Technical Report

Nondisruptive Operations with SMB File Shares

ONTAP 9.x

John Lantz, NetApp
November 2016 | TR-4100

Abstract

This technical report details NetApp® ONTAP® support for nondisruptive operations (NDO) when using the SMB protocol. NDO capabilities are covered for each of the primary SMB dialects, including SMB 1, SMB 2, and SMB 3.

Information Classification

Public

Version History

Version	Date	Document Version History
Version 3	November 2016	John Lantz Updated for ONTAP 9.x
Version 2.1	June 2015	Marc Waldrop Updated for NetApp clustered Data ONTAP® 8.3
Version 2	April 2013	Meghan Liese Updated for clustered Data ONTAP 8.2
Version 1	August 2012	Meghan Liese Initial release

TABLE OF CONTENTS

1 Overview

Nondisruptive operations (NDO) are a fundamental capability of ONTAP 9. However, the stateful nature of the SMB protocol can prevent continuous data availability during some operations. These interruptions to data availability are a result of the protocol, and have the same impact regardless of the underlying storage infrastructure.

In particular, logical interface (LIF) migrations, volume moves (aka NetApp DataMotion™ for Volumes), aggregate relocations (ARLs), controller upgrades, and storage failovers may or may not be disruptive to SMB clients depending on which version of SMB that they are using.

2 Requirements

Nondisruptive operations for SMB file shares in ONTAP depend largely on the version of SMB that is being used.

2.1 SMB Features That Support NDO

Each SMB version includes many features, but for the purposes of nondisruptive operations in SMB file shares, the most important versions and features are those shown in Table 1.

Table 1) SMB features that support NDO.

Version	Feature	ONTAP 9 NDO Capabilities
SMB 1	None	Volume move
SMB 2	Durable handles	Volume move LIF migration
SMB 3	Durable handles Persistent handles*	Volume move LIF migration Aggregate relocation* Storage failover* Nondisruptive upgrades*
* Requires continuously available SMB file shares		

A comprehensive list of all ONTAP supported SMB features, including capabilities, requirements, and best practices for implementation, can be found in TR-4543: SMB Protocol Best Practices.

2.2 SMB Versions That Are Supported by OS

ONTAP 9 supports all modern versions of the SMB protocol. However, it is important to note that SMB functionality that is essential to nondisruptive operations must be supported by both the client (Windows or macOS) and the server (ONTAP). Table 2 indicates which SMB versions are supported by ONTAP, Windows, OS X, and macOS.

Table 2) OS support for SMB.

OS	SMB 1	SMB 2	SMB 2.1	SMB 3	SMB 3.1.1
ONTAP 9	✓	✓	✓	✓	✓
Windows XP	✓				
Windows 2000	✓				
Windows Server 2003	✓				
Windows Vista SP1	✓	✓			
Windows Server 2008	✓	✓			
Windows 7	✓	✓	✓		
Windows Server 2008 R2	✓	✓	✓		
Windows 8	✓	✓	✓	✓	
Windows Server 2012	✓	✓	✓	✓	
Windows 10	✓	✓	✓	✓	✓
Windows Server 2016	✓	✓	✓	✓	✓
OS X 10.7 Lion	✓	✓			
OS X 10.8 Mountain Lion	✓	✓			
OS X 10.9 Mavericks	✓	✓	✓		
OS X 10.10 Yosemite	✓	✓	✓	✓	
OS X 10.11 El Capitan	✓	✓	✓	✓	
macOS 10.12 Sierra	✓	✓	✓	✓	

3 SMB 1

Compared to more modern SMB releases, SMB 1 is slow, is not suited to nondisruptive operations, and is a security risk because of its use of the now-compromised MD5 hashing algorithm.

Extended support for Windows Server 2003, the last version of Windows Server to support only SMB 1, ended in 2015. All currently supported versions of Windows Server support SMB 2 or later.

3.1 Client SMB 1 Connections

ONTAP 9 enables SMB 1 support on SMB servers by default. SMB 1 cannot be disabled in ONTAP 9.1.

3.2 Domain Controller SMB 1 Connections

Beginning in ONTAP 9.1, storage virtual machines (SVMs) can be configured to use specific SMB protocols (SMB 1 and SMB 2) when connecting to domain controllers. ONTAP 9.1 uses SMB 1 as the default SMB dialect for domain controllers.

```
vserver cifs security modify –vserver <vserver name> -smb1-enabled-for-dc-connections {true|false|system-default}
```
(default: `system-default`)

(The `system-default` for ONTAP 9.1 domain controller SMB 1 connections is `true`.)

3.3 Planning for SMB 1 Deprecation in ONTAP

SMB 1 support will be disabled by default in a future version of ONTAP. NetApp is following Microsoft's lead in this regard because Microsoft has announced that it is deprecating SMB 1 and will be disabling it, by default, in a future version of Windows Server.

Customers should migrate to a more recent version of SMB as soon as possible.

Consider using `statistics show -object smb1` to audit your shares for SMB 1 traffic.

3.4 Volume Moves

Although SMB 1 has no native functionality that promotes nondisruptive operations, thanks to ONTAP, volume moves are nondisruptive to all SMB sessions.

Volume move nondisruptively moves a volume to another physical location in the cluster. The volume may move to an aggregate that is owned by its partner node—or any other node in the cluster—without affecting the availability of the data that is contained in the volume. Volume move changes the physical location of the volume while leaving the logical location (position in the namespace or paths to LUNs) unchanged.

Note: Direct connections to a LIF on the same node that is hosting the SMB share may see better performance than indirect connections (hence SMB automatic node referrals functionality). Direct connections have been highlighted in green in this technical report's visualizations. Please be aware that as network bandwidth increases and cluster architectures evolve (multinode, etc.) that performance gains associated with direct versus indirect connections may not be as significant today as they have been in the past.

Simplified Volume Move Visualization

Figure 1) Volume A is on node 1. The client has a direct connection to volume A by using LIF 1.

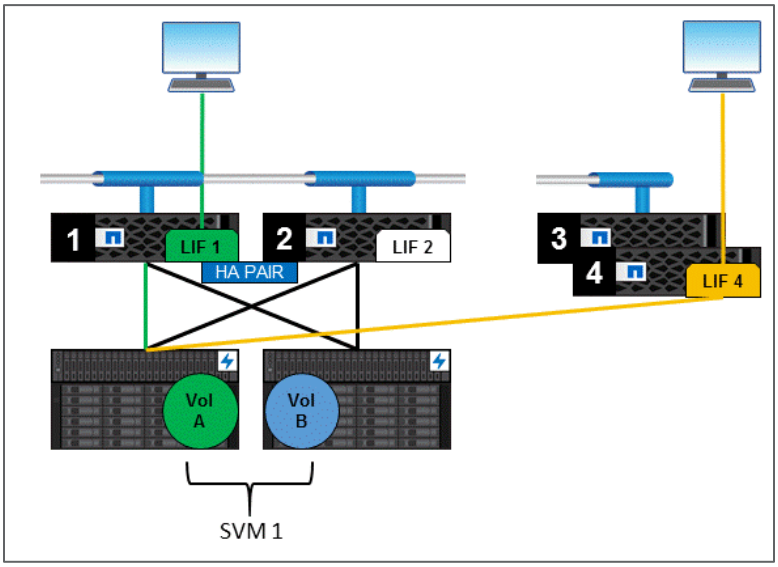
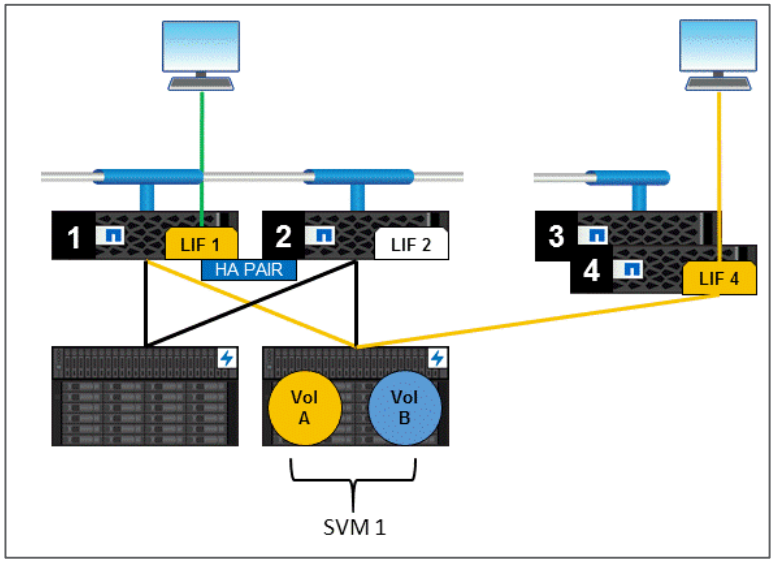


Figure 2) Volume A is moved to node 2. The client now has an indirect connection to volume A, but there is no session disruption. Volume moves are nondisruptive for all SMB traffic.



The primary use cases for volume move include cluster expansion, shelf upgrades, and service-level objective management.

For more information about volume moves, see TR-4075: DataMotion for Volumes.

4 SMB 2

First available in Windows Server 2008 and Windows Vista, SMB 2 was a major rewrite of the SMB protocol. It modernized the protocol and brought many enhancements—including simplified commands, asynchronous and compound operations, and increased read/write sizes—that resulted in significant performance improvements over SMB 1.

The durable handles of SMB 2 are essential for nondisruptive LIF migrations of SMB file shares.

4.1 Client SMB 2 Connections

ONTAP 9 enables SMB 2 support on SMB servers by default. To disable (or reenable) SMB 2, run:

```
vserver cifs options modify –vserver <vserver name> -smb2-enabled
{true|false}
```

(default: `true`)

(An advanced privilege level is required.)

`-smb2-enabled` enables or disables both SMB 2 and SMB 2.1.

4.2 Domain Controller SMB 2 Connections

Beginning in ONTAP 9.1, SVMs can be configured to connect to domain controllers using SMB 2. ONTAP 9.1 uses SMB 1 as the default SMB dialect for domain controllers.

```
vserver cifs security modify –vserver <vserver name> -smb2-enabled-for-dc-
connections {true|false|system-default}
```
(default: `system-default`)

(The `system-default` for ONTAP 9.1 domain controller SMB 2 connections is `false`.)

4.3 Volume Moves

Volume moves are nondisruptive to all SMB sessions in ONTAP. See section 3.4, Volume Moves.

4.4 LIF Migrations

LIF migrations can be used to balance or optimize network traffic across the cluster by moving network resources to storage resources.

When a LIF is migrated between nodes, client sessions become disconnected. SMB 1 is unable to prevent the lock state from being lost during these network outages, but SMB 2+ clients, by using durable handles, can reestablish disconnected sessions and maintain nondisruptive operations.

When an SMB 2 client opens a file and establishes a lock, a durable handle is created that records session details, the most important of which is the server that is associated with it. If the session is interrupted, the file remains open on the client, and the durable handle is used to reconnect the session.

Note: Durable handles are passive and require the client to reestablish disconnected sessions. The durable handle is lost if a client does not reestablish the session within 15 minutes.

Durable handles do not guarantee non-disruptive LIF migrations. If a client loses its durable handle, or if multiuser access to a file downgrades the oplock or lease level, then the connection to the open file may be disrupted.

Simplified LIF Migration Visualization

Figure 3) Volume A is on node 1. The client has a direct connection to volume A by using LIF 1.

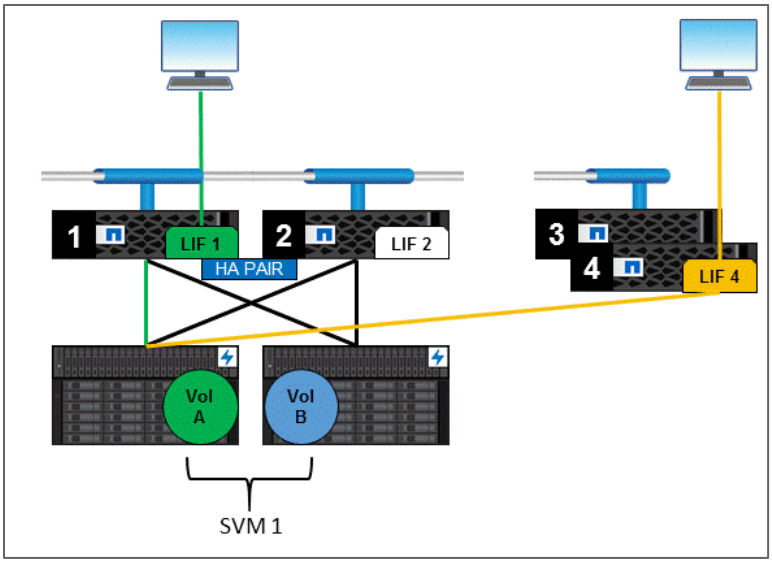
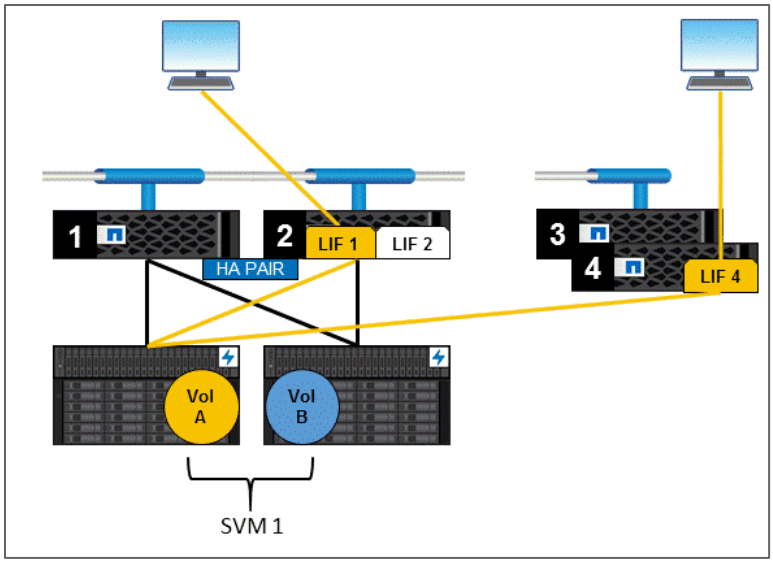


Figure 4) LIF 1 is moved to node 2. The client now has an indirect connection to volume A by using LIF 1. LIF migrations are nondisruptive for SMB 2+ sessions. SMB 1 sessions are disrupted because they lack the durable handles that are necessary to preserve the lock state during the LIF migration.



5 SMB 3

First available in Windows Server 2012 and Windows 8, SMB 3 was a major improvement to the SMB protocol. It saw significant enhancements, such as scale-out, transparent failover, persistent handles, and witness, that were designed to support continuously available shares on scale-out application file servers.

Although SMB 2's durable handles are nondisruptive for LIF and volume moves, they are not mirrored across nodes. Therefore, a failover (or an upgrade) of a node results in the lock state being lost. SMB 3's persistent handles improve upon durable handles by mirroring the lock state across nodes in a cluster.

SMB 3's persistent handles are essential for nondisruptive aggregate relocation, controller upgrades, and storage failover.

Note: ONTAP 9 does not support persistent handles for use cases other than Microsoft Hyper-V over SMB and SQL Server over SMB.

5.1 Client SMB 3 Connections

ONTAP enables SMB 3 and SMB 3.1.1 support on SMB servers by default. To disable (or reenable) SMB 3 or SMB 3.1.1, run:

```
vserver cifs options modify –vserver <vserver name>
-smb3-enabled {true|false}
-smb31-enabled {true|false}
```

(default: `true`)

(An advanced privilege level is required.)

5.2 Domain Controller SMB 3 Connections

ONTAP SVMs cannot be configured to connect to domain controllers using SMB 3. ONTAP 9.1 uses SMB 1 as the default SMB dialect for domain controllers (SMB 2 domain controller are optional).

5.3 Continuously Available Shares

Continuous availability (CA) is a share property that, by using SMB 3 scale-out, persistent handles, witness, and transparent failover, allows file shares to be accessible during otherwise disruptive scenarios such as controller upgrades or failures.

ONTAP supports CA shares for both Hyper-V over SMB and SQL Server over SMB use cases. Each use case has unique dependencies and configuration requirements that have been documented in the following reports:

- TR-4172: Hyper-V over SMB
- TR-4247: SQL Server over SMB

CA shares can be enabled by using:

```
-share-properties continuously-available
```

Note: Clients that do not support SMB 3 can connect and access data on a file share that has the continuously available property set, but they cannot take advantage of persistent handles.

5.4 Volume Moves

Volume moves are nondisruptive to all SMB sessions in ONTAP. See section 3.4, Volume Moves.

5.5 LIF Migrations

LIF migrations are nondisruptive to SMB 2+ sessions in ONTAP. See section 4.4, LIF Migrations.

5.6 Aggregate Relocation and Controller Upgrades

Aggregate relocate (ARL) is a nondisruptive process that moves ownership of aggregates between nodes that share storage (high-availability [HA] pair controllers). The node that originally owns the aggregate is referred to as the *source node*. The node that takes ownership of the aggregate is referred to as the *destination node*.

During the ARL, aggregates that are owned by the HA pair continue to serve data by using the destination node that is not being upgraded. Data migration does not require any data to be physically copied because ARL does not depend on the HA interconnect when reassigning aggregate ownership.

In practice, ARL is frequently used during controller upgrades—relocating aggregates (and LIFs) between HA pair controllers while controllers are upgraded with new hardware.

For more information about ARL, see the High-Availability Configuration Guide.

ARL for Continuously Available Shares

ARL is nondisruptive to SMB 3 CA file shares because of their support of persistent handles, which mirror the lock state on the destination node. As soon as an SMB 3 client loses connection to the source node, witness and transparent failover functionality work together to connect the SMB client to the destination node, reestablish the session, and allow the client to reclaim the file.

A new session is established on the destination node, the lock state is preserved, and the operation is nondisruptive to the client.

ARL for Non-CA Shares

ARL is a disruptive operation for non-CA SMB file shares. Because there is no mechanism—such as persistent handles— in these environments that can share the lock state across nodes, the relocation of an aggregate to a partner node disrupts the SMB session between the client and the SVM on the source node. Because persistent handles, witness, and transparent failover are available only on CA shares, the client must manually reestablish the session with the SVM, which is now on the destination node.

Simplified ARL Visualization

Figure 5) Aggregate A is on node 1. The client has a direct connection to aggregate A by using LIF 1.

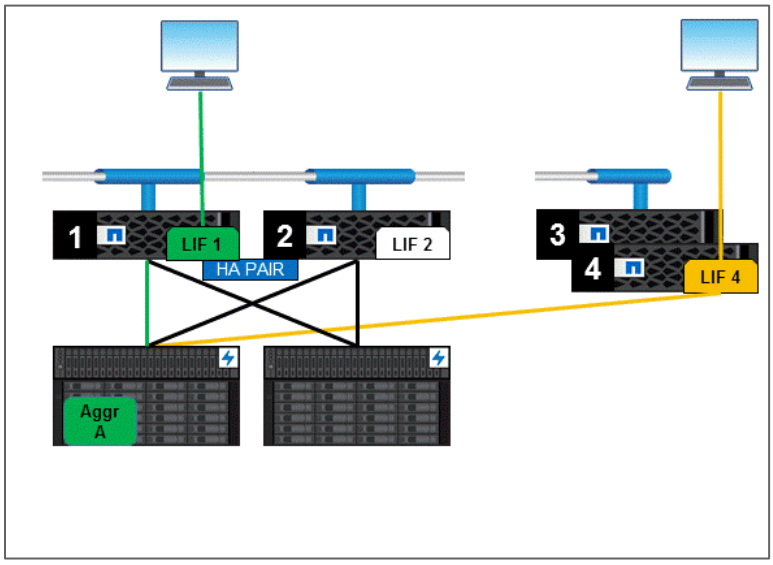
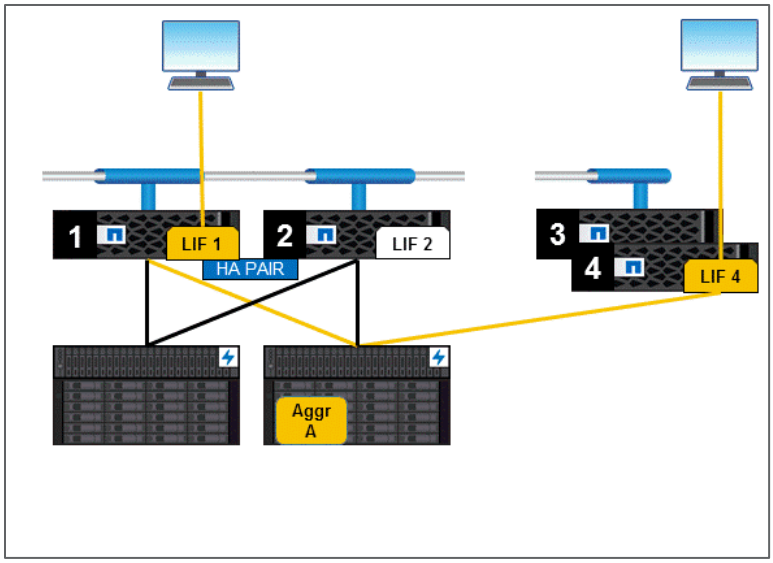


Figure 6) Aggregate A is moved to node 2. The client now has an indirect connection to aggregate A by using LIF 1. ARL relocation is nondisruptive to sessions that use SMB 3 CA shares. All SMB sessions to non-CA shares are disrupted because they lack the persistent handles that are necessary to preserve the lock state during LIF migration.



5.7 Storage Failover

A high-availability (HA) pair consists of two partner nodes, and a cluster is made up of numerous HA pairs. HA pairs are the physical components that provide storage resiliency at the system level for improved overall availability of the cluster.

If a NetApp controller fails, ONTAP takes advantage of the ability of an HA pair controller to fail over aggregates to partner nodes. In many ways, a storage failover in ONTAP acts like an automated albeit unplanned ARL operation.

For more information about storage failover, see the High-Availability Configuration Guide.

Storage Failover for CA Shares

Storage failover is nondisruptive to CA SMB file shares because of their support of persistent handles, which mirror the lock state on the destination node. As soon as a client loses connection to the source node, the witness and transparent failover functionalities work together to connect the client to the destination node. They then reestablish the session and allow the client to reclaim the file.

A new session is established on the destination node, the lock state is preserved, and the operation is nondisruptive to the client.

Storage Failover for Non-CA Shares

Storage failover is a disruptive operation for non-CA SMB file shares. There is no mechanism, such as persistent handles, in these environments that can share the lock state across nodes. Therefore, the relocation of the aggregates to the partner node disrupts the SMB session between the client and the SVM on the source node. The session lock state is lost and must be reestablished on the partner node after a failover or a giveback is complete.

Simplified Storage Failover Visualization

Figure 7) Storage failover (1). Aggregate A is on node 1. The client has a direct connection to aggregate A by using LIF 1.

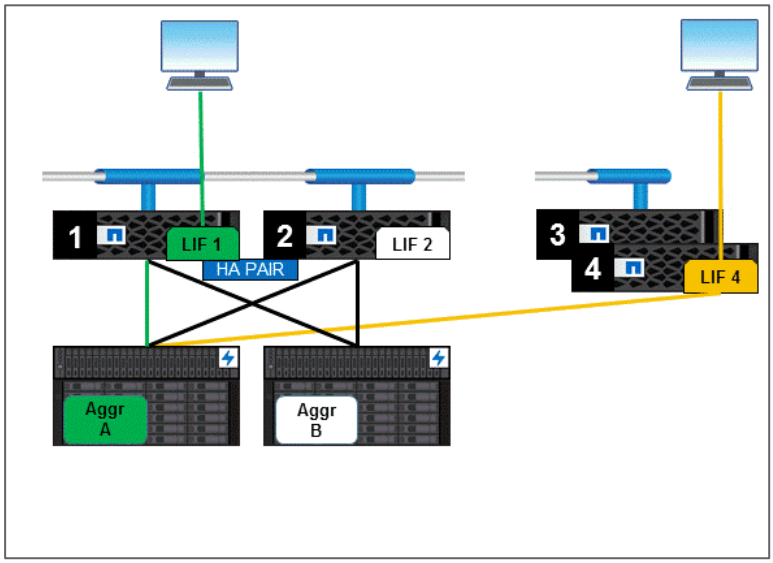
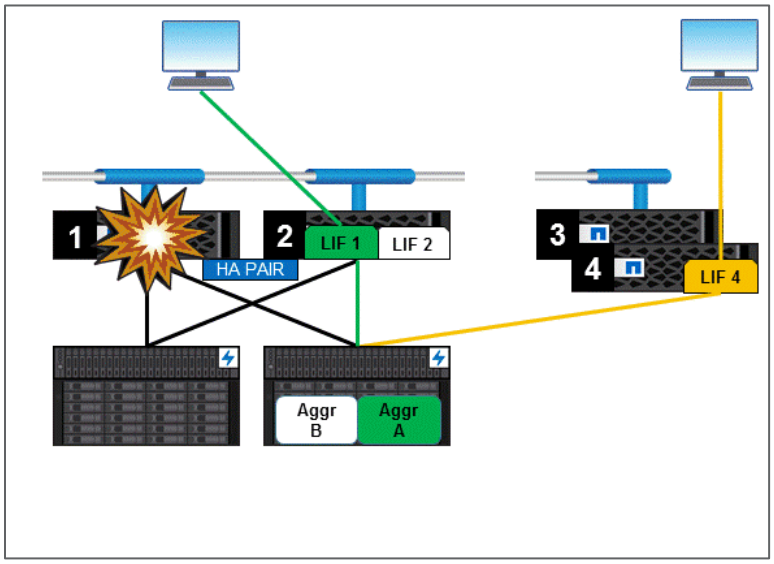


Figure 8) Storage failover (2). Aggregate A and LIF 1 are moved to node 2. The client continues to use a direct connection to aggregate A by using LIF 1. Storage failover is nondisruptive to sessions that use SMB 3 CA shares. All SMB sessions to non-CA shares are disrupted because they lack the persistent handles that are necessary to preserve the lock state during the storage failover.



5.8 Nondisruptive Upgrades

The term nondisruptive upgrades (NDU) refers to ONTAP's ability to nondisruptively upgrade system software and firmware without having to take downtime. NDU allows each node of an HA pair to be upgraded individually to a later version of ONTAP.

During an NDU, a takeover of the data service process that belongs to the node that is being upgraded is performed before reboot, thereby minimizing disruption of client I/O. Following the reboot, a giveback is initiated, returning the data service to the newly upgraded node.

NDU can be performed manually or it can be performed automatically by using automated nondisruptive upgrades (ANDU). ANDU validates cluster components to confirm that the cluster can be upgraded nondisruptively. It installs the target ONTAP image on each node and, based on the number of nodes in the cluster, executes either a rolling or a batch upgrade in the background. All core NDU commands and routines that are required to nondisruptively upgrade the cluster are imbedded in the ANDU process. Also, the administrator can monitor the progress, pause or resume an upgrade, and see the cluster upgrade history.

For more information, see the Upgrade Express Guide and the Upgrade and Revert/Downgrade Guide.

NDU with CA Shares

NDU is nondisruptive to CA SMB file shares because of their support of persistent handles, which mirror the lock state across nodes in the cluster. As soon as an SMB 3 client loses connection to the source node, the witness and transparent failover functionalities work together to connect the SMB client to the destination node. They then reestablish the session and allow the client to reclaim the file.

A new session is established on the destination node, the lock state is preserved, and the operation is nondisruptive to the client.

NDU for Non-CA Shares

NDU—in the traditional sense—is disruptive to all SMB sessions other than those that use continuously available shares. Clients that use SMB 1, 2, or 3 (without CA shares) lose their session lock state during controller takeover and reboot, and the information that the client requires about the open file is lost.

If disruption is acceptable, NetApp recommends using maintenance windows to upgrade controllers that host non-CA shares to new versions of ONTAP.

If disruption is unacceptable, NetApp recommends using rolling upgrades, a nondisruptive solution that makes use of volume move and LIF migration and that can maintain the session lock state for SMB 2+ clients during ONTAP upgrades.

Rolling Upgrade

A rolling upgrade consists of one or more nondisruptive upgrade procedures that are being executed on the nodes of the cluster. Rolling upgrades allow the cluster to continue serving data while the nodes in the cluster are running mixed versions of ONTAP.

The number of NDU procedures depends on the number of HA pairs in the cluster. For example, a rolling upgrade for a two-node cluster would be complete when a single nondisruptive upgrade is completed on the HA pair that makes up the two-node cluster. Alternatively, a rolling upgrade on a four-node cluster would not be complete until a nondisruptive upgrade was first executed on one HA pair and then the second HA pair, in a serialized fashion. A standard rolling upgrade allows only one HA pair to be upgraded at a time.

For nodes that are moving volumes in the cluster, a LIF migration might be necessary to maintain direct access to a file after a volume has been moved. There might be some variance in performance

throughput for a file that was accessed directly and subsequently received indirect access after a volume move. LIF migration is nondisruptive for SMB 2+.

Note: The time that it takes for the volume to be moved is the product of several variables, but it can be a lengthy process for larger volumes on busy systems. This time needs to be considered when planning the rolling upgrade. Although disruptive, upgrades that use maintenance windows take significantly less time.

Clusters with mixed versions of ONTAP are not intended to run for long periods. NetApp recommends completing a rolling upgrade as soon as possible.

Simplified Rolling Upgrade Visualization

Figure 9) Volume A is on node 1. The client has a direct connection to volume A by using LIF 1.

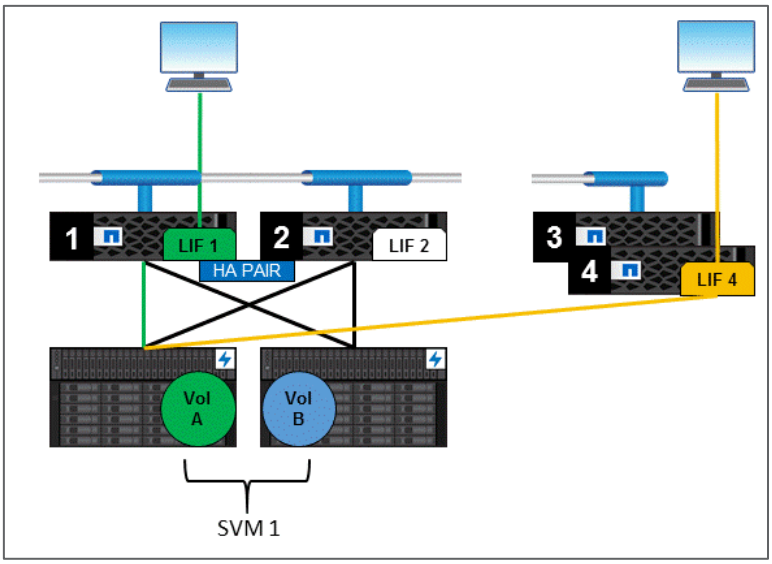


Figure 10) Volume A is copied to node 2. Cutover takes place, and session traffic is reestablished on LIF 2. ONTAP is upgraded on Node 1.

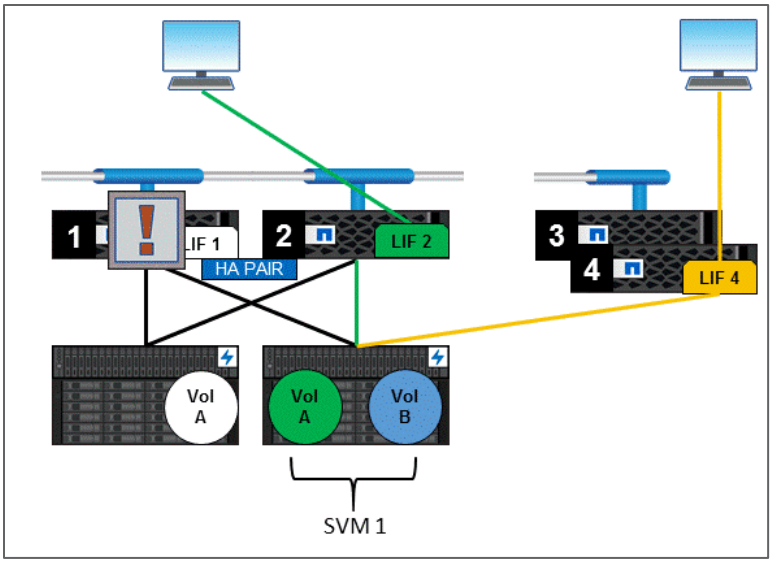


Figure 11) Node 1 is now using the latest version of ONTAP. Volume A and volume B are copied to node 1. Cutover takes place, and session traffic is reestablished on LIF 1. Node 2 is upgraded.

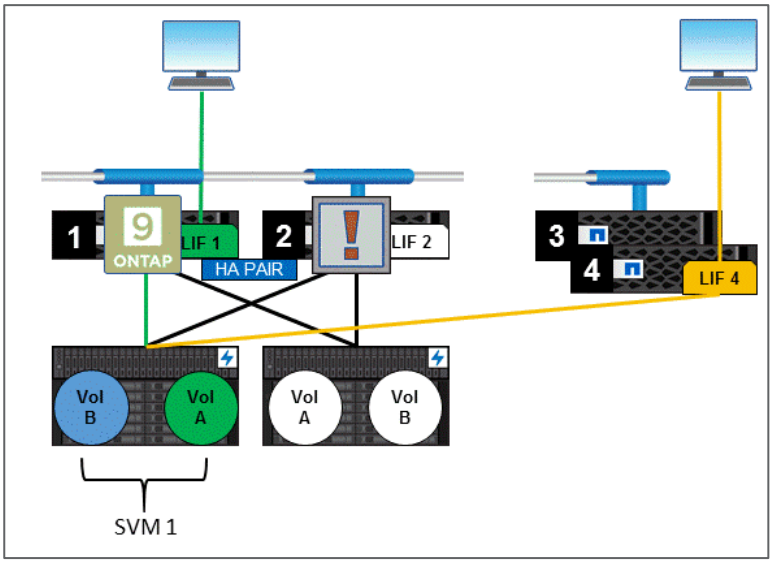
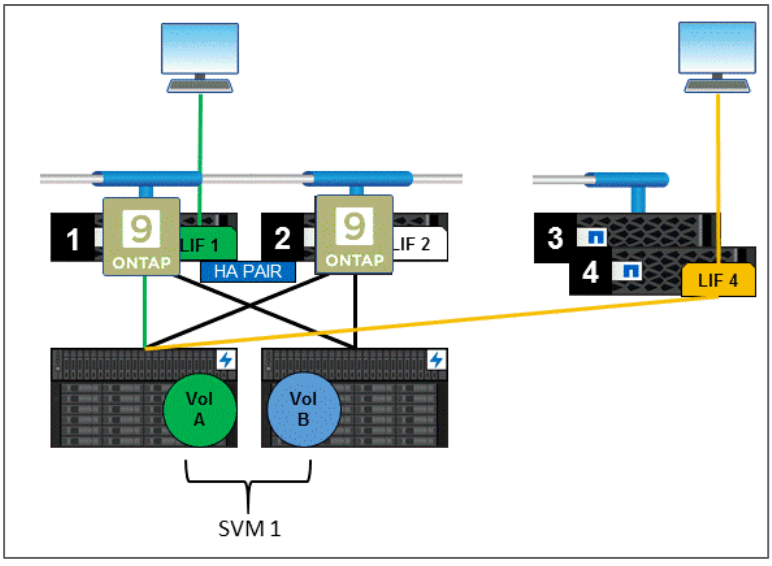


Figure 9) Node 1 and node 2 now both use the latest version of ONTAP. Volume B is moved back to node 2. This process continues, one HA pair at a time, until the entire cluster has been upgraded. Rolling upgrades take considerably longer but can be nondisruptive for SMB 2+ sessions.



Maintenance Windows

Although disruptive, the most straightforward and fastest process for upgrading ONTAP is to plan downtime by using maintenance windows. Maintenance windows disrupt SMB sessions while the NetApp controller fails over to its partner node and then again when a giveback occurs after the upgrade.

To minimize disruptions, maintenance windows should be scheduled during periods in which little activity or stress is placed on the storage system.

The entire procedure can be found in the Upgrade and Revert/Downgrade Guide of the ONTAP product documentation or in the NetApp My AutoSupport® Upgrade Advisor.

Note: To prevent data loss, SMB sessions not using CA shares must be terminated before the upgrade procedures. SMB 3 connections that use CA shares do not need to be terminated. They remain nondisruptive throughout the process.

Additional Resources

- TR-4075: DataMotion for Volumes
  http://www.netapp.com/us/media/tr-4075.pdf
- High-Availability (HA) Pair Controller Configuration Overview and Best Practices
  http://www.netapp.com/us/media/tr-3450.pdf
- Upgrade and Revert/Downgrade Guide
  https://library.netapp.com/ecm/ecm_download_file/ECMLP2492712
- CIFS Reference
  https://library.netapp.com/ecm/ecm_download_file/ECMLP2494081
- High-Availability Configuration Guide
  https://library.netapp.com/ecm/ecm_download_file/ECMLP2492713
- TR-4543: SMB Protocol Best Practices
  http://www.netapp.com/us/media/tr-4543.pdf
- TR-4172: Hyper-V over SMB
  http://www.netapp.com/us/media/tr-4543.pdf
- TR-4247: SQL Server over SMB
  http://www.netapp.com/us/media/tr-4543.pdf
- Upgrade Express Guide
  https://library.netapp.com/ecm/ecm_download_file/ECMLP2507747

Contact Us

Let us know how we can improve this technical report.

Contact us at docfeedback@netapp.com.

Include TECHNICAL REPORT TR-4100 in the subject line.

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.