Technical Report

SnapManager for SQL Server on Data ONTAP 8.1 Operating in Cluster-Mode

Abhishek Basu, NetApp

ABSTRACT

This technical report describes guidelines and best practices for NetApp® SnapManager® enabled Microsoft® SQL Server® on Data ONTAP® 8.1 operating in Cluster-Mode. The NetApp technologies discussed in this report are vital to achieving an integrated storage solution that is cost effective, operationally efficient, flexible, and environmentally friendly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1   INTRODUCTION

This guide covers best practices to set up a NetApp SnapManager enabled Microsoft SQL Server on Data ONTAP 8.1 operating in Cluster-Mode. Cluster-Mode allows intercluster asynchronous volume replication (replication between volumes hosted on different clusters) for the purpose of enabling disaster recovery and replication to a cluster in a remote site.

Figure 1 explains the architecture of Data ONTAP 8.1 operating in Cluster-Mode.

Figure 1) Cluster-Mode architecture.
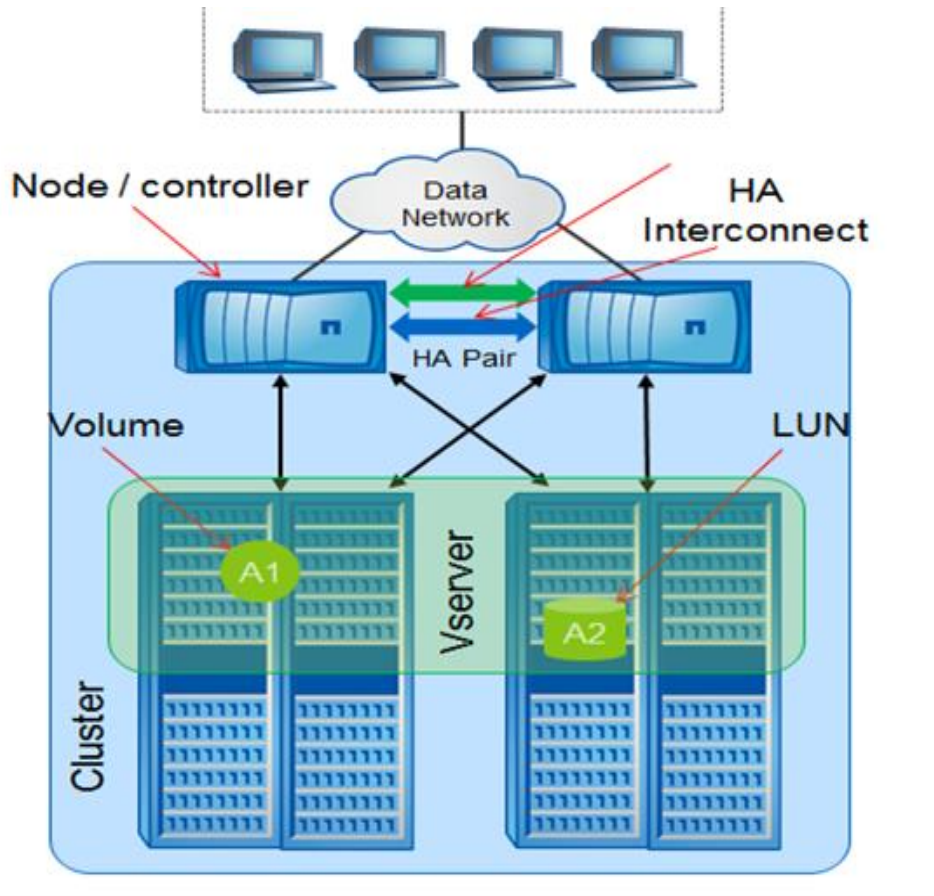
2   BENEFITS OF CLUSTER-MODE

Key benefits of Data ONTAP 8.1 Cluster-Mode include:

- Scalability
- High availability
- High performance
- Flexibility
- Multi-tenancy

2.1 SCALABILITY

Data ONTAP 8.1 operating in Cluster-Mode has scalability provisions for a single system image for two to four nodes. Its throughput scales linearly to multiple GB/sec, and it has linear scale performance for a single volume with striping. There is linearly scaling read performance with load sharing mirrors and collective read/write performance in a single namespace. For these reasons Cluster-Mode:

- Helps in rapid and seamless deployment of new storage
- Is built for continuous scale-out operation with no downtime
- Scales out transparently to compute farm with unchanged namespace

Figure 2) Cluster-Mode scalability.
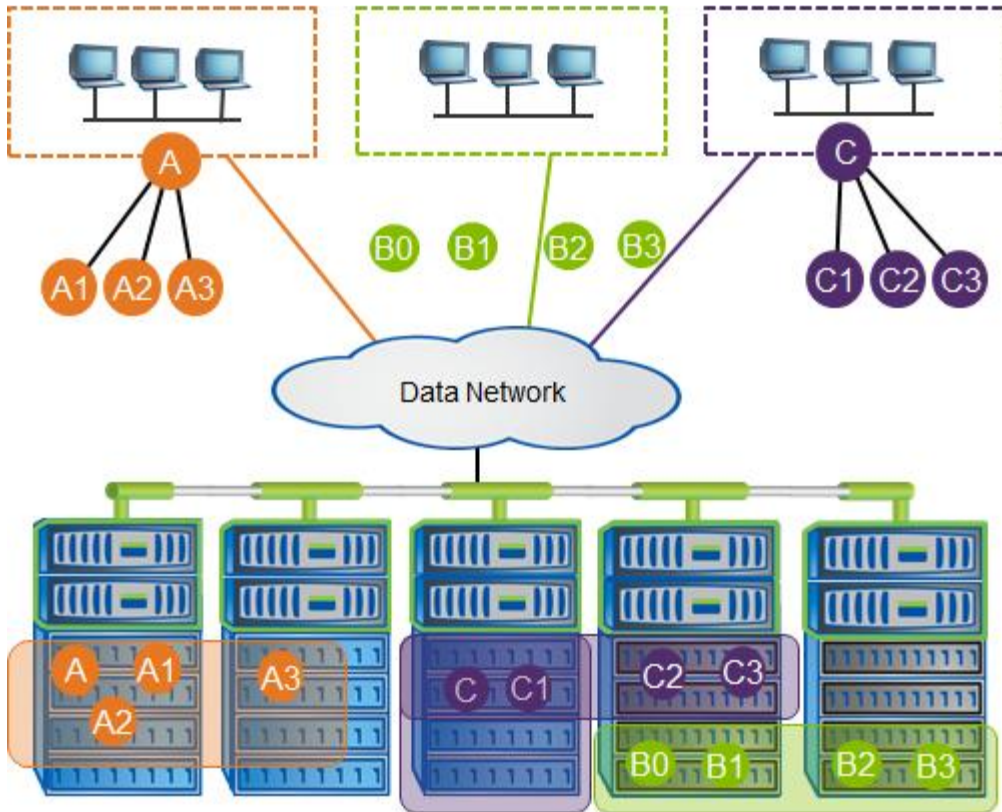
2.2 HIGH AVAILABILITY

Data ONTAP 8.1 operating in Cluster-Mode is architected with the high-availability requirements of today's businesses in mind. Its recovery capability is provided by a pair of nodes (storage systems), called an HA pair, that are configured to serve data for each other if either node stops functioning. Operations like volume movement, node failover, switch failure, and site recovery can be performed without any disruption to the storage and applications. IT organizations can nondisruptively meet the needs of SQL Server data growth and increased or changing SQL Server application workload by adding more controllers or storage without disrupting current applications. These organizations also enjoy increased SQL Server application uptime and failover protection during infrastructure (hardware and software) maintenance and upgrades by means of a highly available storage back end.

2.3 HIGH PERFORMANCE

Data ONTAP 8.1 operating in Cluster-Mode also provides high performance. It can scale up linearly as the number of nodes increases and can deliver increased throughput as well.

Table 1) High-performance features.

Component	Features
Sessions from the host to the storage subsystem	Windows MPIO/Data ONTAP DSM 3.5
Host Ethernet port	NIC teaming (supported systems only)
Server	Windows failover cluster manager
Ethernet	2 Ethernet switches
FC	2 Fabric switches
Storage systems	2-node Data ONTAP Cluster-Mode system (minimum)
LIFs in the storage systems	Multiple ports mapped to the same LIF

2.4 FLEXIBILITY

Cluster-Mode offers a high degree of flexibility for storage operations because it can:

- Accommodate any mix of FC and SATA drives
- Allow nondisruptive data movement between tiers
- Provide transparent access to volumes on any node from any node
- Connect many volumes into a single namespace
- Move volumes transparently between nodes

2.5 MULTI-TENANCY

Cluster-Mode provides multi-tenancy, as illustrated in Figure 3. It is important for different customers (in the case of public clouds) and different business units (in the case of private clouds) to be securely isolated at the compute, network, and storage layers. Cluster-Mode facilitates multi-tenancy at the storage layer by segregating storage entities such as aggregates, LIFs, LUNs, and volumes and containing them in a Vserver. Because each Vserver operates in its own namespace, each unit or customer mapped to this Vserver is completely isolated. Each Vserver supports role-based access control (RBAC), and specific protocols like NFS, CIFS, iSCSI, FC, and FCoE can be assigned to it.

Figure 3) Multi-tenancy.
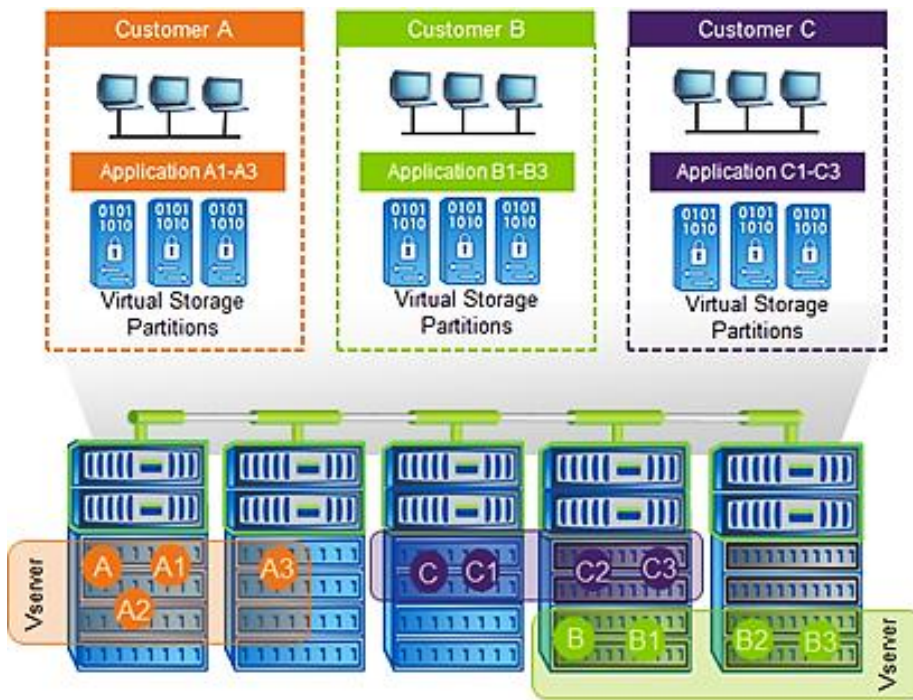
3 CLUSTER TENETS

Key tenets of Cluster-Mode clusters include:

- Nonstop, nondisruptive operation of clusters and their Vservers
- Vservers are the basis for secure multi-tenancy
- Data location and container transparency
- Single system and delegated management
- Unified storage and multiprotocol support

3.1 CLUSTER

Cluster-Mode nodes are configured in fault-tolerant pairs. This enables continuous service to end users if individual nodes go offline or are taken offline. The following characteristics define the Cluster-Mode cluster:

- A collection of colocated nodes
- Domain of nondisruptive volume move
- Home to one or more Vservers
- Boundary of volumes
- HA is contained within the cluster except for
    – NetApp MetroCluster™ HA
    – Attached storage aggregates
    – Private cluster networks
- Single management, shared configuration

- May be tightly or loosely coupled within qualified bounds—for example R4, a collection of 6080s, or a collection of R4s

3.2 VSERVER

A Vserver is a secure, virtualized storage server that has its own administration, IP addresses, exports, storage objects, and namespace, independent of physical location. A Vserver can reside on any node in the cluster that provides a physical resource and has the following attributes:

- Multiprotocol, unified storage
- The scale-out equivalent of a 7G controller or vFiler® unit
- The highest level user-visible entity
- Assigned to a single tenant
  - Administration can be delegated
- Hosted on just one cluster
  - Can aggregate or partition cluster resources

3.3 NODE

A node:

- Is one instance of Data ONTAP
- Runs natively on a NetApp controller
- Acts as a virtual storage array on NetApp or other controller

4 SIMILARITIES BETWEEN 7G AND CLUSTER-MODE

Similarities between Data ONTAP 7G and Data ONTAP 8.1 operating in Cluster-Mode include:

- Controllers and disk shelves
- RAID 4 and RAID-DP®
- (WAFL®) Write Anywhere File Layout
- Snapshot™ copies
- Aggregates
- Flexible volumes
- Active-active pairing for failover
- Asynchronous mirroring

5 DIFFERENCES BETWEEN 7G AND CLUSTER-MODE

Table 2 lists the differences between Data ONTAP 7G and Data ONTAP 8.1 operating in Cluster-Mode:

Table 2) Data ONTAP 7G versus Cluster-Mode.

Data ONTAP 7G	Data ONTAP 8.1 Operating in Cluster-Mode
Cluster = 2 nodes	Cluster = 2 to 4 storage area network (SAN) nodes
CFO (cluster failover)	SFO (storage failover)
VIF = link aggregation	LIF = Logical interface

Data ONTAP 7G	Data ONTAP 8.1 Operating in Cluster-Mode
Volumes = traditional, flex	Volumes = flex, striped
Root volume = vol0	Root volume = mroot, vsroot
vFiler units	Vserver
Flat command-line interface (CLI) structure	Hierarchical CLI command structure
Generally uses Remote Shell (rsh) to connect to CLI	Generally uses Secure Shell (SSH) to connect to CLI

6   LIMITATIONS OF CLUSTER-MODE AND RECOMMENDATIONS

6.1   CURRENT LIMITATIONS OF CLUSTER-MODE

- Large environments need multiple clusters.
- Multisite: Clusters don't stretch beyond the data center.
- Limits to the supported scale of clusters.
- Cluster-Mode does not support IPv6.
- Cluster-Mode based storage systems do not support NetApp SnapVault® software because of the SnapDrive® for Windows® (SDW) limitations.
- File-based thin provisioning in Cluster-Mode is not supported.
- Currently no support exists for VMDK on NFS and VMFS datastores residing on Cluster-Mode systems.
- Limits to hardware and software heterogeneity allowed in a cluster:
    - Customer-imposed limits
    - Organizational and administrative boundaries

6.2   CLUSTER-MODE BEST PRACTICES

- Users must configure user credentials for all cluster server and Vserver management endpoints from the storage system that is to be accessed by SDW and SMSQL.
- For a highly available connection to the storage system, NetApp requires installing the supported version of multipathing software such as the Data ONTAP DSM for Windows MPIO.
- ALUA support is available on all Cluster-Mode configurations with Data ONTAP DSM as well as with MSDSM.
- Performing LUN management tasks requires at least one iSCSI LIF and one management LIF per Vserver.
- NetApp recommends adding the Data ONTAP Cluster IP and credentials (cluster server credentials and Vserver credentials) along with the management LIF in SnapDrive > Transport Protocols Settings > Storage Systems.
- NetApp recommends having the Vserver and cluster IP mapped in DNS or explicitly specified in the Windows Host file in `C:\WINDOWS\system32\drivers\etc`.
- The cutover window defined for a SAN volume should not exceed the expected timeout value on the host side. During the cutover phase of the volume move, all I/O access is queued and requests are blocked to the source volume. SnapDrive sets a timeout value of 120 seconds on the host during the volume move. Also, when a SAN volume is moved, ALUA is used for optimized access to the volume.
- Each node must have a data LIF for optimized access to the volume.

- The FlexClone® license must be enabled on the storage system to enable Snapshot copies. No separate license to enable SIS cloning exists on Cluster-Mode systems.
- Snapshot backup scheduling can be performed only if the cluster server credentials are provided, because it is the responsibility of the cluster server administrator to manage space allocation on the Vservers.
- For Data ONTAP Cluster-Mode systems, guest virtual machines residing on ESX hosts are supported. However, this support is restricted to RDM LUNs only. VMDK on NFS and VMFS datastores residing on Cluster-Mode systems are not currently supported.
- Storing the RBAC configuration file in the storage system root volume of the Vserver is not supported.
- To get ASUP™ alerts for Data ONTAP 8.1 operating in Cluster-Mode, cluster server credentials must be provided in the SDW storage settings.
- Users cannot establish a SnapMirror® relationship between a 7-Mode source volume and a Cluster-Mode destination volume.
- Intercluster replication requires at least one intercluster logical interface (intercluster LIF) per node. The intercluster LIF can be assigned to a data port or to a dedicated intercluster port.
- Configure the Windows firewall to allow SnapDrive services for Windows communications.

7 STEP-BY-STEP SETUP WORKFLOW FOR CLUSTER-MODE

Here are the sequential steps (commands) to set up a Cluster-Mode environment in Data ONTAP 8.1:

1. `cluster create/join`
2. `aggr create`
3. `vserver create`
4. `volume create`
5. `vserver iscsi create`
6. `vserver fcp create`
7. `network interface create –data-protocol iscsi`
8. `network interface create –data-protocol fcp`
9. `lun create`
10. `igroup create; [portset create]`
11. `lun map`
12. `setup iscsi initiator`
13. `setup fcp initiator`
14. `config SAN Fabric Zones (FCP) / VLANs (iSCSI)`

Figure 4) Cluster-Mode details.

```
APPCL::> cluster show
Node                    Health  Eligibility
--------------------    ------- ------------
APPCL-01                true    true
APPCL-02                false   true
2 entries were displayed.

APPCL::> vserver show
                        Admin    Root                          Name     Name
Vserver      Type       State    Volume      Aggregate         Service  Mapping
-----------  ---------  -------- ----------- -----------       -------  -------
APPCL        admin      -        -           -                 -        -
APPCL-01     node       -        -           -                 -        -
APPCL-02     node       -        -           -                 -        -
excserver    cluster    running  root_vol    excaggr           file     file
infraserver  cluster    running  infraroot   infraggr          file     ldap
sdw          cluster    running  sdw_root    infraggr          file     file
shpserver    cluster    running  shproot     shpaggr           file     file
sqlserver    cluster    running  sqlroot     sqlaggr           ldap     file
vmshp        cluster    running  vmshp_root  vmshpaggr1        ldap     ldap
vmsql        cluster    running  vmsql_root  vmsqlaggr1       ldap     ldap
10 entries were displayed.


APPCL::>
APPCL::> aggr show
Aggregate        Size  Available Used% State     #Vols Nodes              RAID Status
---------  --------- ---------- ----- -------   ------ ------------------ ------------
aggr0       119.5GB    5.57GB   95%  online        1 APPCL-01           raid_dp,
                                                                         normal
aggr0_APPCL_02_0 -         -       - unknown       - APPCL-01           -
excaggr          -         -       - unknown       - APPCL-01           -
infraggr         -         -       - unknown       - APPCL-01           -
shpaggr          -         -       - unknown       - APPCL-01           -
sqlaggr          -         -       - unknown       - APPCL-01           -
vmshpaggr1       -         -       - unknown       - APPCL-01           -

Info: Node APPCL-02 that hosts aggregate vmsqlaggr1 is offline
vmsqlaggr1       -         -      - -              - APPCL-02           -
8 entries were displayed.

APPCL::> vol show
  (volume show)
Vserver   Volume       Aggregate      State    Type      Size  Available Used%
--------- ------------ -------------- -------- ---- ---------- ---------- -----
APPCL-01  vol0         aggr0          online   RW     113.3GB   88.79GB   21%
APPCL-02  vol0         aggr0_APPCL_02_0
                                      -        -          -         -      -
```

8 SQL SERVER AND SMSQL LAYOUT ON CLUSTER-MODE

Here are the prerequisites for deploying SQL Server and SMSQL on Cluster-Mode:

1. Install Data ONTAP DSM 3.5.
2. Install SnapDrive 6.4 for Windows.
3. Enter the Vserver name and data LIFs in the etc hosts file.
4. In SnapDrive 6.4 for Windows, add the management LIF that you created while creating the Vserver in "Add Storage Systems" under the Transport Protocol option. All the data is managed through the management LIFs belonging to this Vserver in SnapDrive.
5. For managing Cluster-Mode LUNs, select http or https as the transport protocol. The RPC protocol is not supported.
6. In SnapDrive 6.4 for Windows, create the LUN in the selected Vserver.

7. To establish a SnapMirror relationship, add the cluster server credentials of the source and destination volumes in the primary host.

For SQL Server and SMSQL layout on Cluster-Mode, NetApp recommends the following:

1. There should be a dedicated Vserver for each individual physical host server.
2. There should be an aggregate for each instance of a virtual machine.
3. There should be a dedicated volume for each instance of a SQL Server.
4. Respective DB LUNs for the instance are to be created in the dedicated volume for the instance. You can also use separate volumes according to business requirements.
5. SQL Server can be installed on either the local C: drive or in a LUN provisioned from SnapDrive.
6. SnapManager for SQL can be installed on either the local drive or in a LUN provisioned from SnapDrive.
7. The SnapManager components must be installed on the host machines or VMs where the SQL Server instances are to be managed.
8. The user databases should be created directly on the NetApp-provisioned LUNs.

Figure 5 explains SQL Server file layout in a Cluster-Mode environment.

Figure 5) SQL Server layout.
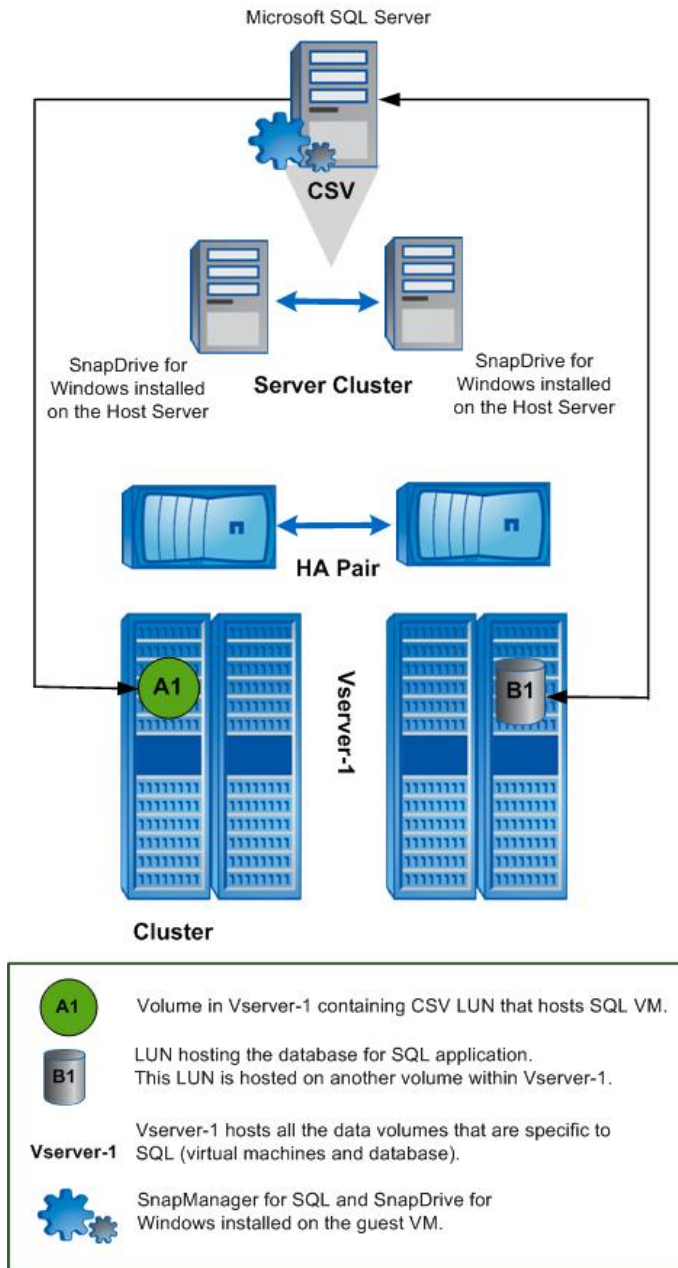
9 SMSQL 5.2 OPERATIONS GUIDANCE AND NEW FEATURES

Snap Manager for SQL Server 5.2 is designed to support Data ONTAP 8.1 operating in Cluster-Mode. It offers enterprises stability, high performance, storage efficiency, functionality, and interoperability. Previous versions of Snap Manager for SQL Server do not support Cluster-Mode. All features from the previous releases of SMSQL are part of this release. This release will continue to support management of 7-Mode storage systems. There is also a dependency on SnapDrive for Windows 6.4. Previous versions may not be optimal for use with Cluster-Mode.

The snapmanager_suite license must be provided on all nodes of the Cluster-Mode system.

Table 3 specifies the minimum software requirements for Cluster-Mode support.

Table 3) Minimum software requirements for Data ONTAP 8.1 operating in Cluster-Mode.

| Solution Component	Minimum Revision
Primary storage	
Data ONTAP	8.1
NFS, ASIS, FlexClone, SnapMirror, SnapRestore®, and NearStore® licenses	N/A
Backup storage	
Data ONTAP	8.1
NFS, ASIS, SnapMirror, and NearStore licenses	N/A
NetApp management software	
NetApp SnapManager for SQL Server	5.2
NetApp SnapDrive	6.4
Microsoft SQL Server 2008 R2, Microsoft SQL Server 2008, Microsoft SQL Server 2005 Enterprise Edition 64-bit version	
Windows Server® 2008	x64, Enterprise Edition, R2

9.1   NEW FEATURES OF SNAPMANAGER FOR SQL SERVER 5.2

This section describes the new features of SnapManager for SQL Server 5.2.

FEDERATED BACKUP

Federated backup in SMSQL 5.2 offers a new federation feature so that users can specify databases not only from different SQL Server instances on the same host, but also from remote hosts on different or on the same storage controllers. Users can connect to the remote servers, select databases, add them to the backup list, and make Snapshot copies on remote servers as well as local backups at the same time.

Figure 6) Federated backup options.
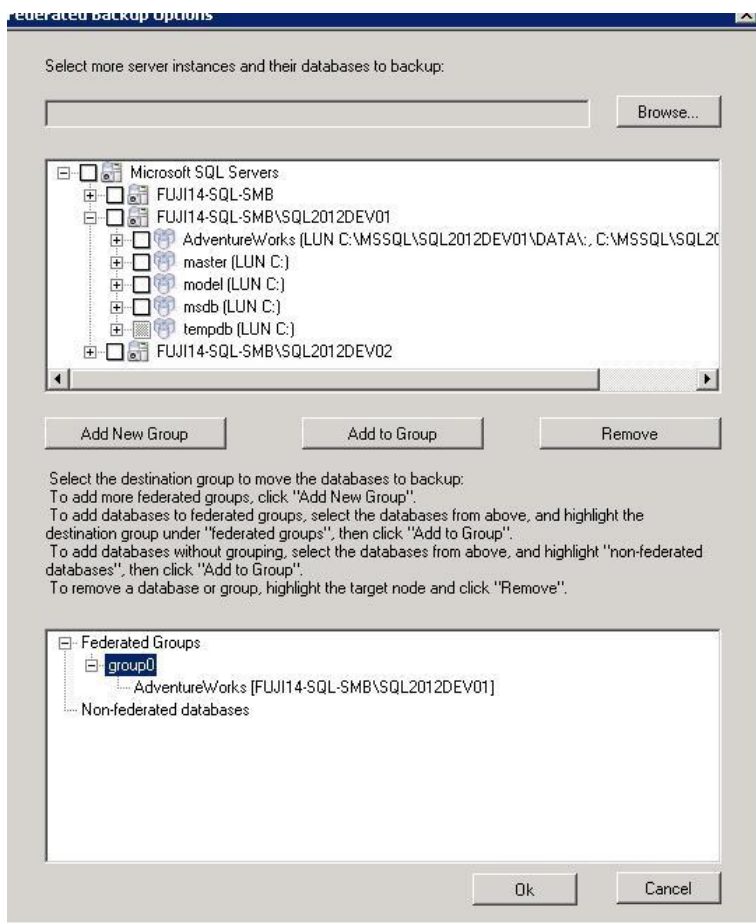
FEDERATED LOG RESTORE AND RESTORE LOG TO MARK

"Restore to marked transaction" operations allow you to restore a database to a marked transaction. Using the marks created during a federated full backup, you can restore a backup to a marked transaction across multiple databases for a synchronous restoration.

**Note:** You can restore either to a marked transaction or to a point in time.

Figure 7) Federated log restore and restore log to mark.
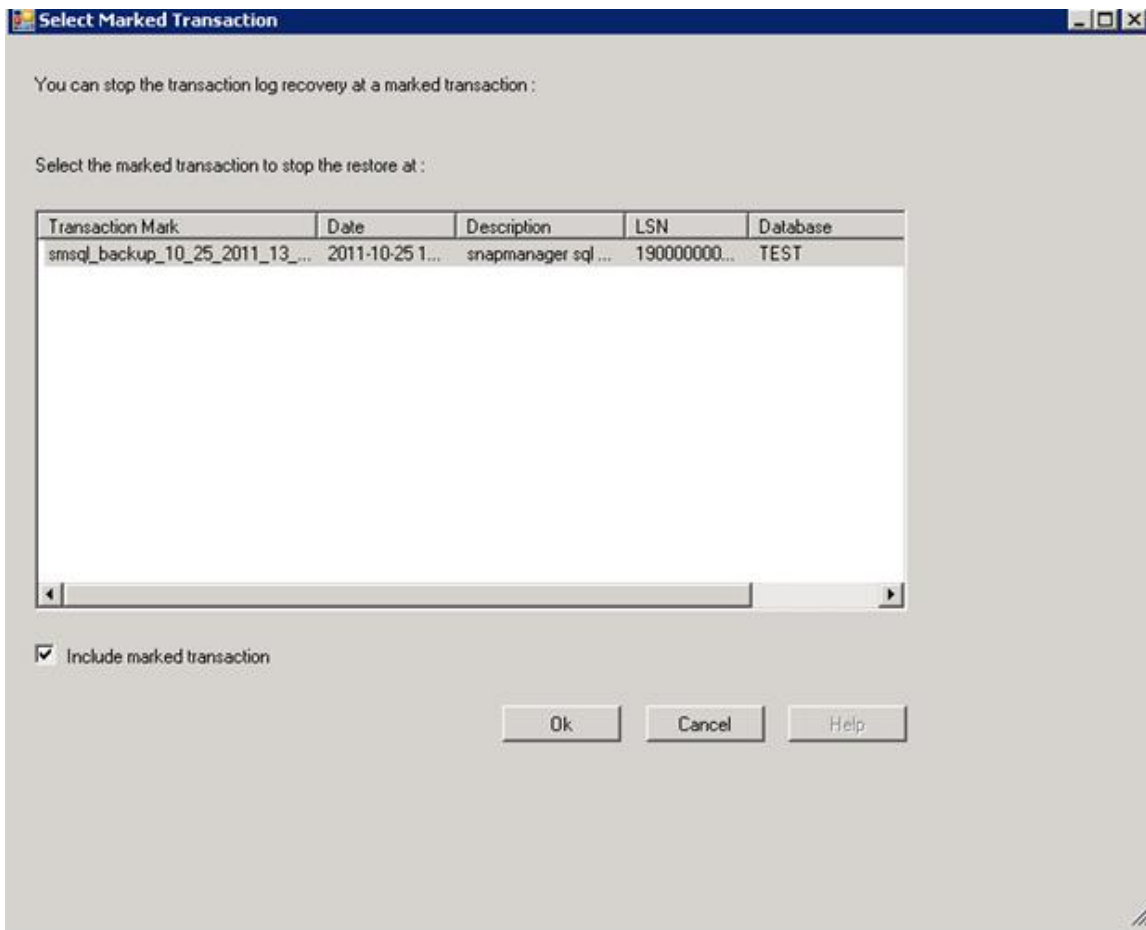
BACKUP RETENTION MANAGEMENT

To efficiently manage the retention of transaction log backups, it is necessary to limit the number of old transaction logs that are saved. When the new backup runs, it marks the previous backups as being point-in-time only and then deletes the logs as well as the Snapshot copies to reduce log traffic.

Figure 8) Backup retention management.
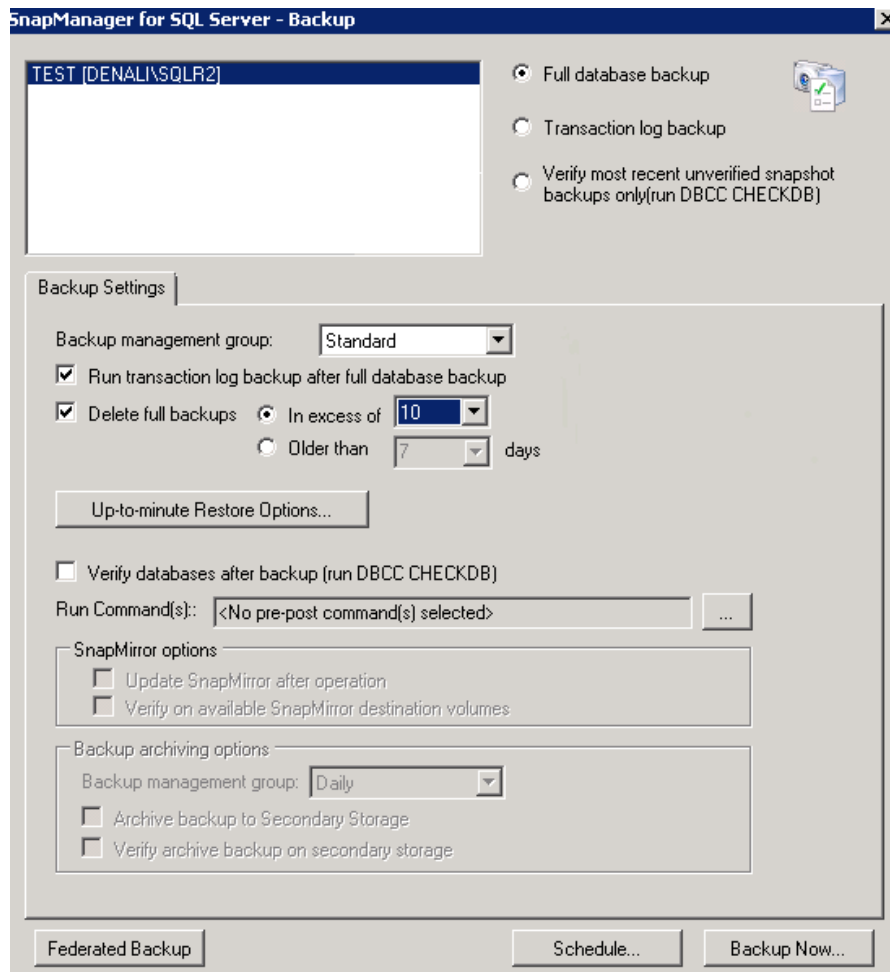
PERFORMANCE IMPROVEMENT

The following key improvements have been made to SMSQL in this release:

- SMSQL Snapshot copy cache
- On-demand lazy load enumeration of SMSQL backups
- Grouping databases into larger Snapshot groups
- Batching I/O into larger transfer requests
- SnapDrive cache for mount and dismount operations

SMSQL SNAPSHOT CACHE PERFORMANCE

When enumerating Snapshot copies for each LUN, SMSQL makes a relatively expensive call to SnapDrive and ultimately to the NetApp storage appliance. This feature caches the results of previous queries to the NetApp storage for the existing Snapshot copies. The cache does not persist for the lifetime of the SMSQL service.

The operations that use cache are:

- Delete phase of backup

- Restore enumeration of existing backups
- Delete individual dialog load
- Cmdlet enumeration of existing backups

Use this registry key to enable or disable the SMSQL cache on the SMSQL server:

```
 Key: HKEY_LOCAL_MACHINE\SOFTWARE\Network Appliance\SnapManager for SQL Server\Server
ValueName: EnableEnumSnapshotCache
REG_DWORD:
0 = disable cache
1 = enable SMSQL lun snapshot cache (default
```

ON-DEMAND LAZY LOAD RESTORE

This feature helps to enumerate backups for restore on an as-needed basis. It preemptively stages the enumeration so that the entire set of backups is not loaded at once.

Use the following registry key to enable or disable on-demand (deferred or lazy) restore in the GUI:

```
Key: HKLM\SOFTWARE\Network Appliance\SnapManager for SQL Server\Client
ValueName: EnableOnDemandRestoreTree
REG_DWORD:
0 = disable and load all backups all the time (sme 5.0-6.0 behavior),
1 = enable loading of backup detail only on as needed basis, lazy load on demand
(default)
```

BACKUP GROUPING PERFORMANCE

This feature groups databases that share storage into larger groups to be processed as a single backup set (or subjob) of the entire backup job.

Note: Snapshot copies are currently done at the storage volume level, and a volume can have many qtrees, LUNs, and, of course, file system objects.

The advantages of this grouping include:

- By default reduces the per-group overhead (of VDI, NetApp, Windows, and so on) for each backup subjob
- Allows more NetApp volumes to have Snapshot copies made concurrently, using increased parallel processing
- Improves the performance of databases that are not already on a single NetApp storage volume

Default Limits

Minimum databases per backup set = 35

Maximum databases per backup set = 255

This feature groups a minimum of 35 databases into a backup set. If each database is located on a separate NetApp volume, then after 35 databases it commits the backup set to execute. If combining databases causes it to exceed the maximum number of databases, this feature commits the combined database as a backup set of the overall current backup.

The registry key can be configured to modify the minimum and maximum number of volumes to combine:

```
KEY: HKEY_LOCAL_MACHINE\SOFTWARE\Network Appliance\SnapManager for SQL Server\Server
TYPE: REG_DWORD
VALUE: "MaxDbConcurrentBackup" = default is 255
VALIE: "MinDbConcurrentBackup" = default is 35
```

A registry key can also be used to enable or disable the SMSQL grouping:

```
Key: HKEY_LOCAL_MACHINE\SOFTWARE\Network Appliance\SnapManager for SQL Server\Server
ValueName: EnableBackupSetGrouping
REG_DWORD:
0 = disable the multi-volume backup set grouping, reverts backup sets based on volume
boundaries.
1 = enable the multi-volume backup set grouping, (default)
```

BATCHING I/O

This feature helps to batch more enumeration requests into a single call to SnapDrive for Windows.

This improvement is significant in configurations with a high number of LUNs and databases. It is also effective when there are many outstanding backups on the hard drives. (If only one or two backups exist, the improvement is less noticeable.) This feature reduces the per SnapDrive for Windows call and batches the controller I/O for several LUNs into a single larger call, thus reducing the per-call overhead.

Note: This happens only on the delete all backups dialog load in the SMSQL MMC client.

SNAPDRIVE FOR WINDOWS CACHE

This feature is in the server side only. Improvement is noticeable in enumeration of LUNs within SnapDrive and is intended primarily to improve verification speeds.

```
The registry key to enable or disable the SnapDrive cache on the SMSQL server calls
into snapdrive:
Key: HKEY_LOCAL_MACHINE\SOFTWARE\Network Appliance\SnapManager for SQL Server\Server
ValueName: DisableSDLunEnumCache
REG_DWORD:
0 = enable the SnapDrive cache (default)
1 = disable the SnapDrive cache
```

10 APPENDIXES

10.1 CLUSTER-MODE TERMINOLOGY

This section defines Cluster-Mode terms.

cluster. In Data ONTAP 8.x operating in Cluster-Mode, a group of connected nodes (storage systems) that share a global namespace and that can be managed as a single virtual server or multiple virtual servers that offer performance, reliability, and scalability benefits.

command line interface (CLI). The Data ONTAP Cluster-Mode CLI is a command-based mechanism similar to the UNIX® tcsh shell in that it provides tab completion, advanced queries, and UNIX-style patterns and wildcards.

epsilon. Epsilon is an extra partial weight configured to one node. It does not determine the master but helps to form a majority. The epsilon node is epsilon for the entire cluster and not just for the individual RDB units. It is manually configurable but can change automatically. To add an epsilon to a node, use the **cluster modify** command with the −node and −epsilon parameters.

high availability (HA). In Data ONTAP 8.x, the recovery capability provided by a pair of nodes (storage systems), known as an HA pair.

high-availability (HA) pair. In Data ONTAP 8.x, a pair of nodes (storage systems) configured to serve data for each other if either of the nodes stops functioning.

logical interface (LIF). A LIF is a logical interface that is mapped to a physical port. A physical port can have up to eight LIFs in Cluster-Mode. A LIF is required to access a Vserver. The three types of LIFs are cluster management, node management, and data management (Vserver).

network. The cluster network connects nodes to form a cluster, and the data network connects the cluster to the client.

quorum. A quorum is formed when a majority of the eligible nodes in a cluster are healthy and in contact with one another. There is one quorum per RDB ring at any given time .The node with the lowest SiteID that is online is elected master; the rest are secondary RDB members.

replicated database (RDB). A replicated database that stores and maintains the data that manages the cluster. The operations in the RDB are transactional in nature. There are four RDB apps: VLDB, VifMgr, Management, and SpinAuth. This database is key to maintaining high-performance consistency in a distributed environment. Each RDB unit has its own replication ring.

ring. A ring is made up of one master, which is a read/write database, and other read-only databases. The writes go to the master and are then replicated to others in the ring through the cluster network.

storage failover (SFO). When two nodes are connected together, they make an SFO pair. The SFO pair should be the same controller model and can be enabled from either node.

virtual interface (VIF). A VIF is a virtual interface, whereas network ports are physical. There are three types of VIFs and ports: management, cluster, and data. Four physical ports can be grouped into a single VIF.

Volume Location Database (VLDB). A VLDB contains information about the index of the D-blade that currently owns a volume and currently serves an aggregate. VLDB content is cached on each N-blade to speed up the data path.

Vserver. A Vserver is a secure, virtualized storage server that is assigned to a single tenant.

Web interface. The Cluster-Mode Web interface is a model with which the user interacts through a Web browser.

11 ACKNOWLEDGMENTS

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®