Best Practices Guide for Exchange 2010 and Tegile Systems' Zebi Hybrid Storage Array

Version 2.0: May 2013

Contents

The Exchange Story

Unquestionably, e-mail and enterprise messaging are business-critical applications. More often than not, these communications run on Microsoft Exchange Server. Although Microsoft has released Exchange 2013, Exchange 2010 continues to dominate the on-premises e-mail and enterprise collaboration market due to the product's streamlined nature. With Exchange 2010 Microsoft added significant features over previous versions, which enables the product to offer improved availability, search, and recovery, but it also presents challenges in terms of storage efficiency, management complexity and resilience, particularly due to changes that Microsoft has made in Exchange architecture over the years. This paper discusses these architectural changes and the ways by which the Tegile Zebi line of arrays can address them.

The Tegile Systems' Zebi hybrid array provides the most balanced storage solution for Microsoft Exchange Server. Using Tegile's patent pending Metadata Accelerated Storage System (MASS) technology, Zebi accelerates Exchange Server performance up to seven times, delivers 14 times more mailboxes, and occupies up to 75 percent less storage space than legacy storage arrays.

In addition, Zebi simplifies storage decision and management, supports resilience with application consistent data protection, and offers iSCSI, Fibre Channel (FC), NAS, and CIFS capability in the same box. Zebi provides unified storage for a wide variety of enterprise applications, lowering the total cost of ownership (TCO), improving business productivity, continuity and service levels, and optimizing recovery point objective time (RPO) and reducing recovery time objective (RTO) to minutes instead of hours.

In order to achieve adequate performance, efficiency and resilience, Exchange administrators need to perform storage planning and verification tasks prior to deployment. This Best Practices Guide provides a brief introduction to the Zebi solid-state-disk based(SSD-based) hybrid array and provides recommendations on:

- Storage planning and verification
- Storage-related configuration best practices for hosts, virtual machines (VMs), host bus adaptors (HBAs), switches, and storage arrays
- Achieving optimal performance, efficiency and resilience using Tegile Zebi storage arrays

Zebi Metadata Accelerated Storage System (MASS)—The Ultimate in Efficiency

The Tegile Zebi hybrid array is powered by Zebi Metadata Accelerated Storage System (MASS) technology and delivers the trifectaof storage—high performance, high capacity and high reliability at a low cost.

Traditional disk array storage systems store data and metadata together, with metadata interspersed with data on hard disks. Over time, due to incidences of modified, deleted and rewritten data, the metadata becomes fragmented. Furthermore, features such as de-duplication can lead to the rapid multiplication of metadata, causing significant performance deterioration. Utilizing MASS, the Zebi network storage array system organizes and stores metadata on high-speed devices with optimized retrieval paths. This accelerates all storage functions within the system, raising the performance of near-line SAS hard disk drives to the level of very costly high-RPM SAS or Fibre Channel (FC) drives. As a result, even an entry level Zebi array can provide customers with a storage solution delivering tens of thousands of IOPS, but at a fraction of the cost of a comparably performing traditional storage array or an all flash array.

Because near-line SAS hard disks offer much more capacity than high-RPM SAS/FC disks, Zebi performance is similar to SSDs and provides the high capacity of near-line hard disk drives.

A NOTE TO EXPERIENCED EXCHANGE ADMINISTRATORS

Microsoft Exchange 2010 no longer provides any of its own data de-duplication, as it did in previous versions of the product. In Exchange 2003, Exchange provided fully de-duplicated e-mail messages and attachments through a feature Microsoft calls single-instance storage. In Exchange 2007, Microsoft eliminated single-instancing for e-mail messages, but still provided the service for message attachments, which can be quite large. Eliminating single-instance storage for messages in Exchange 2007 enabled Microsoft to reduce the system's I/O footprint by orders of magnitude, greatly simplifying storage design for Exchange planners. In Exchange 2010, Microsoft has taken the final step and fully eliminated single-instance storage. As was the case with Exchange 2007, this decision by Microsoft has again significantly reduced Exchange's I/O footprint, making it a much nimbler and more manageable product.

Eliminating de-duplication from Exchange has also created new capacity challenges. Now, duplicated e-mails and subsequent attachments are duplicated for each recipient. If a message is sent to 100 recipients, that message and any attachments to it will be stored 100 times in the Exchange database.

By utilizing Zebi de-duplication, duplicated e-mail content is only stored once. In essence, Tegile customers have gotten back the great capacity benefits that were inherent in Exchange 2003 while still getting the great performance gains that are reality with Exchange 2010. It truly is a "best of both worlds" scenario.

Inline Compression and De-Duplication

Integrated inline compression and de-duplication reduces real-world capacity requirements by 75 percent with only minimal performance impact. This makes data de-duplication truly operable for mission-critical applications like Microsoft Exchange, even those that require high levels of performance and low latency (Figure 1).

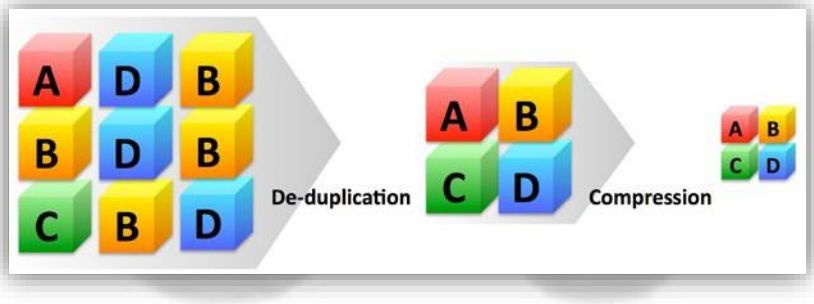


FIGURE 1: A REPRESENTATION OF DE-DUPLICATION AND COMPRESSION

Exchange Storage Reference Architecture

Figure 2 gives an example of Microsoft Exchange 2010 database availability group (DAG) high availability (HA) reference architecture.

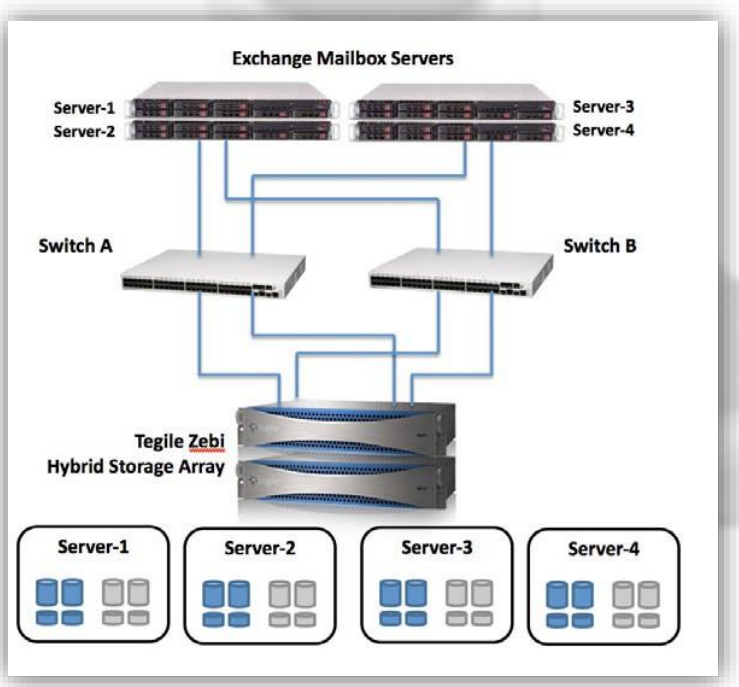


FIGURE 2: EXCHANGE 2010 DATABASE AVAILABILITY GROUP (DAG) HIGH AVAILABILITY REFERENCE ARCHITECTURE

In this example, the DAG contains four mailbox servers, as well as eight active and eight passive databases. Each Exchange server hosts an equal number of active databases and passive databases. Additionally, each Zebi controller hosts an equal number of active and passive databases and log volumes. A no-single-point-of-failure design is applied to servers, switches, storage and cabling, and the database, log and I/O load are balanced across all hardware components.

In addition, Zebi has the ability to host virtual servers for hub transport and client access servers, enabling the additional HA and disaster recovery (DR) benefits that are a result of the underlying server virtualization infrastructure.

Storage Performance Sizing

Performance is one of most important factors to evaluate while searching for a suitable Microsoft Exchange storage solution. Understanding Exchange's I/O characteristics helps buyers select the right storage solution and storage configuration.

During normal Exchange operations, there are two primary types of I/O:

- **Database** that is primarily 32KB random read and write. Database IOPS and latency are the two most important performance factors to measure and verify prior to deploying a Microsoft Exchange Server solution.
- **Transaction logs** that are 4KB and up to I/O size sequential read and write.

There could also be additional background database maintenance and log replication IOs.

Microsoft Exchange mailbox database needs are the primary driver when it comes to storage IOPS requirements in an Exchange environment. The total necessary IOPS can be determined based on the following factors:

- Estimated daily message volume
- Number of mailboxes
- Different mailbox accounts tiers
- DAG configuration

As displayed in Figure 3, the recommended general formula for calculating the total required IOPS of Microsoft Exchange databases is:

FIGURE 3: EXCHANGE IOPS CALCULATION FORMULA

Keep in mind the following:

- The number of IOPS per mailbox depends on the daily outgoing and incoming message volume and DAG or non-DAG configuration. For example, a mailbox with a daily message volume of around 100 messages requires 0.10 IOPS per mailbox for a DAG configuration.
- The number of database copies depends on a DAG or non-DAG configuration. For non-DAG configuration, there is only one database copy; for DAG configuration, there could be two or more database copies.
- There is some wiggle room. This includes 20 percent overhead to accommodate an unexpected load to the Microsoft Exchange Server that could occur during peak times.

As an example, consider the following sample Exchange characteristics:

- 30,000 mailboxes
- 100 daily messages each
- Two DAG copies

30,000 × 0.10 × 2 × 120% = 7,200 IOPS

Some deployments will have multiple tiers of mailboxes. Organizations often separate mailboxes into separate tiers in order to, for example, establish different quotas for different people, and plan for different mailbox traffic characteristics. In this case, Exchange storage planners should use the same formula above for each mailbox tier and then add the resulting IOPS figures. By applying this formula, IT managers and administrators can determine the total required IOPS for entire environments.

Exchange is an application that carries some latency requirements. Microsoft mandates the following requirements for the following transactions:

- **Database read.** Average latency ≤ 20 milliseconds, maximum latency ≤ 200 milliseconds
- **Transaction log write.** Average latency ≤ 10 milliseconds, maximum latency ≤ 100 milliseconds, but it's highly recommended that the database write average latency remain at under 20 milliseconds.

Jetstress tests should be done to verify that these performance requirements are met for the target exchange profile before product phase.

Tegile offers a range of products and solutions that meet both IOPS and latency requirements of Microsoft Exchange Server and support the deployment of tens of thousands of mailboxes on a single entry-level array.

Bear in mind that database and transaction logs carry different I/O characters in term of I/O size, intensity and randomness (random versus sequential I/O). The configuration for database and transaction logs needs to consider potential difference in I/O characteristics. Achieving the best performance usually requires the proper

WHAT IS JETSTRESS?

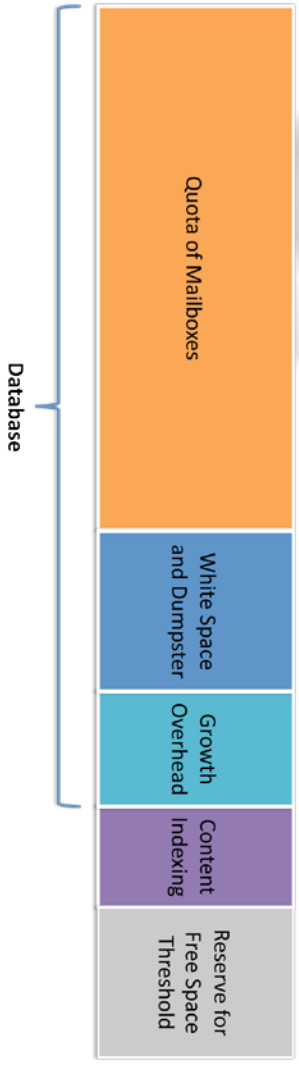
Jetstress 2010 is a tool which enables Exchange administrators to verify the performance and stability of a disk subsystem prior to putting a Microsoft Exchange Server 2003, 2007 and 2010 server into production. Jetstress helps verify disk performance by simulating Exchange disk I/O load. Specifically, Jetstress simulates the Exchange database and log file loads produced by a specific number of users. Use Performance Monitor, Event Viewer, and ESEUTIL in conjunction with Jetstress to verify that storage meets or exceeds established performance criteria. After a successful completion of the Jetstress Disk Performance and Stress Tests in a non-production environment, an administrator has ensured that the Exchange disk subsystem is adequately sized for the user count and user profiles.

configuration deployment of the host hypervisor (for environments in which Exchange is virtualized), HBA storage switch and storage systems. In the End-to-End Best Practices Section later in this document, detailed recommendations are provided to help IT managers and administrators achieve optimal Microsoft Exchange performance with Tegile Zebi storage arrays.

To assist in planning, Microsoft has developed the Exchange 2010 Mailbox Server Role Requirements Calculator, which can be used to calculate storage sizing and performance requirements, even for complex Exchange implementations. For more information about Exchange storage I/O requirements, please refer to Microsoft's Understanding Database and Log Performance Factors resource page on TechNet.

Storage Capacity Sizing

Perhaps the most important question that Exchange administrators ask themselves is, "How much space do I need to allocate for the mailbox databases and logs?" The answer is not as simple as tallying the mailbox quota and multiplying it by the number of mailboxes in the database. Microsoft Exchange Server requires capacity overhead for some functions, such as outgoing and incoming mails, single item recovery, search and more.



If the database and log files are not configured with sufficient space, issues will arise when the mailbox database nears capacity and the available space on the database or log volumes approaches zero. If space runs to zero, Exchange no longer allows new messages to be sent or received and unpredictable service issues can occur. As such, capacity sizing is a crucial step in Microsoft Exchange Server planning.

On the database side, the aforementioned required overhead space is for:

- **White space**. As items are removed from a database, the location at which that content resided may not be immediately usable until database maintenance is performed. During this time period, those locations are unavailable and known as white space.
- **Dumpster**. The Dumpster is the location at which "soft-deleted" email items reside. By default, soft-deleted items are stored for 14 days and calendar items are stored for 120 days in Exchange 2010. After these time periods, items are permanently deleted. While items exist in the Dumpster, they can be recovered, if necessary.
- **Content indexing**. Content indexing is a feature which accelerates Exchange's ability to find content based on a user query. In order to provide this service, the Exchange database must store a content index.

On log side, the extra space is needed for log retention in order to meet backup strategy and backup window requirements. In addition, factoring in 20 percent overhead is needed to handle unexpected growth conditions. Also, an additional database restore volume should be reserved for offline database maintenance and database restore purposes. Finally, extra free space should be buffered since many operating systems issue a warning alert when capacity utilization reaches certain thresholds (for example 80 percent).

After applying these various factors, the required capacity for each mailbox could become almost twice as large as the mailbox quota.

Database and log volume configuration has a significant on impact Exchange transaction I/O performance. In the End-to-End Exchange Storage Best Practices section, detailed configuration best practices are provided to achieve the optimal performance and efficient capacity provisioning.

High Availability (HA) and Site Resilience

Because Exchange is considered mission-critical software in many places, HA and site resilience are often implemented in enterprise Microsoft Exchange Server deployments. To provide HA and non-interrupted service, a no single point of failure (NSPF) architecture is necessary, which entails achieving server, switch, and storage redundancy, as well as the interconnections among them. Tegile Zebi arrays have built-in HA support and NSPF storage design, as well as fully redundant storage controllers, power supplies, fans, and RAID support to protect data from possible drive failures. This makes Tegile Zebi arrays a perfect fit for even the most critical Exchange implementations. Plus, it enables full hardware fault tolerance from the top to the bottom of the Exchange stack.

On the database side, Microsoft Exchange Server 2010 includes the use of DAGs as one approach to achieving HA at the mailbox server and database level. A DAG is simply a combination of an active Exchange database coupled with a number of passive database copies. A DAG can include up to 16 mailbox servers that host a set of databases and provide automatic database-level recovery from failures that affect individual servers or databases. In addition to protecting databases, DAGs can be used to achieve site resilience which is critical in disaster recovery scenarios. Each database has one active copy and one or more synchronous and asynchronous passive copies. In the case of an active database storage failure or mailbox server failure, the service can continue through other mailbox servers and passive database copies protected in the DAG, even in remote locations.

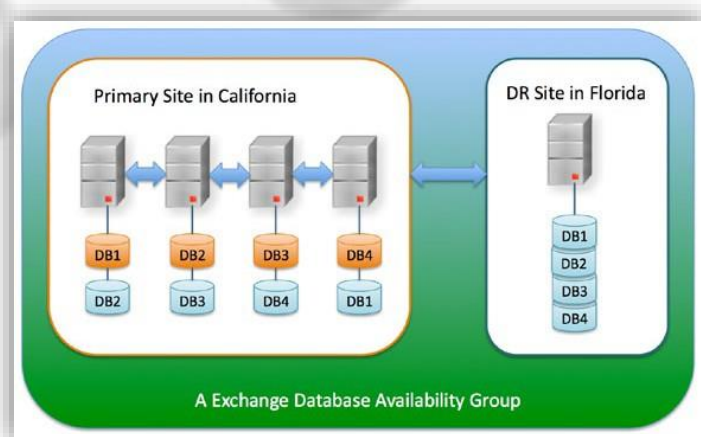


FIGURE 4: EXCHANGE 2010 DAGS IN ACTION

Figure 4 shows a DAG configuration that provides HA in a single site and disaster recovery across sites. The DAG contains four active mailbox servers on the primary site which hosts two database copies—one active (orange) and one passive (blue). It also shows a passive mailbox server in the remote disaster recovery (DR) site, which hosts a passive copy of each database.

Depending on the resiliency requirements and the distribution of the active mailbox users, multiple DAG groups can be configured to better suit an organization's needs.

DAGs add additional I/O load to data storage because of operations such as database synchronization and log replication. Since DAGs work on the concept of database copies and because Exchange no longer does any single-instancing, DAGs also significantly increase storage capacity requirements by two times or more.

With Tegile Zebi, there is a better way. Even while providing unparalleled high performance and sizeable capacity, Tegile's Zebi take it a step farther by providing inline compression. Its de-duplication technology can further lower storage investment costs by reducing DAG database and log capacity by up to 75 percent.

For resiliency, and as alternative to DAG, Tegile Zebi provides remote database and log replication. Zebi remote replication is based on an application consistent snapshot, which means that the databases and logs on the remote Zebi array are always transactionally consistent. When the primary site is down, the Microsoft Exchange Server in a DR site can continue providing mail services, which minimizes service interruption during disaster recovery.

Backup Strategy

There is no one-size-fit-all backup strategy; every organization has unique needs. Depending on the RTO and RPO service level agreement, a combination of the following can be used to implement a backup strategy:

- Hardware based VSS integrated VSS snapshots built into the storage array
- Remote replication built into the storage array
- Microsoft Exchange 2010 DAG for HA and resiliency
- Microsoft Exchange 2010 single item recovery
- Microsoft Exchange mailbox archiving
- Third-party backup and archiving software

The Tegile Zebi array supports RAID and NSPF designs to provide redundancy at hardware level in order to alleviate the need for backup should a storage hardware component fail.

The array also offers snapshot support, which is integrated with Microsoft VSS to provide application aware point-in-time data protection. This can be an effective strategy against data

loss due to a malicious attack or human error, such as accidental deletion of Exchange objects. It is an efficient solution, which conserves capacity by only storing changes instead of duplicating the entire databases, and by allowing recovery within minutes rather than hours.

When choosing hardware VSS-based snapshots to backup data, it can provide additional capacity for the data due to Zebi's built-in data compression and in-line de-duplication.

Besides the mailbox databases and logs, the logs and configuration data on the hub transport and client access servers should be included as part of a backup strategy. By storing Microsoft Exchange Server roles on the Zebi array, similar VSS integrated snapshots and remote replication can also be employed to protect support server data.

Virtualized Infrastructure

More and more organizations are deploying Microsoft Exchange Servers on virtualized infrastructure. Tegile Zebi arrays support VMware vSphere, Citrix XenServer and Microsoft Hyper-V; all are certified by the respective hypervisor providers.

Zebi arrays also provide:

- Hypervisor-aware backup
- Accelerated VM provisioning and installation of Microsoft Exchange Server roles through integration with the VMware vStorage API of Data Protection, VAAI and Microsoft VSS.

To best use Zebi in conjunction with virtualized Microsoft Exchange Server deployments, follow the general Zebi best practices for vSphere, XenServer or Hyper-V, plus the additional Microsoft Exchange specific best practices in this document.

Optimizing Storage Configuration for Exchange

In this section, detailed configuration best practices are provided to achieve optimal performance, reliability and resiliency when utilizing Zebi storage arrays in Microsoft Exchange Server deployments.

These best practices apply to Microsoft Windows Server 2008 and Microsoft Exchange Server 2010 in iSCSI and FC environments.

For general Zebi array best practices and a configuration guide, please refer to http://www.tegile.com/support.

iSCSI

Zebi works well with the Microsoft iSCSI initiator on a host Microsoft Windows Server.

In a virtualized environment, it's advisable to use Microsoft's iSCSI Initiator in Windows VMs. In addition, a dedicated storage network that accesses Zebi through iSCSI is recommended between Zebi arrays and the Microsoft Exchange servers.

Jumbo Frames

When using iSCSI as the storage interconnect, enable end-to-end Jumbo Frames between Microsoft Exchange servers and switches connected to Zebi storage arrays.

On the Microsoft Exchange servers, enable jumbo frame by setting the maximum transmission unit (MTU) to 9,000 on the network interface card (NIC) of the Microsoft Exchange Servers. In a virtualized environment, jumbo frame should be enabled on the network interfaces both at the host OS level, as well as at the guest OS level.

Please note: on VMware vCenter 5.0, when using VM cloning to replicate Windows Server VMs, the Jumbo Frames configuration on the guest OS may not be cloned during the process. The configuration must be configured and verified on each Windows VM after the clone operation is done.

On certain gigabit or 10GbE switches, Jumbo Frames are disabled by default. In that case, an administrator must manually enable it on the ports that are connected to Zebi storage, as well as on those ports that are related to the Microsoft Exchange Servers.

On Zebi, the jumbo frames feature is enabled on the iSCSI ports by setting the MTU to 9,000.

Fibre Channel

The queue depth or execution throttle on the FC HBAs should be configured as 128, 255 or to the maximum the HBAs can support.

If using QLogic QLE2562 or equivalent FC 8G HBA, the default execution throttle is already 255. If using QLogic ALE2462 or equivalent FC 4G HBA, the default execution throttle is only 16. It must be manually set to 255 in order to achieve optimal Microsoft Exchange Server performance. On a Windows host OS, this can be configured using QLogic SANSurfer software.

If using Emulex FC HBA, the default queue depth is set to 32. On Windows host OS, use Emulex HBA utility LPUTILNT to manually configure the driver parameter "QueueDepth" to maximum 254.

In a virtualized environment, the FC HBA queue depth or execution throttle should be configured properly at host OS level. Please refer to the host OS administration guide to configure the HBA parameters properly.

Database and Log LUN Configuration

A two LUNs per database architecture is recommended:

- One dedicated database LUN
- One dedicated log LUN are necessary to achieve optimized performance according to Microsoft Exchange IO characteristics.

Zebi arrays support an unlimited number of LUNs so there is no need to be concerned about using a two LUNs per database architecture. The boxes to the right provide detailed recommendations for configuring storage for Microsoft Exchange volumes on Tegile Zebi arrays.

EXCHANGE AND ZEBI RECOMMENDED BEST PRACTICES

Consider larger and fewer LUNs, while keeping the database itself within 2TB. The database LUN can be larger than 2TB.

Configure the database LUN and the related log LUN inside the same Zebi Pool and the same Zebi Project for transactionally consistent backup. Multiple database LUNs and related log LUNs, of the same backup schedule, can be grouped together in the same Zebi Pool and Project.

Choose a database application type.

For log LUNs, use an 8K block size and enable compression and disable de-duplication.

For database and log LUNs, select the basic disk type and GPT partition table type when initializing the disks on Microsoft Windows 2008 Servers.

If using Zebi to host the Windows VM for Microsoft Exchange Server roles, opt for the virtualization application type when configuring the LUN for the VM.

Group the mailboxes with similar service level requirements into the same database.

Do not place multiple databases into the same LUN, and do not locate the logs of multiple databases into the same LUN. This enables flexible backup and restore, as well as optimized performance.

Use the mirrored profile when creating pools for both database LUNs and transactional log LUNs.

For database LUNs, use a 32K block size and enable compression and de-duplication.

For database and log LUNs, select the 64KB allocation size when formatting NTFS partition on the disks on Microsoft Windows 2008 Servers.

Balance the database and log LUNs access between the Tegile array controllers, Ethernet or FC paths, switches and Microsoft Exchange Servers.

Multipath

Zebi is compatible with the built-in multipath framework on Microsoft Windows, VMware ESX and Citrix XenServer.

Asymmetric logical unit access (ALUA) is supported on the FC interconnect on Zebi arrays. If using the FC, ALUA should be configured in the multipath configuration on the host side. With ALUA, both storage processors in an array are able to see all volumes that exist on the array even though only a single processor may have "ownership" or an individual volume. As such, although there are multiple pathways to the storage, one is optimized and one is not. ALUA can alleviate issues of path thrashing, which can result in poor performance.

Snapshot and Replication

Install and configure the Tegile VSS agent properly on each of the Microsoft Exchange Servers to provide transactionally consistent backup and restore.

On Zebi, configure snapshot at the project level to schedule the snapshot backup at the desired time and frequency. The snapshot can be scheduled to take place as frequently as every 15 minutes, providing organizations with excellent RPO capability.

If using Zebi replication for resiliency or remote backup, configure replication at the project level to schedule snapshot-based replication at the desired frequency.

Conclusion

Tegile Zebi hybrid storage offers the most balanced storage solution for enterprise applications like Microsoft Exchange Server. Powered by Tegile MASS technology, with in-line compression and de-duplication, Zebi delivers seven times the performance, 14 times more mailboxes and a capacity savings of 75 percent. Microsoft Exchange Server deployments require proper planning and verification. Following best practices recommended by both Microsoft and Tegile will help accelerate this process to achieve optimal performance, resiliency and reliability of Microsoft Exchange deployments that employ Tegile Zebi arrays.

About Tegile Zebi Storage Arrays

Tegile Zebi arrays leverage the performance of solid state disks (SSD) and the low cost per TB of high-capacity disk drives to deliver up to seven times the performance of legacy arrays while, at the same time using up to 75 percent less storage thanks to Tegile's powerful data reduction technologies. Tegile's arrays don't simply use SSDs as a tier of storage; instead, Tegile has infused the entire data path with the performance benefits of SSDs, giving every application a performance boost.

Tegile Zebi arrays feature:

- 7 times the performance of competing arrays.
- Up to 75 percent capacity savings through data reduction technologies.
- NAS and SAN connectivity from the same array.
- Built-in business continuity.
- Simplified storage management.
- Purpose-built for virtualization with VM-aware management tools.

To learn more about Tegile Zebi storage solutions, please visit or call us at:

http://www.tegile.com/

Toll Free 1-855-583-4453

Follow us on Twitter: http://www.twitter.com/tegile