

SQL Server FCI and File Server on VMware vSAN 6.7 using iSCSI Service



SQL Server FCI and File Server on VMware vSAN 6.7 using iSCSI Service

# **Table of contents**

SQL Server FCI and File Server on VMware vSAN 6.7 using iSCSI Service
Executive Summary
Business Case 3
iSCSI Introduction
Solution Overview
Key Highlights
Introduction
Purpose5
Scope5
Audience
Technology Overview
Solution Configuration
Solution Architecture
Hardware Resources
Software Resources
Network Configuration
Test Tools and Configurations    9
SQL Server and File Server Virtual Machine Configuration
Solution Validation
Overview
Configuring the iSCSI Service on vSAN
Configuring the WSFC and iSCSI Disks for SQL Server and Scale-Out File Server
Setting Up the SQL Server Cluster
Setting up Performance Monitoring
SQL Server and File Server Performance on All-Flash vSAN
iSCSI Transparent Failover
Application Role Failover and vMotion for WSFC Nodes
Best Practices
Conclusion
References
About the Author and Contributors



# SQL Server FCI and File Server on VMware vSAN 6.7 using iSCSI Service

# Executive Summary

This section covers the executive summary of this guide.

# **Business Case**

The modern-day CIO is concerned with high performance, high availability, and low cost when planning their Database Management Systems (DBMS). Over the years, Clustering Databases have become the mainstream choice over standalone databases in production environments. Clustering improves the availability of SQL Server instances by providing a failover mechanism to a new node in a cluster in the case of physical or operating system failures. The most used model is the activepassive nodes that have been in existence for decades creating an embedded base of clustering legacy systems that could benefit from today's new technologies such as HyperConverged Infrastructure (HCI).

VMware vSAN<sup>™</sup> has been widely adopted as an HCI solution for business-critical applications like Microsoft SQL Server. vSAN aims at providing a highly scalable, available, reliable, and high-performance HCI solution using cost-effective hardware, specifically direct-attached disks in VMware ESXi<sup>™</sup> hosts. vSAN adheres to a policy-based storage management paradigm, which simplifies and automates complex management workflows that exist in traditional enterprise storage systems with respect to configuration and clustering.

With the introduction of vSAN 6.7, the vSAN iSCSI target service now supports **Windows Server Failover Cluster** (WSFC) with shared disks for SQL Server Failover Cluster Instance(FCI), File Server and so on. This allows data center administrators to run workloads using legacy clustering technologies on vSAN. vSAN 6.7 can now support the shared target storage locations when the storage target is exposed using the vSAN iSCSI target service for either physical servers or virtualized servers, such as SQL servers, using guest iSCSI initiators.

The most significant benefit to using the vSAN iSCSI target service is that iSCSI service guarantees transparent failover, in case of hardware failures, operating system failures, application failures, service failures or planned downtime. This means the shared disk, which conventionally was deemed as a single point of failure, can now be protected by vSAN high availability, providing undisruptive target failover on the backend.

# iSCSI Introduction

iSCSI targets allow your Windows Server to share block storage remotely. iSCSI leverages the Ethernet network without requiring any specialized hardware. iSCSI target is a service available in Windows Server 2016 and can be enabled from the Add Roles and Features Wizard tabs, or through the VMware PowerCLI<sup>™</sup> command line automatically. The concepts of an iSCSI service are introduced as follows:

- **Target** : Targets are created to manage the connections between the iSCSI target server and the servers that need to access them. You assign logical unit numbers (LUNs) to a target, and all servers that log on to that target will have access to the LUNs assigned to it.
- **iSCSI target server:** iSCSI target server is the server where the iSCSI target service is running. In Windows Server 2016, there is a service called iSCSI service that you can install to configure the iSCSI target server.
- **iSCSI virtual disk** : iSCSI virtual disks are created on the iSCSI target server and associated with the iSCSI target. iSCSI virtual disk represents an iSCSI LUN, which is connected to the clients using an iSCSI initiator.
- **iSCSI initiator** : The Microsoft iSCSI Initiator Service enables you to connect a host computer that is running Windows 7/Windows Server 2008 R2 or later versions to an external iSCSI-based storage array through an Ethernet network adapter. The iSCSI initiator service runs on the client and is used to make a connection to the iSCSI Target by logging onto a target server. [1]

# Solution Overview

This reference architecture validates the solution of a Microsoft SQL Server Failover Clustering Instance and Microsoft Scale-Out File Server using shared disks backed by vSAN iSCSI target service.

# Key Highlights

This solution provides:

• Validation of running an SQL Server Failover Clustering Instance on vSAN iSCSI target service.



- Performance results of the mixed workloads including SQL Server online transaction processing (OLTP) and Scale-Out File Server file copies on a shared All-Flash vSAN using the vSAN iSCSI service.
- Validation of transparent failover of iSCSI vSAN against component failures, including the planned shutdown for maintenance of the host, and unexpected host shutdown.
- Validation of application role failover for both the SQL Server and the Scale-Out File Server using the vSAN iSCSI target service.
- Showcase of VMware PowerCLI<sup>™</sup> and PowerShell to illustrate the management efficiency of deploying and running Windows clustered service like SQL Server and Scale-Out File Server by using the vSAN iSCSI service.



# Introduction

This section provides the purpose, scope, intended audience and technology overview of this document.

## Purpose

This reference architecture validates the ability of vSAN iSCSI target service to provide the capability of SQL Server Failover Clustering Instances and ScaleOut File Servers to support OLTP workloads and Windows file share on the same HCI platform. And this solution validates the transparent failover on an iSCSI network, unexpected host shutdown and planned shutdown for maintenance as a proof solution for real-world applications on a VMware HCI platform.

# Scope

This reference architecture:

- Provides the steps and scripts for configuring an SQL Server Failover Clustering Instance (FCI) on shared disks, backed by the vSAN iSCSI target service.
- Provides best practices to configure FCI on the vSAN iSCSI service.
- Provides performance metrics of the SQL Server OLTP workload and file copying of Scale-Out File Server.
- Demonstrates the transparent failover capability of the vSAN iSCSI target service for an SQL Server running an OLTP workload and Scale-Out File Server running a copy load.
- Demonstrates the application role failover capacity and advantages of the vSAN iSCSI target service.

# Audience

This solution is intended for SQL Server database administrators, virtualization and storage architects involved in planning, architecting, and administering a virtualized, and Failover Cluster Instance SQL Server, with a complementary clustered file server on VMware vSAN iSCSI target service.

# Technology Overview

This section provides an overview of the technologies used in this solution:

- VMware vSphere <sup>®</sup> 6.7
- VMware vSAN 6.7
- vSAN iSCSI Target Service
- VMware PowerCLI
- Microsoft SQL Server 2016
- Windows 2016 Scale-Out File Server

## VMware vSphere 6.7

VMware vSphere 6.7 is the next-generation infrastructure for next-generation applications. It provides a powerful, flexible, and secure foundation for business agility that accelerates the digital transformation to cloud computing and promotes success in the digital economy.

vSphere 6.7 supports both existing and next-generation applications through its:

- Simplified customer experience for automation and management at scale
- Comprehensive built-in security for protecting data, infrastructure, and access
- Universal application platform for running any application anywhere

With vSphere 6.7, customers can run, manage, connect, and secure their applications in a common operating environment, across clouds and devices.

# VMware vSAN 6.7

VMware vSAN, the market leader HCI, enables low-cost and high-performance next-generation HCI solutions, converges traditional IT infrastructure silos onto industry-standard servers and virtualizes physical infrastructure to help customers easily evolve their infrastructure without risk, improve TCO over traditional resource silos, and scale to tomorrow with support for new hardware, applications, and cloud strategies. The natively integrated VMware infrastructure combines radically simple VMware vSAN storage,



the market-leading VMware vSphere Hypervisor, and the VMware vCenter Server® unified management solution all on the broadest and deepest set of HCI deployment options.

vSAN 6.7 introduces further performance and space efficiencies. Adaptive Resync ensures fair-share of resources are available for VM IOs and Resync IOs during dynamic changes in load on the system providing optimal use of resources. Optimization of the destaging mechanism has resulted in data that drains more quickly from the write buffer to the capacity tier. The swap object for each VM is now thin provisioned by default, and will also match the storage policy attributes assigned to the VM introducing the potential for significant space efficiency.

# iSCSI Target Service on vSAN 6.7

The vSAN iSCSI target service allows physical servers to access vSAN storage using the iSCSI protocol, extending workload support to physical servers. iSCSI targets on vSAN are managed the same as other objects with Storage Policy Based Management (SPBM). This implies that the iSCSI LUNs can be configured as RAID-1, RAID-5 or RAID-6, and have other characteristics, such as IOPS limits. vSAN functionality such as deduplication, compression, and encryption can also be utilized with the iSCSI target service to provide space savings and security for the iSCSI LUNS. In addition, authorization of initiator with initiator group, CHAP and Mutual CHAP authentication are also supported for greater security.

vSAN 6.7 expands the flexibility of the vSAN iSCSI service to support WSFC. When an organization has WSFC servers in a physical or virtual configuration, vSAN can support the shared target storage locations when the storage target is exposed using the vSAN iSCSI target service.

# VMware PowerCLI 10.0

VMware PowerCLI provides a Windows PowerShell interface to the VMware vSphere, VMware vCloud ®, vRealize® Operations Manager<sup>™</sup>, and VMware Horizon® APIs. PowerCLI includes numerous cmdlets, sample scripts, and a function library. VMware PowerCLI provides more than 500 cmdlets for managing and automating vSphere, vCloud, vRealize Operations Manager, and VMware Horizon environments. See VMware PowerCLI Documentation for more information.

VMware PowerCLI 10.0.0 is the latest version, for installation and updating information see VMware PowerCLI Blog .

## **Microsoft SQL Server 2016**

SQL Server consistently leads in performance benchmarks, such as TPC-E and TPC-H, and in real-world application performance. Gartner recently rated SQL Server as having the most complete vision of any operational database management system. With SQL Server 2016, performance is enhanced with a number of new technologies, including many **new features**, and enhancements including multisite clustering across subnets, faster and more predictable failover of instances, supportability of Windows Server Cluster Shared Volumes, and so on.

The following versions are supported by this solution:

- Microsoft SQL Server 2012, 2014, 2016 and 2017
- Windows Server 2012, 2012 R2 and 2016

# Windows 2016 Scale-Out File Server

Scale-Out File Server provides scale-out file shares that are continuously available for file-based server application storage. Scaleout file shares provide the ability to share the same folder from multiple hosts on the same cluster.

Benefits provided by Scale-Out File Server in include:

- Active-Active file shares: All cluster nodes can accept and serve SMB client requests.
- Increased bandwidth: The maximum share bandwidth is the total bandwidth of all file server cluster nodes.
- CHKDSK with zero downtime: A CSV File System (CSVFS) can use CHKDSK without impacting applications with open handles on the file system.
- Clustered Shared Volume cache: CSVs in Windows Server 2012 introduces support for a Read cache, which can significantly improve performance in certain scenarios, such as in Virtual Desktop Infrastructure (VDI).
- Simpler management: With Scale-Out File Server, you create the scale-out file servers, and then add the necessary CSVs and file shares. It is no longer necessary to create multiple clustered file servers, each with separate cluster disks. [2]



# Solution Configuration

This section describes the solution architecture, hardware and software resources, network configuration, test tools, and the SQL Server and File

Server virtual machine configuration in this RA.

# Solution Architecture

We recommend a minimum of four nodes for the vSAN Cluster to ensure continued data protection and availability during planned or unplanned downtime. Four ESXi hosts with vSAN 6.7 were used to validate this solution.

The configuration consisted of five Windows 2016 virtual machines on a vSAN Cluster as shown in Figure 1:

- SQL Server FCI Active Node
- SQL Server FCI Passive Node
- Scale-Out File Server Active Node
- Scale-Out File Server Passive Node
- Domain Controller (DC)/DNS Server

The File Server test client is on a separate vSAN cluster with one 10GbE network connection on the same subnet as the vSAN Cluster housing the 5 VMs listed above. The SQL Server test client is installed on the same server as the SQL Server Active Node.



## Figure 1. Solution Architecture

# Hardware Resources

In this solution, we used two low-latency 1.5 TB Samsung NVMe SSDs as the cache tier.



Each ESXi host in the vSAN Cluster had the following configuration as shown in Table 1.

## Table 1. Hardware Resources per ESXi Host

Property	Specification
Server	Dell PowerEdge R630
CPU	2 sockets, 24 cores each of 2 6 GHz with hyper-threading enabled
RAM	256 GB DDR4 RDIMM
Network adapter	2 x Intel 10 Gigabit X540-AT2, + I350 1Gb Ethernet
Storage adapter	2 x 12Gbps SAS PCI-Express (Dell PowerEdge RAID H730 mini)
Disks	Samsung <u>NVMe</u> : 2 x 1.5 TB drives as cache SSD SSD: 4 x 400 GB drives as capacity SSD (ATA Intel SSD SC2BX40)

## Software Resources

Table 2 shows the software resources used in this solution.

Table 2. Software Resources

Software	Version	Purpose
VMware vCenter Server and ESXi	6.7	ESXi Cluster to host virtual machines and provide the vSAN Cluster. VMware vCenter Server provides a centralized platform for managing VMware vSphere environments
VMware vSAN	6.7	Software-defined storage solution for hyper-converged infrastructure
Microsoft SQL Server	2016 Enterprise Edition, SP1	Database software
Microsoft Windows Server	2016 x64 Standard Edition	Operating System for the VMS: • SQL Server database virtual machines • Load generation virtual machines • Domain controller
DVD Store	3	OLTP workload generator

# Network Configuration

We created a vSphere Distributed Switch<sup>™</sup> to act as a single virtual switch across all associated hosts in the data cluster.

The vSphere Distributed Switch used two 10GbE adapters for the teaming and failover. A port group defines properties regarding security, traffic shaping, and NIC teaming. To isolate traffic, we used two port groups, one for vSAN, VM and vMotion traffic and the other for the iSCSI service traffic. The settings for both port groups are the default.

We assigned one dedicated NIC as the active link and assigned another NIC as the standby link for the iSCSI service. For vSAN, VM and vMotion traffic, the uplink order is reversed. See Table 3.

## Table 3. Network Configuration



Distributed Port Group	Active Uplink	Standby Uplink
vSAN/VM/vMotion	Uplink2	Uplink1
iscsi	Uplink1	Uplink2

Note : For better performance, we enabled Jumbo Frames on the physical switch and the vSphere Distributed Switch.

\$proc=New-Object System.Diagnostics.Process
\$procellinfo=New-Object System.Diagnostics.ProcessStartInfo ("xcopy", "C:\backup\dbbackup.bak \\sofs\fs01 /Y")

## **Performance Metrics Data Collection Tools**

We measured three critical metrics in the performance tests:

- The OLTP workload throughput measured in orders per minute (OPM)
- The copy rate for the File Server
- The average latency of each IO operation on the guest OS

We also used the vSAN Performance Service and the esxtop utility on the ESXi hosts to monitor vSAN performance.

- vSAN Performance Service : vSAN Performance Service is used to monitor the performance of the vSAN environment, using the web client. The performance service collects and analyzes performance statistics and displays the data in a graphical format. You can use the performance charts to manage your workload and determine the cause of problems.
- esxtop : esxtop is a command-line utility that provides a detailed view of the ESXi resource usage.

## Test Tools and Configurations

The following test tools were used in validating this solution.

## **DVD Store Version 3**

DVD Store 3 (DS3) is an open source test/benchmark tool that simulates an online store that sells DVDs. Customers can log in, browse DVDs, browse reviews of DVDs, create new reviews, rate reviews, become premium members, and purchase DVDs. Everything needed to create, load, and stress this online store is included in the DVD Store project. DS3 comes in three standard sizes: small, medium and large, with each having a different standard database size, customer number, orders, and products as shown in Table 4.

#### Table 4. DVD Store 3 Standard Database Sizes

Database	Size	Customers	Orders	Products
Small	10 MB	20,000	1,000/month	10,000
Medium	1 GB	2,000,000	100,000/month	100,000
Large	100 GB	200,000,000	10,000,000/month	1,000,000

## xcopy and Multiple Sessions

xcopy is a command -line tool used for copying multiple files or entire directory trees from one directory to another and for copying files across a network. We used PowerShell scripts to continuously copy a file from the client to the three shared folders exposed



from the clustered File Server. The following is an example to issue the copy session as one thread.

## SQL Server and File Server Virtual Machine Configuration

#### DB Size and File Server Space Usage

We configured two SQL Server virtual machines for the performance tests. Given the test goals, a 200 GB sized database was created. It contains 400 million customers, an order history of 20 million DVDs per month, and a catalog of 2 million DVDs. The 200 GB database consumed approximately 306 GB of disk space.

For the File Server, we shared 3 folders with provisioned space of 245 GB, 245 GB and 62 TB respectively to demonstrate small and large shared folders. We tested copying an 80 GB database backup file to the three folders simultaneously, so the actual File Server space consumption is around 240 GB.

#### VM Configuration

We assigned 24 vCPUs and 128 GB of memory to the VM hosting SQL Server Failover Cluster Instance and assigned 4 vCPUs and 16 GB of memory to the VM hosting File Server. We set the maximum server memory to 96 GB for the SQL Server instance. Table 5 lists the SQL Server 2016 VM and File Server VM configurations.

#### Table 5. SQL Server 2016 VM and File Server VM Configurations

VM Role	vCPU	Memory (GB)	SQL Server Version	Operating System
VM1/VM2— 306 GB DB	24	128	2016 Enterprise Edition, SP1	Windows Server 2016 Standard Edition 64-bit
VM3/VM4— Scale-Out File Server	4	16	N/A	Windows Server 2016 Standard edition 64-bit

Windows Server 2016 was used as the guest operating system (OS) for the SQL Server and File Server virtual machines. Each SQL Server and File Server virtual machine in the cluster had VMware Tools installed, and each virtual hard disk for data and log was connected to separate VMware Paravirtual SCSI (PVSCSI) adapter to ensure optimal throughput and lower CPU utilization [1]. The virtual network adapters used the VMXNET 3 adapter, which was designed for optimal performance. [2]

## **OS and SQL Server Instance Settings**

The settings which can benefit from the DVD Store workload are summarized below:

- Operating System Settings:
  - $\circ~$  Set the power plan to High Performance in the guest OS.
  - Enable SQL Server to use large pages by enabling the Lock pages in memory user right assignment for the account that will run the SQL Server in Group Policy.
- SQL Server Startup Parameters:
  - -x Disable SQL Server performance monitor counters.
  - $\circ~$  -T661 Disable the ghost record removal process.

## VM Disk Layout

For the OLTP workload, the database size is based on the actual disk space requirement and additional space for the database growth:

- The iSCSI virtual disk configuration for the SQL Server database is:
  - $\circ$  2 x 100 GB OS disks, 2 x 250 GB data file disks, 1 x 100 GB log file disk, and 1 x 80 GB tempdb disk.

For the file server, we created one Scale-Out file server with three file shared folders:

• The iSCSI virtual disk configuration for the File Server is:



 $\circ~$  2 x 100 GB OS disks, 2 x 245 GB file share disks, and 1 x 62 TB file share disk.

The OS virtual disks were from native vSAN, and the others were created from the iSCSI service. In total, we have seven LUNs under seven targets created as the shared disks for SQL Server FCI (four disks) and File Server (three disks).

See Configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters (1010398).

See Choosing a network adapter for your virtual machine (1001805).



# Solution Validation

This section describes the solution validation.

## Overview

We validated this solution running Microsoft SQL Server on the vSAN iSCSI service and Scale-Out File Server using shared disks. We accomplished this by:

- Configuring the vSAN iSCSI Service
- Configuring the Windows Cluster and iSCSI disks
- Setting up the SQL Server Cluster
- Setting up Performance Monitoring
- Measuring the performance of the SQL Server and Scale-Out File Server on All-Flash vSAN
- Validating the iSCSI Transparent failover for both the SQL Server and Scale-Out File Server
- Validating Application Role Failover

We used PowerShell scripts throughout our configuration and validation.

# Configuring the iSCSI Service on vSAN

vSAN 6.7 enables guest iSCSI initiators on both physical or virtual servers to access a VM either on the same vSAN Cluster or another one. This feature enables an iSCSI initiator to transport block-level data to an iSCSI target on a storage device in the vSAN cluster.

The high-level steps to configure iSCSI service on vSAN are:

- Enable the iSCSI Target Service on the vSAN cluster
- Create an iSCSI Target and its associated LUN.
- Add one or more LUNs to an iSCSI Target, or edit an existing LUN
- Create an iSCSI Initiator Group to provide access control for iSCSI targets. Only iSCSI initiators that are members of the initiator group can access the iSCSI targets. Without the definition of the Initiator group, any initiator can see any targets
- Assign a Target to an iSCSI Initiator Group

See Using the vSAN iSCSI Target Service for the detailed information.

After configuring the vSAN iSCSI target service and LUNs, we discovered the vSAN iSCSI targets from the physical or virtual servers via software initiator. To discover vSAN iSCSI targets, we used the IP address of any host in the vSAN cluster and the TCP port of the iSCSI target.

**Recommendation** : To ensure high availability of the vSAN iSCSI target, configure multipath support for your iSCSI application. You can use the IP addresses of two or more hosts to configure the multipath.

# Configuring the WSFC and iSCSI Disks for SQL Server and Scale-Out File Server

There are several steps to configure the Windows Cluster and the iSCSI shared disks. These steps must be done before installing the SQL Server FCI and before configuring the Scale-Out File Server on the guest OS in the VM.

We used PowerShell and PowerCLI scripts to automate these steps. The high-level steps to configure the required features for Windows Cluster, MPIO, and to configure the disks are as follows:

- Enable the cluster feature on a Windows server to be used as an SQL Server Failover Clustering host.
- Enable the Windows Multipath IO (MPIO) feature and explicitly set the failover policy to " Fail Over Only. "
- Reboot the VM.
- Discover the target on each host of the Windows cluster and connect them using multiple connections for failover purposes.
- Create the Windows cluster.
- Format the shared disks.



• Add disks to the Windows cluster.

At this point, the Windows cluster is ready for a cluster application or service such as SQL Server FCI.

**Recommendation** : Use PowerShell and PowerCLI to automate the process and minimize potential setup errors and time spent on configuration steps. We have provided sample scripts, see Configuration in VM to set up iSCSI initiator and cluster shared disks for SQL Server and File Server.

**NOTE:** We used the default vSAN SPBM policy for the virtual disks of VMs of SQL Server and File Server.

## Setting Up the SQL Server Cluster

A Windows Server Failover Cluster (WSFC) is a group of independent servers that work together to increase the availability of applications and services. A Failover Cluster Instance (FCI) is a SQL Server instance that is installed across nodes in a WSFC. In the event of a failover, the WSFC service transfers ownership of instance's resources to a designated failover node. The SQL Server instance is then restarted on the failover node, and databases are recovered as usual. At any given moment, only a single node in the cluster can host the FCI and underlying resources. That means the FCI is active/passive high availability solution.

We followed the details in Create a New SQL Server Failover Cluster (Setup) to set up our clustered SQL Server Failover Cluster Instance.

## Setting up Performance Monitoring

To calibrate our performance results, we needed to setup monitoring of both the vSAN Datastore capacity and the location of the VMs.

## Managing the Quota for File Server from vSAN

In our test environment, we created one 62 TB disk and exposed it as a scale-out shared folder. By default, vSAN's space allocation policy is thin provisioned. Thus the raw capacity on the vSAN datastore was less than 10 TB.

We needed to monitor the space usage from vSAN and take proactive actions such as stopping new space allocation from vSAN when the space usage is above 80 percent. We recommend using the vSAN Health Service to monitor the space usage and for triggering an alert when the space usage exceeds certain thresholds. See KB 2108740 for details.

For more finely controlling the accessibility of the shared folder using File Server from OS level after the space usage of vSAN is above the threshold especially in a testing environment, you may also use the sample PowerShell and PowerCLI script we created. This script is available for download, see **Report the space usage of a vSAN datastore and set the file share attribute to Read-only** to avoid over-commitment using PowerShell.

## **Rebalancing VM Location for a WSFC node**

VMware HA and vMotion permit the VMs running Windows Cluster services with iSCSI shared disks to be on the same host. This may hurt performance for a mission-critical virtual machine which is intended to be on a host independently. To avoid this from happening, use the DRS VM-Host Affinity Rule to separate the location of the VMs on the WSFC nodes.

You may also use the sample code we created to migrate the VMs which can move to a host that should only have one Windows Cluster VM on it. This script is available for download, see Automatically assign a VM to a host with 1:1 mapping to avoid performance contention on vSAN.

## SQL Server and File Server Performance on All-Flash vSAN

The purpose of this test was to verify that the vSAN iSCSI target service can support mixed clustered services simultaneously including SQL Server FCI and Scale-Out File Server.

## **Test Scenarios**

This test measured the impact of stressing an SQL Server 2016 on a Windows Server 2016 guest VM using DVD Store 3 to generate the OLTP workload. The number of driver threads, which generate OLTP activity, were adjusted as needed to saturate the host. Repeatability was ensured by restoring a backup of the database before each run. Steady state was achieved by ensuring throughput gradually increased during the 10-minute warmup period, and reached and maintained steady state during the 20-minute run time for each test.

In addition, we simultaneously stressed the network bandwidth with the Scale-Out File Server. Our test generated multiple file copy sessions from a separate vSAN Cluster to the shared disks where the Scale-Out File Server Windows Server VM resides over a 10GbE network connection. We used multiple xcopy sessions to copy a large backup file to the shared folder to saturate the network bandwidth. For the three shared file folders, we ran a continuous loop of three simultaneous file copy sessions.

The VM Disk layout was described earlier. In total, seven LUNs under seven targets were created as the shared disks, with each



target having one LUN. The target owners were distributed on the four hosts. This was to make sure no target owner contention existed.

# **Testing Results**

We measured the key performance indicators include aggregate DVD orders per minute (OPM), file copy rate, and virtual disk latency in the guest OS. The total IOPS, bandwidth, and latency on the vSAN level were also measured.

**Application Performance:** The overall OPM was 30,112 in the DVD Store 3 output report and the average file copy rate was 382 MB/sec over the 10GbE network. See Table 6.

#### Table 6. Application Performance

Orders Per Minute (OPM)	Average Copy Rate (Mbytes/sec)
30,112	382

**vSAN Performance:** When the OLTP workload and concurrent file copy session were running, the overall IOPS on vSAN was around 30K, with around 29K write IOPS and 1K read IOPS. Using Failure to Tolerance Method RAID-1 (Mirroring), the average bandwidth was around 880 MBs with around 870 MBs write and 10 MBs read bandwidth. The average write latency was less than 0.7 milliseconds and read latency less than 0.6 milliseconds.



esxtop CPU Utilization: The ESXi host, where the SQL Server active node ran on, consumed most of the physical core CPU resources at around 85%. The CPU utilization of the other ESXi hosts was each less than 16%.

## Figure 2. Physical Core Utilization per ESXi Host

Virtual Machine Performance : The average Scale-Out File Server disk write latency was 16 milliseconds. The network bandwidth of the Scale-Out File Server Active Node over the 10GbE network with one physical NIC was 406 MB/sec received and 387 MB/sec sent while receiving file copies from the test client and sending writes to virtual disks through iSCSI service. This load saturated the 10GbE NIC bandwidth.

The average read and write disk latency of SQL Server data files was less than or equal two milliseconds, and the write latency of log disk was 5 milliseconds on average.

## Table 7. Virtual Machine Performance



Scale-Out File Server (Active Node)			
Average File Server shared disk latency in millisecond (write)	Network Adapter (MB Received/sec)	Network Adapter (MB Sent/sec)	
16	406	387	
SQL Server Failover Instance (Active Node)			
Average data disk latency in millisecond (read)	Average data disk latency in millisecond (write)	Average log disk latency in millisecond (write)	
1	2	5	

## **Summary**

- iSCSI vSAN can support mixed workloads on clustered services including SQL Server Failover Cluster Instance and File Server simultaneously and efficiently with high transaction numbers (OPM was more than 30K) on a single virtual machine, and high copy rate (over 300 MB/sec over the 10GbE network).
- Given the backend data service of vSAN, the test results showed that physical core CPU utilization of the SQL Server active node was high at around 85 percent, while the CPU utilization of the File Server was low. That demonstrated the DVD Store 3 workload is CPU bound, while the file copy was network capacity bound.
- The disk performance was excellent; the NVMe SSD provided high bandwidth and low latency.

## iSCSI Transparent Failover

vSAN can handle several failure scenarios including disk, disk group, host failure, disk controller, or network. We have already verified three major failure scenarios in our published reference architecture - Microsoft SQL Server 2014 on VMware vSAN 6.2 All-Flash .

On vSAN iSCSI service, all LUNs are visible to initiators on all ESXi hosts. For any given time, every iSCSI target has its owner host and only the owner host handles the IOs for the iSCSI target. However, the initiator can create multiple iSCSI sessions to any ESXi hosts through multipath. If the initiator connects to the non-target owner host, it will be redirected to the target owner host automatically. If the owner host fails, a new owner is elected to serve the IO for the target. This process is called owner transfer. The transfer can happen along with planned or unplanned host failures.

vSAN 6.7 expands the flexibility of vSAN iSCSI service for Windows Server Failover Cluster (WSFC) with transparent failover capability. That means the underlying target owner transfer is transparent for an application like SQL Server Failover Cluster on WSFC.

This section focuses on the validation of this functionality, transparent failover.

# **iSCSI Transparent Failover for SQL Server FCI**

In this scenario, we validated the transparent failover capability of the vSAN iSCSI service for SQL Server Failover Cluster Instance with a running workload for the following failures:

- Unplanned host failures
- Planned host downtime

## **Test Procedures**

The high-level steps for this test are:

- Start a DVD Store 3 workload on the SQL Server Failover Instance.
- Verify all shared disks are on one host, this host will be the point of failure.
- Verify the host has no running VMs at the time. We do not want to have the failure related to a compute resource interruption.
- For an unexpected host failure scenario, power off the host directly.
- For a planned host downtime scenario, enter the host into maintenance mode (EMM) using the options of no VM migration



#### and no data evacuation.

## **Testing Results**

For both test scenarios, there were no SQL Server IO exceptions in the error log and no running exceptions. The application had no outages during the target owner transfer.

And for the target owner transfer process, the target owner transfer finished in few seconds for unplanned host failures and planned host downtime.

From an application behavior after failover perspective, the transactions per second of the user database of DVD Store 3 recovered to the level before the failure. The unplanned host failure caused the IO to be suspended for about 45 seconds. Planned host downtime had no IOs suspended.

The failure impact duration was short – 1 minute or less. See Table 8.

#### Table 8. iSCSI Transparent Failover Validation Results

Failure Scenarios	Temporary IO Suspending in Guest OS?	Workload duration (MM:SS)	Target Owner transfer after failure
Unplanned host failures	Yes	00:45	In few seconds
Planned host downtime	No	00:22	Almost Immediate after EMM finished

iSCSI Transparent Failover for Scale-Out File Server

In this scenario, we validated the transparent failover capability of the vSAN iSCSI service for Scale-Out File Server while running multiple file copy sessions:

- Unplanned host failures
- Planned host downtime

## **Test Procedures**

The high-level steps for this test are:

- Start a continuous loop of 3 xcopy sessions to copy a large backup file to a shared folder.
- Verify all shared disks are on one host, this host will be the point of failure.
- Verify the host has no running VMs at the time. We do not want to have the failure related to a compute resource interruption.
- For an unexpected host failure scenario, power off the host directly.
- For a planned host downtime scenerio, enter the host into maintenance mode (EMM) using the options of no VM migration and no data evacuation.

# Testing Results

For both failure scenarios, no copy session outage or cluster role failover was observed.

We observed no obvious IO impact or copy rate downgrade in planned host failure and short-time I/O pause in the unplanned host shutdown from the Performance Monitor on Windows.



## Summary

Both the SQL Server and Scale-out File Server clustered applications on shared disk backed by the vSAN iSCSI service did not experience any outages or service failures both failure scenarios.

We observed the short time and temporary IO suspending on unexpected host shutdown for SQL Server FCI. Enabling MPIO is very important for the initiator to find additional paths to connect the targets which previously connected.

## Application Role Failover and vMotion for WSFC Nodes

Application Roles are clustered services like Failover Cluster Instance (FCI) or generic file server. An FCI is a SQL Server instance that is installed across nodes in a WSFC; this approach is applied to the File Server in a WSFC too. This type of SQL Server Instance depends on resources for storage and virtual network name. The storage can use Fibre Channel, iSCSI, FCoE, SAS, or use locally attached storage for shared disk storage. Those storages should provide SCSI-3 Persistent Reservations and the requirement applies to other service running on WSFC too, like File Server. And this is supported on vSAN native from 6.7 U3 or later only.

Starting from vSAN 6.7, iSCSI service provides the opportunity for those applications to run clustered services like FCI and File Server on vSAN.

We verified the following functions and demonstrated the capability of Failover Cluster and the advantages of the vSAN iSCSI service.

# **Application Role Failover**

We verified the application role failover for both SQL Server FCI and File Server. The application role failover time was less than 20 seconds for the SQL Server FCI and less than 5 seconds for the File Server. The failover for the SQL Server FCI was disruptive because the FCI was restarted on the SQL Server FCI passive node.

# Non-interruptive File Copy During Role of Scale-Out File Server

Though the file shared is built on the WSFC, unlike the FCI on WSFC, the Scale-Out File Server of Windows Server 2016 provides the capability of non-disruptive role failover, according to our test result. In our validation, we failed over the node of the File Server while starting file copy session, the role failover was completed in 4-5 seconds and the copy session continued running without any interruption.

## vMotion WSFC Nodes to One Host

You can run multiple server node VMs of WSFC on the same host, which is useful when a host fails. In that situation, you may have multiple server nodes VMs available to support Windows level high availability. We verified this by vMotioning one service node VM (for example: SQL Server Active Node) to the host which hosted another service node VM (for example: SQL Server Passive Node). Multiple server node VMs is not supported on the SAN configuration [3].

However, letting the VMs on the same host may cause the compute resource contention for mission-critical applications like SQL Server. To avoid resource contention, we kept a 1-1 mapping of a server node VM to a single host. We recommend using DRS VM-Host Affinity Rule to separate the location of the VMs on different WSFC nodes. For testing purposes, we used PowerShell scripts to automatically reallocate the resources to keep this 1-1 mapping to guarantee the mission-critical applications had enough compute resource during testing. This script is available for download, see **Automatically assign a VM to a host with 1:1 mapping to avoid performance contention on vSAN**.

**Recommendation** : Use DRS VM-Host Affinity Rule to separate the location of the VMs on different WSFC nodes.



# **Best Practices**

The following best practices are recommended for the configuring the vSAN iSCSI service for SQL Server Failover Cluster Instance and File Server on a vSAN Cluster.

- Do not have more than 16 targets in one vSAN cluster
- Do not have more than 128 LUNS in one vSAN cluster
- Do not have more than 128 sessions on each host in a vSAN cluster
- Separate the vSAN traffic and iSCSI traffic across different physical network cards
- For a mission-critical application, we recommend using one target with one LUN mapping, even though multiple LUNs under one target can simplify management.
- We recommend setting multiple paths for each target on the guest OS. And set the following recommended values using PowerShell. See the Set-MPIOSetting for the default settings. We highly recommend using the scripts found at Configuration in VM to set up iSCSI initiator and cluster shared disks for SQL Server and File Server to set the values found in Table 9.

#### **Table 9. MPIO Setting Recommendations**

Settings	Description	Recommended Value
Custom Path Recovery	The value to specify whether MPIO performs custom path recovery.	Enabled
NewPathRecoveryInterval	The value to specify a customer path recovery time in seconds.	20 seconds
NewRetryCount	The value to specify the number of times to retry an IO request.	60
NewPDORemovePeriod	The value to specify the physical device object (PDO) removal period in seconds.	60 seconds
NewPathVerificationPeriod	The value to specify the length of time to verify a path in seconds.	30 seconds
NewDiskTimeout	The value to specify the length of time the server waits before it marks an IO request as timed out.	60 seconds

- We only support active/passive high availability of iSCSI service; the only supported failover policy is Failover Only (FOO). We highly recommend using the script found at Configuration in VM to set up iSCSI initiator and cluster shared disks for SQL Server and File Server to set the policy so that the newly discovered disks can identify and leverage the policy automatically
- Do not set the multiple path number of each virtual disk to more than the value of FTT+1. This setting is to limit the session number while keeping multiple paths alive if there is host failure
- iSCSI vSAN allows multiple clustered Windows nodes to reside on the same host; this is one of the benefits of the solution. However, it might hurt the mission-critical application workloads when multiple nodes are on the same host. To avoid this, we suggest using DRS VM-Host Affinity Rule or using PowerShell scripting to balance the VM placement automatically.
- We recommend enabling Jumbo Frames on both the physical switch and distributed switch by setting the MTU to 9,000 for better performance.



# Conclusion

VMware vSAN 6.7 iSCSI Target Service supports both SQL Server Failover Instances and Scale-out File Server seamlessly with transparent iSCSI target failover.

In addition:

- NVMe SSDs can be used as a high-performance caching device for write-intensive workloads. Building on vSAN iSCSI service with backend vSAN storage, customers can see immediate benefits for applications like OLTP workload and File Server with low latency. The mixed workload can run simultaneously with high transaction rate and file copy rate.
- With PowerShell and PowerCLI scripts, the configuration of iSCSI storage for shared storage services like SQL Server FCI and File Server in guest OS can be automatic and less error-prone.
- vSAN Health Service and vSphere advanced options like DRS VM-Host Affinity Rule can be used to monitor and manage the usage of FCI and File Server on vSAN. Meanwhile, PowerShell and PowerCLI scripts provide powerful management capability and more precise control and automation in testing environments. We have provided links to sample code for this use.
- Application role failover using the shared disk on the vSAN iSCSI Target Service is verified in this solution. Scale-Out File Server provides non-interruptive file copy capability during role failover, and vMotion can be leveraged for iSCSI shared disk for SQL Server or Scale-out File Server on a single host and provide the capability of running Windows clustered services on the same host.



# References

- Installing and Configuring target iSCSI server on Windows Server 2012 [Online]. Available: https://secureinfra.blog/2012/03/30/installing-and-configuring-target-iscsi-server-on-windows-server-2012-server -8-beta/
- 2. M. Senel, "Target, Configure Windows 2012 R2 As Iscsi," 01 April 2018. [Online]. Available: https://muratsenelblog.wordpress.com/2016/02/25/microsoft-scale-out-file-server/
- "vSphere MSCS Setup Limitations," [Online]. Available: https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.mscs.doc/GUID-6BD834AE-69BB-4D0E-B0B6-7E176907E0C7.html
- "Performance characterization of Microsoft SQL Server on VMware vSphere 6.5," [Online]. Available: https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/performance/sql-server-vsph ere65-perf.pdf



SQL Server FCI and File Server on VMware vSAN 6.7 using iSCSI Service

# About the Author and Contributors

Tony Wu, Senior Solution Architect in the Product Enablement team of the Storage and Availability Business Unit, wrote the original version of this paper.

**Ellen Herrick, Technical Writer** in the Product Enablement team of the Storage and Availability Business Unit, edited this paper to ensure that the contents conform to the VMware writing style.



