Windows Server 2012 R2 Storage

Technical Scenarios and Solutions

Microsoft

Table of contents

Copyright information

Rethinking Storage with Windows Server 2012 R2

Overview

According to recent studies, more than 40 percent of information technology (IT) staffs in the world today are encountering significant increases in data volume.[1] The exponential growth of available data has created significant challenges for IT. Traditional methods such as database storage no longer prove as effective for query and analysis. Multiple data sources make finding and extracting useful insights even more difficult. Effectively maintaining compliance has become a logistical nightmare.

With Windows Server 2012 R2 however, you can take advantage of a proven, enterprise-class datacenter and cloud platform that can scale to run the largest workloads. Delivered as a dynamic, available, and cost-effective cloud solution, Windows Server 2012 R2 provides automated protection and cost-effective business continuity that enables you to keep your business up and running while reducing downtime risks.

---

[1] Source: Agile BI, Complementing Traditional BI to Address the Shrinking Business-Decision Window, November 2011, Aberdeen Group, Inc.

The following diagram provides a quick view of some of the new and enhanced storage features of Windows Server 2012 R2:

Layer	Feature	Description
Storage	Storage Spaces (enhanced in Windows Server 2012 R2) **	Delivers deduplication, storage tiering, and scalability with just a bunch of disks (JBOD) as well as storage virtualization and resiliency. **Enhanced in Windows Server 2012 R2 to provide Policy-based automatic data tiering to maximize cost and Input/Output Operations Per Second (IOPS).*
Storage	Offloaded Data Transfer	Moves large chunks of data or entire virtual machines between storage devices with speeds comparable to local storage — without taxing valuable server resources.
Storage	Storage Quality of Service (new in Windows Server 2012 R2)	Throttles storage assigned to VHD/VHDX to prevent a single virtual machine from consuming all I/O bandwidth (including incoming and outgoing IOPS). You can use Microsoft System Center Operation Manager for monitors and alerts.
File System	Resilient File System ReFS	Helps protect business-critical data from errors and keeps data available when needed. The architecture provides scalability and performance for constantly growing dataset sizes and dynamic workloads.
File System	Data Duplication (enhanced in Windows Server 2012 R2) **	Uses sophisticated data deduplication functionality to extend branch office storage capabilities by reducing storage size on a disk. Data duplication can be used on any server, by itself, or with Microsoft BranchCache to extend branch office storage capabilities. **Enhanced in Windows Server 2012 R2 to provide better performance and reliability, plus support for live virtual machines in a VDI environment*
File System	Storage Spaces Thin Provisioning and Trim	Offers new native support for thin provisioning and trim that provisions storage as needed and reclaims storage when no longer needed.
File and Block Access	SMB Scale-Out for Applications (enhanced in Windows Server 2012 R2) **	Enables the deployment of Microsoft SQL Server databases and Hyper-V storage on flexible, high-performance Server Message Block (SMB) file shares, improving the simplicity of deployment and management. ** *Enhanced in Windows Server 2012 R2 to offer increased performance for small IO's, improved logging and diagnosing capabilities, and more efficient load balancing across difference nodes in a file server cluster.*

| Networking	SMB Direct (enhanced in Windows Server 2012 R2) **	Provides high-performance storage capabilities that won't burden server central processing unit (CPU). SMB Direct utilizes Remote Direct Memory Access (RDMA) hardware support for SMB to deliver enterprise-class performance without relying on expensive Fibre Channel storage area networks (SANs). *** Enhanced in Windows Server 2012 R2 to deliver more than 30 percent performance improvement with small IO's.*
Virtualization	Live Storage Migration	Maintains maximum service-level availability even while moving virtual machine storage between host servers (independent of the actual virtual machine).
Virtualization	Online VHDX Resize (enhanced in Windows Server 2012 R2) **	Grows or shrinks a VHDX virtual disk dynamically, while the virtual machine is running, with no downtime.
Management	Comprehensive Storage Management (enhanced in Windows Server 2012 R2) **	Offers easy–to-use comprehensive storage management capabilities that enable administrators to manage remote storage on a single management interface, automate storage management tasks with new Windows PowerShell command-line interface capabilities, and reach beyond Windows Server storage to manage storage arrays and SANs with Storage Management Initiative Specifications (SMI-S). *** Enhanced in Windows Server 2012 R2 to provide more System Management Application Program Interface (SM-API) support in cluster tools and much stronger integrations with System Center 2012 Virtual Machine Manager.*
Data Protection	System Center Data Protection Manager	Offers a centralized, item level backup/recovery tool for the entire enterprise. * Requires System Center 2012
Disaster Recover	Hyper-V Replica (enhanced in Windows Server 2012 R2) **	Provides a workload-agnostic storage solution that replicates efficiently, periodically, and asynchronously over Internet protocol (IP)-based networks, typically to a remote site. Hyper-V Replica also enables an administrator to easily test the replica virtual machine without disrupting the ongoing replication. (*Note: Hyper-V Replica does not rely on storage arrays or other software replication technologies*). *** Enhanced in Windows Server 2012 R2 to support tertiary replication and configurable replication schedules.*
Data Protection	Hyper-V Recovery Manager (new in Windows Server 2012 R2) **	Combines Windows Azure, System Center VMM, and Hyper-V Replica to deliver planned and cost-effective business continuity of workloads. With Windows Azure Hyper-V Recovery Manager, you can protect services by automating the replication of the virtual machines that composes them at a secondary location. You also receive continuous health monitoring of the primary site and coordination of the orderly recovery of services in the event of a site outage. * Requires System Center 2012 SP1

For more information on Windows Server 2012 R2 storage features see: www.microsoft.com/storage

In this whitepaper, we will discuss Windows Server 2012 R2 features for an enterprise-class storage platform built on Windows. This scenario defines Windows Server as a collection of capabilities that deliver extremely high levels of storage performance, resiliency, efficiency, and management. Some of these features include the following:

- **Storage Spaces with automated data tiering:** Storage Spaces were one of the groundbreaking features of Windows Server 2012. Storage Spaces with automated data tiering enables you to make better use of expensive solid-state devices (SSD) and serial-attached SCSI (SAS) devices by dynamically allocating the SSD storage devices to the busiest workloads. This feature had previously been available only in expensive, proprietary SAN solutions.
- **Data Deduplication:** Data Deduplication makes storage more efficient by minimizing redundant data on a disk. In certain workloads, such as virtual hard disk (VHD) stores, you can reduce by up to 90 percent the space used on disk.
- **Storage Quality of Service (QoS):** With Storage QoS you can throttle incoming and outgoing IOPS used by a VHD/VHDX file to prevent a single virtual machine from consuming all I/O bandwidth. Storage QoS comes with a sophisticated monitoring policy to alert you when the system fails to meet a minimum bandwidth.
- **SMB Multichannel:** With SMB Multichannel, you gain Load Balancing/Failover (LBFO) protection against path failures on SMB Shares. When combined with the enhanced application storage support with SMB 3.0 shares, you can scale out workloads such as SQL Server and Hyper-V to high-availability storage spaces running on clustered file servers.
- **Comprehensive storage management:** Extensive support for Storage Management Initiative - Specification (SMS-I) enables administrators to manage remote storage and automate storage management tasks with new Windows PowerShell command-line interface capabilities. Administrators can also reach beyond Windows Server storage to manage storage arrays and SANs with Windows Server 2012 R2 SMI-S based tools. With Windows Server 2012 R2, more SMI-S support has also been integrated into Failover Cluster Manager and System Center 2012 R2 Virtual Machine Manager to give you easy provisioning and management of both Microsoft and non-Microsoft storage solutions.
- **Disaster Recovery (DR):** Windows Server 2012 R2 offers enhanced DR tools that provide rich protection and recovery to fit the massive scale of today's multi-site, virtualized datacenters. In addition to the existing two-site replication topology that exists now, Hyper-V Replica in Windows Server 2012 R2 supports tertiary replication to a third DR site for added protection of mission-critical virtual machines. Hyper-V Recovery Manager, which is available through the Windows Azure Administration Portal, provides a Windows Azure-based, offsite console used to facilitate an orchestrated site recovery of Hyper-V virtual machines from a replica site. Additionally, System Center 2012 Data Protection Manager is the cornerstone of enterprise backup and recovery for enterprise applications, operating systems, virtual machines, and many other enterprise assets.
- **Reduced storage costs:** Offloaded Data Transfer (ODX) enables the rapid movement of large chunks of data between SAN drives by offloading the CPU and memory used during copy operations to the SAN controllers. ODX integrates into core operating system (OS) operations and tools such as Windows File Explorer, Hyper-V Manager, and System Center Virtual Machine Manager to reduce the time needed to complete large file copies between SAN disks. This capability is extremely useful with live storage migration operations. Thin Provisioning enables customers to purchase only the storage resources that they need at the time, thus improving efficiency of existing storage capacity. With Virtual Machine Fibre Channel, you gain unmediated access to Fibre Channel Protocol (FCP) fabric directly from a virtual machine.

This whitepaper also provides a side-by-side comparison, where possible, of performance, resiliency, and management expectations after implementing a storage solution built on the Windows Storage Stack, versus using a SAN and FCP fabric. At the end of this whitepaper, you'll find some resources to set up the Windows Storage Stack and test it for yourself.

Enterprise-Class Storage Built On Windows
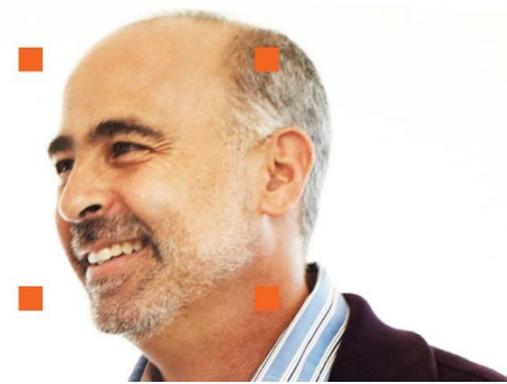
SAN-style solutions built on the Microsoft Stack

Customers today rarely use direct attached storage (DAS) for server workloads. Instead, they have made significant investments in SANs and remote storage solutions. At Microsoft, we continue to design solutions that simplify storage management while protecting the investments you have already made in traditional storage.

At the same time however, with the amount of data exponentially growing, SAN solutions have become more expensive to deploy and expand. As more and more customers look to add additional storage, the use of industry-standard hardware is becoming more common. In addition, the increased adoption of 10 Gbps and faster network interface cards (NICs) has led companies to use their existing network infrastructure to remotely access shared storage running in JBOD enclosures.

Microsoft transforms SAN solutions by integrating SAN features with the power and familiarity of Windows Server. Windows Server 2012 R2 gives you the ability to buy low-cost industry-standard hardware that you can easily scale up as the need arises.

The ability to expand your storage infrastructure using the Microsoft Stack can help you keep up with growing storage needs, while delivering required performance and reliability. For example, not every workload needs a SAN. Storage Spaces combined with Scale-Out File Servers can deliver much of the same performance and reliability as a SAN, while reducing storage system costs by up to 80 percent per terabyte (TB).

Using high-availability Storage Spaces and management capabilities built in Windows Server 2012 and System Center 2012, you can expand your existing storage infrastructure at a fraction of the $/IOPS and $/TB cost that it would take to expand a proprietary SAN without sacrificing flexibility, reliability, or performance.

Some of these features include the following:

- Block storage made available through Storage Spaces and provisioned using industry-standard disks and JBOD enclosures
- Storage exposed through Windows-based Scale-Out File Servers
- Hyper-V workloads and SQL Server databases that access storage through existing networking infrastructure over the enhanced SMB 3 protocol

Figure 1: A diagram of the Microsoft storage stack
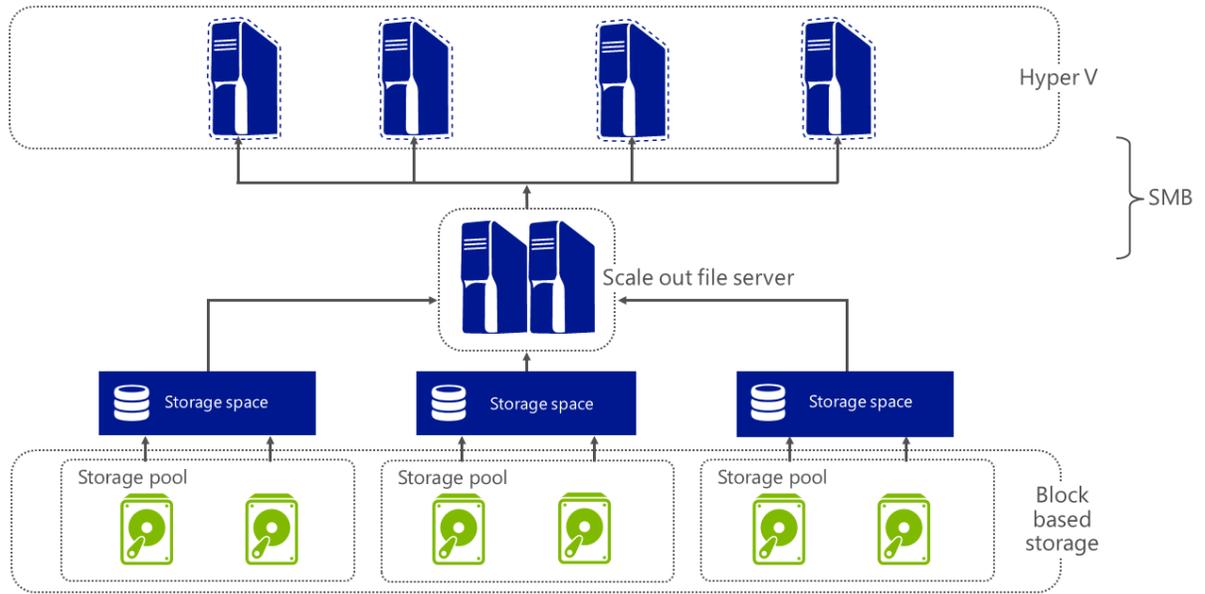
Price Point and Performance Comparison

To illustrate the cost and performance differences between a traditional SAN storage solution and one built on the Microsoft stack, the [Enterprise Strategy Group](#) (ESG) tested response times of virtual machines using different storage topologies (Fibre Channel-SAN, Internet Small Computer System Interface (iSCSI)-SAN, Storage Spaces over SMB with RDMA NICs). The results (see Figure 2) measure SQL Server response times for transactions in milliseconds when running 2, 4, 6, and 8 instances of SQL Server with each type of remote storage.

SQL Server Response Time Comparison
(OLTP Workload, Windows Server 2012, SQL Server 2012)
(Less is better)

(Horizontal bar chart. Y-axis: Number of Hyper-V virtual machines — 2, 4, 6, 8. X-axis: Average Transaction Response (s), from 0.00 to 0.10. Series: Storage Spaces over SMB, FC SAN, iSCSI SAN, Storage Spaces over SMB with RDMA.)
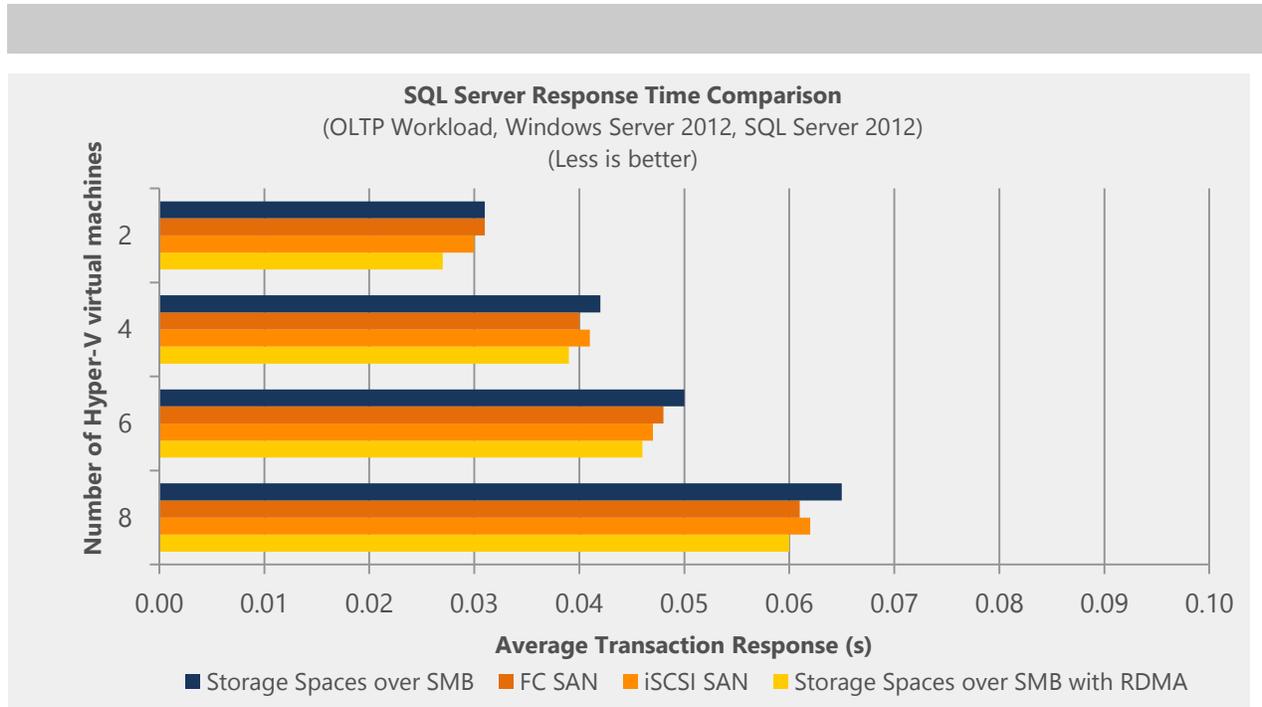
Figure 2: Response Time Comparison

The time comparison between the different instances illustrates several interesting conclusions:

- Performance scales linearly as the number of virtual machines increases from two to eight
- The performance differences between the storage configuration test scenarios are negligable
- Storage Spaces over SMB with RDMA is slightly higher (1-4 percent) than iSCSI/FC SAN test scenarios.
- Storage Spaces over SMB performance is nearly identical with iSCSI/FC SAN (averaging 3 percent slower)

Despite the negligible results, the acquisition cost of the Microsoft file-server cluster with Storage Spaces and RDMA support was 50 percent lower than the two SAN solutions. The major reason for the price difference was due to Microsoft being able to utilize inexpensive, industry-standard, commodity hardware and eliminate the cost of proprietary hardware and software from SAN solutions.

To better illustrate these results on a financial scale, ESG Lab converted the raw capacity of TBs to gigabytes (GBs) and divided the total cost of acquisition by the GB, resulting in the commonly-used metric of $/GB (see Figure 3). Similarly, to the overall cost of acquisition, iSCSI- and Fibre Channel SANs proved to be almost twice the cost of the Microsoft file-based storage offering with Storage Spaces and RDMA.

$/GB Cost of Acquisition Comparison

$/GB Cost of Acquisition Analysis
(14.4TB of raw capacity from 24 10K 600GB SAS drives)
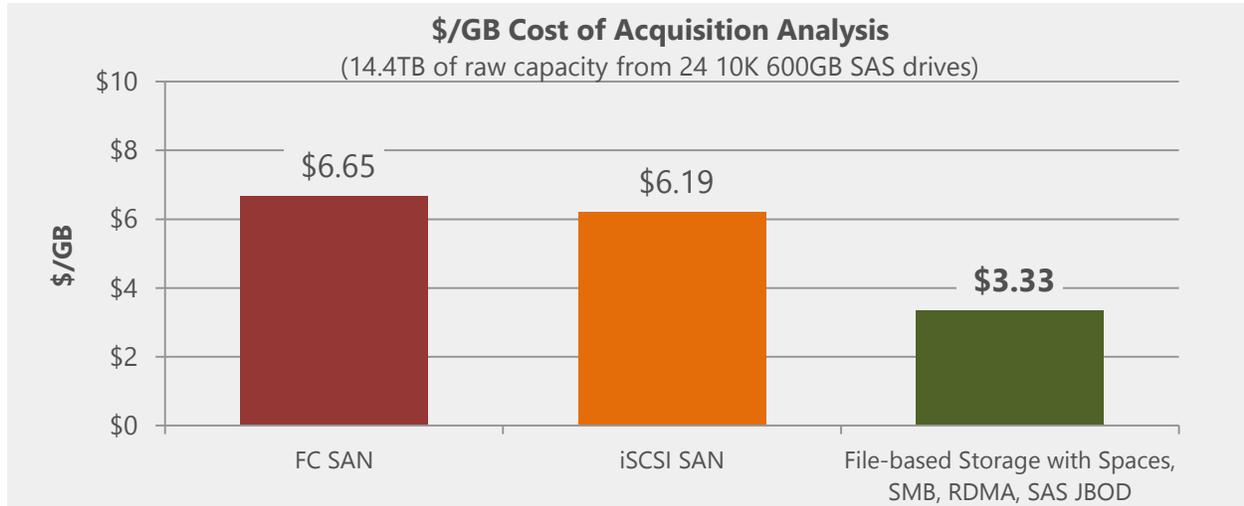
Figure 3: A comparison of cost acquisition between model scenarios

ESG Lab did not factor in the cost of management and maintenance. Because many organizations have Microsoft and Windows experts, as well as storage experts, they can manage the Windows storage without the need for additional training. Many SAN IT vendors require vendor-specific storage specialists to provision, manage, and monitor the storage infrastructure.

Storage Spaces

Challenges

Customers want great performance, reliability, and simplified management at a lower cost.

Solution

Storage Spaces in Windows Server 2012 R2 give you the ability to consolidate all of your Serial Attached SCSIS (SAS) and Serial Advanced Technology Attachment (SATA) connected disks regardless of whether they are Solid State Devices (SSDs) or traditional hard disk drives (HDDs) into Storage Pools. These Storage Pools help you keep specific data isolated, simplifying administration. You can then assign these pools to different departments within your enterprise or to customers. Once you have created these pools, you can then create logical storage devices from them, called Storage Spaces.
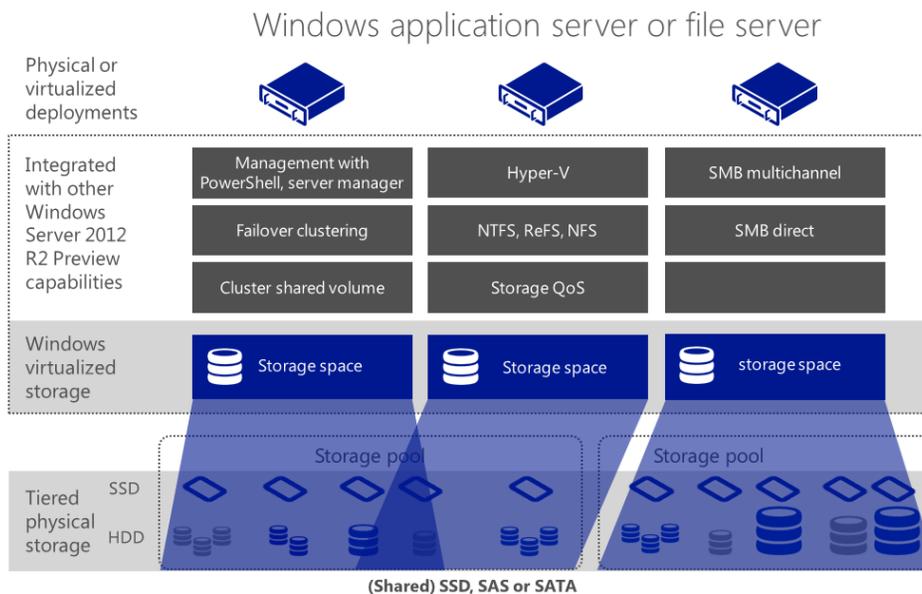

Figure 4: Storage pools in Windows Server 2012 R2

Storage Space virtual disks work the same as regular Windows disks. However, they can be configured for different resiliency schemes such as the following:

- Mirroring
- Two-way and three-way mirroring - for added protection
- Parity (RAID 5)

When you create a Storage Space, you can either choose thin or fixed provisioning. This capability lets you increase your storage investments according to need. With thin provisioning, you can create virtual disks larger than the current storage pool capacity and then add disks later to support needed growth. As a result, administrators only need to purchase the physical storage needed. They can expand dynamically when necessary.

For example, a company stores its Hyper-V virtual machines on logical disks created with Storage Spaces. With trim provisioning, when a virtual machine deletes a large file, it communicates the deletion to the host, which then communicates it to the Storage Space. As a result, the Storage Space automatically reclaims this space, making it available for use again.

Storage Spaces are compatible with other Windows Server 2012 R2 storage features, including SMB Direct and Failover Clustering, so you can use simple inexpensive storage devices to create powerful and resilient storage infrastructures on a limited budget. At the same time, you can maximize your operations by utilizing industry-standard hardware to supply high-performance and feature-rich storage to servers, clusters, and applications alike.

Storage Space Customer Benefits

- SAN-like functionality
- Several levels of resiliency
- Optimized capacity utilization with thin and trim provisioning
- SMB Direct (RDMA) and Failover Cluster support

Hybrid Storage Pools and Tiered Storage

Windows Server 2012 R2 introduces a new, policy-based, auto-tiering mechanism for Storage Spaces. Auto tiering provides more flexibility for hot and cold workloads and supports industry standard enclosures with SATA, SAS, and SSD devices.

Challenges

- Customers and datacenters need to be able to allocate and de-allocate SSD and SAS storage to the workloads that need it "now" (hot and cold data).
- Hosters and IT departments want to lower their cost/IOPS ratio.

Solution

Tiered storage assigns different categories of data to different types of storage media to reduce total storage cost. Categories can be based on levels of protection needed, performance requirements, frequency of use (hot and cold data), and other considerations. Since assigning data to particular media may be an ongoing and complex activity, Windows Server 2012 R2 Storage Spaces provide for automatically managing the process based on a defined policy.

The auto-tiering mechanism within Storage Spaces stores hot data (data that changes frequently) on the faster, but more expensive, SSD media. All data starts as hot data.

The mechanism stores cold data (data that changes infrequently) on the slower, less expensive, hard disk drives. If cold data becomes hot, it automatically moves to the

SSD media. If hot data becomes cold, it moves to the hard disk drives.

Tiered Storage Customer Benefits

- Higher utilization of expensive SSD media
- Policy-based provisioning
- Cost savings with better efficiency
- Built on Storage Spaces, which enhances flexibility

Application Storage Support for SMB 3

Challenges

- Customer needs to separate storage and compute elements of virtual machines, and move the virtual machines without moving the storage.
- Customer needs to be able to achieve all of the above using existing storage resources.

Solution

With Windows Server 2012 R2, SMB 3 shares support continuous availability. This capability opens the door for the Scale-Out File Server roles on cluster-shared volumes version 2 (CSV 2). You can easily create CSV's in Windows Server 2012 R2 clusters by converting regular cluster storage into Cluster Share Volumes, which provide continuous availability to the shares during failover.

📝 *Note*

> *For more information on CSV 2: http://technet.microsoft.com/en-us/library/jj612868.aspx*
>
> *For more information on using CSV 2/Scale-Out File Server storage for high availability virtual machines: http://channel9.msdn.com/Shows/TechNet+Radio/TechNet-Radio-Windows-Server-2012-How-to-Scale-Out-a-File-Server-using-Hyper-V*
>
> *For more information on Hyper-V end-to-end performance when using Scale-Out File Servers: http://channel9.msdn.com/Events/TechEd/NorthAmerica/2013/MDC-B335#fbid=3UQcOaX2s5n*

Cluster Shared Volumes differ from traditional cluster storage because every node in the cluster can act as a SMB entry-point for the session. Since any node can host the Scale-out File Server share session, the share (application store) is always available as long as any node is up. Through the use of Scale-out File Server shares, applications such as SQL Server 2012 and Windows Server 2012 Hyper-V can use the shares as durable data storage for databases and virtual hard drives. This functionality enables a resource to be moved from one cluster node to another without having to take storage offline.

With SQL Server 2012 (see Figure 5), when the database engine starts, it establishes the session to \\foo1\Share1 for database files (user database logs and data files) (see Figure 5). The database engine uses the share as if it was local storage. The required performance comes from the use of SMB Direct and RDMA cards.
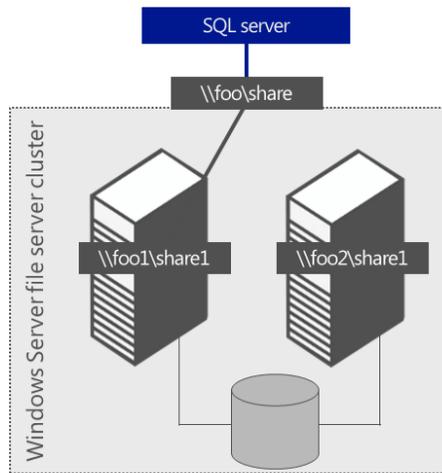
Figure 5: A SQL Server database using a scale-out share for user database files

If the active SMB session is lost because of a node hardware failure, the SQL Server SMB session gets moved to any surviving node without a service interruption. This capability keeps the databases online during a failover or planned move from one node to another. Windows Server 2012 Hyper-V supports this same type of storage for virtual machine hard disk files, enabling a virtual machine to be live migrated from one host or node to another while the virtual machines stays running — thus providing continuous availability to application clients using the virtual machine.
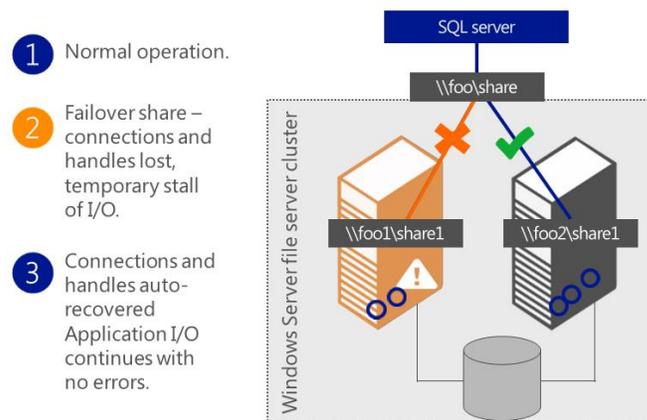
Figure 6: An active SMB session is lost

Continuous Availability Shares using SMB-Direct

Hyper-V servers can enhance Scale-Out File Server connections by taking advantage of Remote Direct Memory Access (RDMA) calls to the storage services. This feature is called SMB-Direct.

SMB-Direct uses RDMA to access the remote server storage directly, instead of going through the layers and redirectors needed to establish a non-RDMA SMB session. SMB-Direct is not a requirement for Scale-Out File Server deployments, but it is highly recommended when storing virtual machine hard drive files on Window Server 2012 Scale-Out File Server shares.

SMB Encryption and VSS Support

Windows Server 2012 R12 now supports encryption in SMB 3, enabling data transferred over wire to remain safe and secure. In order to support backups and snapshots of important data, Windows Server 2012 R2 also now supports Volume Shadow Copy Service (VSS) on remote SMB 3 stores. This functionality enables any third party back-up software that plugs into VSS to back-up your SMB file shares (including documents, virtual machines, and databases) using VSS.

Since SMB 3 shares can use your existing network infrastructure – there is no need for a dedicated storage network. This capability reduces management costs and hardware costs. In general, it is highly recommended that customers upgrade to RDMA network cards for SMB 3 application shares.

SMB 3 Scale-Out Benefits

- Continuous availability file stores for Hyper-V and SQL Server
- Load-balanced IO across all nodes
- Distributed access across all nodes
- VSS support
- Transparent failover and client redirection
- Continuous availability at a share level versus a server level

Scale-Out File Servers

Scale-Out File Server shares are for more than just continuous availability. They also provide a mechanism for load balancing and intelligent session state that determines which node is the best connection option based on recent IO stats to the storage. Below is a summary of the additional benefits for using Scale-Out File Server storage for SQL Server and Hyper-V:

Per-Share SMB Scale-Out

- Single node access to multiple nodes simultaneously
- Finer-grained load distribution
- SMB sessions managed per share
    - Instead of per-file server

Optimized Rebalancing of Scale-Out File Server

- Scale-Out File Server clients automatically redirect to best node
- Redirection traffic minimized
- SMB sessions redirected to node with best storage connectivity
- Scale-Out File Server dynamically maintained as volume ownership changes

Intelligent CSV Placement Policies

- CSV volumes distributed across the cluster
- Awareness of other CSV volume placement
- Optimized failover placement policies spread CSV ownership across the cluster

When to use Scale-Out File Server Shares

The two most common applications that take advantage of SMB 3 Scale-Out File Server (application) shares are Hyper-V to store VHD files remotely and SQL Server 2008 R2 / SQL Server 2012 to store user database files remotely. This compatibility provides continuous availability failover support of SQL Server

instances and Hyper-V high-availability virtual machines without disruption of service. The use of RDMA gives the network path up to 40 Gbps or higher throughput, depending on the NICs used. This keeps you from sacrificing performance with Scale-Out File Server when using RDMA NICs.

SMB 3 application shares provide you with an additional file-based storage option for your mission-critical applications. Continuous availability shares are very easy to provision and deploy by using highly available Storage Spaces with JBOD back-ends.
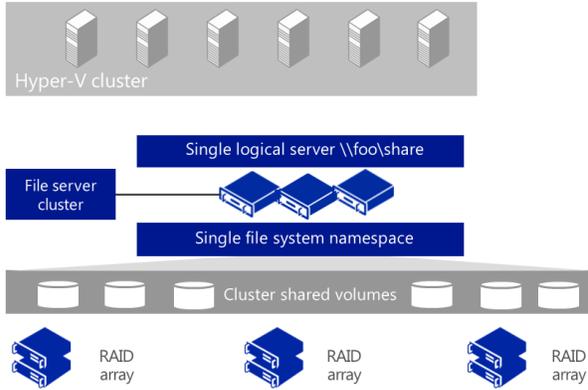
Figure 7: A SMB 3 File Server Cluster

Data Deduplication

Challenges

- Bringing down the cost per IOPS.
- Bringing down the cost/GB of storage.

Solution

Data Deduplication, a storage efficiency feature first introduced in Windows Server 2012, helps address the ever-growing demand for file storage. Instead of expanding the storage used to host data, Windows Server automatically scans through your disks, identifying duplicate chunks and saving space by storing these chunks only once. This functionality saves you money by optimizing your existing storage infrastructure. In addition, deduplication offers even greater savings by extending the lifespan of current storage investments.

From both internal testing and that performed by ESG Lab, Data Deduplication has shown storage savings of 25-60 percent for general file shares and 90 percent for OS VHDs. This is far above what was possible with Single Instance Storage (SIS) or New Technology File System (NTFS) compression.

Source: ESG Lab and Microsoft Internal Testing

Average savings with data deduplication by workload type
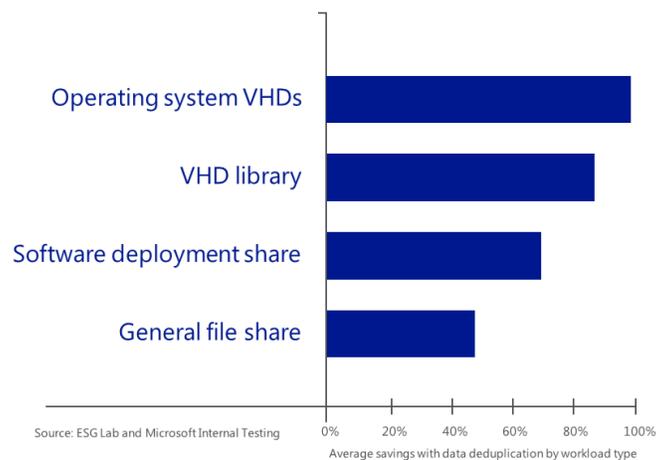
Figure 8: Data Deduplication cost savings

Data Deduplication also throttles CPU and memory usage, enabling implementation of large volumes without impacting server performance. In addition, metadata and preview redundancy help prevent data loss due to unexpected power outages. Checksums, along with data integrity and consistency checks, help prevent corruption for volumes configured to use Data Deduplication.

Routine compression run times can also be scheduled for off-peak times to reduce any impact those operations might have on data access.

Network Fault Tolerance with SMB Multichannel

Challenge

Achieve reliability and performance with existing hardware.

Solution

With SMB Multichannel, you can automatically and transparently handle network path failures without application service disruption. Windows Server 2012 R2 scans, isolates, and responds to unexpected server problems, enabling network fault tolerance (if multiple paths are available between the SMB client and the SMB server).

SMB Multichannel also provides aggregation of network bandwidth from multiple network interfaces when multiple paths exist. Server applications can then take full advantage of all available network bandwidth, making them more resistant to network failure.

SMB Direct (RDMA)

Challenges

- Limited network channel performance based on how fast the CPU can handle traffic as opposed to the line rate of the NIC

- The ability to offload data transfer to the underlying hardware

Solution

The SMB protocol in Windows Server 2012 and Windows Server 2012 R2 includes support for RDMA network adapters, which enables storage performance capabilities that rival Fibre Channels. RDMA network adapters offer this performance capability by operating at full speed with very low latency due to CPU offloading and the ability to write directly to the memory of the remote storage node using RPCs. Reliable transport protocols, implemented on the adapter hardware, enable zero-copy networking with kernel bypass.

With SMB Direct, applications, including SMB, can perform data transfers directly from memory, through the adapter, to the network, and over to the memory of the application requesting data from the file share. This capability is especially useful for read/write intensive workloads such as Hyper-V or SQL Server, resulting in remote file server performance comparable to local storage.

In traditional networking, a request from an application to a remote storage location must go through numerous stages and buffers on both the client and server side, such as the SMB client or server buffers, the transport protocol drivers in the networking stack, and the network card drivers.

With SMB Direct, the application makes a data transfer request straight from the SMB client buffer, through the client NIC to the server NIC, and up to the SMB server. This change provides the application with access to remote storage at the same speed as local storage.

The following scenarios provide examples of a solution without SMB-Direct and RDMA and one that works with it:

Scenario 1: Without SMB-Direct and RDMA

1. A client application makes a request for 500K of data. This request can be a client-to-server request or a server-to-server request (as with scale-out shares for Hyper-V storage).
2. The client formulates the SMB request and sends it over the TCP/IP stack to the server.
3. The server processes the 500K of data into packets. This process takes up CPU cycles on the server.
4. The client receives the packets and assimilates them back into the original 500k of data, then sends them to the stack. This process takes up CPU cycles on the client.

Note: The server cannot send the entire 500k at once because the data must be chopped into smaller packets.
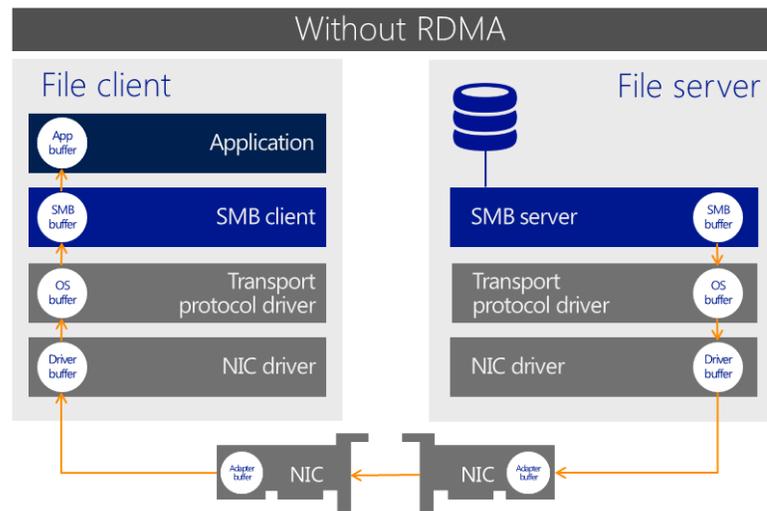

Figure 9: A client and server example without SMB-Direct and RDMA

When using multiple NICs to handle many IO intensive operations, the CPU can become bottlenecked while repeatedly performing the kernel mode activity. The more CPU time spent on processing network packets, the less CPU time is available for other tasks such as database requests and virtual machine workloads.

Scenario 2: With SMB-Direct and RDMA

Through work with partners and vendors, Microsoft has brought to Windows Server 2012 a class of specialized NICs to support high-speed data transfers. These NICs have faster CPUs and support RDMA so that they can transfer data between them without involving the host CPU.

In the SMB-Direct (RDMA) example below:

1. The client requests 500K of data.
2. The SMB-Direct transport driver locates a place in memory where that data should reside and registers it on behalf of the NIC.
3. An RDMA request is sent to the server requesting to read the 500K of data directly into client memory.
4. The client application and server transfer data directly, memory-to-memory, outside the packet level constraints incurred when normal peer-to-peer networking happens over TCP/IP.
5. The two NICs (or NIC teams) perform the transfer without any outside CPU dependencies and overhead.

SMB Direct is extremely fast. In testing, transfers had 1-2 millisecond latencies. This speed almost equals that of using local storage. Support for Infiniband, RoCE, and iWARP network interfaces also help further increase speed.
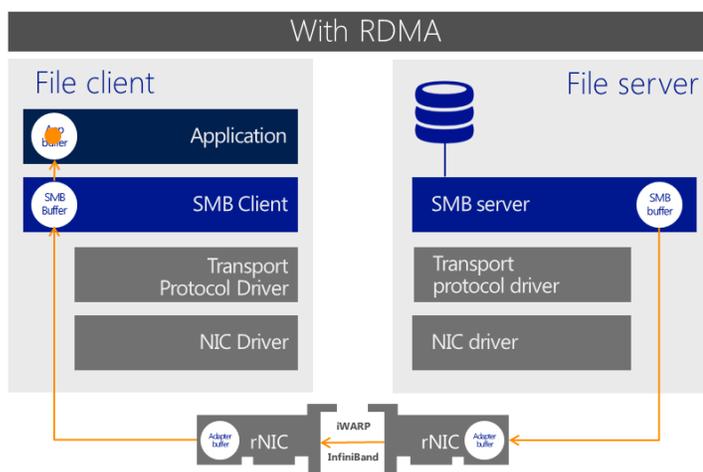
Figure 10: A client and server example with SMB-Direct and RDMA

With iWarp and Infiniband networking technologies, SMB Direct can deliver transfer rates equivalent to 50Gbps on a single 56GB NIC. SMB Direct can also enhance transfer rates and increase the reliability of network paths through direct compatibility with SMB Multichannel as part of SMB 3.0.

Did You Know?

- Mellanox and Intel are some of the vendors that make RDMA capable NICs

- RDMA will become de facto in 10GB and 40GB NICs

SMB-Direct Benefits

- Higher performance and lower latency through CPU offloading
- High-speed network utilization (including InfiniBand and iWARP)
- Remote storage at the speed of local storage
- A transfer rate of approximately 50Gbps on a single NIC port
- Compatibility with SMB Multichannel for load balancing and failover

Online VHDX Resize

Hyper-V in Windows Server 2012 R2 now offers a way to perform online VHDX resize operations (both expanding and trimming). Virtual machines that host applications with service-level agreements (SLA's) must be able to increase and decrease the size of dynamic disks while the virtual machine is running versus having to take the virtual machine and applications offline to do it.

Storage QoS

Another new storage feature with Hyper-V in Windows Server 2012 R2 is Storage Quality of Service (QoS). With this feature, you can control incoming, outgoing, and maximum IOPS.

Challenge

When you virtualize databases or certain virtual machine workloads that are storage heavy, you need to be sure that the databases get the IO bandwidth they need and that you have the ability to monitor storage bandwidth usage.

Solution

With storage QoS, you can set the maximum IOPS that the virtual machine can use. This can prove useful if you have an extraction, transform, load (ETL) job that runs during normal business hours but that you want to prevent from using all available IO on the virtual switch. QoS can throttle the available bandwidth to the guest network adapter when it reaches the limit that you set. You can also set a minimum amount of bandwidth so that it's available even if other processes need resources.

Hyper-V Live Storage Migration

Challenges

- Consolidating virtual machine storage without downtime
- Getting virtual machines off of DAS storage to a shared storage environment

Solution

Before Windows Server 2012, a virtual machine's storage could be moved only while the virtual machine was shut down. Hyper-V in Windows Server 2012 introduced live storage migration, which lets you move virtual hard disks attached to a running virtual machine. Through this feature, you can transfer virtual hard disks, with no downtime, to a new location for upgrading or migrating storage, performing backend storage maintenance, or redistributing your storage load. A new wizard in Hyper-V Manager or the new Hyper-V cmdlets for Windows PowerShell take you through the process. Live storage migration is available for both SAN-based and file-based storage.

When you move a running virtual machine's virtual hard disks, Hyper-V moves storage the following way:

1. Throughout most of the move operation, disk reads and writes go to the source virtual hard disk.
2. While reads and writes occur on the source virtual hard disk, the new destination virtual hard disk copies the disk's contents.
3. After the initial disk copy is complete, disk writes are mirrored to both the source and destination virtual hard disks while replicating outstanding disk changes.
4. After synchronization of the source and destination virtual hard disks, the virtual machine switches over to using the destination virtual hard disk.
5. The source virtual hard disk then gets deleted.
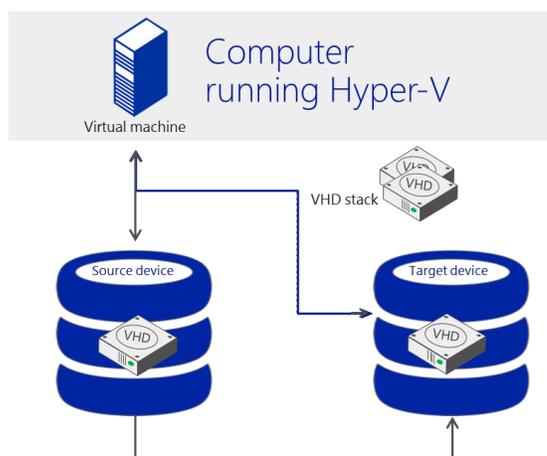
Figure 11: An example of Hyper-V Live Storage Migration

Just as virtual machines might need to be dynamically moved in datacenters, allocated storage for running virtual hard disks sometimes must be moved for storage load distribution, storage device services, or other reasons.

Updating the physical storage available to Hyper-V is the most common reason for moving a virtual machine's storage. You also may want to move virtual machine storage between physical storage devices, at runtime, to take advantage of new, lower-cost storage supported in Hyper-V (such as SMB-based storage), or to respond to reduced performance that results from bottlenecks in the storage throughput. Windows Server 2012 R2 provides the flexibility to move storage along virtual machines to the cluster without any downtime or loss of availability.

Hyper-V Live Storage Migration Benefits

- Easy migration of virtual machine into a cluster while the virtual machine is running
- Improved virtual machine mobility
- Flexible placement of virtual machine storage based on demand
- Migration of virtual machine storage to shared storage without downtime

Simplified Storage Management

In addition to providing a great new feature set with Windows Server 2012 R2, Microsoft has also provided a great management layer that helps you manage the storage solutions whether you have a single server or multiple servers.

Storage Management API (SM-API) is the storage industry standard framework for management tools. No matter how large or small your storage needs are, Microsoft has made rapid strides with Windows Server 2012 R2 and System Center 2012 to provide a platform that helps you build reliable, repeatable, and consistent automation frameworks, so that you can automate your regular tasks and work on what is more important for your career and business.

SM-API

At a Glance: SM-API Improvements

- A new architecture provides ~10x faster disk/partition operations
- The new Storage Spaces provider offers remote and cluster-awareness capabilities
- New support management of Storage Spaces using System Center 2012 Virtual Machine Manager

Challenges

- A need for SM-API support for System Center managed storage pools and clouds and high-availability storage spaces in the cluster
- Faster and easier provisioning

Solution

To help tie the storage capabilities of Windows Server 2012 R2 together, Windows Server 2012 R2 takes advantage of a unified interface that uses Windows Management Instrumentation (WMI) for comprehensive management of physical and virtual storage, including third-party intelligent storage subsystems. The unified interface uses WMI to deliver a rich experience to IT pros and developers using Windows PowerShell scripting that provides a diverse set of solutions. Windows Server 2012 R2 integrates these features with high-performance computing (HPC) and Microsoft System Center for better performance. In addition, management

applications can use a single Windows application programming interface (API) to manage different storage types by using simple management protocol (SMP) or standards-based protocols such as SMI-S.

The unified interface for storage management provides a core set of defined WMI and Windows PowerShell interfaces as well as features for more advanced management.
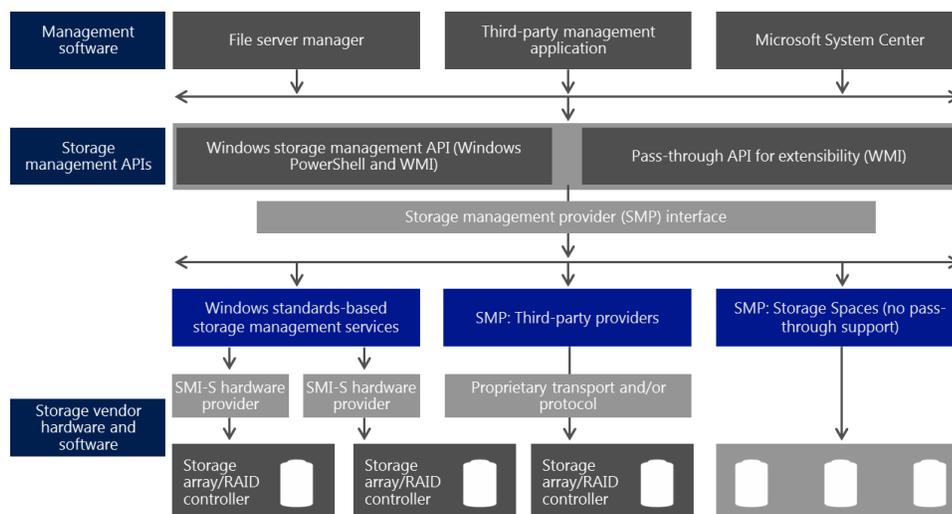
Figure 12: The unified storage management architecture

The unified interface provides a powerful and consistent mechanism for managing storage, which can reduce complexity and operational costs. The storage interface also offers capabilities for advanced management of storage in addition to the core set of defined WMI and Windows PowerShell interfaces.

Much of the Windows Server 2012 SM-API functionality was implemented with the addition of the File and Storage Services manager role and tools. What was missing was the ability to remotely manage cluster-based storage spaces. This feature comes as part of Windows Server 2012 R2. With the File and Storage Services management interface in Server Manager, you can accomplish the following tasks:

- Create storage pools by grouping many physical disks into easily-managed volumes, and then manage those storage pools as a unit.
- Add and share volumes anywhere on the network by using SMB and NFS.
- Create and manage iSCSI virtual disks and targets.

Storage management with Windows Server 2012 R2 has been expanded to provide better support from System Center 2012 R2 Virtual Machine Manager. Some of the benefits of System Center 2012 R2 VMM 2012 include the following:

- Storage Discovery
  - SAN discovery
    - FC, iSCSI, SAS
  - NAS discovery
    - Self-contained NAS, NAS head
  - File server discovery:

Storage Management Benefits:

- A new architecture provides ~10x faster disk/partition enumeration operations
- Remote and cluster-awareness capabilities
- SM-API exposes new Windows Server 2012 R2 features (Tiering, Write-back cache, and so on)
- SM-API features added to System Center VMM
- End-to-end storage high availability space provisioning in minutes in VMM console
- More Windows PowerShell

    - Windows Scale-Out File Server
    - SAN/NAS/file server discovery levels.
  - Level 0: Provider, subsystem (array), file server and all relevant associations
  - Level 1: Storage pool, file share, and all relevant associations
  - Level 2: LUN, masking sets, initiator ports (array knows about), target ports, iSCSI portal, file share ACLs, and all relevant associations
  - Level 3: Physical disks and all relevant associations
- Capacity Management
- Scale-Out File Server deployment
- Storage Spaces Provisioning

Improved Storage Management UI with Server Manager

The Server Manager UI enables you to view the storage solution across your entire datacenter. Whether you have a solution based on SAN, Storage Spaces, or redundant array of independent disks (RAID) controller, you can view the status of all your disks, logical unit numbers (LUNs) and so on. This functionality is very useful when trying to associate a specific physical disk with a virtual disk created in a storage space.

In addition, every storage feature is manageable with a Windows PowerShell module. Every operation that can be performed using Server Manager can also be done using Windows PowerShell, enabling you to automate provisioning of new storage spaces and virtual disks whether you have one server or multiple servers.

System Center Operations Manager also has management packs created for various storage features, which enable you to monitor the performance and status of each of these features.

Did You Know?

With System Center 2012 R2 VMM and Windows Server 2012 R2, you can provision a Scale-Out File Server with a high availability storage space in minutes from either bare metal servers or existing Windows Server 2012 R2 Servers, which are racked, stacked, and ready to go. The steps only take a couple of minutes and involve using the VMM console in the following way:

1. Select a Run-As account and Baseboard Management Controller (BMC) protocol.
2. Select the servers (the servers cannot already have the Hyper-V role installed) using an IP Range or Server Names.
3. System Center 2012 R2 VMM will execute a Wake-on-LAN (WOL) to start the servers and begin the Scale-Out File Server provisioning process, as well as generate the Windows PowerShell code to repeat for other provisions.
4. You can use JBOD storage for back-end storage running a Storage Pool.
5. Once the provisioning is done, you can expand the front-end file servers by adding more servers to the Scale-Out File Server (similar to the process above), and you can add additional disks to the storage pool - all while the Scale-Out File Server is running.

Protecting Your Data

For data protection of core operating system and applications, Windows Server 2012 R2 uses a layered approach that incorporates the new Resilient File System (ReFS) and NTFS check disk (CHKDSK) utility. Both ReFS and CHKDSK have been engineered to provide the highest level of availability.

With ReFS, even if there is a bad area on the disk, ReFS volumes will stay online while repairs are made to the affected areas of the volume. CHKDSK on NTFS now only runs against the affected areas of an NTFS volume, which similarly keeps the volume online, even if there are errors.

Other layers include disk-level protection and fault tolerance (implemented through high availability and resilient storage spaces) and workload and datacenter protection and availability (implemented through System Center 2012 SP1 Data Protection Manager, Windows Azure Backup, StorSimple, and Hyper-V Replica).

Protecting File Systems and Directories

Resilient File System (ReFS) targets maximizing data availability, scaling efficiently to very large data sets across diverse workloads, and guaranteeing data integrity by means of resiliency to corruption. By utilizing an integrated storage stack comprising ReFS and Storage Spaces, you can deploy the most cost-effective platform for available and scalable data access using commodity storage.

At a Glance: ReFS

- Maintains the highest levels of system availability
- Provides a full end-to-end resilient architecture
- Uses the Copy on Write method to update files thereby avoiding corruptions caused during disk power outage
- Offers Checksums on all metadata

Challenges

- Increased volumes
- Increased resiliency to power failures
- High availability

Solution

Windows Server 2012 introduced an engineered file system, Resilient File System (ReFS), meant to maintain compatibility with the popular NTFS file system and support a new generation of storage technologies and scenarios. Some of these scenarios include the following:

- Maintain the highest levels of system availability and reliability under the assumption that underlying storage can be inherently unreliable.
- Provide a full end-to-end resilient architecture when used in conjunction with Storage Spaces so that these two features magnify the capabilities and reliability of the other when used together.
- Use the Copy on Write method to update files and save them to a new location after every update, thereby avoiding corruptions caused during disk power outage.

- Perform Checksums on all metadata in ReFS at the tree level and store the Checksums independently from the tree page itself. This feature enables the detection of all forms of disk corruption, including degradation of data on media.
- Forms the storage foundation, along with Storage Spaces, on Windows for the next decade and beyond with features that enable significant storage stability, flexibility, scalability, and availability.

CHKDSK in Windows Server 2012 R2 has been improved to run in seconds versus minutes or hours. Scanning of volumes is now done online, dramatically reducing the time that volume is offline for repairs.

Resilient File System (ReFS)

- Rapid recovery from file system corruption without affecting availability.
- Resilient against power outage corruption.
- Periodic checksum validation of file system metadata.
- Improved data integrity protection.
- ReFS remains online during subdirectory reconstruction and knows where orphaned subdirectories exist and automatically reconstructs them.
- Ideal for file server volumes.

Source: "ESG Lab"

CHKDSK

- Seconds to fix corrupted data.
- No offline time when used with CSV.
- Disk scanning process separated from repair process.
- Online scanning with volume and offline repair.

Hours (chart: 20, 15, 10, 5, 0 — categories 1, 2, 3; Series1, Series2)
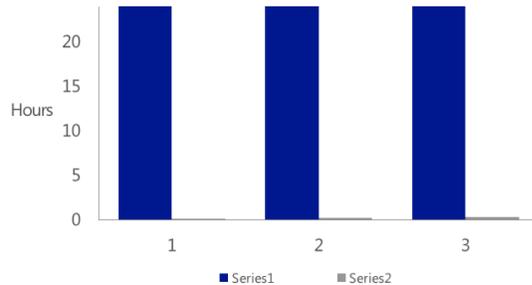
Figure 13: Features of ReFS and CHKDSK

ReFS Benefits

- More resilience to power failures
- Highest levels of system availability
- Larger volumes with better durability
- Scalable to petabyte size volumes

Windows Azure Backup

Windows Azure Backup Service is a cloud-based storage service managed by Microsoft. To transfer data between servers running Windows Server 2012 R2 and Windows Azure Backup you can use either the Windows Azure Backup Service Agent or the Online Backup cmdlets for Windows PowerShell.

- Ability to leverage Windows Azure cloud services to back up data.
- Backs up data to Azure on a user set schedule.
- Reduced cost for backup storage and management.
- Ideal for small businesses, branch offices, and departmental business needs.
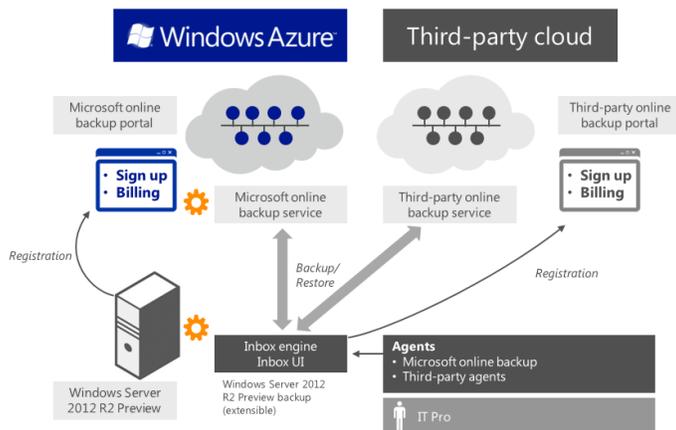
Figure 14: Windows Azure Backup

Target Segments

Windows Azure Backup Service offers a low-cost backup and recovery solution for small businesses, departmental backups, and remote office backups that lack a local backup infrastructure. The Windows Azure Backup Service is not meant to be a replacement for enterprise-level backup solutions such as System Center 2012 R2 DPM.

> **At a Glance: Windows Azure Backup:**
> - Cloud-based storage service for online backups
> - Windows PowerShell instrumented
> - Simple, reliable Disaster Recovery solution for applications and data
> - Supports System Center 2012 R2 DPM

System Center 2012 R2 Data Protection Manager (DPM)

System Center 2012 R2 DPM enables disk-based and tape-based data protection and recovery for servers such as SQL Server, Exchange Server, SharePoint, virtual servers, file servers, and support for Windows desktops and laptops. System Center 2012 R2 DPM can also centrally manage system state and Bare Metal Recovery (BMR). New features include offsite replication and backup on Windows Azure Cloud Services and applications. In particular, System Center 2012 R2 DPM targets the following areas:

- File data from volumes, shares, and folders
- Application data, such as Exchange Server storage groups, SQL Server databases, Windows SharePoint Services farms, and Microsoft Virtual Server and its virtual machines
- Files for workstations running Windows XP Professional SP2 and all Windows Vista editions except Home
- Files and application data on clustered servers
- System state for protected file and application servers
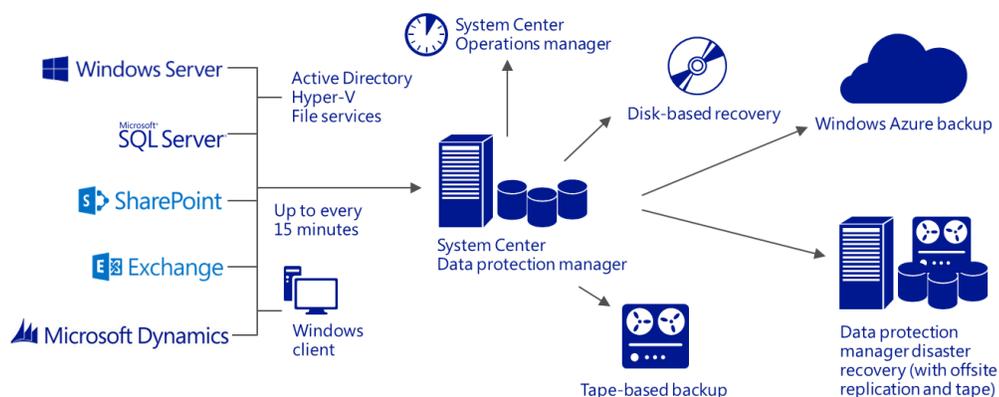
Figure 15: System Center 2012 DPM

System Center 2012 R2 DPM includes the following features from System Center DPM 2012:

- Disk-based data protection and recovery
- Command-line scripting using Windows PowerShell
- Enterprise deployment methods for distributing the System Center 2012 R2 DPM agent
- Enterprise monitoring with Operations Manager
- Tape-based backup and archive solutions
- Disaster Recovery solutions, which provide BMR of servers running Windows

Windows Azure Backup Benefits

- Reduced total cost of ownership (TCO) by providing scalability, elasticity, and simplified storage management
- Peace of mind by providing a reliable and secure short-term offsite backup and restore solution
- Simplicity through seamless integration into the existing System Center 2012 R2 DPM backup, recovery, and monitoring workflows

Windows Azure Backup

As an alternative to backing up System Center 2012 R2 DPM to disk or to a secondary on-premise System Center 2012 R2 DPM server, you can back up System Center 2012 R2 DPM servers and data protected by those servers to the cloud, using Windows Azure Backup.

Windows Azure Backup works with the System Center 2012 R2 DPM disk-based protection feature. When you enable online protection, the disk-based replicas are backed up to an online location, providing an added level of protection and peace of mind.
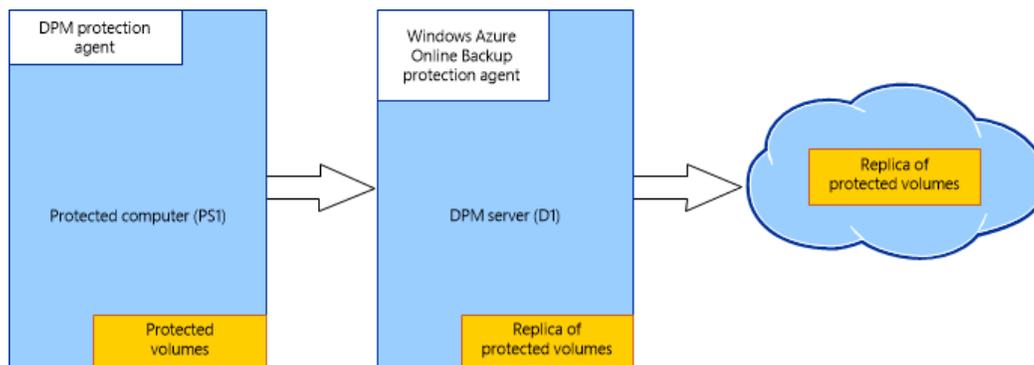
Figure 16: How System Center 2012 R2 DPM works with the Windows Azure Backup

Virtualized deployment

System Center 2012 R2 DPM can be deployed in a virtual environment on a virtual machine. You can also configure storage using VHDX storage pool disks that are shared through the VMM library.

Linux virtual machine backup

System Center 2012 R2 DPM provides support for the protection and backup of Linux virtual machines, in addition to the support already provided for Hyper-V virtual machines. Linux backups only support file-consistent snapshots, not application-consistent ones. In addition, protection of Linux virtual machines is not supported using Windows Azure Backup Service.

SQL Server cluster support

System Center 2012 R2 DPM now supports the use of clustered SQL Server nodes, removing the standalone limitation that existed in System Center 2012 and System Center 2012 SP1. This new functionality enables you to use a remote SQL Server cluster in addition to a standalone server. Other benefits include the following:

- Reliability
  - Support for a SQL Server cluster mitigates the single point of failure when a standalone SQL server is used.
- Scalability
  - As your System Center 2012 R2 DPM deployment grows, you can add another clustered SQL Server database or instance.
- Consistency
  - Support for a SQL Server cluster makes System Center 2012 R2 DPM consistent with other System Center 2012 components.

In addition, the System Center 2012 R2 DPM reporting server can be installed on the same standalone or clustered SQL Server used for the System Center 2012 R2 DPM database or on a different SQL Server.

System Center 2012 R2 Data Protection Management: Enterprise Ready

Microsoft IT uses 80 System Center DPM servers running back up for 3.5 petabytes (PBs) of data. For machines that lie outside the domain, which is a very common scenario, System Center DPM supports certificate-based protection. System Center DPM also has protection of clustered workloads and integrates with System Center Operations Manager for monitoring backup and recovery operations the same way you monitor your applications, fabric, and infrastructure.

System Center DPM offers customers a single backup and recovery solution that scales to any size and backs up any Microsoft or non-Microsoft application with ease. Clustered and virtualized workloads are completely supported for online backups whether those jobs are simple Volume Snapshot Service (VSS) copies of CSV volumes or complex backup operations for Exchange, SQL Server, SharePoint, customer relationship management (CRM), or any other

enterprise application. System Center 2012 R2 DPM can also integrate with Operations Manager and System Center VMM for efficient backup of cloud-based applications.

System Center 2012 R2 DPM Benefits

- Offers single backup and recovery solution that protects all enterprise data, infrastructure, and apps
- Provides offsite replicas with Windows Azure Backup
- Scales to any size and tightly integrates with System Center Operations Manager and System Center 2012 R2 VMM
- Supports clustered SQL nodes and a virtualized deployment

Hyper-V Replica

Windows Server 2012 introduced Hyper-V Replica as a built-in feature that provides asynchronous replication of virtual machines for the purposes of business continuity and disaster recovery. In the event of failures (such as power failure, fire, or natural disaster) at the primary site, the administrator can manually failover the production virtual machines to the Hyper-V server at the recovery site. During failover, the virtual machines recover to a consistent point-in-time within minutes. The virtual machines running in the recovery site are accessible to the rest of the network in near real-time. Once the primary site comes back, the administrators can manually convert the virtual machines to the Hyper-V server at the primary site.

At a Glance: Hyper-V Replica:

- Synchronous virtual machine replication to recovery site with point-in-time recovery
- Three-site group available with Primary, Recovery, and DR sites.
- Failure recovery in minutes
- Closely integrated with System Center 2012 R2 VMM and Failover Clustering to support live migrations

Challenges

- Quick recovery of business functions from a downtime event with minimal or no data loss.
- Protection against power failure, IT hardware failure, network outage, human errors, IT software failures, and natural disasters.
- Affordable and reliable business continuity that helps companies recover from a failure.

Solution

Once a virtual machine or group of virtual machines has been configured for replication, Hyper-V Replica lets you replicate your Hyper-V virtual machines over a network link from one Hyper-V host at a primary site to another Hyper-V host at the Hyper-V Replica site. This solution has no reliance on storage arrays or other software replication technologies. In Windows Server 2012 R2, you have the ability to further chain the replication topology to a 3rd site for an added layer of DR protection.
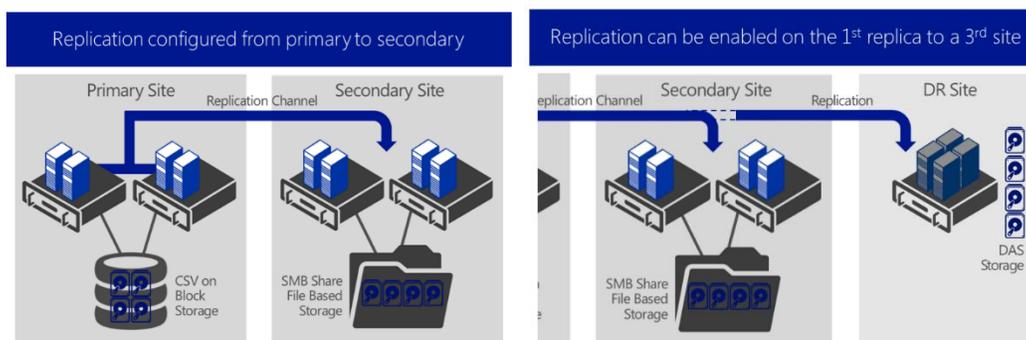

Figure 17: Secure replication of virtual machines from different systems and clusters to a remote site over a WAN

Failover Cluster Support

Hyper-V and Failover Clustering can be used together to make a virtual machine highly available. Administrators can seamlessly migrate their virtual machines to a different host in the cluster in the event of an outage. They can also load balance their virtual machines without impacting virtualized applications. This functionality requires the use of Scale-Out File Server and CSV-2 for the storage of the VHDX files on continuous availability shares.

While this capability can protect virtualized workloads from a local host failure or scheduled maintenance of a host in a cluster, it does not protect businesses from outage of an entire datacenter. While Failover Clustering can be used with hardware-based SAN replication across data centers, these are typically expensive. Hyper-V Replica fills an important gap in the Windows Server Hyper-V offering by providing an affordable in-box disaster recovery solution.

Hyper-V Replica Architecture

Hyper-V Replica architecture consists of the following components:

Hyper-V Replica Replication Tracker (RT)

The Hyper-V Replica Replication Tracker periodically replicates updates to the virtual hard disks associated with a virtual machine to a Hyper-V Replica server. This replication is configurable to 30 seconds, 5 minutes, or 30 minutes in Windows Server 2012 R2. Once you enable replication for a virtual machine/VHD, the replication log (*.HRL) tracks all changes made to the virtual machine/VHD. The replication log is stored in the same folder as the VHD file being replicated. Each VHD on a virtual machine being replicated has an associated log. Logs are transmitted from primary server to the Hyper-V Replica server.

Recovery history controls how far back a virtual machine can be recovered. History for a virtual machine is determined by the settings made by the Hyper-V administrator on each virtual machine. The default option is to store only the latest point for recovery. This option causes all the replicated log data to be merged to the original virtual machine VHD(s). Once acknowledgement is sent by the Hyper-V Replica server that changes have been merged, the process repeats itself. Replication metadata, needing to be preserved during virtual machine mobility scenarios such as cluster failover, live migration, and so on, is stored inside the virtual machine configuration.

Assuming that replication acknowledgements are received by the primary site on schedule, the following scenarios dictate how changes are merged.

- **Scenario 1: Storage of the latest recovery point**
    - At the configured interval of 30 seconds, 5 minutes, or 30 minutes, the replication log is frozen on the primary server and sent to the Hyper-V Replica server. A new log is started on the primary server and the replicated log is merged into the VHD on the Hyper-V Replica server.
- **Scenario 2: Standard Replicas**
    - This scenario is the same as scenario 1, with the addition of the Hyper-V Replica server creating a recovery snapshot every hour to match the number of replicas to keep. Once the limit is reached, the oldest snapshots merge to the base replica VHD, removing stale snapshots.
- **Scenario 3: Standard Replica with Application-Consistent recovery points**
    - This scenario is the same as scenario 2, with the addition of a VSS snapshot being taken at the desired interval sent to the replica site. When an app-consistent copy arrives, it gets applied and the entire replica snapshot becomes app consistent.

    -

Hyper-V Replica Broker Manager

The Hyper-V Replica Broker Manager operates in a similar way to the Hyper-V Replica Server, except that instead of a primary and replica server being in the mix, the Replica Broker Manager maintains a network listener in the Failover Cluster. As a result, all virtual machines in the cluster, up to 8000 of them, can be synchronized with replicas in another cluster, even in another datacenter or city. This capability raises the scale of Hyper-V replication to up to 64 nodes. When setting up the Hyper-V Replica Broker Manager in the cluster, the only difference is that the two replication partners are clusters instead of standalone hosts.
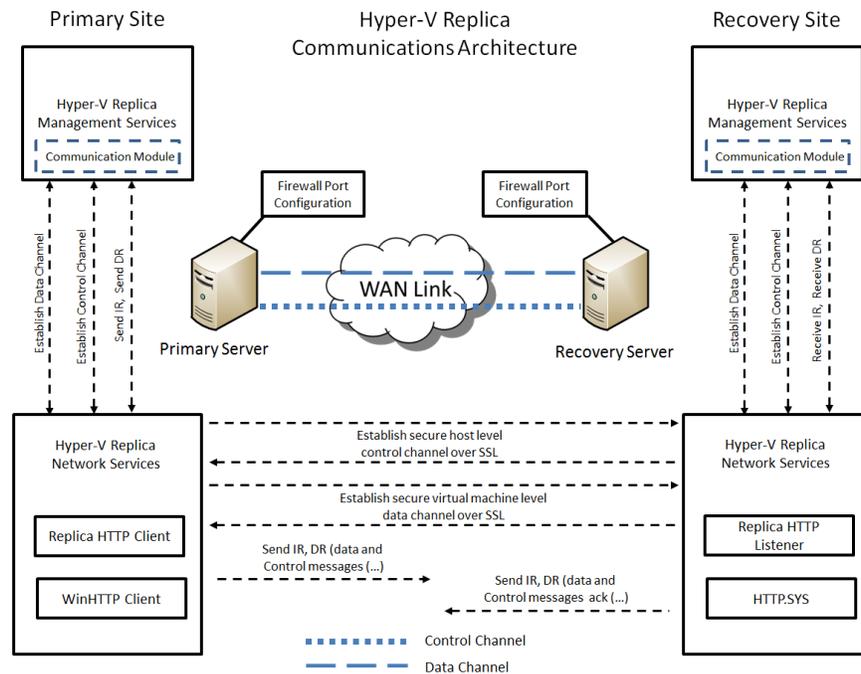

Figure 18: Hyper-V Replica Communications Architecture

The Hyper-V Replica architecture implements a transport layer, which enables bi-directional communication for data and control. Access control and authentication for Hyper-V Replica pass over this transport as well as the replica changes. It's possible to encrypt replication traffic using a public-key infrastructure (PKI).

Hyper-V Replica Benefits

- Offsite warm standby of some or all virtual machines
- Scalability up to 8000 virtual machines
- Configurable at virtual machine level
- Windows PowerShell Automation
- Replication metadata moves with the virtual machine in all mobility scenarios

Hyper-V Recovery Manager

Hyper-V Recovery Manager is a new cloud-scale disaster recovery solution for Hyper-V. It's built on the foundation of Windows Server 2012 R2 and System Center VMM 2012 SP1. This solution is part of the new Microsoft Disaster Recovery as a Service (DRaaS) architecture.

Challenges

- To reduce costs related to downtime.
- DR solutions with synchronous replication are expensive
- Constant monitoring of services
- Manual recovery of many virtual machines
- Cost and complexity of workloads leave companies unprotected

Solution

With more and more enterprises moving to public and private cloud architectures, Hyper-V Replica, first introduced in Windows Server 2012, needed to evolve into a new cloud-based solution. The primary design goals for Hyper-V Recovery Manager needed to deliver the following:

- Simplicity
- Reliability
- Extensibility
- A consistent user experience across all services and apps

Hyper-V Recovery Manager

Hyper-V Recovery Manager extends Hyper-V Replica, adding an emphasis on simplicity and consistency of the user experience across clouds. Most enterprise applications generally have several built-in layers of reliability, which are specific to the application. The more applications that you need to protect, the further you get away from consistency and simplicity. Combine that with the cost and complexity of traditional DR solutions, which require expensive deployment of agents and software, and only the most mission-critical enterprise applications usually make the grade for an offsite DR solution.

Hyper-V Recovery Manager levels the playing field for DR. All workloads, regardless of application, vendor, and version, can be protected the same way by delivering business continuity in hours, not days. Data and control channels for a good offsite DR solution are recommended to be encrypted and firewall friendly.

Hyper-V Recovery Manager is built on Windows Server 2012 R2, System Center 2012 VMM SP1, and Windows Azure. At no time, does Windows Azure ever connect to your internal network. Encrypted communications only flow between primary and secondary sites, never to or from Windows Azure. Therefore, you never have to worry about inbound channels from Windows Azure to your corporate datacenter. The two arrows in Figure 19 are DR Orchestration, which point into Windows Azure from the primary and secondary datacenters. DR Orchestration is metadata only. Since the data only flows on datacenter-owned channels, the data preserves its confidentiality.
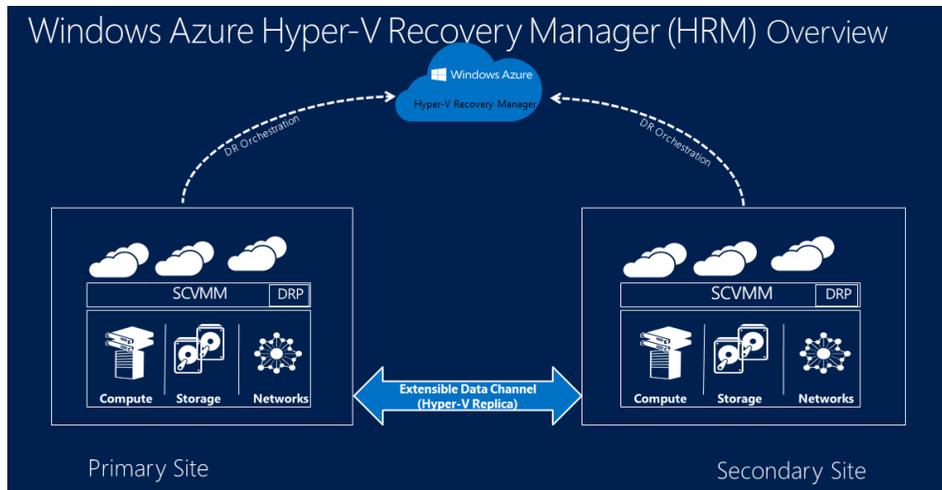
Figure 19: The major components of Hyper-V Recovery Manager

Figure 19 illustrates two sites, which act as independent private clouds (primary site and secondary site). The cloud-based DR orchestration/management of Hyper-V Recovery Manager is available through the Windows Azure Management Portal. The extensible data channel currently supports Hyper-V Replica control and data with a network-level transport. However, with the channel being extensible, it is possible for Microsoft and partners to develop extensions for this layer to better support custom applications and DR requirements.

The communications channels from your datacenter to Windows Azure (see Figure 20) are outbound-only (using an encrypted session) or optional (using an HTTPS proxy, which can use your own PKI infrastructure). Since HTTPS is the application layer protocol, the connection to Windows Azure is simple, inexpensive, and firewall friendly. No additional changes are needed to your existing Hyper-V Replica host settings. Only System Center VMM needs to talk to Windows Azure, not the Hyper-V hosts themselves.
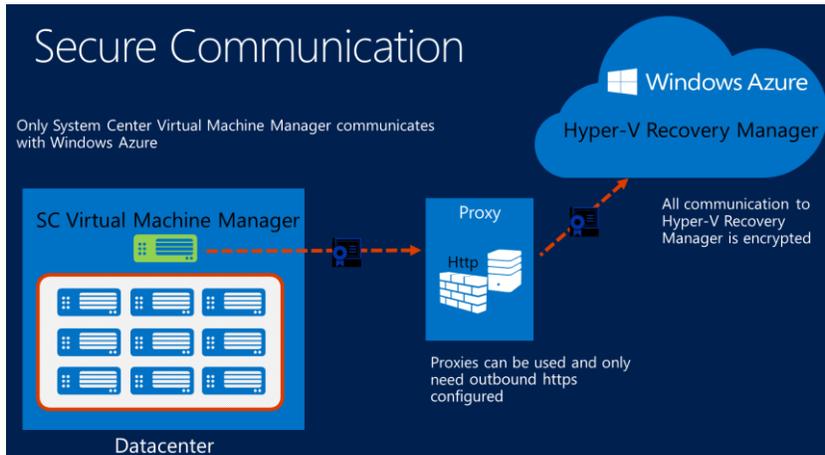
Figure 20: Communications channels from a datacenter to Windows Azure.

Hyper-V Recovery Manager offers a single console for protection and recovery of multiple clouds in one or more sites. Where Hyper-V Recovery Manager goes above and beyond the concept for virtual machine replication is in the protection and recovery of virtual networks as well as virtual machines. Figure 21 depicts how the Gold recovery site not only stores the virtual machines that need to be protected, but also the virtual networks for those virtual machines. Hyper-V Recovery Manager takes DR to a cloud level versus a virtual machine level with Hyper-V Replica alone.
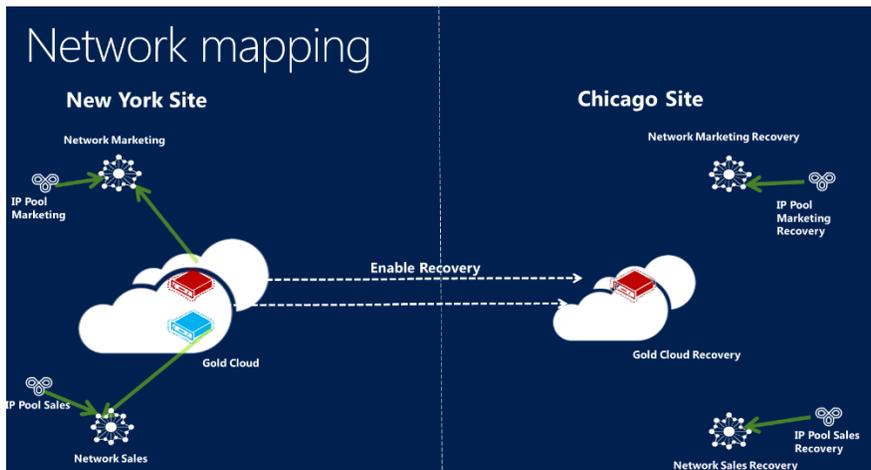

Figure 21: Gold recovery site stores virtual machines and virtual networks

Maximizing Existing Storage Investments

Windows Server 2012 R2 helps businesses and datacenters utilize their existing investments in SAN hardware by taking advantage of the speed and power of standard SAN features whenever possible in the operating system.

Core operating system functions and certain Hyper-V functions can now benefit from SAN functionality to reduce CPU and network bandwidth. This functionality enables you to offload to the SAN whenever possible with offloaded data transfer (ODX) and connect virtual machine workloads directly to FCP storage for unmediated access the SAN from any virtual machine.

To take better advantage of your existing SAN investments with Windows Server 2012 R2, you can incorporate the following features:

- ODX
- Virtual Machine Fibre Channel

Offloaded Data Transfer (ODX)

Offloaded Data Transfer (ODX) provides CPU and network offloading to SAN hardware during file copy and move operations between SAN drives. This feature is especially significant in the cloud space when you must provision new virtual machines from virtual machine template libraries or when virtual hard disk operations require large blocks of data to be copied, such as virtual hard disk merges, storage migration, and live migration. The storage device handles the copy operations and must be able to perform offloads (such as an offload-capable iSCSI, Fibre Channel SAN, or a file server based in Windows Server 2012 R2). This capability frees up the Hyper-V host processors to carry more virtual machine workloads.

At a Glance: ODX

- Token-based copy operations
- File copy and move operations offloaded to SAN
- Support for most iSCSI and FCP environments
- Automatic offloading of network and CPU

Challenges

- High CPU and network bandwidth utilization during file copy and move operations
- Poor use of SAN hardware for copy and move disk operations across SAN drives

Solution

ODX in Windows Server 2012 R2 enables you to accomplish more with your existing external storage arrays by letting you quickly move large files and virtual machines directly between storage arrays, reducing host CPU and network resource consumption.

ODX support is a feature of the storage stack of Hyper-V in Windows Server 2012 R2. When used with offload-capable SAN storage hardware, ODX lets a storage device perform a file copy operation without the main processor of the Hyper-V host reading the content from one storage place and writing it to another.

ODX uses a token-based mechanism for reading and writing data within or between intelligent storage arrays. Instead of routing the data through the host, ODX copies a small token between the source and destination. The token simply serves as a point-in-time representation of the data.

For example, when you copy a file or migrate a virtual machine between storage locations (either within or between storage arrays), ODX copies a token that represents the virtual machine file. This removes the need to copy the underlying data through the servers.
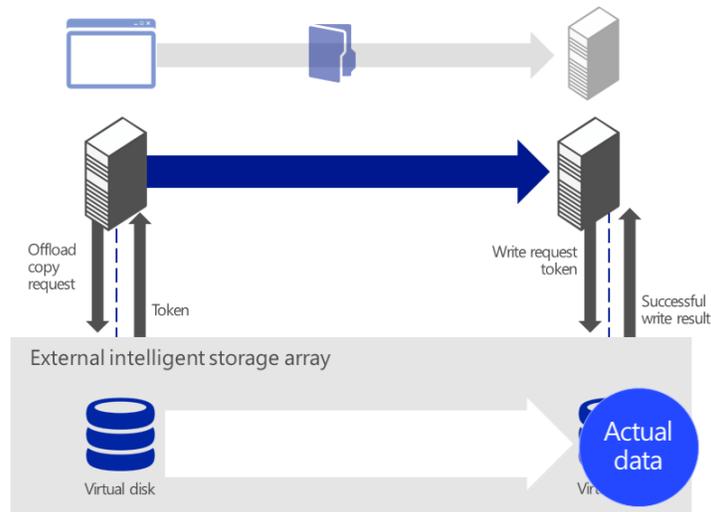
Figure 22: AN ODX token-based copy operation

A user attempts to copy a file or move in Windows Explorer, a command-line interface, or a virtual machine migration from one volume to another that is residing on a SAN which supports ODX. The following happens automatically (even if copying and pasting through Windows Explorer):

1. Windows Server automatically translates this transfer request into an ODX (if supported by the storage array) and receives a token representation of the data.
2. The token is copied between the source and destination systems.
3. The token is delivered to the storage array.
4. The storage array performs the copy internally and returns progress status.

ODX Benefits

- Frees up CPU and network bandwidth during file operations
- Greatly reduces the time needed to copy large amounts of data
- Offers extremely fast live migration and live storage migration in Hyper-V

Virtual Machine Fibre Channel

At a Glance: Virtual Machine Fibre Channel

- Unmediated access to a SAN
- A hardware-based I/O path to any virtual machine
- NPIV and MPIO support
- A single Hyper-V host connected to different SANs

Challenges

- The need for direct access to FCP networks from Hyper-V virtual machines
- True SAN line speed from virtual machines

Solution

Virtual Fibre Channel for Hyper-V, a new feature of Windows Server 2012, provides Fibre Channel ports within the guest operating system, letting you connect to Fibre Channel directly from within virtual machines.

Virtual Fibre Channel support includes the following:

- **Unmediated access to a SAN**
    - Virtual Fibre Channel for Hyper-V provides the guest operating system with unmediated access to a SAN by using a standard World Wide Name (WWN) associated with a virtual machine. Hyper-V lets you use Fibre Channel SANs to virtualize workloads that require direct access to SAN LUNs. Fibre Channel SANs also enable you to operate in new scenarios, such as running the Windows Failover Cluster Management feature inside the guest operating system of a virtual machine connected to shared Fibre Channel storage.
- **A hardware-based I/O path to the Windows software virtual hard disk stack**
    - Mid-range and high-end storage arrays include advanced storage functionality that helps offload certain management tasks from the hosts to the SANs. Virtual Fibre Channel presents an alternative, hardware-based I/O path to the Windows software virtual hard disk stack. This path lets you use the advanced functionality of your SANs directly from Hyper-V virtual machines. For example, Hyper-V users can offload storage functionality (such as taking a snapshot of a LUN) to the SAN hardware simply by using a hardware VSS provider from within a Hyper-V virtual machine.
- **N-Port ID Virtualization (NPIV)**
    - NPIV is a Fibre Channel facility that enables multiple N-Port IDs to share a single physical N-Port. With NPIV, multiple Fibre Channel initiators can occupy a single physical port, easing hardware requirements in SAN design, especially where virtual SANs are called for. Virtual Fibre Channel for Hyper-V guests uses NPIV (T11 standard) to create multiple NPIV ports on top of the host's physical Fibre Channel ports. A new NPIV port is created on the host each time a virtual host bus adapter (HBA) is created inside a virtual machine. When the virtual machine stops running on the host, the NPIV port is removed.

- **A single Hyper-V host connected to different SANs with multiple Fibre Channel ports**
    - Hyper-V enables you to define virtual SANs on the host to accommodate scenarios where a single Hyper-V host connects to different SANs with multiple Fibre Channel ports. A virtual SAN defines a named group of physical Fibre Channel ports that connect to the same physical SAN. For example, assume a Hyper-V host connects to two SANs—a production SAN and a test SAN. The host connects to each SAN through two physical Fibre Channel ports. In this example, you might configure two virtual SANs—one named "Production SAN" that has two physical Fibre Channel ports connected to the production SAN and one named "Test SAN" that has two physical Fibre Channel ports connected to the test SAN. You can use the same technique to name two separate paths to a single storage target.
- **Up to four virtual Fibre Channel adapters on a virtual machine**
    - You can configure as many as four virtual Fibre Channel adapters on a virtual machine and associate each one with a virtual SAN. Each virtual Fibre Channel adapter associates with one WWN address, or two WWN addresses to support live migration. Each WWN address can be set automatically or manually.
- **MPIO functionality**
    - Hyper-V in Windows Server 2012 can use the multipath I/O (MPIO) functionality to help ensure optimal connectivity to Fibre Channel storage from within a virtual machine. You can use MPIO functionality with Fibre Channel in the following ways:

        - Virtualize workloads that use MPIO. Install multiple Fibre Channel ports in a virtual machine and use MPIO to provide highly available connectivity to the LUNs accessible by the host.
        - Configure multiple virtual Fibre Channel adapters inside a virtual machine and use a separate copy of MPIO within the guest operating system of the virtual machine to connect to the LUNs that the virtual machine can access. This configuration can co-exist with a host MPIO setup.
        - Use different device-specific modules (DSMs) for the host or each virtual machine. This approach enables live migration of the virtual machine configuration, including the configuration of DSM and connectivity between hosts and compatibility with existing server configurations and DSMs.
- **Live migration support with virtual Fibre Channel in Hyper-V**
    - To support live migration of virtual machines across hosts running Hyper-V while maintaining Fibre Channel connectivity, you configure two WWNs for each virtual Fibre Channel adapter: Set A and Set B. Hyper-V automatically alternates between the Set A and Set B WWN addresses during a live migration. This functionality helps to ensure that all LUNs are available on the destination host before the migration and that minimal downtime occurs during the migration.

Requirements for Virtual Fibre Channel in Hyper-V include the following:

- One or more installations of Windows Server 2012 with the Hyper-V role installed. Hyper-V requires a computer with processor support for hardware virtualization.

- A computer with one or more Fibre Channel HBAs, each with an updated HBA driver that supports Virtual Fibre Channel.
- Windows Server 2008, Windows Server 2008 R2, or Windows Server 2012 as the guest operating system.
- Connection only to data LUNs
  - Storage accessed through a Virtual Fibre Channel connected to a LUN can't be used as boot media

Virtual Fibre Channel Benefits

- Connect virtual machine storage directly to FCP
- Give the guest operating system unmediated access to a SAN
- Offer unmatched performance when connecting up to four FCP adapters to a virtual machine

Case Study:
Windows Build Team

Situation

- The Windows Build Team needed 2PBs of storage to perform up to 40,000 Windows installations a day. They assembled a mix of traditional datacenter storage technologies for the job.
- 2PBs only stores 5 days of data. They wanted to increase storage, but without increasing budget.
- Typically, a disk failure occurs each month requiring another build and installation before the break can be triaged. This single failure can affect dozens of teams.

Solution

By using Windows Server 2012 R2 features such as Storage Spaces and Data Deduplication, the Windows Build Team gets far more performance and resilience for its storage budget including triple storage capacity. Storage Spaces reduces storage costs by using JBOD devices. Data Deduplication reduces costs by reducing storage requirements. Continuously available clusters provide resilience.

With the new solution, the Windows Build Team now uses 20 file servers, connected to 20, 60-bay JBOD arrays, for a total of 4 Petabytes of resilient storage. Huge volumes of servers hit source repositories during build operations. Each operation generates debug symbol data and stores code. Engineers and partners use the build symbols in their debuggers. These build servers and users connect through the corporate network infrastructure to a switching fabric with an aggregate of 80Gbps network connections to the file servers.

Benefits

By using Storage Spaces with the Windows File Server Cluster, the team was able to gain the following benefits:

- Double their total storage throughput overnight.
- Triple their raw capacity, with less overall cost and more flexibility.
- Reduced cost per TB by 1/3. With the budget they were given to replace 1/3 of the existing storage, they replaced all storage by moving to the Windows File Server Cluster.
- Through the increase in raw capacity and with Data Deduplication, the team realized an actual 5x increase in capacity through more efficient space utilization and 45-75 percent deduplication rate for their symbol data. Instead of using four 60-bay JBODs full of 4TB disks for that data set, they now need only one.
- Data retention increased from 5 days to just under a month.
- The team was able to cull 8 full racks of gear, and the power and cooling necessary to drive them, through a 6x reduction in the number of servers they needed to manage.
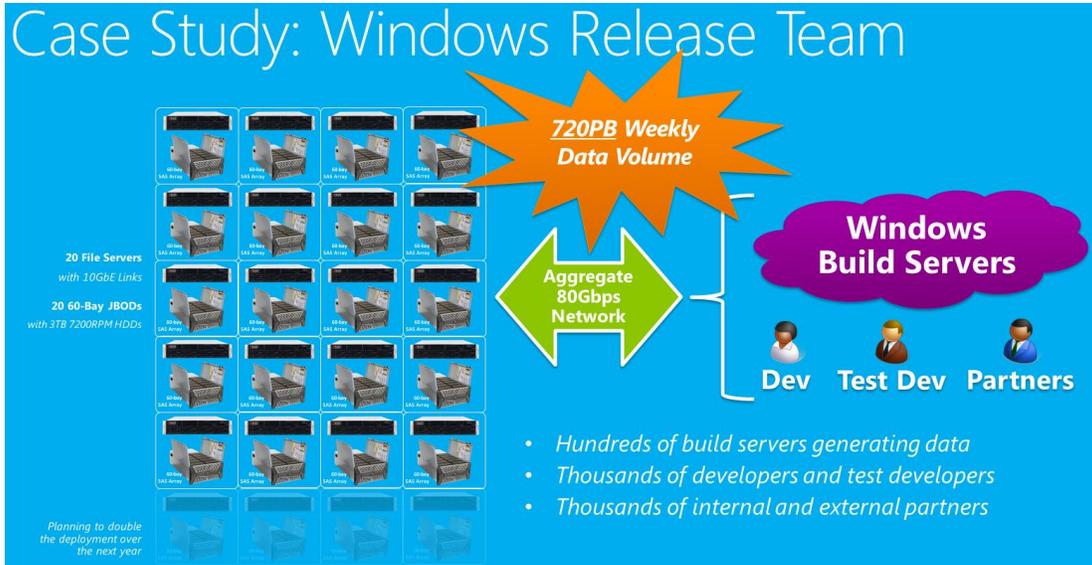
Figure 23: Windows Build Team case study results

Each week these switches see 720PBs of data passing through them. This data is comprised of new data generated by the build servers and data requested by developers, test developers, and partners. All together, the data equals almost 1 Exabyte of data, or a million TBs. Over the course of the next year, the team plans to double its storage and symbol retention, and store other types of Windows Build data.

Summary

The Windows Storage stack can handle most enterprise workloads for a fraction of the price per IOPS and per TB as traditional SAN. With shrinking budgets and increased demand for storage, organizations need to take a fresh look at Windows Server 2012 R2 as a storage solution for workloads that do not require the advanced features and capabilities of traditional SAN storage.

Expanding on-premises storage is an expensive but oftentimes necessary undertaking. A typical organization increases its data storage by approximately 25 percent annually. Only a small portion though gets frequently accessed or used. Such a reality makes traditional on-premises SAN storage solutions for cold data very expensive.

Windows Server 2012 R2 gives IT a new, less-expensive option for providing high-speed, resilient, network storage that can run on industry-standard hardware such as JBOD enclosures, SAS, and SSD disks. Management tools and features that are part of Windows Server 2012 R2, such as Storage Spaces, Automatic Tiering, SMD-Direct, and Scale-Out File Servers, combine to deliver fast, reliable SMB storage to workloads such as Hyper-V virtual machines and SQL Server, thus preserving SAN storage for the most demanding mission-critical applications.

Many companies have already begun to move toward some level of a hybrid cloud infrastructure, which provides customers with a lower TCO and a higher degree of elasticity. Growing your storage infrastructure with the Microsoft Storage Stack gives you the reliability and manageability you need, at an affordable price.