

# White Paper

---

## **Storage Efficiency Strategies for Microsoft Application Environments**

*By Brian Babineau*

December, 2009

---

## Contents

Executive Summary .....	3
Current Microsoft Application Infrastructure Challenges .....	3
Resource Chaos .....	3
Short-sighted Problem Solving .....	3
Native Management Tools do not Scale.....	4
Copies Drive Data Growth .....	4
NetApp Delivers Storage Efficiency Options .....	5
Moving to a Simple Shared Resource Model .....	5
Performance and Protection Without Tradeoffs.....	5
Control Resource Provisioning .....	5
Eliminate Redundant Data.....	6
Frugal, but Useful, Copies.....	6
The End Result .....	7
Getting Started .....	7
Change for the Better .....	7
Hyper-V is a Catalyst.....	8
The Bigger Truth .....	9

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188. This ESG White Paper was developed with the assistance and funding of NetApp.

## Executive Summary

A 2008 ESG survey of nearly 400 enterprises revealed that, on average, organizations spend 28% of their IT budgets on storage inclusive of systems, software, and services.<sup>1</sup> In another ESG study conducted in March 2009, nearly 74% of respondents stated that they were planning on maintaining or increasing storage spending in 2009 compared to 2008.<sup>2</sup> Even with most IT departments operating in “cost reduction” or “cost containment” mode, constant data growth and inefficient operations make it very hard to curb storage-related expenses.

It is unlikely that businesses will stop generating digital data, leaving IT to focus on making storage infrastructures more efficient in an attempt at gaining control of one of the largest budget line items. Rather than guessing about the sources of storage-related issues, ESG suggests that IT leaders start with Microsoft application environments. This paper addresses the many Exchange, SQL Server, and SharePoint implementations running on inefficient storage infrastructures, limiting scalability and complicating what should be fairly simple storage management tasks.

Readers will also learn why recent changes in Microsoft infrastructures, including the introduction of a server operating system and a server virtualization solution (Hyper-V), create a compelling opportunity to improve underlying storage infrastructures right now. NetApp, a leading provider of storage solutions, can help accelerate Microsoft-related storage efficiency initiatives with its portfolio of storage system, storage management, and data protection offerings.

## Current Microsoft Application Infrastructure Challenges

### Resource Chaos

Microsoft’s importance in the data center cannot be questioned—everyone knows how vital Exchange is to workforce communication. Companies are also finding new ways to leverage SQL Server in building database applications and 52% of global enterprises indicated that SharePoint will be a top five IT initiative over the next 24 months.<sup>3</sup> Increased reliance on Microsoft applications is a significant contributor to the growth of Microsoft data center infrastructure, specifically x86 servers running Windows. Adding the expansion of Windows-based servers used for file sharing, small web applications, and other general purposes, most IT departments are currently dealing with Microsoft server sprawl.

Microsoft applications have also contributed to growth in server infrastructure. For example, Exchange 2007’s architecture leverages “server roles” where certain functions are segregated in the application. This separation of duties enables IT to scale functions such as Mailbox and Transport capabilities separately from other application components. It also leads to IT deploying more servers to support each individual function. SharePoint environments will most likely follow a similar trend as the solution is designed with several modules that can scale independently.

### Short-sighted Problem Solving

With application performance in mind, Microsoft developers frequently advise administrators to leverage server attached storage. While some end-users will see a boost in application response times, IT administrators find themselves managing dozens of independent, isolated compute and storage infrastructures. When an Exchange Mailbox server needs more capacity, IT cannot use the free space from the Transport server environments. The net result is poor storage utilization in almost every application environment and unnecessary spending on server resources just to add storage capacity.

---

<sup>1</sup> Source: ESG Research Report, *ESG 2008 Enterprise Storage Systems Survey*, November 2008.

<sup>2</sup> Source: ESG Research Report, *2009 Data Center Spending Intentions Survey*, March 2009.

<sup>3</sup> Source: ESG Research Report, *Microsoft SharePoint Adoption, Market Drivers & IT Impact*, March 2009.

Independent or “siloed” Microsoft server and storage infrastructures also complicate backup and disaster recovery operations. IT has to execute these processes on each individual server and monitor all jobs to ensure completion. These jobs need to be checked for application consistency to ensure that the backup can be used in the event of data loss or corruption. IT also has to make sure that all devices that need to be part of a disaster recovery schema are included to ensure efficient recovery—a device missing from a replication or clustering configuration can increase unplanned downtime. In some instances, organizations omit Microsoft environments from disaster recovery plans simply because it is too hard to manage all of the individual devices.

### **Native Management Tools do not Scale**

It is very easy for organizations to rely on technology such as Volume Shadow Services or native replication capabilities to address server and storage management challenges. These “free” capabilities are great for simple environments with a few servers and minimal storage capacity. For larger infrastructures, these tools create more problems than they solve because administrators have to configure them on a “per device” basis. What’s more, capabilities can vary version to version. As an example, Exchange 2010 has native database replication capabilities whereas Exchange 2007 supports local and log file replication (LCR and CCR) for data protection. Further, not all applications offer the same types of tools, making it difficult for Windows teams to implement scalable Standard Operating Procedures for storage management and data protection operations.

Not only do Microsoft application tools vary by application and by version, but the way each solution deals with storage specifically is drastically different and subject to change. Take Microsoft Exchange 2010: this latest version of the company’s flagship communication application no longer supports single instancing, but Exchange 2007 does. SharePoint saves every version of a file in its content repository, driving up storage requirements. All of this unpredictability makes it extremely difficult to rely on Microsoft-specific tools to optimize storage environments.

### **Copies Drive Data Growth**

Primary Microsoft application environments continue to grow. ESG predicts that SharePoint site capacity will increase at 25% per year and, by supporting unified communications capabilities, Exchange storage needs are likely to expand as well.<sup>4</sup> If primary data growth wasn’t enough, IT must also find a place to store all copies made to safeguard against the risk of data loss and corruption. In some cases, these copies of Exchange, SQL Server, and SharePoint are a combination of snapshots and full volume replicas used for backup; in other instances, IT replicates data to a secondary site to prevent against data loss and application downtime in the event of a disaster. Multiple copies of the same data set are frequently made in the same day, depending on recovery point and recovery time objectives. And IT may save any number of the copies for weeks or months, further increasing storage needs.

Not all copies of primary data sets are for data protection purposes. A 2007 ESG research study uncovered that a majority of organizations created between two and five copies of databases, including SQL Server, per month for both development and test environments and over two-thirds of these copies were full replicas as opposed to subsets.<sup>5</sup> Depending on the size of the primary SQL Server databases, these secondary environments can take a while to create, slowing engineering and quality assurance efforts. And even though IT may delete the copies after development and test efforts are complete, IT still needs excess “standby” storage capacity in case a new one needs to be created.

---

<sup>4</sup> Source: ESG Research Report, *Microsoft SharePoint Adoption, Market Drivers & IT Impact*, March 2009.

<sup>5</sup> Source: ESG Research Report, *Database Archiving Survey*, November 2007.

## NetApp Delivers Storage Efficiency Options

### Moving to a Simple Shared Resource Model

When looking to improve efficiency, two fundamental steps can lead to the right path. The first is moving to a shared storage resource model where capacity is centralized and can be accessed by any Microsoft application. A networked storage model where Microsoft applications are connected to centralized storage resources via SMB (Server Message Block – also known as [SMB, also known as Common Internet File System or “CIFS”]), iSCSI, or Fibre Channel protocols, makes it easier for IT to provision capacity and increase utilization (Note: Exchange, SQL Server, and SharePoint can connect to a networked storage system via iSCSI or Fibre Channel; SQL Server and SharePoint can also connect to a networked storage system via SMB.). IT also doesn’t have to buy new servers if they run out of capacity—more capacity can be added or, if the space already exists, provisioned to the networked storage system.

The second choice is simplifying the networked storage deployment, which is where NetApp’s value begins to manifest. All of NetApp’s storage systems run the same operating system (Data ONTAP) and file system (WAFL) as well as any of NetApp’s data protection and storage management software. And all of NetApp’s storage product lines support traditional file sharing and block-based protocols. Translated, when a customer chooses to deploy any NetApp system to support their Microsoft application environment, the devices will “look and feel” the same. Customers do not have to learn how to configure and provision a new device or figure out how the software will work. NetApp’s unified architecture gives customers choice without complexity and simplifies operations.

### Performance and Protection Without Tradeoffs

When moving to a networked storage infrastructure in a Microsoft application environment, two initial choices have to be made: the types of drives to be used (Fibre Channel, SATA, etc.) and the RAID configuration best suited for the environment. While certainly tactical in nature, these decisions impact the cost, performance, and availability of a storage system. SATA drives are less expensive than Fibre Channel, but are also typically slower. A RAID 1 configuration (mirroring) provides the best performance and protection, but drives up cost as customers must buy double capacity. Alternatively, RAID 5 implementations, while consuming less capacity, do not protect against a double-disk failure.

NetApp eliminates the “tradeoffs” when it comes to making these two very important decisions. First, via its Data ONTAP operating system and WAFL File System, NetApp has optimized data storage on any disk drive. This enables customers to get response times from SATA drives that would normally be seen only with Fibre Channel. NetApp also leverages its own RAID algorithm, RAID-DP, to prevent against double-disk failures. RAID-DP is more closely aligned with RAID5 or RAID6 parity schemes than a RAID 1 or RAID 10 implementation, reducing the amount of extra capacity a customer needs to buy in order to adequately protect their data from disk failures. In all, customers save by leveraging less expensive SATA drives and buying fewer drives altogether because of the minimal RAID penalty—without giving up any performance or protection benefits.

### Control Resource Provisioning

Moving from a server attached storage model to a centralized NetApp system should immediately increase utilization. It does not, however, address the issues and inefficiencies created by over-allocating capacity to an application for performance or to prepare for unknown growth. Many Microsoft application owners cannot predict their future storage requirements, so they request overhead space for fear that resources will run out, slow, or disrupt application access and IT will not be able to provision new capacity fast enough. This occurs frequently in Exchange environments that do not employ mailbox quotas and in SharePoint environments when a site moves from a pilot implementation to production.

In addition to RAID-DP, NetApp also offers other means to solve performance issues without adding more capacity to a device. Customers can choose to use Fibre Channel drives; leverage FlexCache, a solution that enables

customers to dedicate memory resources to a particular server or application for more effective quality of service; or add additional Flash memory with NetApp's Performance Accelerator Module.

Performance is critical in Microsoft application environments, but bigger efficiencies can be gained by cutting allocated yet unused storage capacity. NetApp's thin provisioning capability, referred to as FlexVol, enables IT to allocate capacity to an application without actually committing the resources. As an example, a SQL server application owner may want 1 TB of data for a 200 GB database, with the extra space being used as headroom. With NetApp, IT can allocate the 1 TB, but within the storage system, only the data that has been actually written is committed to the application. If the SQL server application has written 200 GB of data, the balance of the capacity—800 GB—remains in the storage pool for use by another application. In essence, the SQL Server administrator sees 1 TB, but the storage system does not commit all of this capacity until the application actually requires it. NetApp automatically recognizes when it needs to free additional space for an application and does so when the capacity is needed. Minimizing over-allocation significantly drives up utilization as unused capacity remains pooled for use by all Microsoft applications.

FlexVol is complemented by NetApp's ease of provisioning features. Allocating volumes (or LUNs) can happen in minutes, which enables uncommitted capacity to be easily moved to an application. Even if a customer chooses not to use NetApp's thin provisioning, they do not have to worry about downtime as a result of prolonged or disruptive storage provisioning processes. This makes it possible for IT to use "just in time" provisioning, leading to additional capacity savings without impacting Microsoft application environments.

### **Eliminate Redundant Data**

When a customer consolidates Microsoft applications onto a NetApp storage system, they can employ data deduplication to reduce storage capacity requirements. As an example, a customer may run multiple SharePoint farms and connect them all to a central NetApp system to better manage utilization and performance. If deduplication is enabled, NetApp can scan the volumes allocated to each SharePoint server looking for the same "bytes" of data. The scan produces an index of all the bytes stored within each volume—when a duplicate byte is identified, it is removed and replaced with a pointer to the index so that access remains intact.

NetApp customers can schedule when the deduplication process occurs, with most choosing when application activity is the lightest. In addition, deduplication can be enabled at the volume (or LUN) level. This flexibility is extremely useful when two different applications or parts of an application store data on the same NetApp system and one them isn't likely to save the same data multiple times. As an example, an Exchange Mailbox server is likely to save the same attachment multiple times, but the Client Access server only contains configuration data. When running these different Exchange server roles on the same NetApp system, the volumes allocated to the Mailbox server can execute deduplication while the capacity committed to the Client access server does not.

With early deduplication adopters reducing storage requirements tenfold,<sup>6</sup> this is often the easiest way to achieve immediate storage efficiencies. Customers can defer capital purchases by fitting more capacity into a single system and, more importantly, downstream copies of data (made for data protection, test and development, etc.) will be smaller.

### **Frugal, but Useful, Copies**

For local backup purposes, NetApp customers can create Snapshot copies which can be mounted by any backup application server. Using storage-based snapshots is not new for many organizations, but many are under the impression that they need to leave "scratch" or "copy" space to save the snapshots. The size of this space depends on how big the data sets are and how many actual copies are taken. In contrast, NetApp's Snapshot capabilities do not require any additional capacity to be available when they are created. Customers only need to ensure there is enough room to accept new data written to the primary file system (being copied).

---

<sup>6</sup> Source: ESG Research Report, *Data Protection Market Trends*, January 2008.

Without the need for “scratch space,” customers may create several snapshots a day within a system, reducing the risk of data loss and the amount of time needed to complete a recovery. For Microsoft environments, customers can use NetApp’s SnapManager to schedule and manage—a task that includes data restoration if needed—all of the snapshots. The application-specific modules of SnapManager, including solutions for Exchange, SharePoint, and SQL Server, integrate with Microsoft Volume Shadow Services, enabling a consistent point-in-time snapshot of the environment. All SnapManager products validate the backup before the process is complete, ensuring that any given Snapshot can be easily recovered and the application environment restarted if there is a corruption.

Companies can also reduce the size of copies made for disaster recovery purposes with a combination of three NetApp solutions. First, SnapVault and SnapMirror are used to move data between NetApp systems. Rather than sending a full copy of data to a target device every time there is a modification made to the source, Thin Replication lets customers move only the changes. If there is a problem with the source system that requires a partial restore, customers can replicate only the data needed from the target rather than copying all of the information back. Moving less data over a network cuts bandwidth expenses without compromising business continuity plans. And partial data recoveries facilitate faster restore times.

When a company wants to clone a data set for development and testing purposes, they can use FlexClone to create more space efficient copies. Instead of making a full replica of a data set, FlexClone creates the appearance that a clone exists and enables “changes” to be written to this “copy.” The only storage capacity used for the FlexClone is for any changes that are written such as those operations conducted during development and testing purposes. Even though the changes are made, there is no impact or risk to the primary data set as the changes are made. The process of creating usable clones is much faster because data is not actually copied, helping customers create test and development environments more quickly. Ideal for SQL Server implementations, FlexClone enables IT to create as many test and development environments as necessary with no storage penalty for doing so.

## The End Result

Organizations may choose to leverage some or all of NetApp’s solutions to drive storage efficiencies in their Microsoft application environments. The hardest part may be measuring the tangible benefits of these efficiencies. ESG suggests thinking about what the environment would look like if NetApp underpinned Exchange, SQL Server, and SharePoint applications. During such an exercise, ESG suspects companies will realize:

- Storing less data on fewer devices creates several opportunities for operational savings such as lower energy consumption, centralized management of critical storage tasks, and minimal IT burden as all NetApp devices work the same way.
- Increasing utilization cuts capital expenditures as customers can defer storage, and in some cases server, purchases.
- Keeping all Microsoft applications highly available and protected is much simpler when these tasks are centralized and more functional tools are used to execute the operations.
- IT can improve agility and scale as data growth increases by swiftly provisioning storage resources, scheduling and executing data protection operations, and removing redundant data without disrupting application access.

## Getting Started

### Change for the Better

ESG believes storage efficiency projects can have an immediate impact on Microsoft environments due to the changes occurring in Exchange, SQL Server, and SharePoint infrastructures right now. The biggest change is the new release of the Windows Server operating system, which includes an updated version of Hyper-V. Additionally, many companies will be evaluating an upgrade to Exchange 2010 and soon a new version of SharePoint will be available. These changes create a great opportunity to re-architect and simplify the underlying storage

environments. For some, this may mean centralizing storage data protection processes rather than relying on individual application tools or homegrown scripts to back up and replicate data. For others, this may simply mean moving away from server attached storage because Exchange and SharePoint data growth is exceeding expectations.

## **Hyper-V is a Catalyst**

Server virtualization and its benefits are not new to IT departments. In fact, in a February 2009 ESG research study, 71% of companies were found to be using server virtualization in production or test environments. A more telling statistic is that, of those that haven't deployed server virtualization, nearly half are evaluating Microsoft's Hyper-V offering. ESG believes that many of these Hyper-V implementations will kick off Microsoft application infrastructure consolidation projects, paving the way for complementary storage efficiency initiatives.

One of the biggest benefits of using Hyper-V is to improve CPU utilization while maintaining flexibility to move virtual machines from one physical server to another when necessary. In most instances, driving up CPU utilization reduces the number of actual physical servers Windows teams have to manage. If organizations are using server attached storage, a reduction in servers also means a reduction in available storage capacity. Suddenly, a SharePoint farm running on ten physical servers, each with 1 terabyte of storage, may be consolidated to two devices—each with a terabyte of storage—supporting five virtual machines. If the SharePoint application's data needs grow beyond the two terabytes, a customer using a server attached storage strategy has to add another server to get incremental capacity—a move that reverses the benefits of server virtualization. As a result, moving from server attached storage to NetApp networked storage is a logical outcome of Hyper-V projects.

Hyper-V's Live Migration feature represents another reason to use NetApp networked storage with Hyper-V. Live Migration enables virtual machines to move from one physical server to another, but it does not work unless both physical servers (the source and destination for the virtual machine) are connected to a networked storage device. The availability (if a physical server crashes, Live Migration can move the information to any other Windows server running Hyper-V) and performance management (if one virtual machine needs more computer resources, it can be moved to a server that has extra CPU available) benefits of Live Migration—just two of the use cases that streamline Windows application infrastructure operations—make moving to networked storage an even easier decision.

NetApp further eases the move to a Microsoft virtualized server environment with its SnapManager for Hyper-V offering. Similar to NetApp's other Microsoft-integrated data protection offerings, SnapManager for Hyper-V ensures that IT can create a consistent copy of the virtual machine. It also reduces the risk of Hyper-V deployments by discovering and tracking all virtual machines to see which are and are not protected, mitigating the risk that data is left unprotected if Hyper-V environments grow rapidly. Customers can also configure data protection policies with SnapManager and apply them to a single or group of virtual machines. As new virtual machines are added, administrators simply apply the appropriate policy via SnapManager. Additionally, if administrators need to change backup scheduling or retention policies, they can modify the policy within SnapManager as opposed to altering it on each virtual machine. If something happens to a virtual machine, customers can execute the recovery directly from SnapManager. These restoration operations are very straightforward as SnapManager tests a snapshot after it is executed to make sure it is recoverable.

In addition to protecting Hyper-V environments, NetApp also removes obstacles that could impact the use of Live Migration. If a customer wants to move a Virtual Machine from one physical server to another but both physical servers are attached to different NetApp systems, NetApp, with its Data Motion solution, can move the data associated with the Virtual Machine to the other storage system. As a result, customers do not have to overburden one storage system by connecting all of its physical servers to it just for Live Migration purposes.

## The Bigger Truth

Microsoft application owners may simply view NetApp's offerings as functional storage systems that can deliver some back end efficiencies. This perspective may lead to missed opportunities to generate more efficiency out of their NetApp investments. As an example, a customer may be currently using CCR and LCR features in Exchange 2007 to execute local and remote replicas. In Exchange 2010, these capabilities do not exist, but administrators can make up to 16 replicas of Exchange databases. Rather than try to fit new technology into existing processes or vice versa, companies can use NetApp to back up and replicate Exchange. NetApp's copies take up less space than the Exchange replicas and, with Snap Manager for Exchange, all operations can be easily scheduled and managed.

Similar opportunities exist for creating SQL Server test and development environments. Administrators may be used to home grown scripts, but the capacity penalty of creating a full replica every time someone needs a new environment gets very expensive. NetApp centralizes these process, instantiates the environments quickly, and does so without an exorbitant amount of storage resources.

Hyper-V implementations can go much smoother when companies simply add virtual machines into their NetApp-managed data protection schemas. In this case, customers get the benefit of consolidating servers and centralizing data protection. Disaster recovery options for all Microsoft applications become more feasible and affordable with NetApp.

The list could go on and on, but the point is, when speaking from a Microsoft application standpoint, NetApp does more than just store the data—it stores, manages, and protects it. By executing these tasks effectively and, most importantly, efficiently, customers can actually cut out a variety of costs as they begin to scale and upgrade their Microsoft application environments.



Enterprise Strategy Group | **Getting to the bigger truth.**