storage I/O

Exchange has grown from being an I/O-intensive application, in its early days, to being highly efficient, with moderate I/O requirements. Today, it is capable of running on cheap disks that don’t typically offer a very high I/O.

The thought of hosting your data on a simple hard drive makes many people nervous. The fear of losing data, that potentially results from unreliable simple hard drives, is not entirely unwarranted. Although no exact numbers exist, hard drives in general are considered to be the least-reliable parts in any computer system. Hard drives are one of the few mechanical parts in a computer. The platters inside a hard drive spin at a constant rate of 5,400 rpm or more which, logically, leads to an increased wear and tear on the parts of the disk. If not for the expected wear, vibrations and other factors like ambient temperature or temperature swings are also major influencers on a hard disk’s reliability.

In order to overcome some of the negative effects of hard disks, several hard disks can be configured in a Redundant Array of Independent Disks — better known as RAID. The benefits of using RAID is that it increases resiliency, and depending on which type of RAID, also increases the performance by combining the throughput of several disks in the array. The downside is that RAID requires a capable controller, and depending on the type of RAID, you might have to sacrifice one or more disks to guarantee data redundancy.

For instance, RAID 1 combines two disks and presents them to the system as a single disk. Data is written sequentially to each disk, but each disk can independently service separate read requests. When one drive fails, no data is lost as the other drive contains exactly the same data. The failed disk can then be replaced, after which the controller rebuilds the data set by replicating the exact contents of the surviving disk to the replaced one. If an application only has a single copy of its data, RAID is invaluable and a reliable way to ensure data protection and Availability. It’s no surprise that RAID configurations are quite common in the enterprise!

The problem with RAID configurations is they are hardware-based redundancy and application-agnostic. RAID cannot protect data from in-application issues such as logical corruption, it cannot prevent data loss from a controller failure, nor can it protect the application from some OS-level failure. One huge benefit of hardware redundancy is that it buys time for replacing hardware, for instance when a disk fails. This extra layer of protection theoretically allows someone to wait until Monday morning before replacing a disk that failed on Saturday night, for example. This is especially useful in smaller environments that do not have 24/7 IT operations, but do require HA for their applications or systems.

Instead of relying on hardware redundancy to protect its data, Exchange creates and maintains additional database copies on other servers in a cluster, called a Database Availability Group (DAG). A DAG is fully application aware. A variety of processes allow Exchange to maintain several copies of a database and automatically leverage those additional copies to respond to several types of outages. If a server fails, another copy of the database is activated so that service to clients is maintained. It can solve several forms of logical corruption that might have crept into one of the database copies. When a corrupt page is detected in the active database, a non-corrupt version of the page from one of the other database copies is used to replace the corrupt page. By doing so, a DAG protects Exchange and the data it contains from a variety of threats in a more granular way than typical hardware-based redundancy. In a sense, a DAG allows the HA to be handled at the application layer (Exchange) instead of the hardware layer (RAID).
Virtualizing Exchange 2016: the Right Way!

Although technically possible to use JBOD (independent disks) or Storage Spaces Direct, most virtualization systems require some type of shared storage to make many of the virtual platform’s HA features work. The shared storage often comes in the form of a Storage Area Network (SAN) which, when reduced to its essence, is nothing more than a RAID with a lot of disks and supporting hardware such as a controller, enclosure, etc. The software that controls the SAN, along with the built-in hardware redundancy, offers a fairly high level of protection for anything that it stores making it the default choice for most virtual deployments. Nonetheless, even SANs can fail. So, storing all your database copies on one physical array is generally not such a great idea. While SANs may have their own data redundancy solutions (SAN Replication for example), they rarely allow a timely failover of service. Especially when compared to Exchange, which can failover a database often times in less than 15 seconds!

Whatever storage solution you decide on to host Exchange and its data, it must be block-level — except when using SMB 3.0 (code which Microsoft owns and can carefully control) to host VHDX files for Exchange VMs. This means that NFS-based storage is not supported. This point is heavily debated, and many (wild) theories exists about why Microsoft does not support NFS. In the end, the simple truth is that there are many different NFS implementations out there — some better than others. While some solutions might perfectly meet Exchange’s specific requirements and offer a sufficient amount of throughput and reliability for Exchange, the lack of standardization or reliable cross-solution performance guarantees make it very hard to generically ratify NFS as supported by Microsoft. Could Microsoft setup some certification program through which other vendors can validate their NFS solution for Exchange? I believe so. However, the question is whether or not the lack of support for NFS-based storage is truly such a problem. If you have no other option than to use NFS-based storage, carefully assess the risk you expose yourself to and determine how to best deal with it moving forward.

There are plenty of alternatives such as in-guest iSCSI or pass-through disks (Raw Device Mappings) that allow you to present block-level storage directly to the guest. The problem with in-guest iSCSI is that it breaks through the virtualization layer and has functionality implications from a virtualization point of view. Using in-guest iSCSI also requires additional attention with regards to the network design of your virtual platform. Successful iSCSI implementations require a lot of effort and often the costs of doing so aren’t justifiable. Hence why many virtualization vendors do not recommend using in-guest iSCSI. Instead, alternatives like placing virtual hard disks on (shared) block-level storage (SMB 3.0 or host-based iSCSI) are much better.

Validating storage requirements

Microsoft Jetstress is a tool to test and verify the performance of the storage subsystem on which Exchange will be installed. By using the same read/write pattern as a typical Exchange server, the tool stresses the storage subsystem in order to yield I/O performance results that can be compared to the requirements from the Role Requirements Calculator.

Although Jetstress was originally developed to be used on physical deployments only, it can also be used on a variety of virtual platforms. But, not all virtual platforms are supported. So, it is possible that Exchange VMs might be sharing storage with other workloads when virtualizing Exchange. While Microsoft strongly recommends having dedicated spindles for Exchange, it may be unavoidable to share spindles on a SAN with other workloads.


© 2016 Veeam Software
When running Jetstress on shared storage, some organizations choose to run Jetstress when utilization is low. The rationale is to minimize a potential performance impact on other applications running from that same storage. As you can imagine, doing so defeats the purpose of the test and largely voids the results. The intention of Jetstress is to mimic a real-world Exchange server. In order to obtain reliable results, test conditions must match the actual deployment as closely as possible.

**Thin-provisioning**

Another feature of a SAN is the ability to thin-provision disk space. Thin-provisioning is when the storage subsystem dynamically allocates storage, but presents the VM the full storage capacity. For instance, with thin-provisioning enabled, the guest might see a disk of 80 GB, of which only 10 GB is actually consumed on disk. This is very different from fixed-size disks where the full amount of storage is allocated from the start and therefore unavailable to other systems — regardless of whether the disk space is actually used within the guest or not. In such case, you are potentially wasting disk space on the SAN. Undoubtedly, thin-provisioning offers certain advantages over fixed-sized disks. It is one of the ways in which a virtual platform can potentially help drive down costs.

Just like memory and CPU cores, you can over-subscribe disk space, meaning that you assign more (virtual) disk space than physically available in the storage subsystem. In order to benefit from thin-provisioning, you must meticulously keep track of disk space usage and growth over time which enables you to determine at what point you must buy additional storage to avoid running out of disk space. Correctly forecasting future disk space requirements isn't easy and requires a high level of operational maturity.

The storage requirements yielded by the Role Requirement Calculator represents the worst-case scenario in which all mailboxes grow to their specified maximum mailbox size. Even though it is highly unlikely that you ever end up in such a situation, it's much safer to ensure that sufficient disk space is available without continuously having to wonder if and when you must purchase additional storage. This is especially true if you expect the storage to be consumed anyhow. Additionally, you should plan for a sufficiently large buffer in case you cannot add disk space quickly enough. This would, for instance, be the case where the storage vendor cannot deliver disks quickly enough.

The risk in this scenario is an Exchange VM that believes it has an 80 GB volume with plenty of free disk space. However, because the backend storage solution has depleted all available storage, the Exchange databases will likely be dismounted and in a corrupted state due to the inability to write to data blocks.

**Deduplication**

Deduplication can help by eliminating redundant copies of data and thus significantly drive down disk space usage. Despite the perceived benefits for Exchange, there are almost none. That is also one of the reasons why Microsoft does not recommend using SAN vendor deduplication with Exchange and does not support Windows-based file system deduplication.

Deduplication can be most effective when the data you store across multiple guests is very similar. How much disk space you can save varies from one solution to the other and the type of data that is stored.

Imagine having 20 identical VMs, each running Windows Server 2012 R2 and consuming 15 GB of data. Without deduplication, you're using nearly 300 GB of storage, just for running those VMs. With deduplication, you could in theory reduce disk utilization to a single instance of 15 GB.
While the thought of freeing up so much disk space sounds appealing, it is very unlikely you will ever see such results with Exchange. The problem lies in how deduplication works. Often the deduplication process takes places as data is written to disk. Each block of data is inspected and a blueprint of each written block is kept for future reference. If a future block of data is identical to one of the blocks that were written to disk before, the new block is replaced with a pointer to the former block. As you can imagine, the more identical blocks you have, the higher the storage reduction.

If you have multiple databases, you might be tempted to believe that each database contains identical mail data and that deduplication can lower your footprint on disk by many times. While databases may often contain the same set of data, that data isn’t written to each database in the same way, let alone using the exact same data blocks. Each unique database file has a unique checksum, therefore at the block-level it will look different than a different database containing the same data. When working with DAGs, you can have multiple copies of a mailbox database, each with the same checksum and similar at the block level. So, in theory, you could save space by deduplicating each copy of a given mailbox database. It would require each copy be on the same storage solution capable of delivering the deduplication. The downside of this is that you’ve introduced a single point of failure for your multiple database copies and therefore defeated the purpose of the copies altogether. So, even though deduplication will have some effect on your databases, it’s unlikely that the process can drive down storage utilization enough to make it worth the risk.

To restate, the risk associated with deduplication is that you’re undermining the idea of having multiple copies. Even when from an Exchange perspective it looks like you have multiple copies, you are essentially trying to reduce the data to a single instance on the physical disk. Imagine what would happen to all your Exchange database copies if something happened to the SAN.

Most, if not all, SANs offer a very high level of reliability through a series of protection features and redundant hardware. Yet, despite all those safety features it’s not unheard of that a SAN fails, wreaking major havoc in the process. Replicating data from one SAN to another further reduces the risk associated of a failing SAN. But, then aren’t you just mimicking what Exchange does, albeit with much more expensive storage instead?

**Flash storage**

Now that prices are reaching acceptable levels, the use of flash-based storage is on the rise. Using all-flash storage can have a very positive impact on the performance of your I/O-hungry application thanks to the insane performance flash storage offers over conventional hard drives.

Although Exchange used to be an I/O monger, it no longer is today. Despite being supported, there is simply no need for the additional performance that flash drives have to offer. Why invest more money if you can make do with less?

Some storage solutions combine flash drives with traditional hard disks in a so-called tiered storage approach. In such a scenario, data is moved from slower storage to faster, potentially flash-based, storage on an as-needed basis. Similar to an all-flash solution, there is little need for Exchange to host data on the faster storage tier. Because of the various processes that continuously interact with data in the Exchange databases, it is even possible for Exchange data to be permanently moved to the faster storage tier, undermining the entire paradigm of tiering. Just like with all-flash arrays, there is no benefit to tiering Exchange data. All it does is add another layer of complexity to your deployment. Microsoft does not recommend tiered storage for Exchange for these reasons.
Snapshots

Some hypervisor platforms allow taking snapshots of VMs. Snapshots capture the state of a running VM at a given point in time — very similar to taking a picture of a real-life situation. One of the benefits of using snapshots is that you can go back in time to restore one of the snapshots. By doing so, you restore the state of the VM to the point in time the capture was taken. Exchange does not support taking snapshots (or reverting to one), other than for lab environments, because snapshots are not application aware. As such, reverting to a snapshot can have unpredictable consequences on data for which a specific state is maintained — such as Exchange’s databases. Also, some configuration data for Exchange is held within Active Directory, meaning that a discrepancy could be created between Active Directory and Exchange should either be reverted to a snapshot.

VM snapshots should not be confused with VSS-based snapshots which use a supported way to create a VSS-snapshot of a database. This is commonly done when taking a backup of a database and is fully supported, as long as the snapshots are taken through the VSS framework.

High Availability (HA)

Planning for HA means you have to understand what can cause an application to fail. Those threats can be categorized in so-called Failure Domains, such as the network infrastructure, the virtual platform, Active Directory, hardware, witness servers etc.

In order to effectively provide a highly available solution, you must account for each of these failure domains. Solely relying on the ability to execute host-based failovers is generally not a good idea as it only addresses a limited amount of failure domains. However, combining the best of both worlds ensures that you can benefit from the host-based failover technologies as well as the application-aware features within Exchange.

One of the major advantages of a virtual platform is the ability to move VMs from one host to another, for example when a host fails. Most platforms can move a VM which was running on a failed server to another server within a matter of seconds. During that time, the VM on the failed server becomes briefly unresponsive.

For stateless applications, this results in a hiccup of a few seconds. When the VM is then successfully transferred to another host in the cluster, the application typically returns to operating normally. For Exchange, and especially when configured in a DAG, not responding for a few seconds and then returning to normal operation can be a problem. To avoid any issues, following a failover event, the Exchange server should be brought back online by means of a cold boot (a reboot of the machine).

All member servers in a DAG send heartbeats to let each other know they are still there and operating properly. By default, this happens every second. During a failure, for instance when a virtual host goes down, the Exchange VM on that host would stop sending the heartbeat signaling the other members of the DAG that something is wrong. After a few consecutive missed heartbeats (five by default), the other members will start a procedure to evict the failed server from DAG operations. In the meantime, the cluster service on the failed node is restarted in an attempt for it to rejoin the DAG. Once network communications have been reestablished, the evicted node should be able to rejoin the DAG automatically.
There are a few conditions that warrant tweaking the cluster failover timeouts. Although the defaults should work for the vast majority of the deployments, having timeout settings that are too aggressive can lead to problematic DAG operations. Typically, this is observed when a random DAG member is evicted from the DAG — although nothing seems wrong. This could be the case in an environment with an elevated network latency or where the planned move of a VM from one host to another is taking more time than anticipated. Planned moves of a VM, unlike moves following a failure, do not require that VM to be rebooted and are supported by Microsoft. When a VM is moved from one host to another, the memory contents from the guest are gradually copied over to the target host. During the final transition (cutover, if you will), the VM is silenced (paused) for a brief moment to allow the final memory contents to be moved to the alternate host before resuming operations. If this final stage happens to take more than five seconds, you could run into the aforementioned problem.

Recommendations for cluster heartbeat settings might vary from one vendor to another and can be affected by your specific configuration. It’s best to consult with your hypervisor vendor to find out what values to use — if they need to be changed at all. If you decide to modify the timeout settings, do so with precaution and incrementally increase the heartbeat timeouts until you reach a value that no longer affects DAG operations.

**Virtual Machine Affinity Rules**

Because VMs can roam freely across hypervisor hosts in a cluster, it’s not unthinkable that two or more Exchange guests could end up running on the same host. When this happens, the server hosting multiple Exchange VM now bears a greater responsibility and potentially reduces the resiliency of your deployment.

Using Anti-Affinity Rules, you can define which VMs should not be hosted on the same server. When the hypervisor platform makes a decision to move a VM from one host to another, it will take those rules into account and avoid moving Exchange VMs onto the same host — unless there is no other option.

Needless to say, this only works if you have sufficient other hosts in your virtual cluster capable of hosting your Exchange guests.

**Data Recovery with Veeam**

Microsoft’s Preferred Architecture does not mention configuring backups. After all, having a sufficient number of database copies eliminates the need for a backup — at least to a certain extent. In reality, things aren’t always as simple or as straightforward as that. Not every customer has the ability to deploy at least four Exchange servers including all its requirements and dependencies. While one could argue whether or not those environments are perfect candidates to move to Office 365, is beyond the point. A lot of organizations still need backups. For instance, when they require an out-of-band strategy in order to ensure data recoverability in any situation.

Backups can be considered an insurance in case something catastrophic happens to your infrastructure. If you have deployed a DAG with four members across two geographically distant data centers, it is very unlikely that you will ever lose the data in both data centers at the same time.

However, smaller environments that do not have the luxury of having a second data center are exposed to a greater risk and can benefit from the additional safety a backup provides. Of course, that is only true if the backup data is stored off site and you perform regular restore tests. There is nothing more annoying than to find out you are unable to restore data from a backup when you actually have to.

The long-term preservation of historical data is another reason to keep backups. While today’s technology allows quasi-unlimited mailbox sizes — enabling you to virtually keep everything forever — it’s often the human factor that plays an important role in the decision to embrace this new paradigm or not. Even though many administrators would probably like to, you can’t control the actions of an end user, and as such, it is hard to prevent situations like items accidentally being removed.

Features like In-Place Hold can help to ensure that data is kept indefinitely. When an In-Place Hold is enabled, items that are purged out of the mailbox, for instance because the user purposefully removed items from the Deleted Items or the items expired past the Deleted Items Retention period, they are not physically purged from the database. Instead, the items are moved into a hidden folder in the user’s mailbox called the Recoverable Items folder. Depending on how the items were deleted, they could be moved to different places within the recoverable items folder. Most folders in the recoverable items folder are hidden to the user. However, the deletions folder (where items that have been hard deleted (shift+delete) from the Deleted Items are moved to) is exposed to users through the Recover Deleted Items feature in Outlook.

Restoring an item from the Recoverable Items Folder is not an easy process, because you can only do so through PowerShell. On top of that, when an item is restored from the Recoverable Items folder, it loses its hierarchy. The latter is especially tricky when a user accidentally deleted an entire folder and wants the folder restored.

Having a solution, like Veeam Backup & Replication™, which plugs in onto your existing infrastructure and can help you to easily find and restore items through a GUI, is a major advantage for customers unable to leverage the Preferred Architecture to the fullest. Even if you have deployed Exchange following the guidelines from the Preferred Architectures, you still benefit from the product’s ability to easily restore individual items.

---

Backups and disaster recovery (DR)

Backups can be used in a variety of situations. Depending on the type of backup, you can restore individual items, databases or even entire VMs. The latter might sound like an interesting option for Exchange, but is it really?

Restoring a VM snapshot of the OS including Exchange configuration is, even without databases, very tricky — especially if the server was a member of a DAG. Instead of trying to explain how to restore a full VM, let’s take a look at how you can use Exchange’s built-in recovery options to restore a failed server.

Assuming that you have multiple mailbox database copies, recovering a failed Exchange Server is a relatively straightforward process. From a very high level, the process looks like this:

1. Create a new VM with the same specifications, name, IP address and OS as the failed server.
2. Reset the computer account in Active Directory and join the new server to the domain so it uses the same computer account.
3. Run the Exchange setup program using the /m:recoverserver switch. During the installation, information is pulled from Active Directory to ensure the replacement server assumes the roles and configuration from the previously failed server.
4. Re-join the new server to the DAG and reseed databases for it to resume its duties as before.

**Note:** When using this approach, you assume that the failed server is broken beyond repair. At no point should the failed server be brought back online.

Because the server was part of a DAG and other database copies are available, Exchange itself can create new database copies on the restored server by reading from the remaining database copies. But, what if you don’t have additional database copies or what if they all failed?

In the scenario where you have no other database copies left, Veeam Backup & Replication can complement the built-in recovery process:

1. Create a new VM with the same specifications, name, IP address and OS as the failed server.
2. Reset the computer account in Active Directory and join the new server to the domain so it uses the same computer account.
3. Run the Exchange setup program using the /m:recoverserver switch. During the installation, information is pulled from Active Directory to ensure the replacement server assumes the roles and configuration from the previously failed server.
4. If applicable, re-join the new server to the DAG and reseed databases for it to resume its duties as before.
5. Restore (individual) databases from Veeam’s application-aware backup.
6. If applicable, configure additional database copies for the restored database on other members in the DAG.
Application-aware backups with Veeam Backup & Replication

Earlier in this paper, we discussed that reverting to a snapshot is not supported by Exchange because a typical snapshot is not application-aware and can cause some serious damage to Exchange and its data. Although Veeam Backup & Replication also takes a snapshot of the VM, it leverages the VSS framework to create a consistent and application-aware snapshot.

To enable the feature, you must first select the Enable application-aware processing while creating a new backup job:

After an application-aware backup is initiated, a run-time management service is deployed in the process, allowing it to retrieve data from the Exchange VM. This management service invokes the Exchange VSS Provider which then creates a VSS snapshot of the Exchange databases. Afterward, the VSS snapshot is saved into Veeam Backup Files.

When time comes to restore the data from the backup, you now have multiple options: Restore the entire VM, restore files within the VM or restore Application Items. The third option launches Veeam Explorer™ for Microsoft Exchange through which you can restore individual items from and to a user’s mailbox.

When restoring individual items, Veeam Explorer mounts the Exchange Database it backed up using VSS. This allows you to browse the database, as if it were mounted to an Exchange Server, and restore items from any user’s mailbox or folder to the same mailbox, an alternate mailbox or a PST file.
Conclusion

Virtualizing Exchange is not dramatically difficult or different from a typical physical deployment. Even though Microsoft recommends deploying Exchange on physical hardware, it’s perfectly acceptable to virtualize Exchange. In some situations, virtualizing Exchange makes even more sense.

Microsoft has pretty strong and well-defined support statements and recommendations for virtualizing Exchange. These recommendations aren’t pulled out of thin air and are created to protect you from deploying Exchange in a way that might jeopardize your data. As such, it’s best to follow these guidelines to ensure a smooth deployment.

To recap, here’s some of the key points we discussed throughout this white paper:

• Whether you are virtualizing Exchange or not, always stick as closely as possible to the guidelines from the Preferred Architecture (Except to account for the overhead from the hypervisor — there is no difference in how to size your architecture for a virtualized or physical Exchange deployment)

• When virtualizing Exchange, make sure to respect the support statements from Microsoft and work closely together with your virtualization platform team to ensure that the Exchange architecture is fully supported

• Disconnecting the OS or application workload from the hardware has its benefits, such as portability and mobility. However, the built-in data protection features in Exchange are very powerful and exist to prevent data loss in almost every failure scenario, taking away the need for virtualization features. If you do decide to virtualize, leverage the best of both worlds and take advantage of certain hypervisor features in addition to the built-in Exchange capabilities

• Although features like In-Place Hold allow you to preserve and restore virtually any data inside user’s mailbox, using these features can be cumbersome — especially when trying to restore single items
About the Author

Michael Van Horenbeeck is a Microsoft Certified Solutions Master (MCSM) and Office Servers & Services MVP from Belgium, placing him amongst a handful of experts, world-wide, to hold both the MCSM certification and MVP award at the same time. As an independent consultant he has helped customers all over the world to implement, or migrate to Exchange and Office 365.

He is a dynamic tech enthusiast, specializes in a variety of Microsoft technologies including Office 365, Exchange, and (Azure) Active Directory and is highly skilled in all things hybrid — that is how he got his nickname “Van Hybrid”.

Besides writing about technology for a variety of tech websites, Michael co-authored several books including the recent ‘Office 365 for Exchange Professionals’ e-book along with Tony Redmond and Paul Cunningham. He also loves to talk about technology and regularly speaks at various conferences like Microsoft TechEd, Ignite and IT/DEV Connection.

You can follow Michael via Twitter (@vanhybrid) or his blog vanhybrid.com.

About Veeam Software

Veeam® recognizes the new challenges companies across the globe face in enabling the Always-On Business™, a business that must operate 24/7/365. To address this, Veeam has pioneered a new market of Availability for the Always-On Enterprise™ by helping organizations meet recovery time and point objectives (RTPO™) of less than 15 minutes for all applications and data, through a fundamentally new kind of solution that delivers high-speed recovery, data loss avoidance, verified protection, leveraged data and complete visibility. Veeam Availability Suite™, which includes Veeam Backup & Replication™, leverages virtualization, storage, and cloud technologies that enable the modern data center to help organizations save time, mitigate risks, and dramatically reduce capital and operational costs.

Founded in 2006, Veeam currently has 39,000 ProPartners and more than 193,000 customers worldwide. Veeam’s global headquarters are located in Baar, Switzerland, and the company has offices throughout the world. To learn more, visit http://www.veeam.com.
NEW Veeam® Availability Suite™ v9

RTPO™ <15 minutes for ALL applications and data

Learn more and preview the upcoming v9 release
vee.am/v9