Abstract

This white paper provides an overview of the technologies that are used to perform backup and replication of EMC CLARiiON® on Windows Server 2008 Hyper-V. The paper discusses how EMC NetWorker® and EMC NetWorker Module for Microsoft Applications (NMM) are used as backup technologies and MirrorView™/Cluster Enabler (MV/CE) is used as the replication technology.

April 2010
# Table of Contents

Executive summary ........................................................................................................... 4
Introduction ...................................................................................................................... 4  
  Audience ....................................................................................................................... 4  
  Terminology ................................................................................................................... 4  
    Hyper-V terminology .................................................................................................. 4  
    Backup terminology .................................................................................................. 5  
Technology overview ................................................................................................... 5  
  Microsoft Windows Server 2008 Hyper-V .................................................................. 5  
  EMC NetWorker .......................................................................................................... 6  
    Network backup ......................................................................................................... 6  
    Off-host backup ........................................................................................................ 7  
    LAN-free backup ....................................................................................................... 7  
  Microsoft Volume Shadow Copy Service .................................................................... 8  
    Types of VSS providers ............................................................................................. 8  
  Deduplication .............................................................................................................. 9  
    Deduplication at target ........................................................................................... 9  
    Deduplication at source .......................................................................................... 9  
  EMC MirrorView/Cluster Enabler (MV/CE) ............................................................... 10  
Backup and recovery using NetWorker ...................................................................... 11  
  Hyper-V parent partition backup .............................................................................. 11  
  Child partition backup ............................................................................................... 11  
  Deduplication with NetWorker .................................................................................. 12  
    Target-based deduplication .................................................................................. 12  
    Source-based deduplication .................................................................................. 12  
Backup and recovery using NMM .............................................................................. 13  
  Quick migration ........................................................................................................ 14  
Replication ................................................................................................................... 14  
  Host clustering ........................................................................................................... 14  
  Guest clustering ......................................................................................................... 15  
Conclusion ................................................................................................................... 16  
References .................................................................................................................... 17
Executive summary
Protecting an organization’s data and IT infrastructure in the event of disruptive situations has become an increasingly visible business priority. EMC offers products and solutions to safeguard critical business data and ensure business continuity.

EMC® NetWorker® and NetWorker Module for Microsoft Applications (NMM) provide a single, unified solution that not only protects Hyper-V and Windows Server 2008, but also enables nondisruptive backup of applications such as Microsoft Exchange Server, Microsoft SQL Server, and Microsoft Office SharePoint Server running within the virtual machines.

Disaster recovery is increasingly becoming a requirement for many businesses. System administrators and enterprises are looking for ways to run their business even through natural disasters. MirrorView™/Cluster Enabler (MV/CE) provides high availability and disaster recovery.

Introduction
This white paper explains how EMC NetWorker and NMM deliver snapshot-based protection and recovery for the entire Hyper-V server environment, including parent and child partitions, by leveraging Microsoft Volume Shadow Copy Service (VSS). The persistent snapshot management of NetWorker enables fast and efficient snapshot restore and rollback for quick Hyper-V recovery.

This paper also explains how MirrorView/CE enables the cluster to automatically manage resource and applications failover by integrating with the Microsoft failover cluster to improve the recovery time objective (RTO).

Audience
This white paper is intended for EMC customers, partners, and service personnel involved in planning, architecting, or administering a Hyper-V environment with EMC CLARiiON as the storage and also for those who are planning to implement backup and replication solutions.

Terminology

MirrorView — Software designed for disaster recovery solutions by mirroring local production data to a remote disaster recovery site. It offers two complementary remote mirroring products such as MirrorView/Synchronous and MirrorView/Asynchronous.

Hyper-V terminology

Direct iSCSI disk — Direct iSCSI disk format is configured when a virtual machine leverages the iSCSI software initiator within the guest OS to connect directly to a CLARiiON storage system using the iSCSI protocol. In this topology the disk is not seen by the parent Windows 2008 server but is directly presented to the virtual machine.

Guest operating system — An operating system that runs within a virtual machine or child partition. Hyper-V currently supports Windows and Linux guest operating systems.

Partition — A basic entity that is managed by the hypervisor. It is an abstract container that consists of an isolated processor and memory resources with policies on device access.
Virtual Hard Disk (VHD) — VHD format is the common virtualization file format that captures the entire virtual machine operating system and the application stack in a single file stored on a file system in the parent partition.

Backup terminology

Backup — The duplicated copy of the original data that is stored on a variety of storage media for recovering deleted, broken, or corrupted data on the primary disk.

Deduplication — A method of reducing the storage needs by eliminating redundant data from a device.

Disaster recovery — The process of restoring a previous copy of the data and applying logs or other necessary processes to that copy to bring it to a known point of consistency.

Recovery point objective (RPO) — The point in time to which systems and data must be recovered after an outage. This defines the amount of data loss a business can endure.

Recovery time objective (RTO) — The period of time within which systems, applications, or functions must be recovered after an outage. This defines the amount of downtime that a business can endure and survive.

Snapshot — A point-in-time representation of data.

Technology overview

Microsoft Windows Server 2008 Hyper-V

Microsoft Hyper-V is the virtualization software that provides server consolidation by enabling several instances of similar and dissimilar operating systems to run as virtual machines on one physical machine. This cost-effective, highly scalable virtual machine platform offers advanced resource management capabilities.

Hyper-V minimizes the total cost of ownership (TCO) of computing infrastructure by:

- Increasing resource utilization
- Decreasing the number of servers and all associated costs
- Maximizing server manageability

When Hyper-V is enabled the Windows Server 2008 environment becomes the parent partition where the guest OS is created, device drivers are installed and updated, and the management interface is accessed. The guest OS is created in a child partition, which interacts with the parent partition by using the hypercall API through the VMBus. VMBus is a communication channel between the virtual service provider (VSP) in the parent partition and the virtual service client (VSC) in the child partition.

Before loading the parent and child partitions, the Hyper-V hypervisor is loaded directly on top of the hardware, enabling virtual machines to run above the virtualization layer provided by the hypervisor. **Figure 1** shows the architectural diagram of Microsoft Hyper-V.
NetWorker is EMC’s flagship Enterprise Backup and Recovery application. Thousands of organizations use this application to integrate data protection operations, management, and reporting into one central solution. NetWorker provides the best backup and recovery capabilities with several cutting-edge options that can help you move ahead of data growth, which is a big challenge for most organizations. It also includes options that help to scale up and understand what is happening throughout the environment as it relates to backup and recovery. It also includes advanced backup-to-disk capabilities leveraging disks to their fullest potential for backup using snapshot management and data deduplication.

NetWorker’s snapshot policy includes settings that use copies that are rolled-off or backed up to a secondary media such as tape. Backing up shadow copies of files and applications to a disk, tape or virtual tape can be done in conjunction with the NMM applications in three different ways depending on the configuration of the solution.

Network backup
Data from a shadow copy is moved from the client across the network to the backup server or storage node and to any given backup media. This snapshot-assisted backup captures all open files, provides a consistent application image, and then rolls off the data to a secondary media as shown in Figure 2 on page 7.

Figure 2 on page 7 shows the network backup method.
Off-host backup

In this method, a shadow copy is transported to a secondary client for backup. As shown in Figure 3, this happens across the LAN, but because a proxy server is taking the backup load, the production server is unaffected. This is possible only through the EMC VSS provider or a similar supported hardware provider. The Microsoft system provider does not support transportable snapshots.

Figure 3 shows the off-host backup method.

LAN-free backup

NetWorker not only offloads the backup from the production client, but also from the production LAN. In a LAN-free case, a NetWorker storage node is used to mount and back up the shadow copy to its SAN-attached tape, disk, or virtual tape as shown in Figure 4. Similar to the proxy client implementation, this is only possible in combination with the EMC VSS provider or a similar supported hardware provider.

Figure 4 shows the LAN-free backup method.
**Microsoft Volume Shadow Copy Service**

Microsoft Volume Shadow Copy Service (VSS) is a Windows technology and integration standard that coordinates various components to create consistent point-in-time copies of one or more volumes. VSS also implements a framework that enables volume backups to be performed when applications on a system continue to write to the volumes.

VSS consists of the following three components:

- **Requestor** — The application that requests the creation of a shadow copy.
- **Provider** — The interface that provides the functionality to actually create a shadow copy.
- **Writer** — The application-specific software that ensures that the application data is ready for shadow copy creation.

Requestors, providers, and writers communicate in the VSS framework to create and restore volume shadow copies. A shadow copy of a volume duplicates the data on that volume at a defined instant.

A requestor initiates the backup and restores processes. The provider controls the processes and instructs a writer to prepare a dataset for the backup.

*Figure 5* shows the VSS backup process.

![Figure 5: VSS backup process](image)

**Types of VSS providers**

VSS is used as a Windows integration standard that is widely supported by storage vendors today. However, the compliance to VSS as a hardware provider does not imply the use of Microsoft-authored VSS writers. The three types of VSS providers are hardware provider, software provider, and system provider.

A hardware provider acts as an interface between VSS and a hardware storage adapter or controller. The creation of the shadow copy is performed by the hardware outside the operating system. While a Hardware
VSS provider complies with the VSS standard that defines the point-in-time copy; VSS does not specify the underlying mechanism by which the hardware provider produces and maintains the shadow copies.

A software provider typically intercepts and processes I/O requests in a software layer. Similar to hardware providers, VSS places no restrictions on the technique software providers use to create and maintain shadow copies. However, the most commonly deployed software providers are Windows Server application and File System VSS providers that use the Microsoft-authored Windows VSS writers.

A system provider is supplied as a default part of the Windows Server 2003 and Windows Server 2008 operating systems. The system VSS provider works with New Technology File System (NTFS) volumes. Hence, it is also often referred to as the Windows OS or file system VSS provider, in contrast to the Microsoft Application VSS providers.

**Deduplication**

The new requirements in both business and regulatory require data to be retained for a longer period of time. This results in data growth. Not only does this impact the organization’s physical ability to back up, but it impacts the cost as well. Data center operations, which are 24x7, cannot withstand any downtime during backups. There are also bandwidth limitations with the increased adoption of server virtualization. Deduplication saves the storage used for backups, and thus reduces the hardware requirements.

Deduplication can be accomplished through a number of different deployment alternatives. Backup deduplication can occur at the backup target or at the data source and EMC NetWorker supports both of them.

**Deduplication at target**

In this method, the backup application sends the native data to a target storage device and the data is deduplicated at the device during or after the backup as shown in Figure 6. This is ideal for customers who are satisfied with the backup software without experiencing any bottlenecks in transferring the data to the backup storage device.

*Figure 6* shows the deduplication at the target.

**Deduplication at source**

In this method, the data is deduplicated when the backup process begins and before the data is sent over the network to storage. Hence, it provides the benefits of shorter backup windows and lesser bandwidth requirements. This makes it ideal for remote or WAN-based backup, large file servers, and environments where the backup process is hampered by network or other resource bottlenecks.
Figure 7 shows the deduplication at the source.

**EMC MirrorView/Cluster Enabler (MV/CE)**

EMC MirrorView/CE (MV/CE) is host-based software that integrates with Microsoft Failover Clusters to manage cluster resource failover between storage systems. It enables geographically dispersed Microsoft Failover Clusters across CLARiiON MirrorView/Synchronous and MirrorView/Asynchronous links as shown in Figure 8. Geographically dispersed clusters offer increased levels of high availability, disaster recovery, and automation over non-clustered solutions. MV/CE works seamlessly with applications such as Exchange, SQL Server, and Hyper-V by taking the advantage of failover clusters.

Figure 8 shows the MirrorView/CE cluster configuration.
Backup and recovery using NetWorker

The factors that drive IT managers to consider new approaches to data protection are exponential data growth, regulations, aggressive service level agreements, and optimization of backup windows. EMC NetWorker enterprise backup and recovery software centralizes, automates, and accelerates data protection across heterogeneous IT environments. To protect Hyper-V, NetWorker leverages Microsoft VSS to deliver snapshot-based protection and recovery for the entire Hyper-V service environment, including parent and child partitions (guests).

**Hyper-V parent partition backup**

NetWorker provides a full range of backup and recovery services for Windows Server 2008 Hyper-V. NetWorker ensures complete recovery including the protection of operating system components such as system state, data volumes, and the Hyper-V configuration database.

From the parent partition or virtual server host, NetWorker leverages the VSS writer to capture copies of the Hyper-V child partitions from the parent partition. NetWorker supports all Hyper-V Guest operating systems that may be hosted on the virtual machines. This protection benefits from full NetWorker integration including policies, advanced backup operations, snapshot management workflows, and directed or relocated recoveries. The backup at the Hyper-V parent partition provides the benefit of straightforward protection because it performs the backup of Windows and executes full image protection to ensure complete disaster recovery protection. Off-host backup eliminates performance impacts on key virtual servers and removes backup traffic from the network.

Disks that are direct-attached, Fibre Channel or iSCSI attached to the parent can be protected with the VSS software provider. The EMC VSS hardware provider also supports Fibre Channel and iSCSI on CLARiiON.

*Figure 9* shows how the Windows Server 2008 Hyper-V is backed up using the off-host backup method.

**Child partition backup**

NetWorker software can be used in child partitions to deliver application-aware backup. For Microsoft Applications, application-consistent backup data can be captured by leveraging application-specific VSS writers. This solution provides options such as traditional, off-host, and LAN-free backup similar to VMware VCB backups. It can keep persistent snapshots for Quick Recovery and the recovery is available...
from secondary media and snapshot copies or even from snapshot rollbacks. The benefits of this type of backup are application-specific backups and off-host backups. Application-specific backups provide finer-grain recoveries with application log management and off-host backup eliminates performance impacts on Hyper-V server.

At the child partition level, NetWorker can support direct-attached, passthrough, Fibre Channel, or iSCSI-attached disks using the Windows VSS software provider. CLARiiON at the child supports only the iSCSI storage connect method with the VSS hardware provider due to a current Microsoft limitation.

**Deduplication with NetWorker**

EMC offers backup deduplication with NetWorker at the backup target and at the data source. The apt deduplication technology and strategy depends on several factors, including the use case, service level requirements, and the environment.

**Target-based deduplication**

NetWorker is integrated with Data Domain for target deduplication. The backup application sends data to the target storage device and data is deduplicated at the device. This type of deduplication is found in virtual tape libraries (VTLs) and LAN backup-to-disk appliances or platforms. This approach provides the benefit of being plug-and-play with existing backup applications and infrastructure.

**Source-based deduplication**

NetWorker is integrated with Avamar® for source-based deduplication. As Integrated Client is a standard feature of NetWorker, it is possible to deploy NetWorker and Avamar side-by-side or separately as standalone products. NetWorker represents the Avamar data store as a deduplication node. Deduplication is enabled through the NetWorker management console. The save process gathers metadata and sends it to the NetWorker server and data is sent to the deduplication node. Recovery is accomplished from the NetWorker recovery interface (GUI or CLI). After the metadata associated with the data that needs to be recovered is identified, the required deduplicated data is restored from the deduplication node.

Source-based deduplication is the optimal approach to Hyper-V infrastructure backups.

**Optimal location**

The deduplication process at the source backup can quickly and efficiently protect virtual machines by sending only the changed segments of data on a daily basis. This provides considerable reduction (almost 500 times) in daily reduction of network resource consumption when compared to traditional full backups. It also reduces backup times dramatically. Deduplication can be performed at the guest (individual virtual machine) or at the parent partition.

**Optimal granularity**

A VHD is one large file and any changes in the file cause traditional backup to retransmit the entire file. Changes in VHD that deduplication finds are variable length segments called sub-files. Deduplication sends only the changes. Deduplication at the sub-file is the optimal granularity to find changes anywhere within a virtual hard disk file. The deduplicated backups are all stored as “virtual full backups.” This means that data is ready to restore quickly. Traditional backup, in many cases, takes time to reassemble full and incremental backups to deliver the data for recovery. **Figure 10 on page 13** shows an example of how Hyper-V is deduplicated based on the source.
For example, consider multiple virtual machines and hosts.

When backup is addressed for the first virtual machine, the NetWorker source-based deduplication breaks the virtual machine data into sub-file variable-length data segments such as A, B, C, and D. This is the new data that has never been backed up. As each data segment is new, all segments are backed up.

There is a duplicate instance of the virtual machine. In this case, the deduplication technology keeps track of the segments such as A, B, C, and D that are already backed up. As a result, only a unique ID pointer is stored and it is a considerably small file.

A new virtual machine is created using the same template. During the next backup operation, only the new unique segments associated with the changes to the virtual machine are backed up. In this case, because data segment E is new, only this data segment is transferred during the backup. Deduplicating data at the sub-file level dramatically reduces the amount of storage required to retain backup data for extended periods of time.

**Backup and recovery using NMM**

NMM provides full backup and recovery of Windows Server 2008 with Hyper-V. It supports the backup of the Hyper-V server, child partitions, and applications within each child partition and full-level backup. NMM uses the Hyper-V VSS writer to back up and recover the following:

- Hyper-V initial store configuration file
- Each virtual machine or child partition

When the files are backed up, NMM can recover all the Hyper-V components, the initial store configuration file, or individual virtual machines depending on specifications provided in the backup.

Hyper-v can be recovered in the following ways:

- Offline and online recovery
- Hyper-V recovery to the original machine and location
- Hyper-V recovery to a different machine or location
Quick migration

Quick migration is the ability to rapidly migrate a running virtual machine from one physical host system to another with minimal downtime, taking advantage of the familiar high-availability capabilities of Windows Server and Microsoft System Center management tools. For a planned migration, quick migration saves the state of a running guest virtual machine (memory of original server to disk or shared storage), moves the storage connectivity from one physical server to another, and then restores the guest virtual machine onto the second server (disk/shared storage to memory on the new server). Using Quick Migration multiple virtual machines can be moved simultaneously. Pass-through disks in a cluster provide enhanced I/O performance and requires virtual machine configuration file to be stored separately from the virtual machine file. For VHD disks one LUN per virtual machine is the best practice.

For further information on quick migration, refer to the *Quick Migration with Hyper-V -White Paper* available on the Microsoft website.

Replication

Replication is the process of creating copies either at the file or volume level of data through a variety of methods including cloning, mirroring, and snapshot capabilities. In addition, synchronization and movement of data through files, directories, volumes, or entire file systems come under this Replication segment. MirrorView/CE supports Windows Server 2008 (x64) Hyper-V virtualization for CLARiiON arrays employing either host or guest clustering.

Host clustering

Host clustering allows for hosting and failover of the virtual machines by making the physical machine highly available. In this common method of deployment, services and applications can run in multiple individual virtual machines that are highly available. If a virtual machine becomes unavailable, only that single service in the virtual machine fails over (assuming each virtual machine has its own physical disk).

*Figure 11 on page 15* illustrates Hyper-V host cluster configuration with the Windows Native Cluster service running in the Windows Server 2008 x64 parent partition.
Guest clustering enables highly available services and applications in a virtual layer. In this deployment method, failover clustering is installed on several virtual machines and then clustered as physical nodes. Microsoft Windows Server 2008 failover cluster or Microsoft Windows Server 2003 Microsoft Cluster Server software and MirrorView/CE software are installed and configured on the guest virtual machines. MirrorView/CE only supports direct iSCSI disks for guest clustering as shown in Figure 12. This means that the guest virtual machines must be running the Microsoft iSCSI Software Initiator on each guest operating system, which is connected to the CLARiiON storage array through the iSCSI protocol. Each virtual machine must be registered to the CLARiiON storage subsystem to be configured as cluster nodes in the Failover Cluster wizard. The EMC Cluster Enabler Manager displays the virtual machines as cluster nodes. In this configuration, the virtual machines on the Hyper-V server act as cluster nodes, thus the failover and failback processes are similar to those of a regular Windows server.

Figure 12 on page 16 shows the guest clustering in Windows Server 2008 Hyper-V.
**Conclusion**

EMC NetWorker and NMM provide backup solutions for Windows Server 2008 Hyper-V on EMC CLARiiON storage. These technologies simplify backup operations for Hyper-V to meet both operational and disaster recovery in one solution. VSS integration provides the opportunity to remove the backup from the production servers. It increases the granularity of backup and recovery operations and includes the ability to schedule regular backup operations. Snapshots create backup copies quickly and easily. Proxy-based backup helps to move the copies to secondary media.

The recovery operations are simplified as the recovery access can be given in the common user interface along with other Microsoft applications supported by NetWorker. NetWorker allows you to leverage more than just tape recovery. Recovery from snapshots and rollback give additional options to help improve RTOs and RPOs.

NetWorker reduces the storage space and time for daily backups by using deduplication technology. NetWorker is an ideal solution because it supports both traditional and deduplicated backup methods as shown in Figure 14 on page 17.

Figure 14 on page 17 shows the summary of NetWorker support for Hyper-V.
Microsoft host clustering coupled with MirrorView/CE provides Hyper-V virtual machines protection against software and hardware failures between two or more geographical sites.

References
The following documents and resources are available on Powerlink:

- *EMC MirrorView/Cluster Enabler (MV/CE) — A Detailed Review*
- *Backup Acceleration with EMC NetWorker and Snapshot Management — A Detailed Review*
- *EMC NetWorker: Complete Protection for Microsoft Hyper-V — Solution Overview*

For more information on Microsoft Server 2008 Hyper-V, refer to the following Microsoft website.