TECHNICAL WHITE PAPER
February 2026

Architecting Microsoft SQL Server for High Availability on VMware Cloud Foundation®

BEST PRACTICES GUIDE

FEBRUARY 2026

vmware® by Broadcom

Table of contents

List of Tables

Introduction

We live in a data-driven world where businesses and end users generate and consume information at an unprecedented scale. Expectations for continuous access to this information have never been higher. Much of this data resides in Microsoft SQL Server, one of the most widely deployed database platforms. Ensuring that SQL Server remains highly available is fundamental to any mission critical deployment. In addition to availability, other important considerations include performance, scalability, and operational manageability. Deploying SQL Server correctly allows these requirements to be met consistently. Achieving the right balance between utilization, availability, performance, and manageability can be challenging, but it is attainable with proper planning. Virtualizing SQL Server on VMware Cloud Foundation (VCF) 9 offers a modern and flexible way to achieve these goals and is the preferred way, although some organizations still rely on traditional physical deployments.

Why virtualizing SQL Server deployments?

Physical hardware has long been a reliable method for deploying SQL Server. However, the physical deployment model often includes long procurement cycles, delivery delays, and the substantial work required after equipment arrives, such as racking, cabling, and preparing data center space. These steps can introduce significant delays between ordering a system and delivering it to end users. The resulting lack of agility can hinder business responsiveness and competitiveness.

Virtualization improves agility. Although virtualization has been in use for many years, some database professionals still question whether virtualizing database servers is the right approach. These concerns often originate from earlier periods in computing when the environment was fundamentally different:

- Much of the computing world was still 32-bit, which limited density and scalability.
- Virtual machines had strict limits on memory and processors, which led to poor outcomes for database workloads.
- Virtualization technologies were not as mature as they are today, and supporting tools lacked the capabilities available now.

These limitations no longer exist. Modern server platforms are 64-bit, and as of SQL Server 2016, Microsoft no longer ships a 32-bit edition of SQL Server. Physical servers now support large CPU core counts, multi-socket designs, and large-scale memory configurations. Only a small subset of workloads can fully saturate today's hardware, which means organizations often seek better ways to utilize the available capacity while improving agility, performance, and availability.

As of VCF 9, ESX supports the following compute and cluster maximums relevant for SQL Server virtualization:

- 768 virtual CPUs per virtual machine
- 24 TB of memory per virtual machine
- 960 logical CPUs per host
- 4,096 virtual CPUs per host
- 32 virtual CPUs per physical CPU cores
- Up to 24 TB of memory per host
- 1,024 powered-on virtual machines per host
- 96 hosts per VCF cluster
- 8,000 virtual machines per VCF cluster

See the complete list of maximum configurations for all currently supported versions of VMware products and solutions.

In the context of availability, virtualization provides options that complement the availability features provided by both Windows Server and SQL Server. Whether designing a straightforward availability solution for a single data center or extending the architecture to support disaster recovery across sites, the implementation must match business requirements.

Whether you are a SQL Server DBA, VMware administrator, architect, or IT decision maker, this document will help you understand how to architect highly available and mission critical SQL Server 2025 solutions using VCF 9. Please note that, although performance and agility aspects related to availability are discussed, this is not a performance guide. This document does not attempt to be exhaustive, but it covers the most important concepts and highlights the practices relevant to deployments on VCF 9. This document is intended to be used in tandem with the official VMware's Microsoft SQL Server-focused Performance Best Practices Guide and other references mentioned elsewhere in this document.

Supportability and Lifecycle Management

Achieving availability requires more than enabling features at the hypervisor or within the guest virtual machine. When deploying mission critical systems that must deliver both reliability and availability, it is necessary to ensure that every component in the end-to-end stack is supported by the vendors involved. VMware provides mechanisms to simplify this process through its Compatibility Guides. The VMware Compatibility Guide enables administrators to search for supported combinations of VMware products, versions, and hardware platforms. This process is demonstrated in the images below.

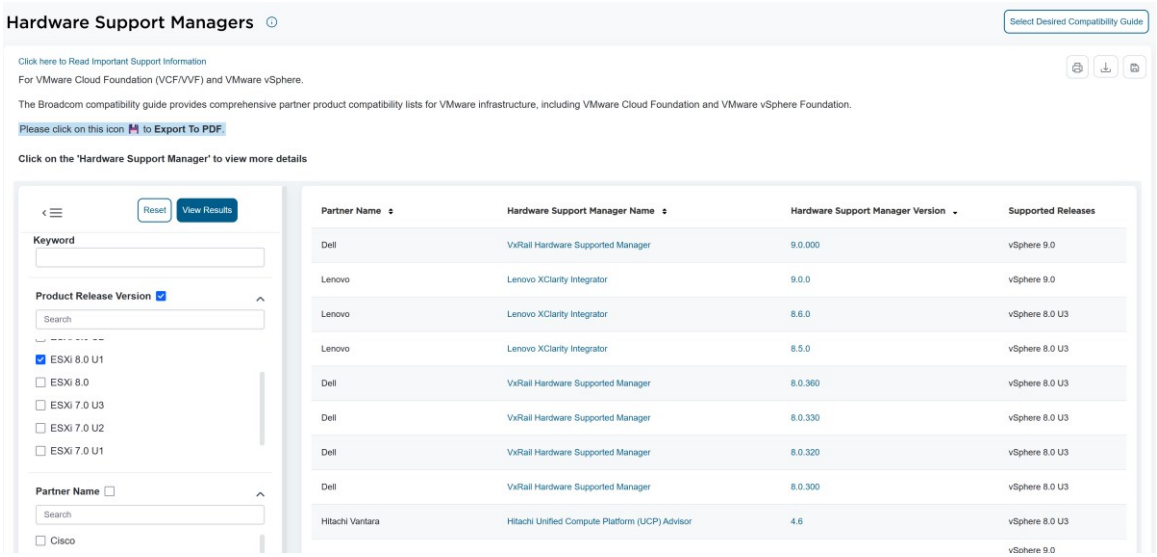


Figure 1 - Querying the VMware Compatibility Guide

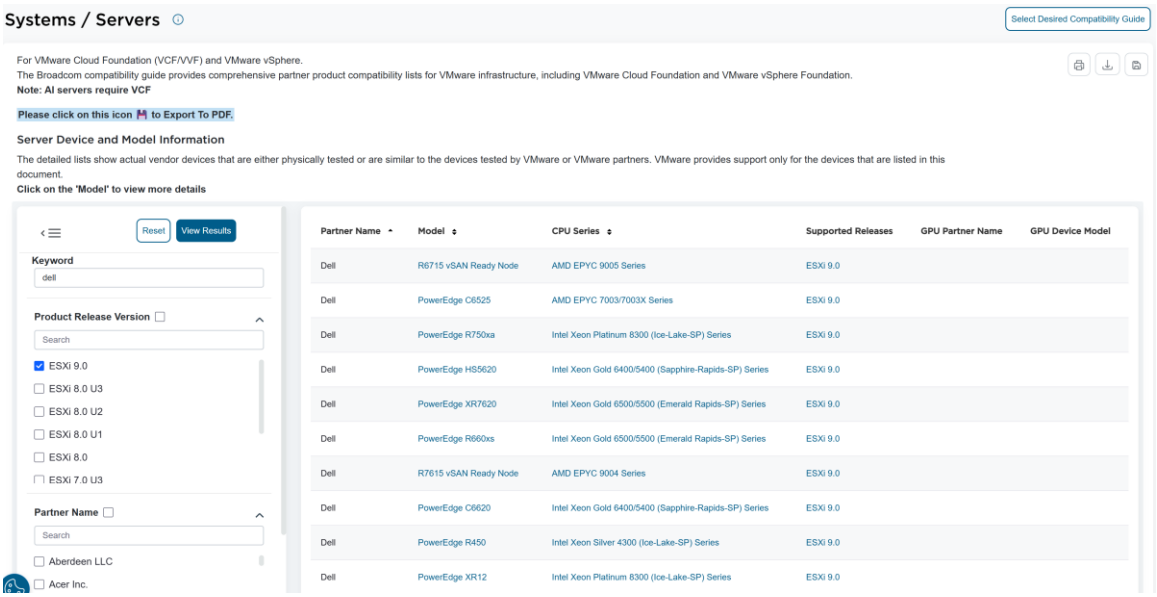


Figure 2 - Results of the Compatibility Guide Query

Supported versions of Windows Server can also be identified in the Compatibility Guide. When searching the Compatibility Guide, choose "**Guest OS**" from the **"Platform & Compute"** tile, then "**OS Family Name"**, as shown in the image below.

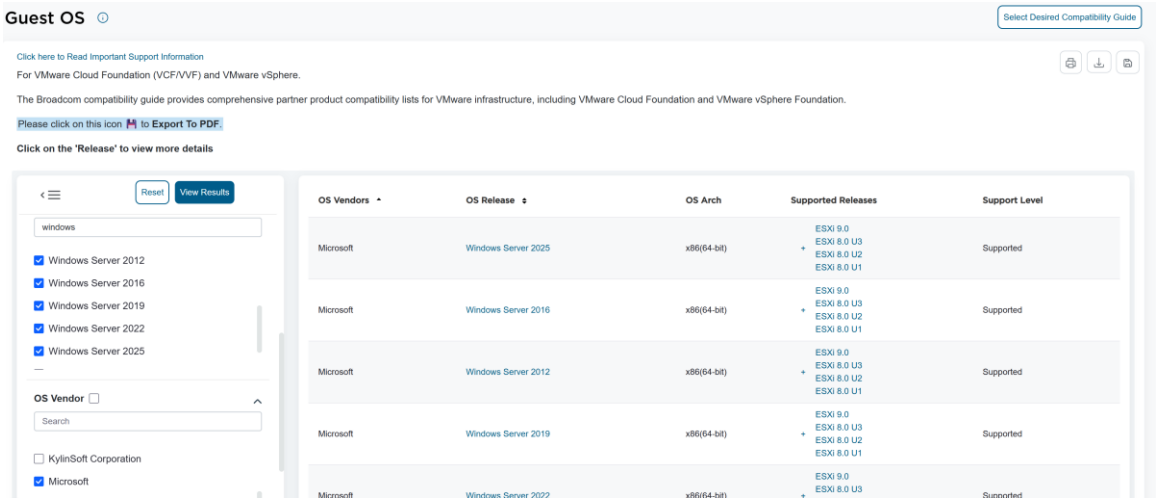


Figure 3 - Querying OS Version Support on VCF

SQL Server must also be supported by the underlying Windows Server version. Microsoft maintains a support matrix which outlines which editions of SQL Server are supported on which versions of Windows Server versions and editions (see "Operating system support" section).

Operating system support

The following table shows which editions of SQL Server 2025 (17.x) are compatible with which versions of Windows. You can also use the support lifecycle information to see if your version of Windows is supported.

Expand table

SQL Server edition:	Enterprise [1]	Standard [1]	Express
Windows Server 2025 (Support lifecycle)			
Windows Server 2025 Datacenter	Yes	Yes	Yes
Windows Server 2025 Datacenter: Azure Edition	Yes	Yes	Yes
Windows Server 2025 Standard	Yes	Yes	Yes
Windows Server 2025 Essentials	Yes	Yes	Yes
Windows Server 2022 (Support lifecycle)			
Windows Server 2022 Datacenter	Yes	Yes	Yes
Windows Server 2022 Datacenter: Azure Edition	Yes	Yes	Yes
Windows Server 2022 Standard	Yes	Yes	Yes
Windows Server 2022 Essentials	Yes	Yes	Yes
Windows Server 2019 (Support lifecycle)			

Figure 4 - SQL Server to Windows Server Support Matrix

Platform choices must be made with long-term supportability in mind. Every component should be supported across its lifecycle. Microsoft publishes lifecycle guidelines for all its products, including SQL Server and Windows Server, at:

As of the release of SQL Server 2025, Windows Server versions earlier than Windows Server 2019 are out of mainstream support. Windows Server 2016 is now in extended support. VMware recommends deploying Windows Server 2025 or Windows Server 2022 where application compatibility permits.

Ensuring that systems remain mission critical requires coordinated lifecycle alignment across the entire stack. Third-party application vendors may dictate the version of SQL Server or Windows Server that can be deployed. This requirement must be considered alongside Microsoft support boundaries. Operating outside mainstream support may introduce security, functional, or compatibility risks.

Organizations should deploy platforms that remain within mainstream support and plan to migrate when versions transition to extended support. Unsupported platforms may introduce operational challenges and security exposures. Another long-term concern is the difficulty of finding administrators skilled in managing deprecated platforms.

Finally, ensure compliance with the guidelines provided by Microsoft in its "Support policy for Microsoft SQL Server products that are running in a hardware virtualization environment" document. The cost of upgrading to newer versions of SQL Server, Windows Server, and VMware products is often offset by improvements that benefit the environment. Structured refresh cycles should be planned for both software and hardware to maintain operational continuity.

Planning for SQL Server Availability Features Under vSphere

This section provides an updated overview and technical understanding of SQL Server availability features as deployed inside virtual machines running on VMware vSphere and VMware Cloud Foundation (VCF) 9. It explains the layers of protection SQL Server offers, how those protections map to modern SQL Server 2025 high availability features, and how to successfully deploy them in a VMware environment that includes vSphere HA, Distributed Resource Scheduler (DRS), and the vMotion Application Notification framework.

The core objective of this section remains unchanged: to ensure that SQL Server workloads achieve predictable, supportable, and resilient high availability when virtualized, while still honoring the operational behaviors of both Windows Server Failover Clustering (WSFC) and the underlying VMware hypervisor platform.

Many of the foundational principles from earlier versions of this guide still apply - for example, the need to properly architect WSFC, manage quorum, and understand the implications of shared storage versus replica-based database protection. However, SQL Server 2025 introduces several enhancements that significantly affect deployment planning, including improvements to synchronous commit latency handling, automatic seeding reliability, backup/restore parallelism, and Contained Availability Groups (AGs), among others.

Similarly, VMware vSphere and VCF 9 have advanced their support for clustered workloads. DRS has evolved to provide more deterministic VM-to-host placement when clustering rules are applied, while vSphere HA includes more granular restart priority and dependency controls that help maintain WSFC stability during host outages. In addition, the introduction of vMotion Application Notification fundamentally changes how SQL Server VMs - particularly AG replicas and Failover Cluster Instance (FCI) nodes - react during live migration events. Proper configuration of this feature can dramatically reduce the risk of unintended cluster failovers or transient downtime during host maintenance or automated migrations.

Finally, VMware's storage capabilities have expanded beyond earlier generations. What were once specialized configurations - such as Clustered VMDKs, vVols with SCSI-3 PR support, and vSAN ESA - are now mature, mainstream options for hosting SQL Server 2025 workloads, including FCIs. These changes, along with updated Microsoft guidance around Windows Server clustering, require revisiting many legacy deployment recommendations to ensure continued alignment with supported practices.

The remainder of this section dives into each protection level, examines the SQL Server features that implement them, and details how to deploy and operate these features successfully in vSphere and VCF 9 environments. Each subsequent subsection retains the original flow and intent of the earlier guide, while fully updating technical expectations, behaviors, and constraints for modern versions of SQL Server, Windows Server, and VMware vSphere.

SQL Server Protection Levels

SQL Server provides three distinct levels of availability protection: instance-level, database-level, and data-level. Each level corresponds to a specific set of native high availability features, and each behaves differently when virtualized under vSphere. Understanding these levels - and how they map to modern SQL Server 2025 capabilities - is essential when designing resilient architectures under VMware vSphere or VMware Cloud Foundation (VCF) 9.

An installation of SQL Server is called an *instance*. Instance-level protection therefore refers to protecting the entire SQL Server installation: binaries, system databases, metadata, SQL Server Agent jobs, and all user databases hosted by that instance. SQL Server 2025 continues to support multiple instances on a single Windows Server Failover Cluster (WSFC) node, although the practical limit is governed not by Microsoft's maximums but by operational concerns such as manageability, performance stability, and recovery expectations. While SQL Server supports a large number of instances per Windows Server installation, deploying only the number required for your business remains the recommended practice.

SQL Server supports two instance types: **default** and **named**.

- A **default instance** on a standalone server inherits the underlying Windows Server name. If the server is named *ALLAN*, then applications connect to the SQL Server instance using simply ALLAN.
- A **named instance** adds a unique identifier that appears after a backslash (for example: ALLAN\SALES or ALLAN\REPORTING). This identifier is internal to SQL Server and not tied to hostname registrations outside the instance itself.

In clustered environments - specifically Failover Cluster Instances (FCIs) - default and named instance behavior differs because an FCI uses a *virtual* network name that is independent of the underlying WSFC node names. Details of this behavior will be discussed in Section 3.2.2.

Database-level protection ensures that a synchronized copy of a database exists on another SQL Server instance, typically on another server, VM, or cluster node. In SQL Server 2025, this functionality is primarily implemented through **Always On Availability Groups (AGs)**, which now include improvements such as:

- More resilient **automatic seeding**, including retry logic
- Enhanced **latency-aware synchronous commit**
- Better **failover readiness health evaluations**
- Improved **read routing** behavior across large replica sets
- Expanded capabilities for **Contained AGs**, allowing metadata, logins, SQL Agent jobs, and credentials to move with the AG during failover

Although AGs provide seamless database-level protection, they do not automatically synchronize all instance-level objects. Any metadata not encapsulated within a Contained AG must be created manually on each replica unless using SQL Server 2025's expanded containment features.

Basic Availability Groups (Basic AGs) - available on SQL Server Standard Edition - remain a reduced-scope version of full AGs. They support one database, two replicas, and no read-scale functionality. These limitations still apply in 2025, though reliability during failover and automatic seeding has improved. Basic AGs are typically recommended only for small or entry-level environments.

Data-level protection covers features that replicate data between SQL Server instances without guaranteeing instance- or database-level high availability. SQL Server Replication (snapshot, merge, or transactional) continues to serve this purpose. However, replication is not intended for high availability and does not guarantee complete database fidelity, metadata synchronization, or transactional failover consistency. It is, therefore, generally unsuitable as an HA mechanism for business-critical applications.

The relationship between these three protection levels and their associated SQL Server high-availability features remains the same conceptually, though SQL Server 2025 significantly enhances the AG-based architecture. Table 1 below maps each protection level to the corresponding SQL Server availability feature, including the most current technologies and their supported configurations.

Protection Level	**SQL Server Availability Feature**
Instance	- Always On Failover Cluster Instance (FCI) - Enhanced TempDB Resiliency (SQL 2022+) - Improved Startup/Recovery & Crash-Resilience (SQL 2025)
Database	- Always On Availability Groups (AGs) - Advanced AG Enhancements: Parallel Redo, Optimized Sync Commit,

	Accelerated Log Transport (SQL 2025) - Basic Availability Groups - Log Shipping - Read-Scale AGs
Data	- Transactional Replication - Snapshot Replication - Merge Replication - Change Tracking & CDC Enhancements (SQL 2022–2025)
Deprecated/Legacy	- Database Mirroring (Removed in SQL Server 2025) - Legacy Failover Partnership Models

Table 1 - Protection Mapping to SQL Server Availability Features

The following subsections continue to focus primarily on modern deployments of Always On Failover Cluster Instances (FCIs) and Always On Availability Groups (AGs), since these remain the dominant methods for deploying SQL Server high availability on VMware vSphere and VCF 9. Both FCIs and AGs require WSFC, meaning every SQL Server HA deployment must begin with proper planning and configuration of Windows Server Failover Clustering inside the VM. Without a functioning WSFC foundation, neither an FCI nor an AG can be deployed.

Figure 6 below illustrates the modern clustering stack as it appears inside a VM when SQL Server 2025 high availability features are deployed under vSphere.

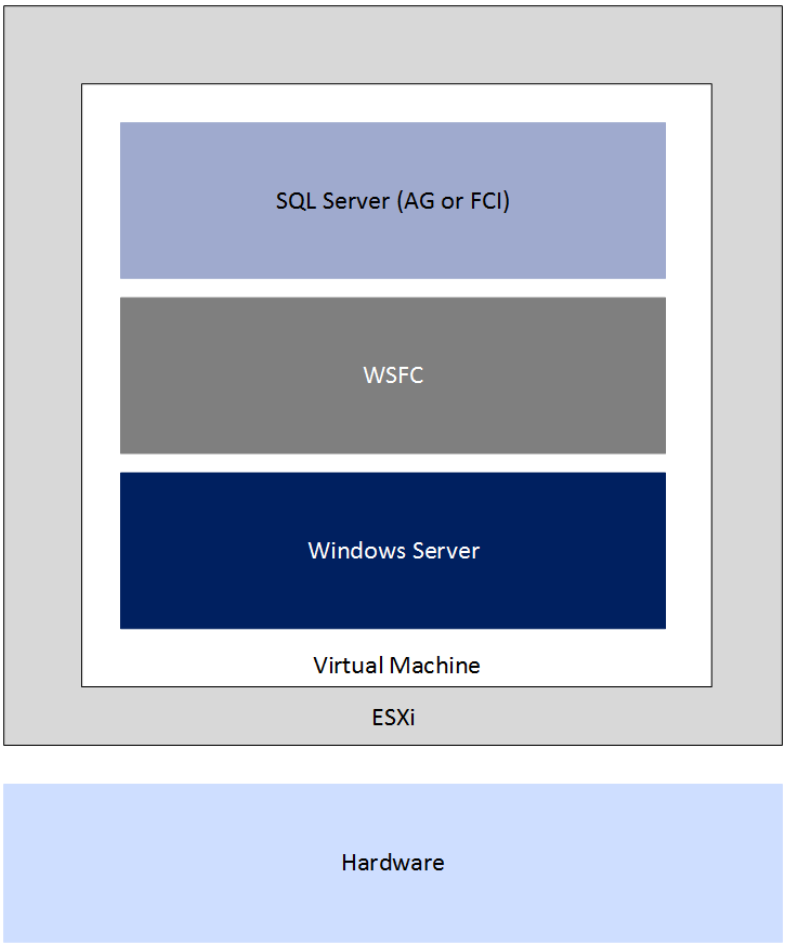


Figure 5 - Clustering Stack for SQL Server as it Relates to vSphere

Database Mirroring, while technically still present in SQL Server for backward-compatibility reasons, remains deprecated and should not be used in any new deployment. Log Shipping continues to function and can still be used as a lightweight DR mechanism. Replication retains its role for data-level distribution. These non-primary features will only be referenced as needed, as they do not materially affect vSphere-based clustering design.

Always On remains Microsoft's current branding umbrella for all clustering-based SQL Server features, including both FCIs and AGs. This naming convention continues to be used in SQL Server 2025. As in previous versions, the terms "SQL Clustering," "AOAG," "AAG," or other incorrect abbreviations should be avoided; this document adheres strictly to Microsoft naming conventions.

Planning Clustered Implementations of SQL Server under vSphere

A Windows Server Failover Cluster (WSFC) establishes the foundational availability layer required for both Always On Failover Cluster Instances (FCIs) and Always On Availability Groups (AGs). Because these SQL Server high-availability features rely on a functioning WSFC, proper planning of the underlying cluster architecture is essential. This applies equally when running SQL Server on VMware vSphere or VMware Cloud Foundation (VCF) 9, where virtualization behaviors - such as DRS migrations, vSphere HA failovers, and storage abstraction - can influence WSFC stability if not designed appropriately.

A WSFC provides availability, not load balancing or horizontal scale-out. While AG readable secondaries may offer scale-out capabilities, WSFC itself does not distribute workload. SQL Server deployments under vSphere must be sized with the assumption that any WSFC node can become active due to maintenance, migrations, or failures.

Nodes must be configured identically: CPU topology, memory, virtual hardware version, VMware Tools, NIC layout, storage controllers, and

reservations. When virtualizing SQL Server 2025 on vSphere or VCF 9, administrators must consider DRS and vSphere HA interactions. Automated balancing, host maintenance evacuations, and anti-affinity enforcement directly affect WSFC stability. vMotion Application Notification substantially reduces transient heartbeat failures during live migration and should be enabled for AG replicas and FCI nodes.

Cluster validation remains necessary, but many warnings are benign in VMware environments (e.g., SET, physical NIC assumptions). Administrators should consult the following VMware's supported clustering guides and documentation for guidance:

- Setup for Windows Server Failover Clustering on VMware vSphere
- Guidelines for Microsoft Clustering on vSphere
- Microsoft Windows Server Failover Clustering (WSFC) with shared disks on VMware vSphere 7.x: Guidelines for supported configurations

The goal remains a WSFC that is stable and predictable across planned and unplanned events, aligned with SQL Server 2025 availability features and VMware hypervisor-level capabilities.

Deploying a Windows Server Failover Cluster

This section provides updated guidance for planning and deploying a Windows Server Failover Cluster (WSFC) to support SQL Server 2025 Always On Availability Groups (AGs) or Failover Cluster Instances (FCIs) under VMware vSphere and VMware Cloud Foundation (VCF) 9. Because AGs and FCIs depend on a stable WSFC, proper planning of cluster architecture, node configuration, networking, and domain integration is critical.

SQL Server 2025 FCIs require Microsoft AD and DNS for virtual network names, while Domain-Independent AGs and AD-detached clusters are supported for AG deployments that do not require a listener.

Cluster validation remains mandatory for supportability; however, several warnings are expected in VMware environments - particularly those related to SET, physical NIC validation, and storage direct-attach assumptions.

Networking must be redundant at the vSphere layer. A single vNIC per VM is fully supported if the backing VMware Distributed Switch (VDS) or NSX-T segment provides genuine physical uplink redundancy. WSFC heartbeats rely on stable, low-latency delivery, and redundancy must be implemented at the hypervisor layer rather than by adding multiple vNICs within the VM.

WSFC quorum configuration should leverage modern options such as dynamic quorum, dynamic witness, and cloud witness. When combined with vSphere HA isolation response behaviors, node placement, and anti-affinity settings, quorum design ensures the cluster remains operational during host failures or isolation events.

Mixed clusters composed of physical and VM nodes remain supported as long as validation is satisfied with VMware-supported exceptions. This model is useful during phased migrations but introduces additional considerations for latency and networking consistency.

Windows Server Edition, Number of Nodes, and Supportability

Windows Server 2008 R2 and earlier require Enterprise or Datacenter Editions to be able to create a WSFC (formerly known as Microsoft Clustering Service (MSCS) in these versions). Windows Server 2012 and later can use either Standard or Datacenter Editions.

Windows Server 2012 and later can support up to 64 servers, known as nodes, in one WSFC. VMware ESXi™ supports up to 16 VMs in a single FCI's configuration.

Current Microsoft documentation treats FCIs and AGs as mutually exclusive and could potentially be interpreted in different ways for the total number of nodes. For a configuration with shared storage, meaning an FCI, the maximum number of nodes depends on the SQL Server edition. For SQL Server 2022 Enterprise edition, it is whatever the Windows Server edition supports. For SQL Server 2022 Standard edition, the maximum is 2 nodes.

See the following applicable references for more details:

- Editions and supported features - SQL Server 2016 | Microsoft Learn
- Editions and supported features - SQL Server 2017 | Microsoft Learn
- Editions and supported features of SQL Server 2019 - SQL Server | Microsoft Learn
- Editions and supported features of SQL Server 2022 - SQL Server | Microsoft Learn
- Editions and supported features of SQL Server 2025 – SQL Server | Microsoft Learn

Since an AG can support up to the maximum number of replicas, the total number of VMs in a WSFC configuration should not be a concern, even if combining one or more FCIs are part of an AG. However, if there is any confusion, you should contact VMware support to ensure that your proposed solution is one that will be supported. Since Microsoft supports up to 64 nodes in a WSFC and SQL Server's number of nodes varies per feature, your limitation will come in based on if you are deploying an FCI and whether you have shared storage.

To ensure that a WSFC configuration is supported by Microsoft introduced a cluster validation wizard in Windows Server 2008. It validates the proposed configuration of the WSFC before creating it. If the validation tests show that the configuration is fine, the VMs can be clustered as a WSFC.

As virtualization became widely adopted, Microsoft introduced the option to skip certain parts of the cluster validation process and interpret certain findings differently, particularly for components that are outside the scope of Microsoft's control or responsibility. The support for these components has shifted to third-party vendors like VMware. Microsoft also updated its guidance on required actions to take when a test fails. In all, Microsoft no longer prevents cluster setup solely based on a failed test report or when tests are skipped.

VMware continues to recommend that customers perform cluster validation before setting up their cluster on VMs hosted on VMware vSphere. When using Raw Device Mapping (RDM) for storage presentation to the nodes, the cluster validation wizard may report an error on the storage test. This is because the Wizard has no notion of this valid storage option. It is ok to skip this test entirely or ignore the warning.

Microsoft deprecated some clustering-related features in Windows Server 2022. However, the cluster validation Wizard still checks for some of these features when validating a cluster setup on Windows Server 2022. For example, Microsoft removed the "LBFO" feature and replaced it with "Switch Enabled Teaming Configurations", a Microsoft hypervisor feature which does not exist in VMware vSphere. Because the wizard checks for this feature on all Windows OS instances, regardless of the underlying hypervisor platform, the wizard generates a **"Switch Enabled Teaming (SET) Configurations"** warning during the validation process.

Although this has been fixed in Windows Server 2025 WSFC (the "SET" test is now skipped and passed), the VMware's cluster support documentation referenced above should be used to interpret validation results to distinguish between true faults and benign virtualization-induced warnings.

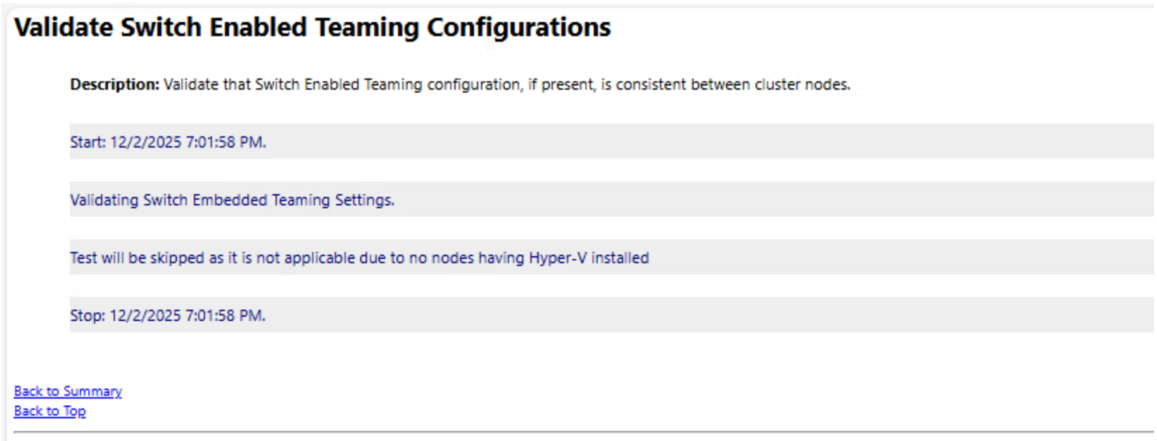


Figure 6 - Validation Wizard Checks for SET

Cluster In a Box, Cluster Across Boxes, Number of Hosts, and Node Placement

VMware vSphere supports two primary placement models for Windows Server Failover Cluster (WSFC) nodes: Cluster In a Box (CIB) and Cluster Across Boxes (CAB). These models have significant implications for SQL Server 2025 Always On Failover Cluster Instances (FCIs) and Availability Groups (AGs) deployed in virtualized environments, particularly when combined with modern vSphere 9/VCF 9 capabilities such as Distributed Resource Scheduler (DRS), vSphere High Availability (HA), and vMotion Application Notification.

CIB places all WSFC nodes on the same ESXi host. This model is not recommended for production because it aggregates all failover risk to a single hypervisor. A host failure results in complete cluster outage, loss of quorum, and interruption of SQL Server services. Although vSphere HA can restart VMs on another host, this does not constitute a high availability configuration since all SQL Server resources become unavailable during such an event. CIB may be acceptable for development or lab scenarios but should not be used for FCIs with shared storage or for production AG deployments.

CAB places WSFC nodes on separate ESXi hosts, ensuring that a single host failure does not impact the entire cluster. CAB is the standard and VMware-supported architecture for production SQL Server clustering under vSphere or VCF. CAB requires the use of DRS VM-to-VM anti-affinity rules to prevent cluster nodes from being placed on the same host inadvertently. With VCF 9, DRS enforces these rules more predictably, even during host maintenance evacuations or automatic balancing operations.

An “N+1” configuration is strongly recommended when clustering workloads in VCF. For example, for a 2-Node CAB-based SQL Server FCIs, a minimum of three ESXi hosts is recommended. This ensures that if one host fails, one node remains available and the failed node can be restarted elsewhere without violating anti-affinity constraints. For AG deployments, the same principle applies: each synchronous replica should reside on a separate host to prevent loss of high availability guarantees.

vSphere HA interacts closely with WSFC. During a host failure, surviving WSFC nodes detect the outage and fail over SQL Server resources at the cluster layer while vSphere HA restarts affected VMs on another host. Consistency between VMware’s HA restart policies and WSFC’s quorum rules is critical. VMware strongly recommends reserving memory for WSFC nodes to avoid ballooning or swapping, which can disrupt WSFC heartbeat stability and AG synchronization.

vMotion Application Notification, introduced in vSphere 8 and integrated fully in vSphere 9, significantly improves the stability of clustered SQL Server workloads during live migration events. Prior to migration, the guest OS receives a pre-migration notification that SQL Server or a custom agent can use to prepare for transient pause conditions. This allows synchronous AG replicas, FCI nodes, and latency-sensitive workloads to maintain cluster stability throughout the vMotion process. For SQL Server 2025 workloads, Application Notification is strongly recommended for all HA nodes.

Proper CAB deployment ensures predictable SQL Server HA behavior by aligning cluster placement with hypervisor operations, DRS rule enforcement, storage independence, and vSphere HA restart logic. This makes CAB the only suitable model for production SQL Server 2025 clustering under vSphere and VCF 9.

Always On Failover Cluster Instances

Always On Failover Cluster Instances (FCIs) remain a central SQL Server high availability technology under SQL Server 2025 and continue to rely heavily on Windows Server Failover Clustering (WSFC). FCIs provide instance-level protection, which means that all databases, system metadata, SQL Server Agent jobs, and instance-wide configurations fail over together as a cohesive unit.

With SQL Server 2025, FCIs benefit from engine improvements such as enhanced startup consistency checks, streamlined master database recovery, improved TempDB initialization, and faster metadata validation during failover operations. These updates reduce failover times and increase the predictability of instance-level recovery events - an important enhancement for virtualized clusters.

VMware vSphere continues to support FCIs through multiple shared-storage models. Historically, Raw Device Mappings (RDMs) were required to support SCSI-3 Persistent Reservations (PRs), but modern alternatives are now preferred. Current VMware recommendations emphasize the use of:

- Clustered VMDKs, supported on vSphere 7.0 and later
- vSAN ESA with native SCSI-3 PR support
- vVols, which provide SCSI-3 PR capability and policy-based management

Note: Although vVols continues to be supported for VCF5.x, VVF5.x and vSphere 8.x (until their respective end-of-support dates), it is no longer supported or marketed for VCF 9 and future releases.

These storage models eliminate the operational complexity of RDMs while providing reliable FCI storage behavior. RDMs remain supported but are increasingly discouraged due to their limited flexibility, lifecycle constraints, and poor alignment with modern vSphere storage features such as Storage vMotion and snapshots.

In VMware environments, FCIs must be deployed using anti-affinity rules to prevent both nodes from residing on the same ESXi host. A Cluster Across Boxes (CAB) design ensures that a host failure does not simultaneously remove both FCI nodes from service. As with AG deployments, vSphere DRS in VCF 9 enforces anti-affinity rules more consistently, ensuring correct node separation even during host maintenance mode evacuations.

vSphere HA plays an important supporting role in FCI architectures. If a host fails, vSphere HA restarts the affected node on another host, allowing the surviving node to control the FCI resource group until the restarted node rejoins the cluster. Proper memory reservation is critical: any ballooning, swapping, or contention on an FCI node can disrupt WSFC heartbeats and increase failover risk.

vMotion Application Notification significantly enhances FCI stability during live migrations. When enabled, the guest OS receives a pre-migration event, giving SQL Server an opportunity to prepare for the brief pause condition associated with vMotion. This reduces the risk of false WSFC heartbeat failures, unintentional failovers, and periods of service unavailability. FCIs hosting synchronous workloads benefit the most from this capability.

In SQL Server 2025, FCIs remain an excellent choice for applications requiring instance-level protection, full metadata portability, and simplified storage semantics. When deployed correctly under vSphere and VCF 9 - with proper storage selection, DRS anti-affinity, memory reservations, and Application Notification - FCIs deliver highly predictable and resilient availability behavior within virtualized environments.

Storage for FCIs

Storage design for Always On Failover Cluster Instances (FCIs) under SQL Server 2025 has evolved significantly, particularly in virtualized environments built on VMware vSphere or VMware Cloud Foundation (VCF) 9. Modern VMware storage platforms now provide multiple fully supported methods for delivering the shared-disk semantics required by FCIs, without relying on legacy Raw Device Mappings (RDMs).

The preferred storage mechanism for FCIs in 2025 is Clustered VMDKs, a mature VMware technology that provides virtual SCSI-3 Persistent Reservation (PR) capability. Clustered VMDKs are supported on vSphere 7.0 and later and eliminate many of the operational constraints that previously limited shared-disk clustering. Clustered VMDKs offer the following advantages:

- Full compatibility with vMotion (subject to cluster configuration)
- Centralized storage management using standard datastore constructs
- Support for Storage vMotion in maintenance workflows
- No reliance on direct LUN mappings or complex RDM lifecycle procedures

For environments using VMware vSAN, the vSAN Express Storage Architecture (ESA) now delivers native SCSI-3 PR support with improved latency, throughput, and resiliency for SQL Server 2025 workloads. vSAN ESA implements distributed storage with advanced caching, compression, and I/O optimization, making it a strong option for clustered SQL Server deployments requiring predictable failover behavior and simplified management.

Virtual Volumes (vVols) are another fully supported storage model for all currently supported versions of vSphere, except VCF 9. vVols offload storage intelligence to the array and provide granular policy-based provisioning while supporting SCSI-3 PR requirements. Administrators can assign performance and availability policies per disk, giving greater control over IOPS guarantees, redundancy, and failover behavior.

Traditional Raw Device Mappings (RDMs) remain technically supported for FCIs; however, they are no longer recommended. RDMs reduce operational flexibility, complicate Storage vMotion workflows, and are increasingly deprecated across VMware design practices. RDMs should only be maintained for legacy clusters pending migration to Clustered VMDKs, vVols, or vSAN ESA.

Storage layout for FCIs running SQL Server 2025 should continue to follow Microsoft and VMware best practices:

- Separate VMDKs for data, log, TempDB, and backup targets when appropriate
- Avoid unnecessary consolidation of workloads on a single datastore
- Ensure consistent storage controller types (e.g., VMware Paravirtual SCSI)
- Distribute VMDKs across multiple SCSI controllers to improve parallelism

TempDB placement remains important for FCI performance. While TempDB can reside on shared cluster storage to preserve consistency across failover, many organizations place TempDB on non-clustered, local VMDKs for performance reasons. SQL Server 2025 improves TempDB startup and allocation behavior, making this configuration more resilient. However, placing TempDB on local disks means TempDB is recreated on failover rather than preserved; applications must tolerate this behavior.

VMware HA interacts with FCI storage during host failures. If a host fails, the surviving node maintains control of the shared disks while vSphere HA restarts the failed FCI node on another host. Storage paths must be fully redundant across hosts, and multipathing best practices should be followed to ensure the restarted node can reacquire the necessary disks.

Selecting the appropriate shared storage option for an FCI - Clustered VMDKs, vVols, or vSAN ESA - ensures SQL Server 2025 clusters maintain predictable failover behavior, simplified lifecycle operations, and alignment with modern VMware platform capabilities.

Using Native VMDKs for FCI on VCF

Considerations and limitations for using Clustered VMDKs are detailed in the "Limitations of Clustered VMDK support for WSFC" section of VMware vSphere Product Documentation.

With a few restrictions, you can enable Clustered VMDK support on existing VMFS Datastore. Because Clustered VMDK-enabled Datastores are not intended to be general-purpose Datastores, VMware recommends that, where possible and practical, customers should create new dedicated LUNs for use when considering Clustered VMDKs.

The most common use envisioned for this feature is the support for shared-disk Windows Server Failover Clustering (WSFC), which is required for creating SQL Server Failover Clustering Instance (FCI).

If you must re-use an existing Datastore for this purpose, VMware highly recommends that you migrate all existing VMDKs away from the target Datastore, especially if those VMDKs will not be participating in an FCI configuration. VMware does not support mixing shared VMDKs and non-shared VMDKs in a Clustered VMDK-enabled Datastore.

You can enable support for Clustered VMDK on a Datastore only after the Datastore has been provisioned.

The process is as shown in the images below:

1. In VMware vCenter® Client, select the datastore for which you want to enable "**Clustered VMDK**".
2. Click on the "**Configure**" tab.
3. Click on the "**General**" section.
4. Click "**Enable**" on the "Clustered VMDK section", as shown below.

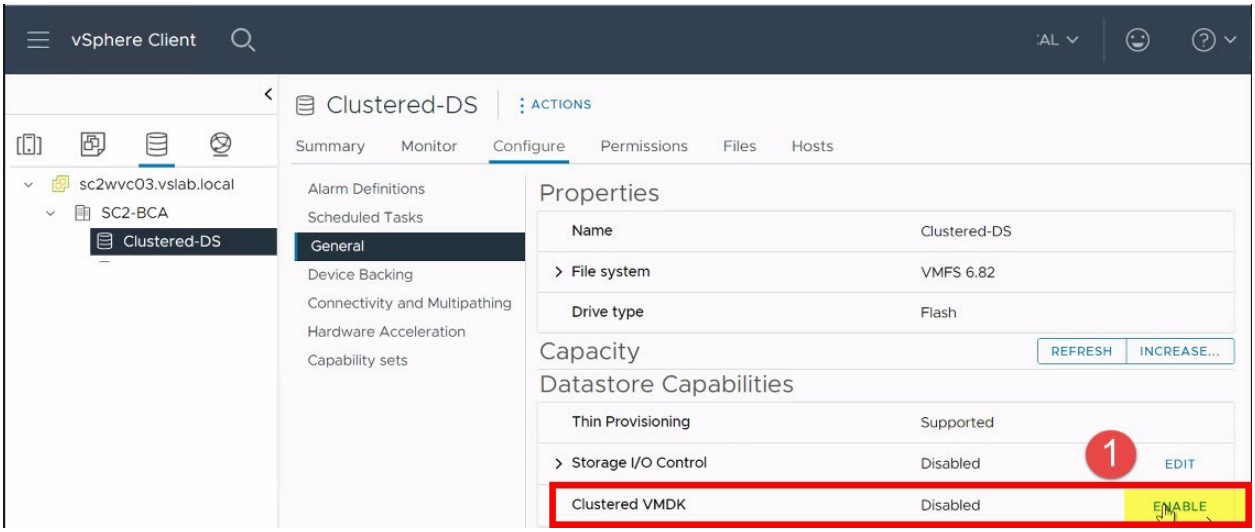


Figure 7 - Enabling Clustered VMDK (Step 1)

5. Read the warning, then Click "**Enable**" to commit the change.

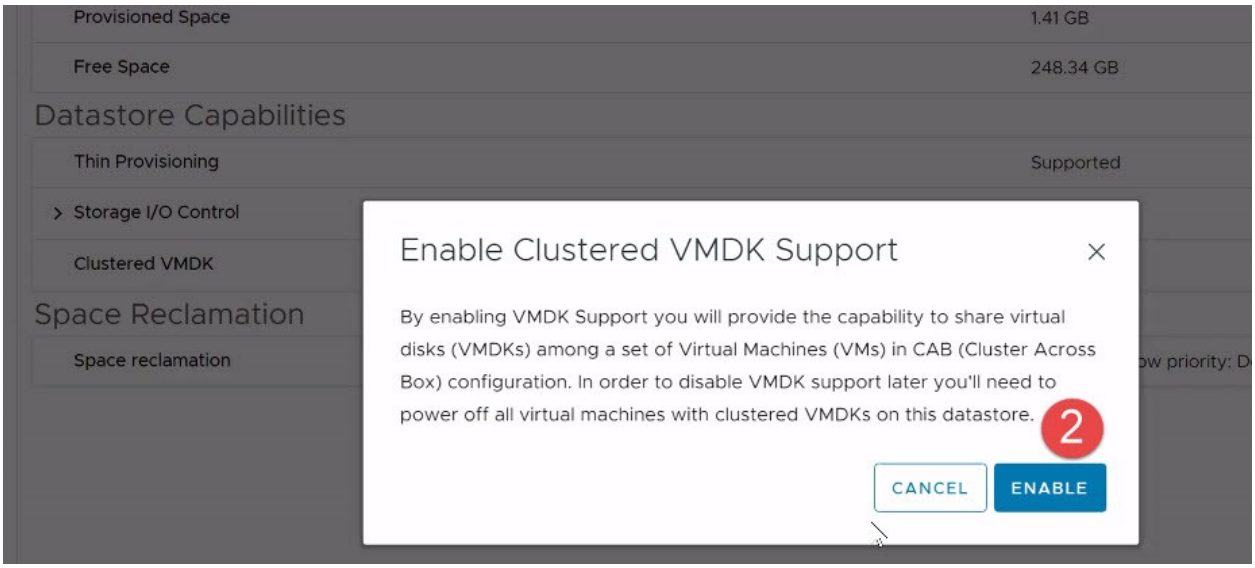


Figure 8 - Enabling Clustered VMDK (Step 2)

6. Verify that the status on "Clustered VMDK" is now set to "Enabled".

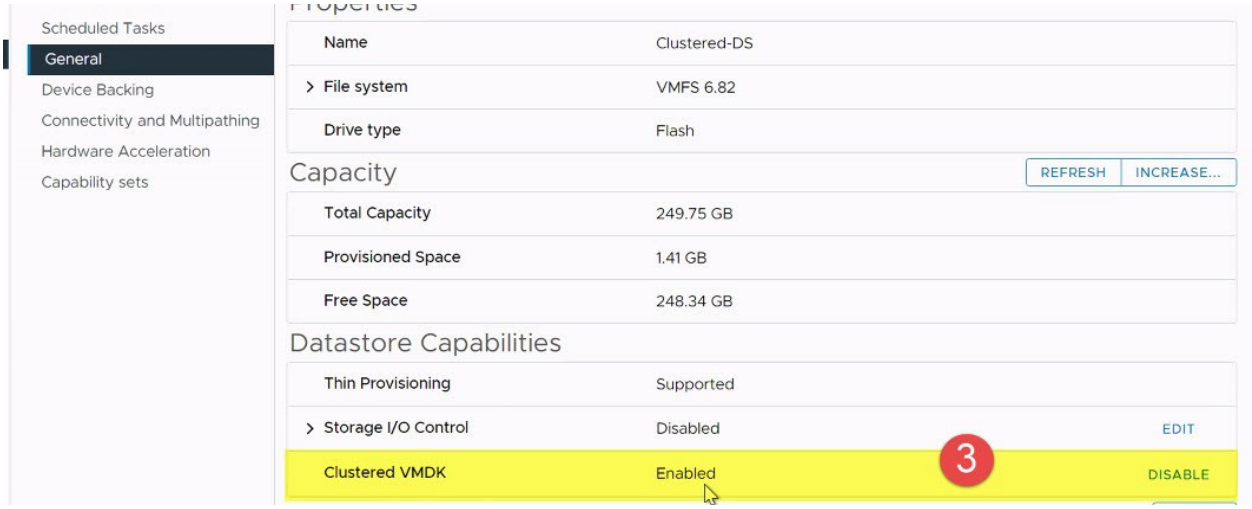


Figure 9 - Enabling Clustered VMDK (Step 3)

Provisioning Shared VMDK on First FCI Node

The process of creating and allocating VMDKs in a "Clustered VMDK" datastore follows the standard vSphere procedures for creating disks. Below are the steps required to create a VMDK and allocate it to multiple VMs in a vSphere environment.

1. Right-click the first VM that will share the disk and select **Edit Settings**.
2. From the **Add New Device** menu, select **Hard Disk**.

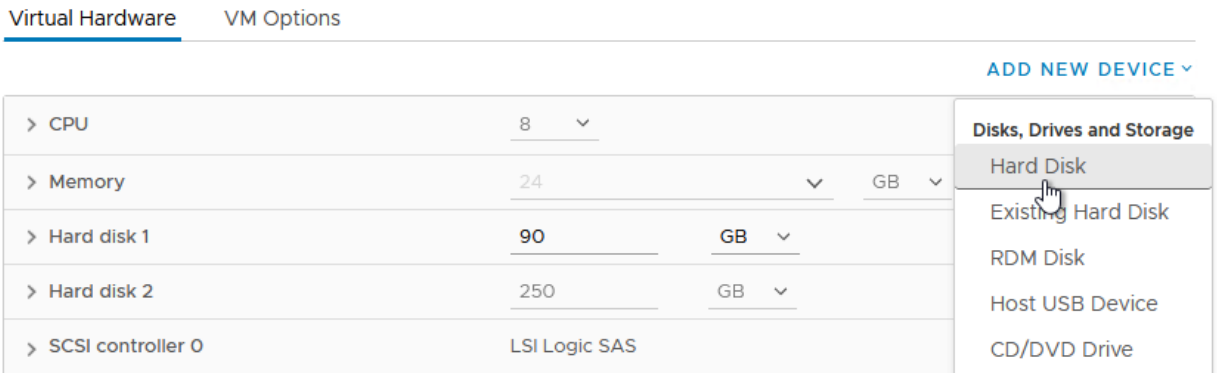


Figure 10 - Add New Shared VMDK Disk to a VM

3. Expand **New Hard disk** and set the **Disk size** to your desired capacity.

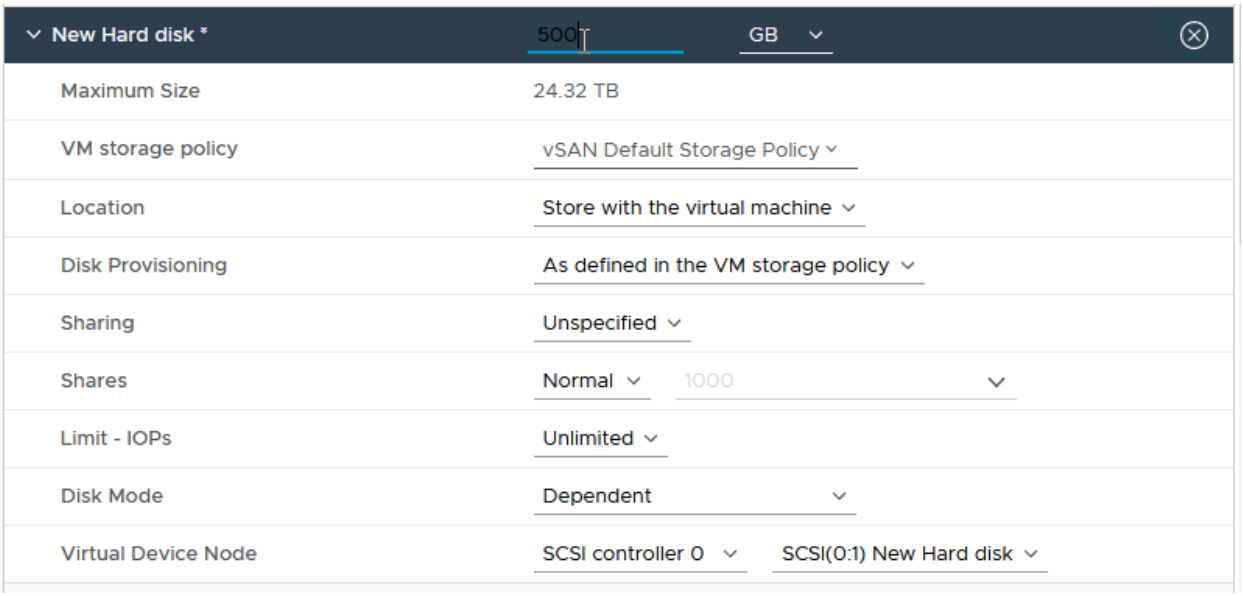


Figure 11 - Specify Desired Disk Size

4. Select **Location**, then click **Browse**.

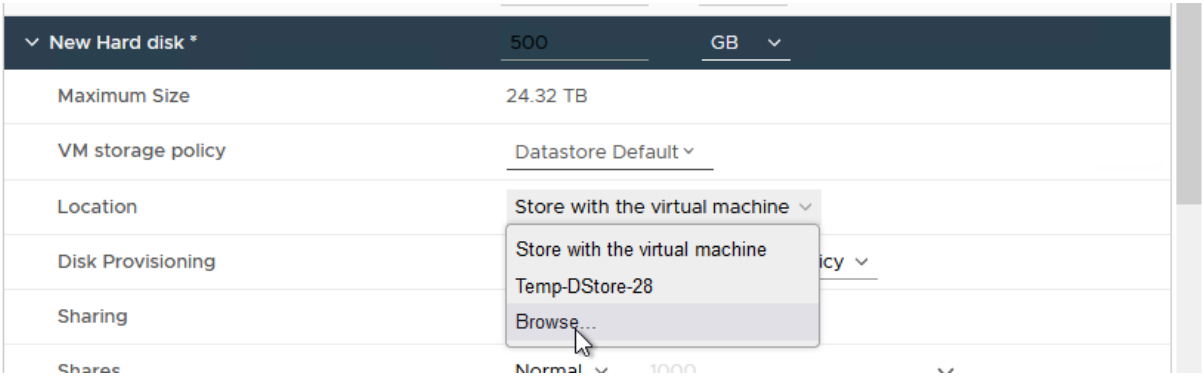


Figure 12 - Browse to Clustered VMDK Datastore

5. Select the "Clustered VMDK" Datastore you created earlier and click **OK**.

   Important: Note this specific file location. You will need to browse to this exact path when attaching this shared disk to other nodes.

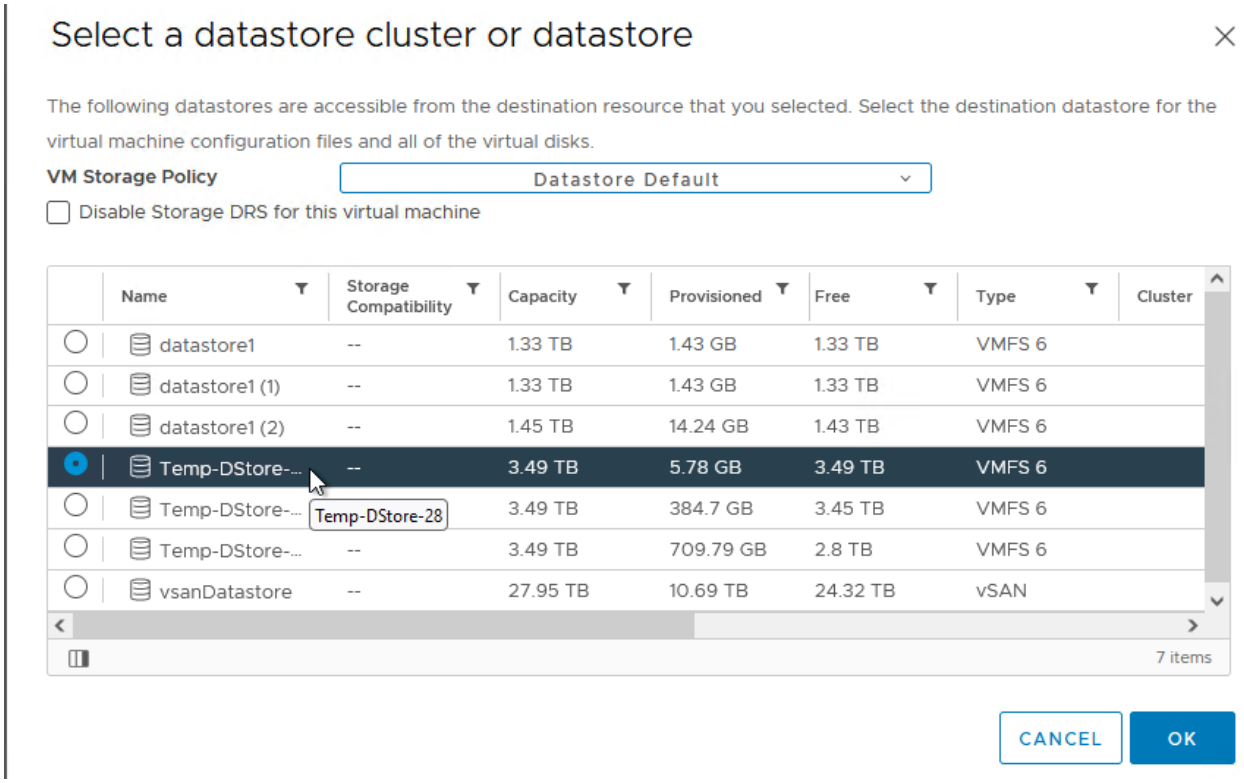


Figure 13 - Select the Clustered VMDK Datastore for the Shared Disk

6. Click **Disk Provisioning** and select **Thick Provision Eager Zeroed**.

   Note: Even if the backing storage array is All-Flash, Eager Zeroed Thick provisioning is recommended for performance and high throughput regarding Microsoft SQL Server data, logs, and TempDB volumes. Thin provisioned disks are supported for this configuration only when using vSAN.

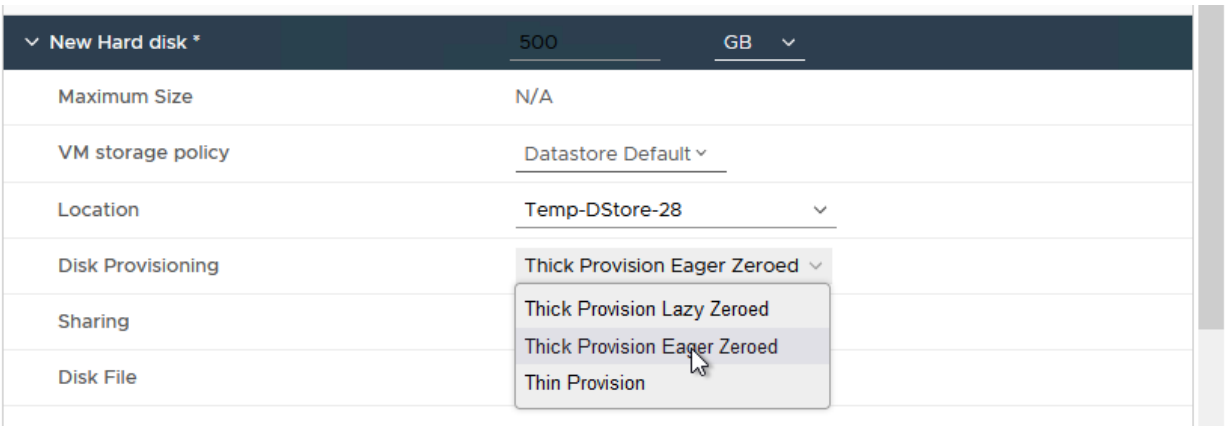


Figure 14 - Select Disk Provisioning Format

7. Click **Disk Mode** and change the setting to **Independent – Persistent**.

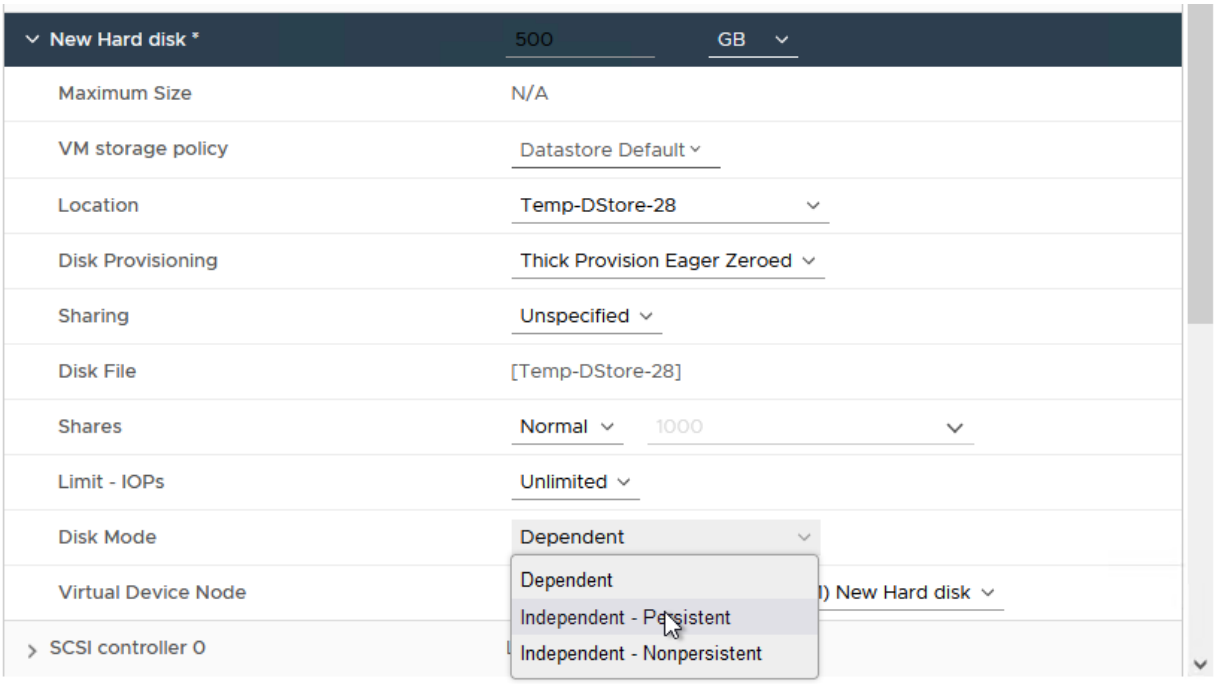


Figure 15 - Specify the Disk Mode

8. From the **Virtual Device Node** menu, select the appropriate **SCSI Controller ID** to which you want to attach the disk.

   Important: Record the SCSI ID used for this disk. You must attach the disk to the exact same SCSI ID on all VMs sharing this disk.

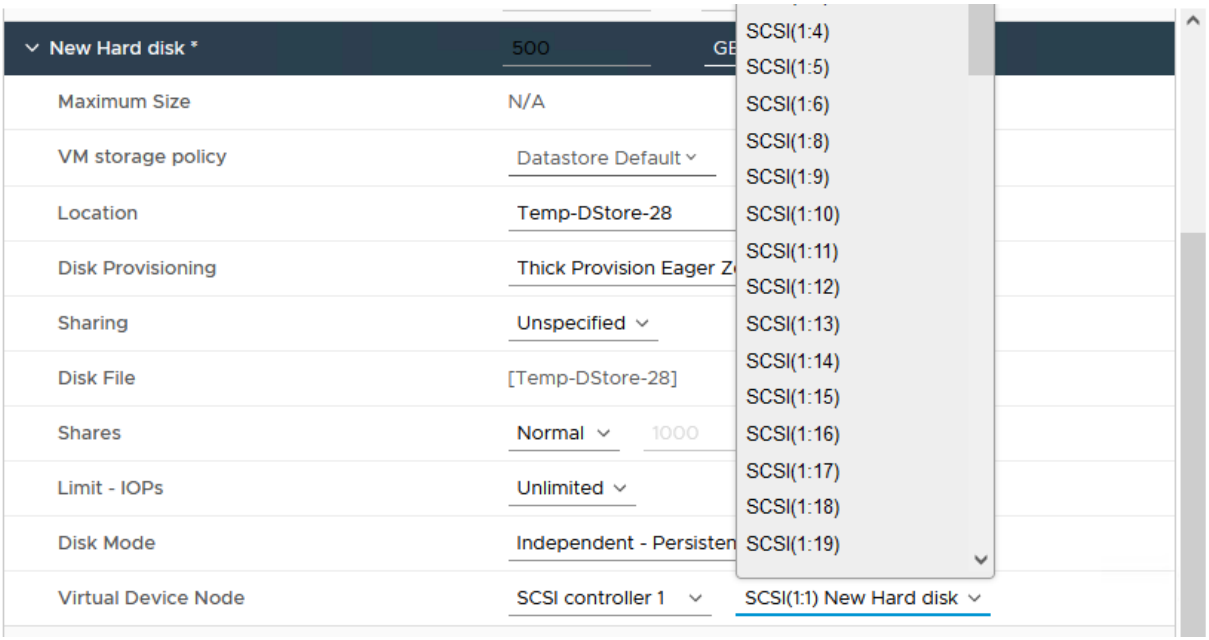


Figure 16 - Attach VMDK to the SCSI Controller

9. Expand the SCSI Controller to which you attached the disk. Click on **SCSI Bus Sharing** and ensure it is set to **Physical**. Click **OK** to commit the changes.

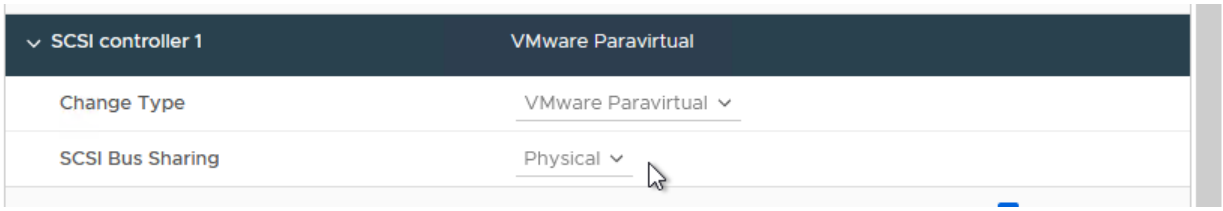


Figure 17 - Select Controller Bus Sharing Option

10. When finished, your configuration should resemble the example below:

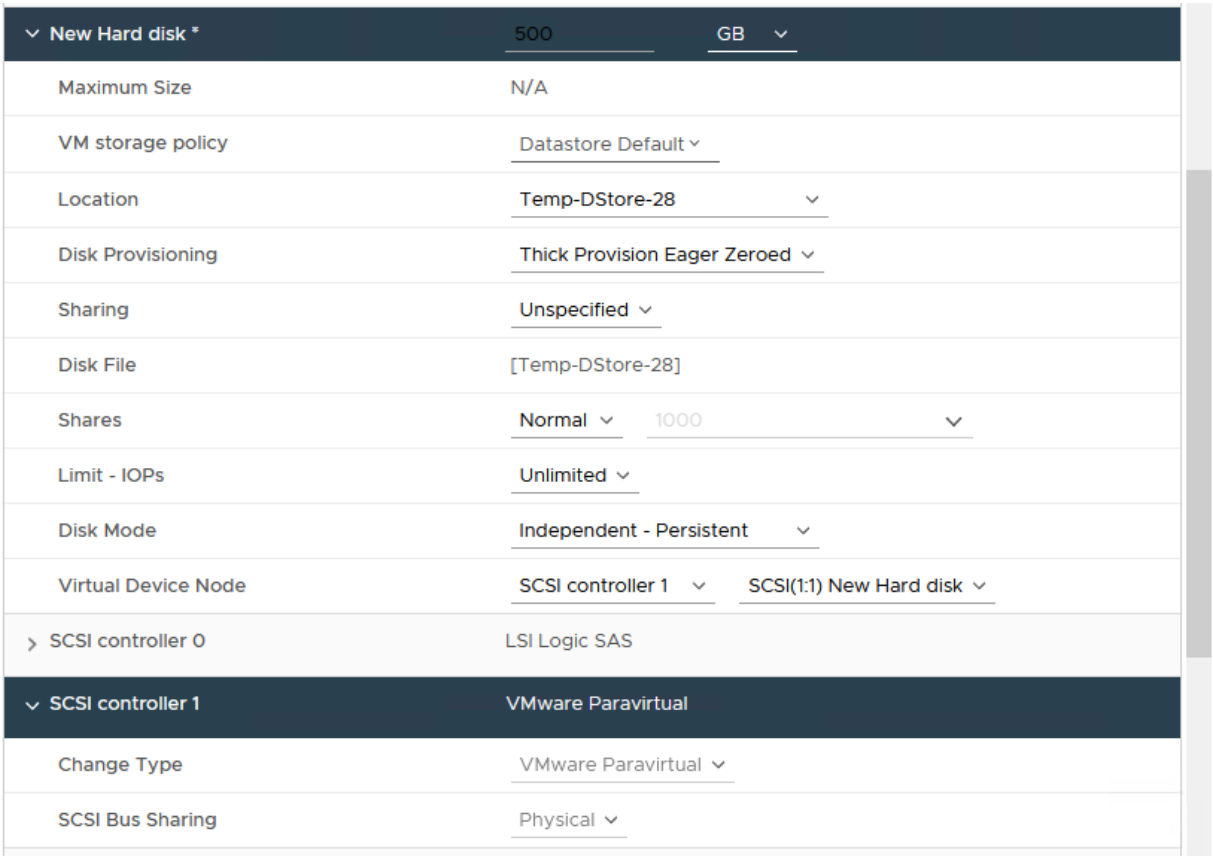


Figure 18 - Sample Shared VMDK Configuration Info

Recommendation: It is recommended that you power on the VM now and format this disk within Windows Server. Once formatted, the disk can be presented to additional VMs.

Adding Shared VMDKs to Additional FCI

1. Right-click the next VM that will share the disk and select **Edit Settings**.
2. From the **Add New Device** menu, select **SCSI Controller**.

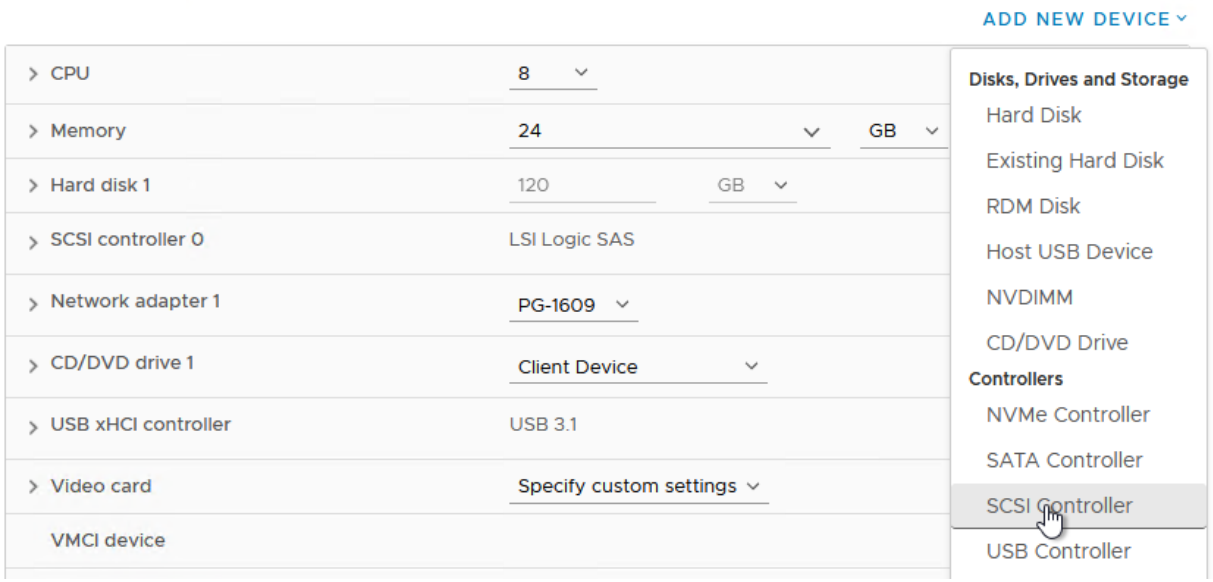


Figure 19 - Add a Controller to be Used by Shared VMDK

3. Expand the newly added SCSI Controller and change the type to **VMware Paravirtual**.

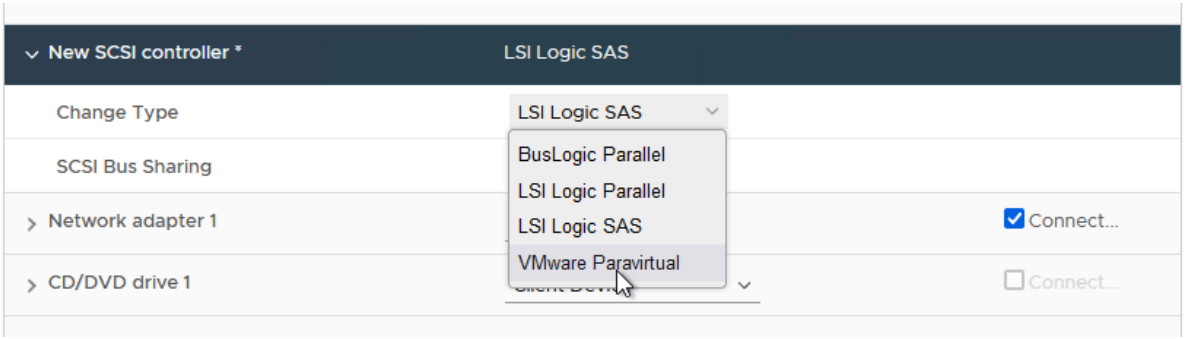


Figure 20 - Ensure that the SCSI Type is "VMware Paravirtual"

4. Change the **SCSI Bus Sharing** setting to **Physical**.

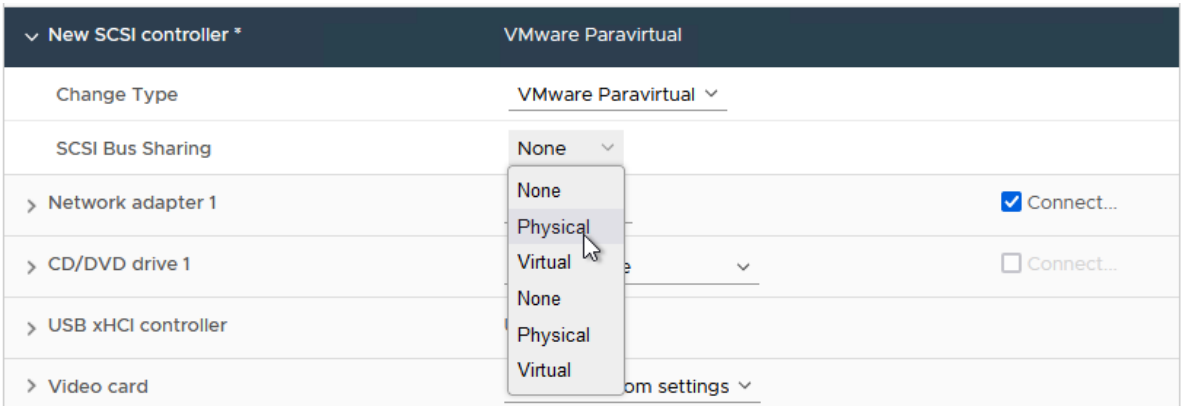


Figure 21 - Change the Bus Sharing Option to Physical

5. Click **Add New Device** again. This time, select **Existing Hard Disk**.

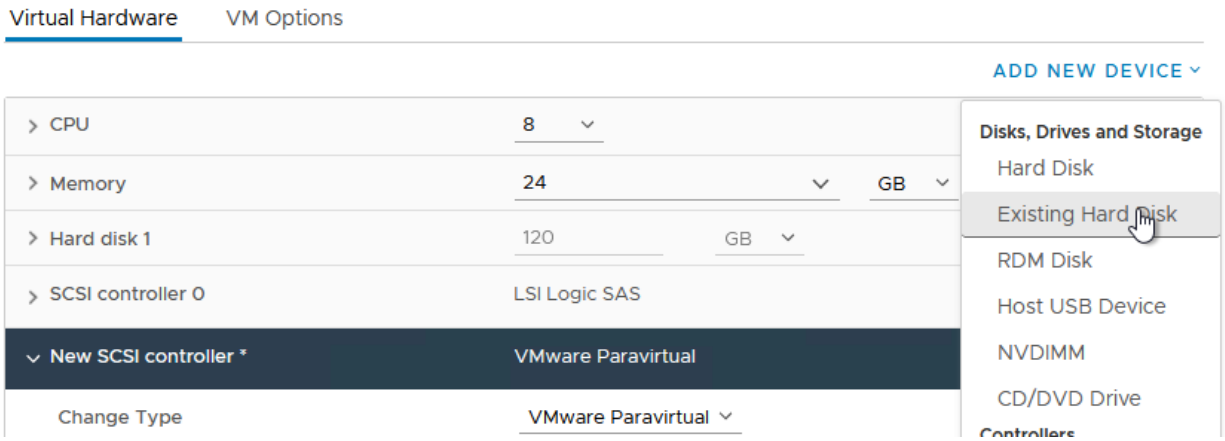


Figure 22 - Select Add New Devices -> Existing Hard Disk

6. Browse to the datastore location containing the disk created in the previous steps. Select the correct disk file.

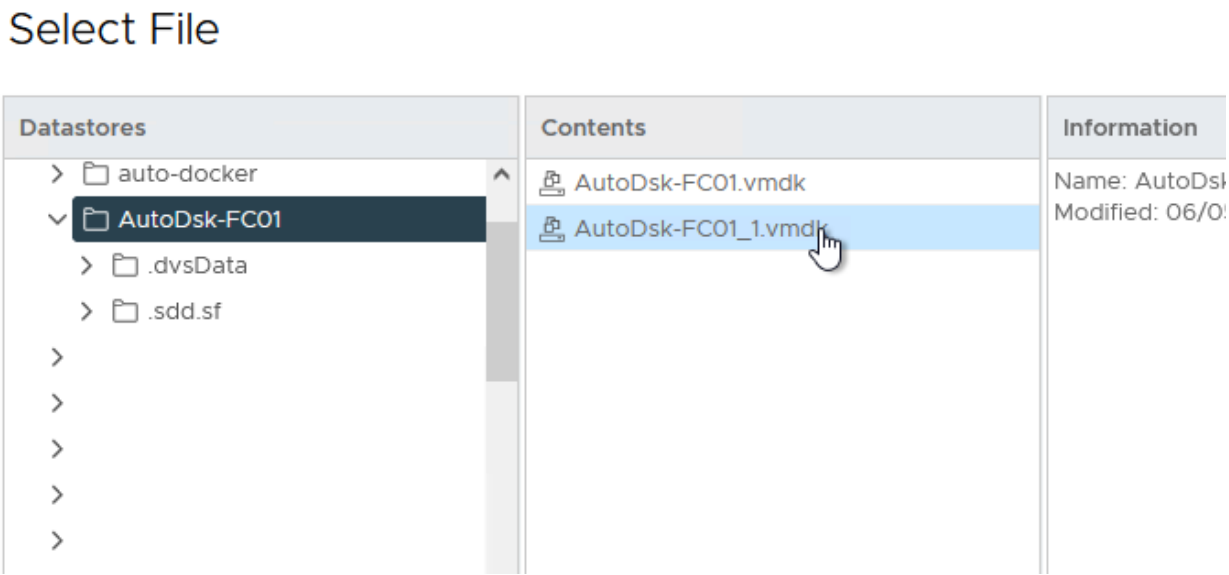


Figure 23 - Select the Disk Created in Previous Steps

7. **Important:** Ensure the disk is connected to the same **SCSI ID** on this node as it was on the original node. The SCSI ID must match on all nodes sharing this disk.

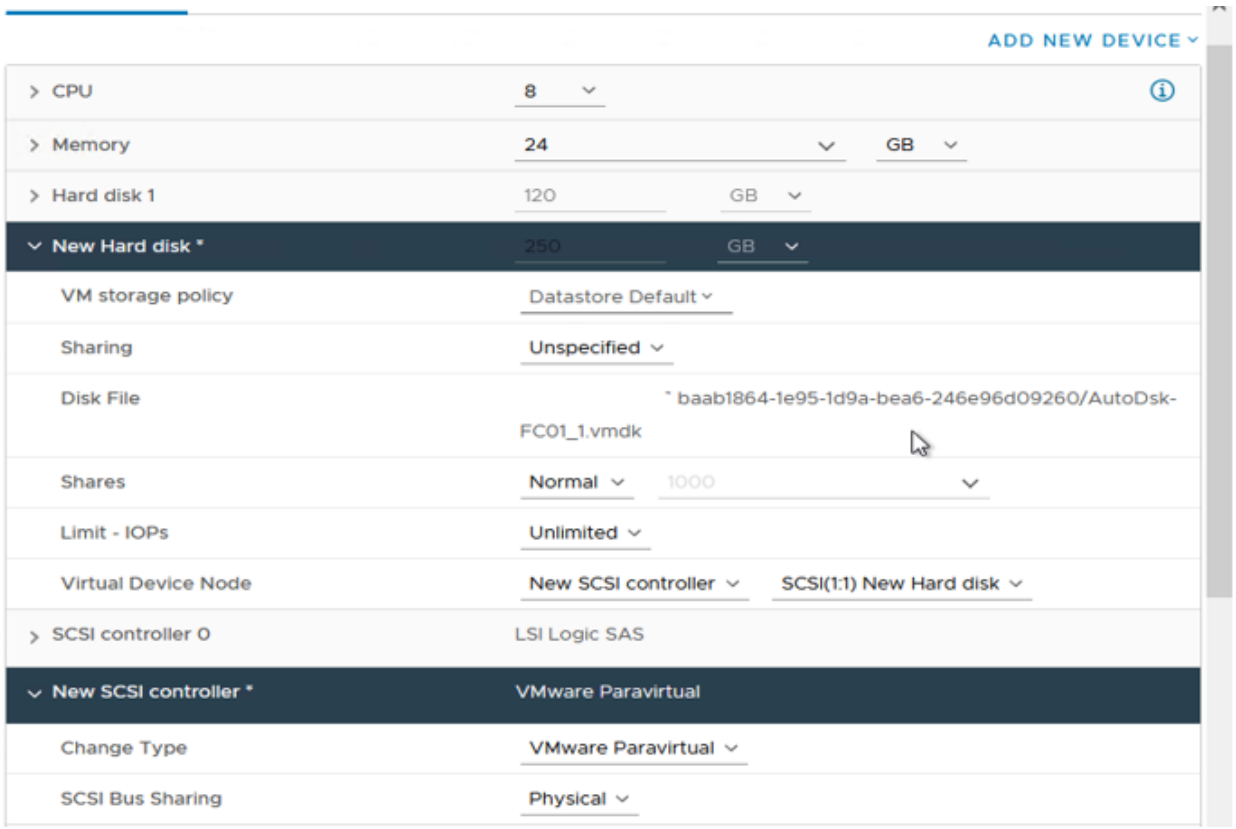


Figure 24 - Review the Settings

8. Click **OK** to complete the configuration.

Note: If you formatted this disk in the previous steps, do not format it again on this node (or any subsequent nodes). The disk is now available for use in Windows Server and WSFC configuration. WSFC will arbitrate access to and ownership of the disk resource from this point forward.

Using Raw Device Mapping (RDM) for FCI

If your architecture requires Raw Device Mappings (RDMs) for a Failover Cluster Instance (FCI), specific configuration steps are required to ensure the disks are recognized correctly by the cluster. The disks must be attached to a **VMware Paravirtual SCSI (PVSCSI)** controller and set to **Physical Compatibility Mode**.

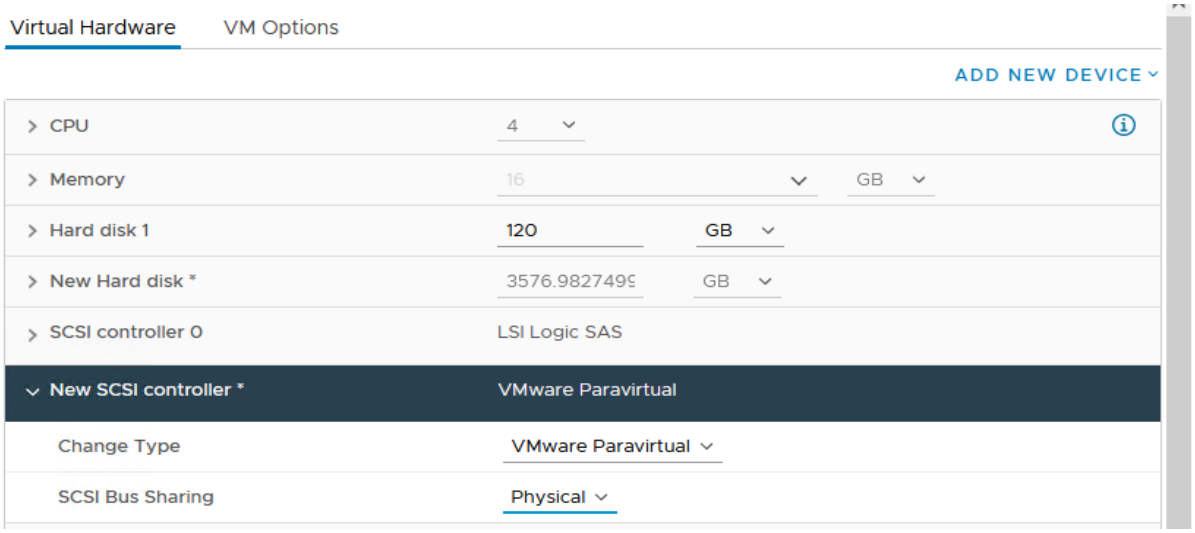


Figure 25 – Add a PVSCSI Controller and Set SCSI Bus Sharing to "Physical"

Adding the RDM to the First Node

1. Edit the settings of the first node (it does not matter which node you start with).
2. From the **Add New Device** menu, select **RDM Disk**.

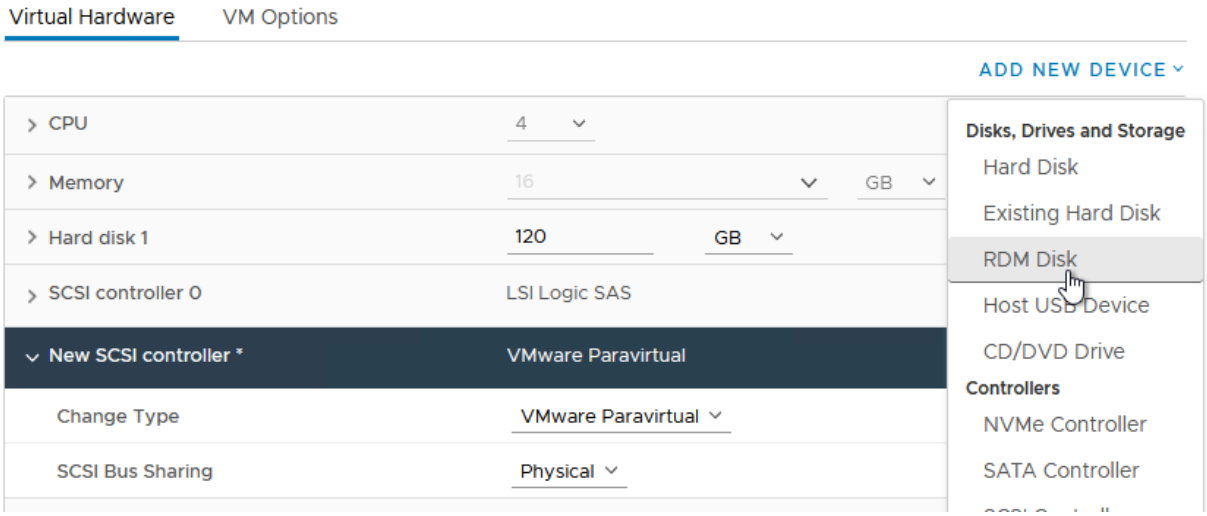


Figure 26 - Add the RDM Disk

3. A list of available LUNs will appear. Select the LUN you wish to map as an RDM and click **OK**.

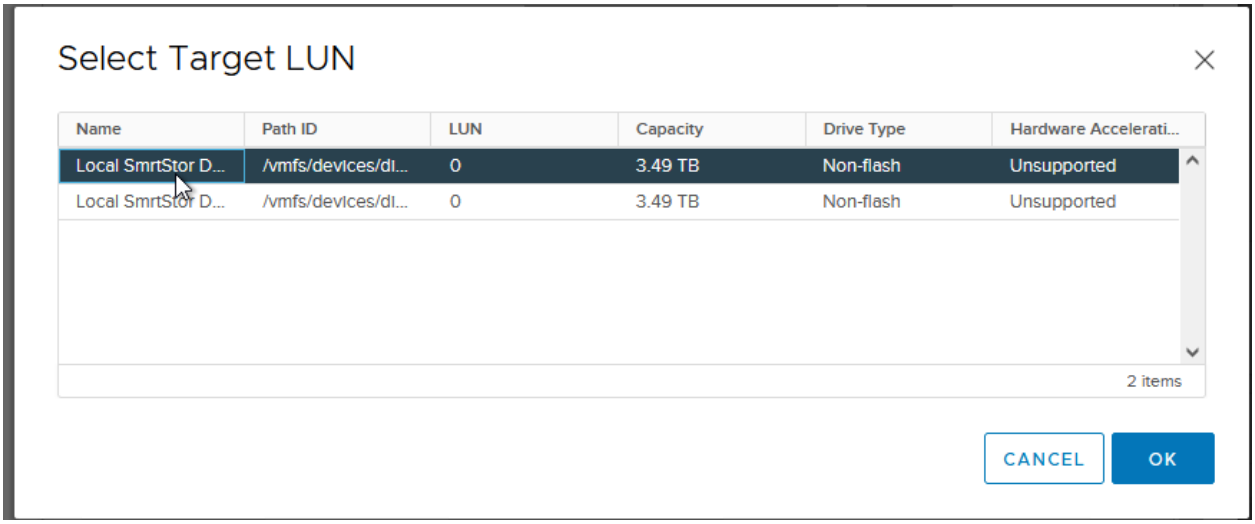


Figure 27 - Select the Target LUN for RDM Disk

4. Expand **New Hard disk** and configure the following settings:
    - **Compatibility Mode:** Physical.
    - **Virtual Device Node:** Assign to the correct PVSCSI controller.
    - **Disk Mode:** Independent – Persistent (required for physical disks).

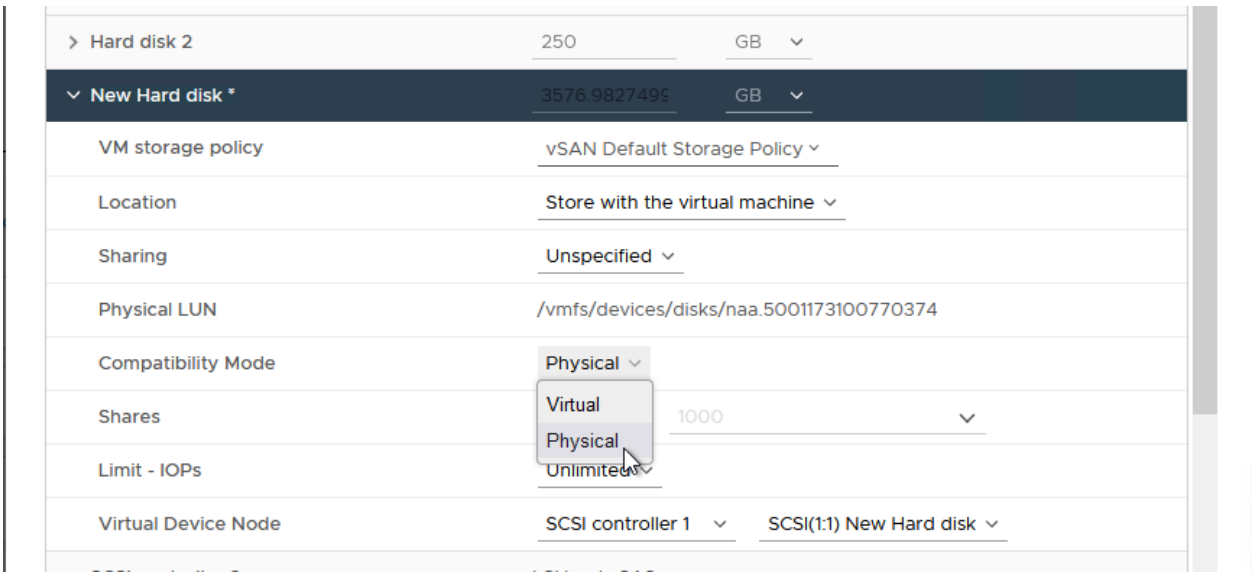


Figure 28 - Configure the RDM Disk as Outlined in Step 4

5. Under **Location**, click **Browse**. You must select a compatible Datastore to hold the RDM's mapping file (the pointer file).

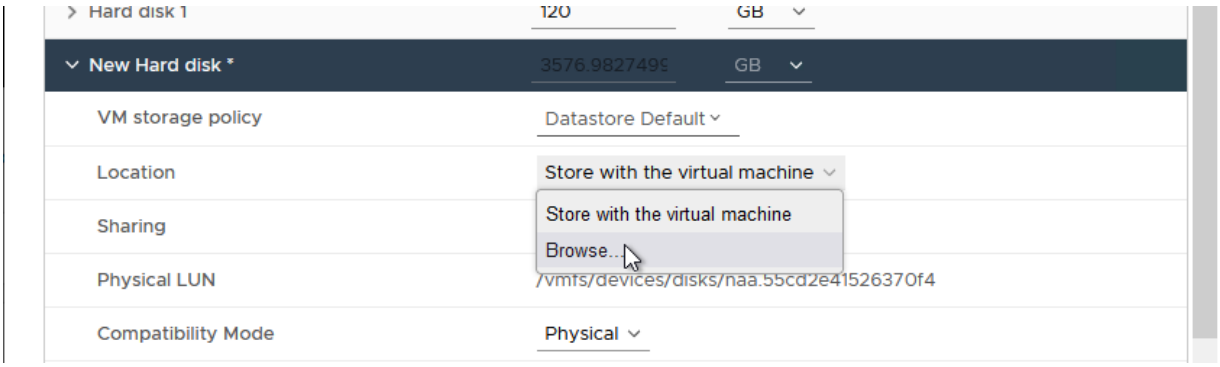


Figure 29 – Browse to the Correct and Compatible Datastore for the Mapping File

6. Select the appropriate datastore for the mapping file.

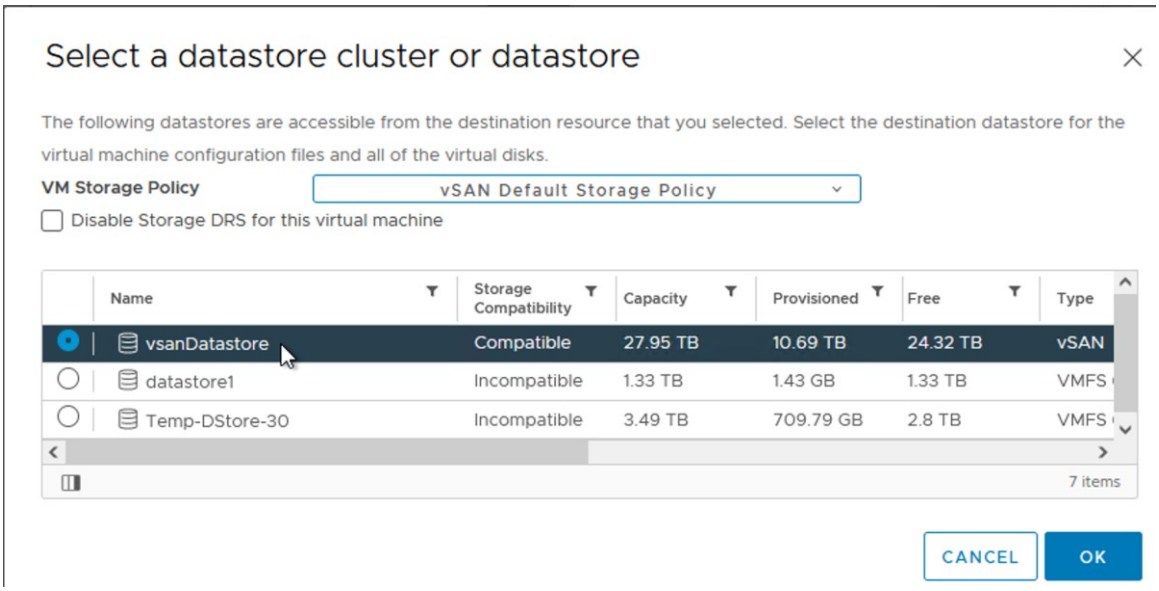


Figure 30 - Select the Datastore

7. Click **OK** to commit the changes.

Understanding RDM Mapping Files

When viewing the datastore in vCenter, you will see a VMDK file that appears to match the size of the full RDM (e.g., several terabytes). This is a "friendly" visual representation provided by the UI.

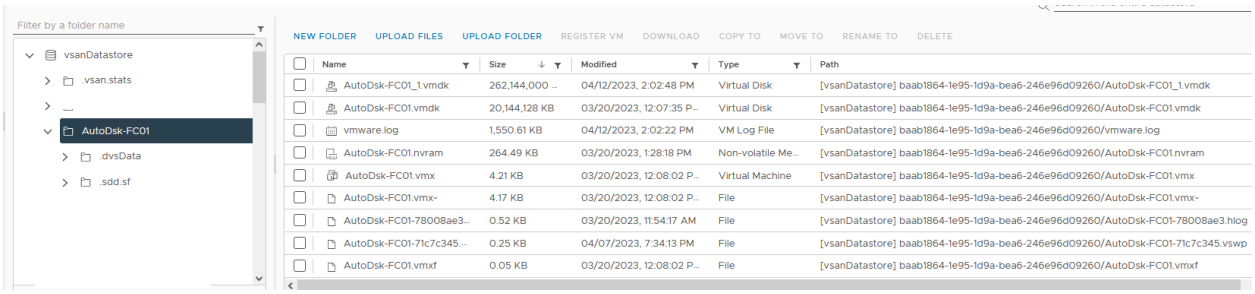


Figure 31 - RDM as Seen in vCenter

Under the covers, the actual VMDK is a tiny mapping file (approximately 520 bytes) with a name format of friendlyname-rdmp.vmdk. This file points to the underlying LUN.

```
[root@sc2esx30:/vmfs/volumes/vsan:525d3597c274c4d0-642c6bb5ba25dd53/baab1864-1e95-1d9a-bea6-246e96d09260] ls -la
total 793624
drwxr-xr-t    1 root     root          3360 Apr 12 21:02 .
drwxr-xr-x    1 root     root           512 Jun  5 02:15 ..
-rw-------    1 root     root             0 Mar 20 18:53 .bbab1864-d244-9dfb-48a5-246e96d09260.lck
drwxr-xr-x    1 root     root           420 Mar 20 18:53 .dvsData
-r--------    1 root     root       1441792 Mar 20 18:53 .fbb.sf
-r--------    1 root     root     267026432 Mar 20 18:53 .fdc.sf
-r--------    1 root     root       1179648 Mar 20 18:53 .pb2.sf
-r--------    1 root     root     268435456 Mar 20 18:53 .pbc.sf
-r--------    1 root     root     262733824 Mar 20 18:53 .sbc.sf
drwx------    1 root     root           280 Mar 20 18:53 .sdd.sf
-r--------    1 root     root       4194304 Mar 20 18:53 .vh.sf
-rw-------    1 root     root           259 Apr  8 02:34 AutoDsk-FC01-71c7c345.vswp
-rw-------    1 root     root             0 Mar 20 19:07 AutoDsk-FC01-71c7c345.vswp.lck
-rw-r--r--    1 root     root           531 Mar 20 18:54 AutoDsk-FC01-78008ae3.hlog
-rw-------    1 root     root        270840 Mar 20 20:28 AutoDsk-FC01.nvram
-rw-------    1 root     root           578 Mar 20 19:07 AutoDsk-FC01.vmdk
-rw-r--r--    1 root     root             0 Mar 20 18:54 AutoDsk-FC01.vmsd
-rwxr-xr-x    1 root     root          4311 Mar 20 19:08 AutoDsk-FC01.vmx
-rw-------    1 root     root             0 Mar 20 19:07 AutoDsk-FC01.vmx.lck
-rw-------    1 root     root            47 Mar 20 19:08 AutoDsk-FC01.vmxf
-rwxr-xr-x    1 root     root          4270 Mar 20 19:08 AutoDsk-FC01.vmx~
-rw-------    1 root     root     268435456000 Mar 20 19:05 AutoDsk-FC01_1-rdmp.vmdk
-rw-------    1 root     root           482 Apr 12 21:02 AutoDsk-FC01_1.vmdk
-rw-r--r--    1 root     root       1587826 Apr 12 21:02 vmware.log
```

Figure 32 - Files Associated with the RDM as Seen in the Command Line

Adding the RDM to Additional Nodes

To configure the remaining nodes in the FCI, you must attach the *existing mapping file* created in Phase 1, rather than creating a new RDM.

1. Edit the settings of the next VM participating in the cluster.
2. Do **not** select "RDM Disk." Instead, select **Existing Hard Disk**.

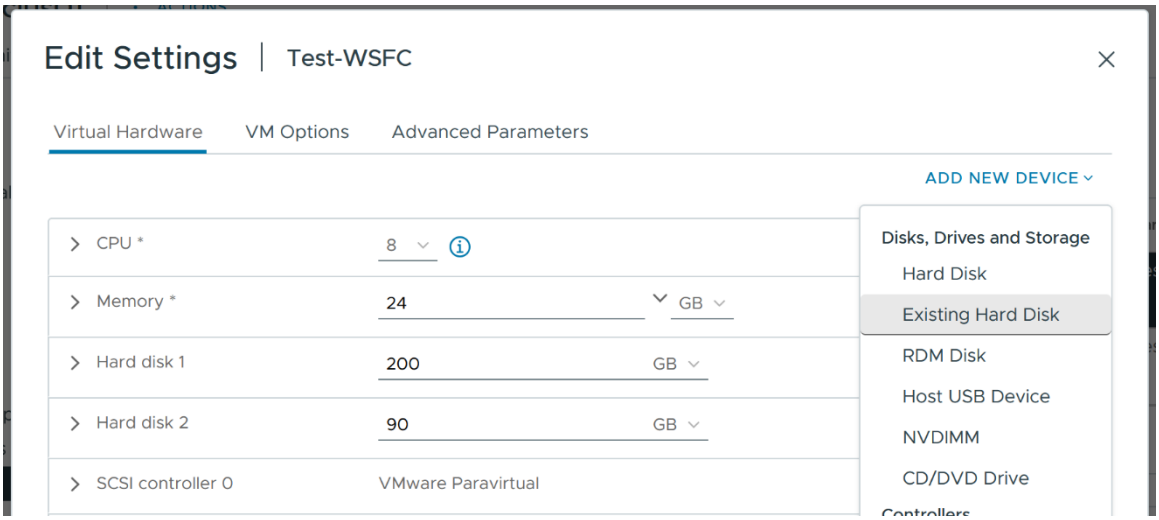


Figure 33 - Adding an Existing RDM Disk

3. Navigate to the datastore location used in the previous steps and select the disk mapping file created earlier. Click **OK**.

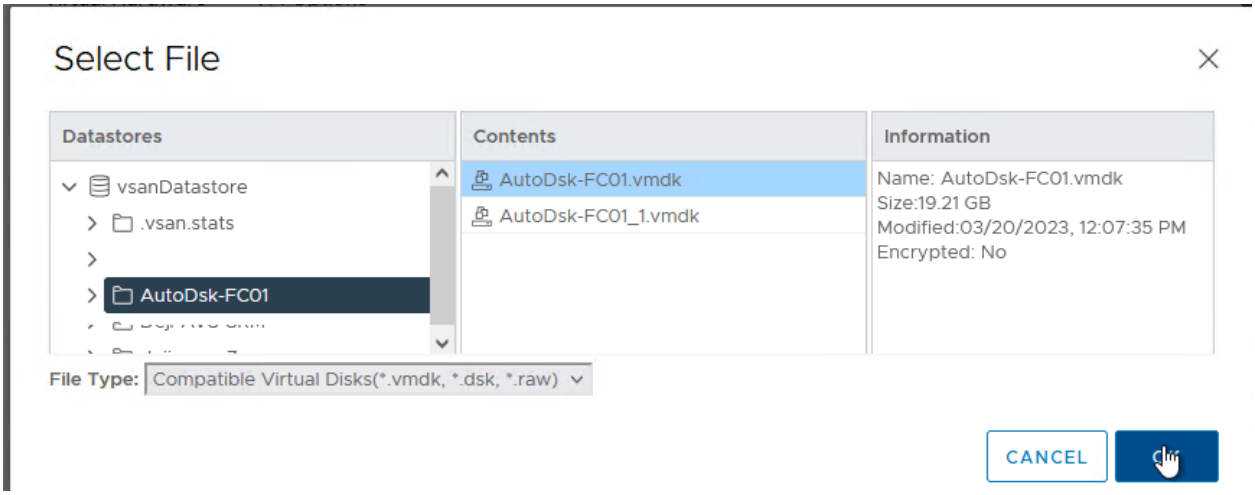


Figure 34 - Selecting Existing RDM Disk

4. Ensure the configuration matches the first node exactly:
    - **Disk Mode:** Independent – Persistent.
    - **Virtual Device Node:** This **must** match the PVSCSI controller ID used on the first node.

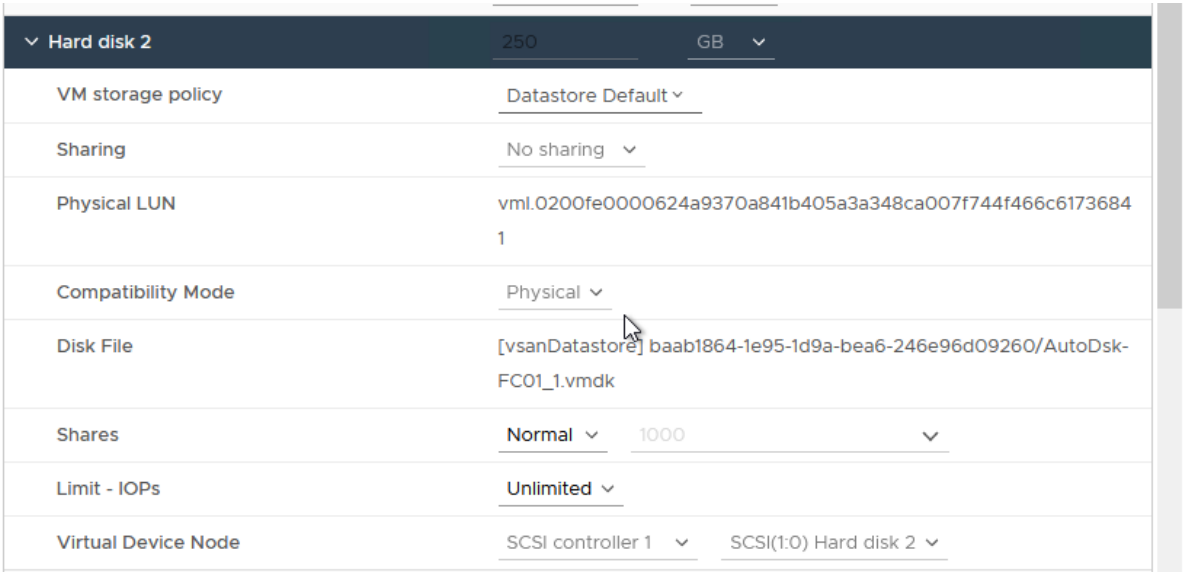


Figure 35 - The Disk Must be Attached to Identical SCSI Port

5. Click **OK** to add the disk.
6. Repeat these steps for any other VMs participating in the FCI.
7. Once completed, the disk will be available to all nodes and can be configured within Windows Server.

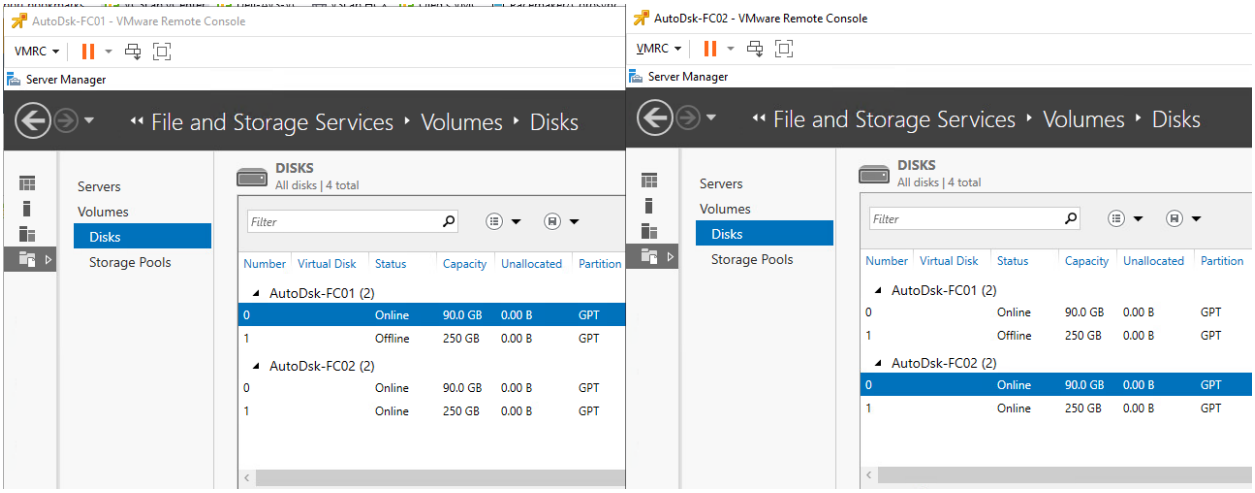


Figure 36 - Shared RDM-backed Disks, As Seen in Windows Server

Formatting Tip: When formatting disks for SQL Server, use a **64KB allocation unit size**. For the file system, use **NTFS**. If using SQL Server 2014 or later, **ReFS** is also supported.

Architectural Considerations and Warnings

Cluster-in-a-Box (CIB) and Validation

If you are using RDMs and the WSFC node VMs reside on the same physical host, the cluster validation wizard will flag an error.

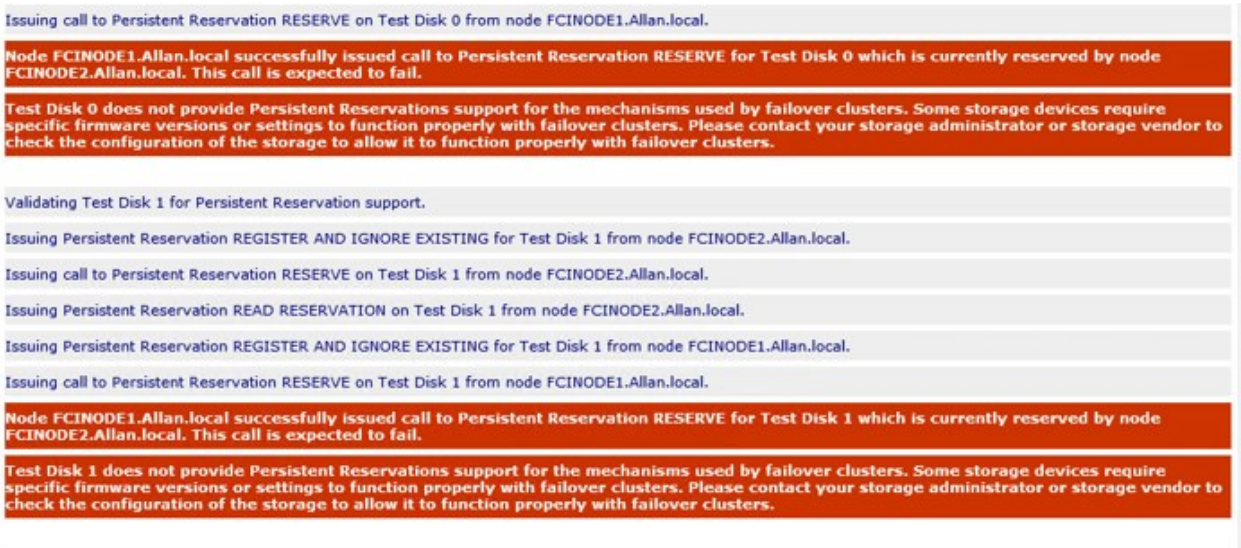


Figure 37 - Disk Errors During WSFC Validation of CIB Configuration

Important: A "Cluster-in-a-Box" (CIB) configuration is **not recommended** for critical SQL Server workloads. If the single physical host fails, all cluster nodes become unavailable. CIB does not provide true High Availability.

Storage Policies and Multipathing

Under the covers, vSphere 5.5 and later supports a Path Selection Policy (PSP) of **Round Robin (PSP_RR)** with shared disks, provided the underlying storage array supports it. PSP is set at the host level; having hosts with different PSP settings running clustered Windows Server VMs is a supported configuration.

SMB 3.0 Support

If you intend to use SMB 3.0 file shares for SQL Server data, logs, and backups, note that VMware only supports this with **Windows Server 2012 or later**. This is a VMware support limitation, not a Microsoft one.

Local Disk Placement

There are three scenarios where "local" (non-shared) disks are used in an FCI configuration:

1. **OS Disks:** The OS disks for each node should not be placed on the same datastores. Where possible, those datastores should utilize different storage units.
2. **Redundancy:** Even though a shared RDM represents a single point of failure, separating the OS disks from the Data/Log/Backup RDMs provides a level of redundancy at the storage layer.
3. **TempDB:** If you desire to create a local VMDK for TempDB (SQL Server's temporary workspace), do so only after confirming that the local storage provides sufficient IOPS. This configuration complicates vSphere HA and vMotion, which will be addressed later in this paper.

Registry Configuration

If using a drive letter, drive letter + mount point, or Cluster Shared Volumes (CSV), VMware recommends setting the following Registry Key' value to **60** (decimal) on each WSFC node: HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Disk\TimeOutValue

Always On Availability Groups

Always On Availability Groups (AGs) provide database-level high availability and disaster recovery for SQL Server 2025. AGs operate by maintaining multiple, continuously synchronized copies of a database - called replicas - across different SQL Server instances. In virtualized environments running on VMware vSphere or VMware Cloud Foundation (VCF) 9, AGs provide a flexible and robust layer of protection that complements hypervisor-level capabilities such as vSphere HA, DRS, and vMotion Application Notification.

SQL Server 2025 introduces several enhancements to AGs, addressing performance, resiliency, and manageability. Among the most important updates are:

- Improved automatic seeding reliability, including retry logic, enhanced throttling for low-bandwidth conditions, and better diagnostic output.
- Latency-aware synchronous commit, which evaluates replica response times and adjusts commit behavior to minimize stalls while maintaining data consistency.
- Enhanced read routing, improving routing decisions across large and geographically distributed replica sets.
- Expanded Contained Availability Groups, allowing metadata objects - such as logins, credentials, and SQL Agent jobs - to fail over with the AG. This reduces cross-node administrative overhead and simplifies AG mobility under vSphere.
- Improved replica health evaluation, enabling better failover readiness assessments and more predictable failover decisions.

AGs require a functioning Windows Server Failover Cluster (WSFC), which provides the underlying cluster membership, quorum management, and resource arbitration. Although AGs do not require shared storage, WSFC remains central to their management and operation.

Under VMware vSphere, AG-based workloads benefit from the abstraction provided by virtualization. Because AGs do not depend on shared disks, they avoid many of the storage complexities associated with FCIs. Storage for AG replicas can be provisioned using standard VMDKs, vVols, or vSAN ESA without requiring SCSI-3 Persistent Reservations.

Placement of AG replicas under vSphere and VCF 9 should follow these guidelines:

- The primary replica and each synchronous secondary replica should run on separate ESXi hosts.
- DRS anti-affinity rules should be used to prevent AG replicas from being co-located on the same host.
- Memory reservations should be applied to SQL Server VMs hosting synchronous replicas to prevent heartbeat or synchronization instability.
- Host groups and VM groups may be used in stretched environments to ensure replicas are distributed across availability zones or fault domains.

Availability Group listeners - the virtual network names that applications use to connect to AGs - require DNS registration and Active Directory integration unless using Domain-Independent AGs. In multi-subnet configurations across vSphere clusters, listener failover behavior depends on DNS refresh cycles and client reconnection logic. Low DNS TTL values and proper multi-subnet configuration are required for predictable behavior.

VMware vSphere HA interacts with AGs by restarting SQL Server VMs on surviving hosts after a host failure. Because AGs rely on WSFC for failover decisions, the cluster determines whether primary or secondary roles need to change following a host outage. In environments with synchronous commit, AG failover is typically automatic; in asynchronous commit configurations, failover may require manual intervention depending on business requirements.

vMotion Application Notification significantly improves AG stability during live migrations. During vMotion, the guest OS receives a pre-migration event, allowing SQL Server to prepare for short pause conditions. This reduces the likelihood of AG failovers triggered by transient heartbeat delays or brief pauses in replica log send/receive activity. Enabling Application Notification is strongly recommended for all AG replicas, particularly those participating in synchronous commit availability modes.

AGs remain one of the most flexible and powerful high availability options for SQL Server 2025, especially in virtualized environments. Their storage independence, granular failover semantics, and improved metadata containment make them an ideal choice for modern cloud and hybrid deployments. When combined with VMware's maturing DRS and HA capabilities, AGs offer predictable and highly resilient protection for mission-critical SQL Server workloads.

ADDS and DNS for AGs

Availability Group listeners depend on name resolution and directory integration unless deploying Domain-Independent AGs. In SQL Server 2025, listener behavior is unchanged in principle but benefits from improved metadata containment when using Contained AGs.

In traditional WSFC-based AG deployments, the listener is implemented as a Virtual Network Name (VNN). This VNN requires:

- Active Directory Domain Services (ADDS) to create and manage the corresponding Virtual Computer Object (VCO).
- DNS registration that enables clients to locate the listener's IP address.

In VMware environments, the dependability of DNS infrastructure becomes especially important. Even when all AG replicas remain healthy, DNS misconfiguration or delayed DNS propagation can disrupt client connections to the listener following failover.

For multi-subnet AG configurations - common in stretched clusters or multi-site vSphere designs - SQL Server 2025 supports multi-subnet listener names with multiple IP addresses registered in DNS. Clients using the MultiSubnetFailover parameter must be configured to handle the failover pattern correctly. DNS time-to-live (TTL) values should remain low to allow rapid reconnection after a failover.

Domain-Independent AGs remove the dependency on ADDS entirely but do not support listeners backed by VNNs. Instead, they use Distributed Network Name (DNN) constructs, which dynamically discover replicas without requiring DNS records or domain objects. While simpler, DNNs lack the failover transparency of traditional listeners and require client-side DNS independence.

- In VMware vSphere and VCF 9 environments, administrators must ensure:
- DNS redundancy across hosts and clusters.
- Proper synchronization of DNS zones across physical and virtual domain controllers.
- Alignment between WSFC network configuration and virtual networking, including NSX-T overlays if used.

Regardless of architecture, AG listener design must be validated thoroughly as part of SQL Server high availability planning, as name resolution failures are a common root cause of connection disruptions after failover.

Storage and Networking for AGs

Availability Groups (AGs) provide database-level high availability without requiring shared storage. Each replica maintains its own copy of the database files, enabling flexible placement across multiple datastores and hosts. SQL Server 2025 continues this architectural model but introduces improvements to I/O resiliency, replica synchronization efficiency, and log transport stability.

General Network Requirements and Validation for WSFC

Microsoft no longer requires more than one network interface card (NIC) for Windows Server Failover Clustering (WSFC). The concept of a "dedicated heartbeat NIC" is obsolete because WSFC will utilize any designated heartbeat NIC for standard "public" traffic during normal operations regardless of its label.

However, if you configure a specific VM with a single vNIC for use in a WSFC, the cluster validation wizard will generate a warning. This is expected behavior. The validation report will highlight this in three stages:

1. **Category View:** At the top of the report, the Network category will display a warning status (see **Figure 38**).

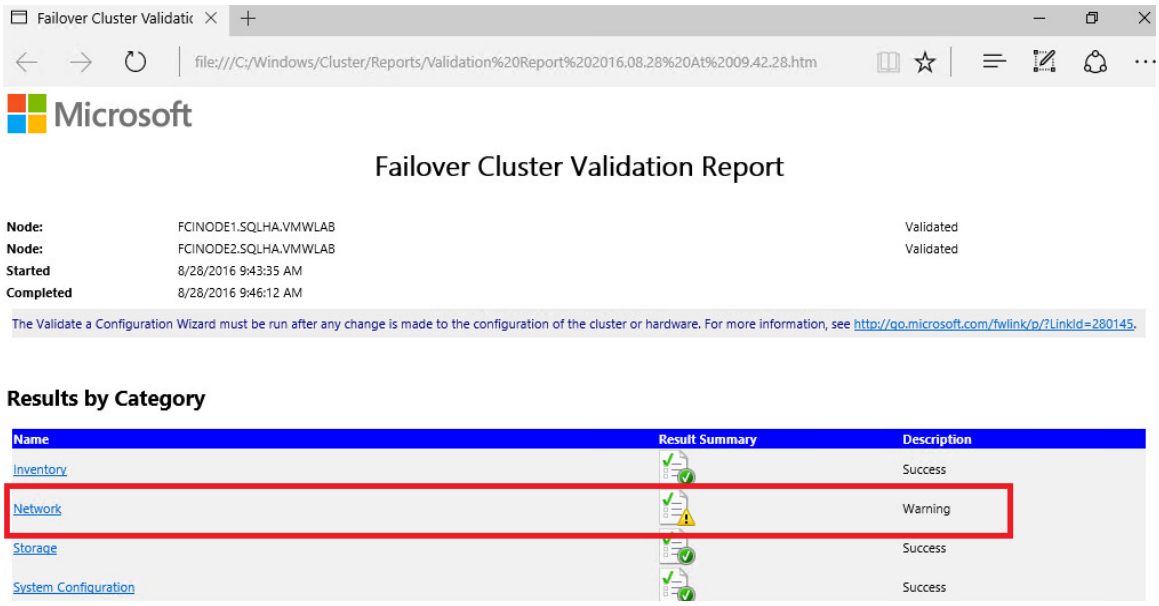


Figure 38 - Network Validation Category Reflecting a Warning

2. **Test List:** Clicking on the Network category reveals that the specific test generating the alert is **Validate Network Configuration** (see **Figure 39**).

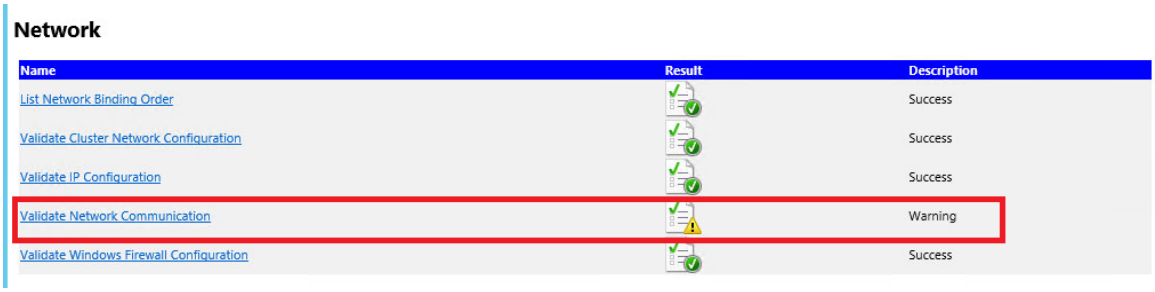
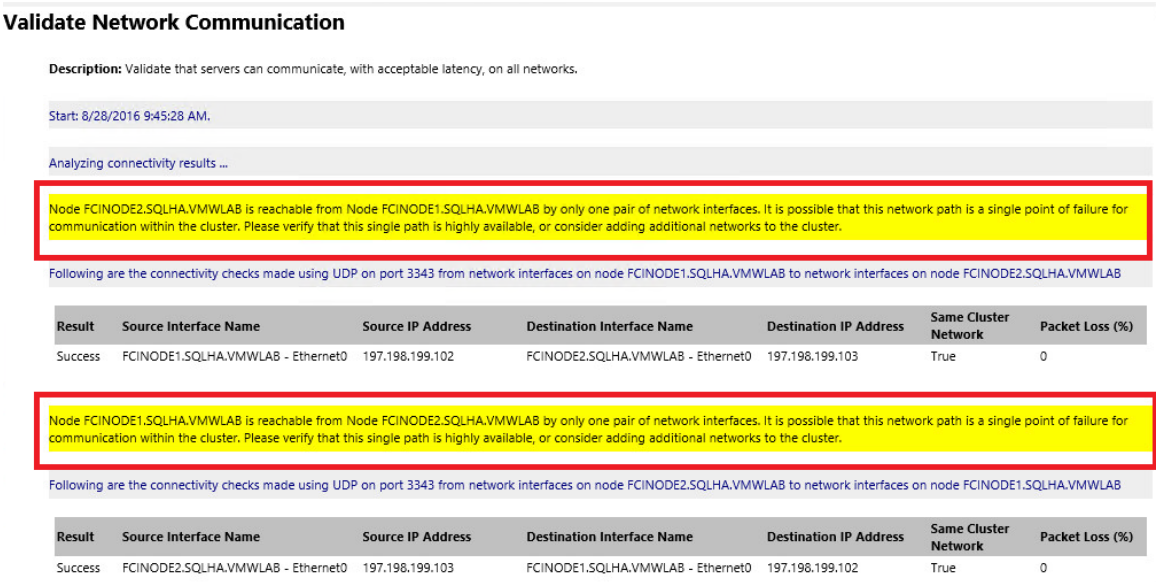


Figure 39 – Specific Network Configuration Check Warning

3. **Detailed Detail:** Drilling down into the details will show specific messages stating that the cluster detects only one NIC, flagging it as a potential single point of failure (see **Figure 40 below**).



Validate Network Communication

Description: Validate that servers can communicate, with acceptable latency, on all networks.

Start: 8/28/2016 9:45:28 AM.

Analyzing connectivity results ...

Node FCINODE2.SQLHA.VMWLAB is reachable from Node FCINODE1.SQLHA.VMWLAB by only one pair of network interfaces. It is possible that this network path is a single point of failure for communication within the cluster. Please verify that this single path is highly available, or consider adding additional networks to the cluster.

Following are the connectivity checks made using UDP on port 3343 from network interfaces on node FCINODE1.SQLHA.VMWLAB to network interfaces on node FCINODE2.SQLHA.VMWLAB

Result	Source Interface Name	Source IP Address	Destination Interface Name	Destination IP Address	Same Cluster Network	Packet Loss (%)
Success	FCINODE1.SQLHA.VMWLAB - Ethernet0	197.198.199.102	FCINODE2.SQLHA.VMWLAB - Ethernet0	197.198.199.103	True	0

Node FCINODE1.SQLHA.VMWLAB is reachable from Node FCINODE2.SQLHA.VMWLAB by only one pair of network interfaces. It is possible that this network path is a single point of failure for communication within the cluster. Please verify that this single path is highly available, or consider adding additional networks to the cluster.

Following are the connectivity checks made using UDP on port 3343 from network interfaces on node FCINODE2.SQLHA.VMWLAB to network interfaces on node FCINODE1.SQLHA.VMWLAB

Result	Source Interface Name	Source IP Address	Destination Interface Name	Destination IP Address	Same Cluster Network	Packet Loss (%)
Success	FCINODE2.SQLHA.VMWLAB - Ethernet0	197.198.199.103	FCINODE1.SQLHA.VMWLAB - Ethernet0	197.198.199.102	True	0

Figure 40 – Benign and Inapplicable Single NIC Warning

Understanding the Risks and Requirements

While the validation wizard flags a single vNIC as a risk, this configuration is technically acceptable **if** you have followed prescriptive guidance for vSphere networking.

If your single vNIC is backed by a fully redundant infrastructure (a vSwitch or vSphere Distributed Switch mapped to redundant physical NICs), the configuration is valid. The redundancy exists at the hypervisor layer, underneath the guest OS.

Critical Warning: If your vNIC connects to a vSwitch mapped to a single physical NIC, a simple network interruption could take down your WSFC and any associated Availability Groups (AG) or Failover Cluster Instances (FCI). If your infrastructure relies on a single physical path, you are not highly available. Do not configure a cluster, AG, or FCI in this environment.

Hardware Considerations and Teaming

True redundancy requires careful hardware planning. A single physical network card with multiple ports connected to different switches still represents a single point of failure (the card itself). Similarly, blade servers share an enclosure and backplane, which concentrates networking risk.

To establish redundancy at the vSphere level, you should team NICs within ESXi. This provides redundancy and can increase bandwidth through load balancing, failover, or a combination of both.

- **vSphere Level:** Team the NICs in ESXi (See How to configure NIC teaming in ESXi and ESX).
- **Guest Level:** Microsoft supports NIC teaming inside the guest VM. If you choose this route, it is recommended to use the built-in teaming feature provided in Windows Server 2012 or later.

Storage for Availability Groups

Because AG replicas do not share disks, each SQL Server VM hosting an AG replica must have independent storage volumes. Under VMware vSphere and VCF 9, the following storage technologies are recommended:

- **Standard VMDKs**: Suitable for most AG workloads, offering full compatibility with vMotion, snapshots (when supported by application requirements), and Storage vMotion.
- **vSAN ESA**: Provides improved latency, increased throughput, and enhanced resiliency for write-intensive workloads. vSAN ESA is an excellent platform for AG replicas due to distributed storage and modern data path optimizations.

- **Virtual Volumes (vVols)**: Provide granular policy-based management, allowing per-disk performance tuning and storage independent failover handling.

Storage best practices for AG workloads under vSphere include:

- Separating data, log, TempDB, and backup VMDKs to optimize I/O distribution.
- Using VMware Paravirtual SCSI (PVSCSI) controllers for all data-intensive disks.
- Spreading VMDKs across multiple SCSI controllers to improve command parallelism.
- Ensuring consistent storage configuration across all AG replicas.

SQL Server 2025 enhances automatic seeding performance and resiliency, making initial synchronization of replicas more efficient, even in bandwidth-constrained environments. Administrators may still prefer manual backups for large databases, but automatic seeding now handles transient errors more gracefully.

An example disk presentation approach for an AG configuration using VMDKs is shown below. There may be more than one data file which may or may not be on the same datastore for a particular VM, but that would be acceptable compared to the alternative.

For a comprehensive prescriptive guidance on optimally provisioning storage for Microsoft SQL Server on vSphere, see the "Section 3.8 - Storage Best practices" section of Architecting Microsoft SQL Server on VMware vSphere.

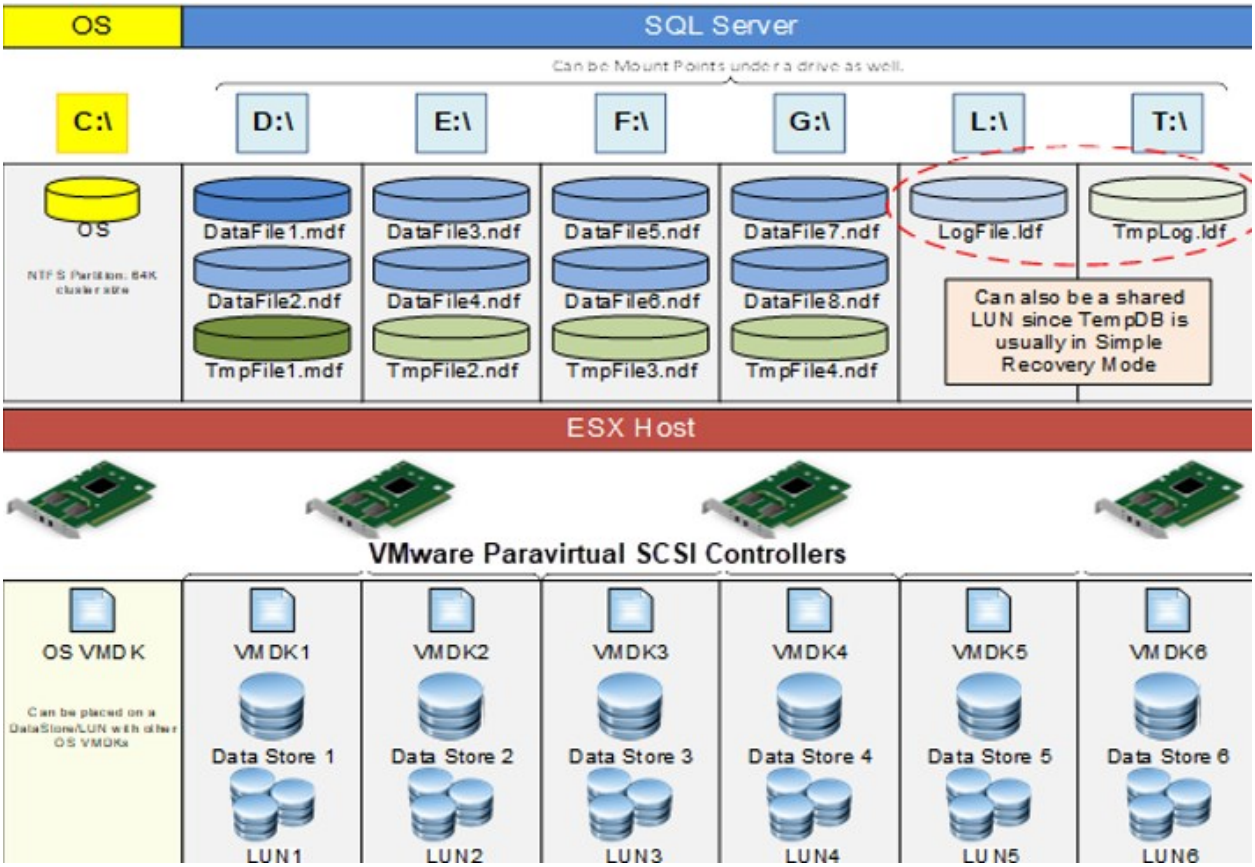


Figure 41 - Sample Disk Presentation for VM Participating in an AG

Networking for Availability Groups

AGs rely heavily on stable network connectivity. Log synchronization traffic between primary and secondary replicas requires low latency, consistent throughput, and minimal packet loss. VMware vSphere and VCF 9 provide several networking capabilities that enhance AG performance:

- **vSphere Distributed Switch (VDS)**: Provides centralized, consistent network configuration across hosts.
- **NSX-T**: Offers overlay-based segmentation, micro-segmentation, and advanced QoS capabilities. AG traffic can benefit from explicit bandwidth guarantees.
- **Jumbo Frames**: Substantially improves performance for heavy log transport workloads, provided that the configuration is end-to-end.

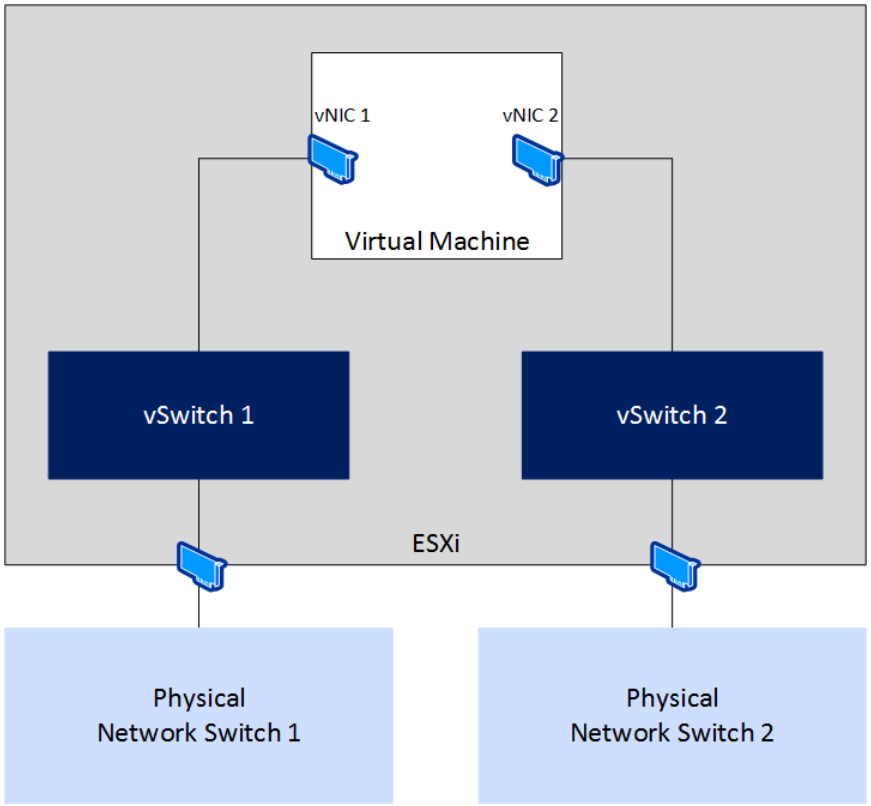


Figure 42 - Example Logical Network Topology for a WSFC Node

The following networking practices ensure AG stability:

- Use redundant uplinks for every port group hosting AG replicas.
- Avoid single-uplink or non-redundant network paths.
- Maintain low network latency between synchronous replicas, especially when synchronous commit is required.
- Place synchronous replicas within the same physical site or fault domain to avoid latency-induced performance degradation.

vSphere Integration (DRS, HA, Fault Tolerance, vMotion)

SQL Server 2025 high availability designs deployed on VMware vSphere or VMware Cloud Foundation (VCF) 9 benefit from a mature and highly reliable virtualization platform. However, successful clustering under vSphere requires a clear understanding of how vSphere High Availability (HA), Distributed Resource Scheduler (DRS), and vMotion interact with Windows Server Failover Clustering (WSFC), Always On Failover Cluster Instances (FCIs), and Always On Availability Groups (AGs). These hypervisor-level mechanisms operate independently of SQL Server and WSFC, but their underlying behaviors directly influence cluster stability, failover timing, and overall availability.

Distributed Resource Scheduler

vSphere's Distributed Resource Scheduler (DRS) is a sophisticated feature designed to reallocate compute resources based on policies configured by the administrator. One of the primary benefits of DRS is its ability to provide dynamic load balancing for Virtual Machines running within a vSphere cluster. DRS operates by moving (via vMotion) a VM to a different ESXi host based on utilization thresholds set for both the host and the vSphere cluster.

A variant of this feature, Storage DRS (SDRS), applies similar logic to storage, moving VMDKs between datastores in a datastore cluster based on IOPS latency or disk space utilization.

DRS continuously monitors resource utilization within a given vSphere Cluster. It responds to observed trends by executing defined actions. The feature provides fine-grained configuration and tuning capabilities, enabling administrators to control exactly when DRS should intervene (thresholds) and what specific actions are acceptable when those thresholds are breached.

The Core Requirement: Anti-Affinity

A fundamental requirement for successfully running clustered SQL Server in a vSphere infrastructure is strictly enforcing physical separation. VMs participating in the same Windows Server Failover Cluster (WSFC) must not run on the same ESXi Host. If they do, a single host failure could take down multiple cluster nodes, negating the availability provided by the cluster.

The mechanism for enforcing this requirement is the **DRS Anti-Affinity Rule**.

It is important to note that DRS is not enabled by default in a vSphere Cluster. However, when the cluster contains WSFC-clustered workloads, enabling DRS becomes a requirement to enforce these separation rules.

When enabled, the default DRS automation level is set to "Manual," as shown in **Figure 43**. This means that even if DRS detects unbalanced host resource utilization within the cluster, it will not automatically move VMs. Instead, it will simply report the imbalance and recommend administrative actions to remedy the situation.

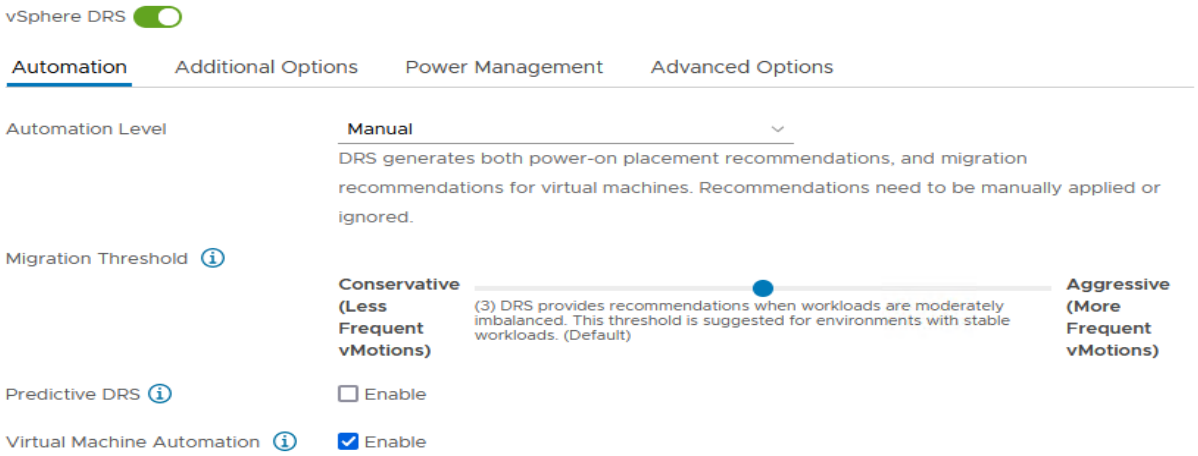


Figure 43 - DRS Default Policies and Behavior

For any VMs participating in the same WSFC - regardless of whether it is supporting a Failover Cluster Instance (FCI) or an Availability Group (AG) configuration - administrators must configure the appropriate DRS rules to ensure that the nodes are always physically separated onto different ESXi hosts.

Configuring DRS Anti-Affinity Rules for Clustered SQL Server Nodes

For detailed information on how to configure the appropriate DRS rule for clustered SQL Server nodes, please refer to Using vSphere DRS Groups and VM-Host Affinity Rules with WSFC Virtual Machines and Create VM-VM Affinity Rules for WSFC Virtual Machines.

The following is a high-level description of the process for configuring DRS anti-affinity rules for a 2-node WSFC:

1. In vCenter, click on the name of the **vSphere Cluster** containing the VMs.
2. Click on **Configure**.
3. Ensure that **vSphere DRS** is turned on. If it is not, toggle it to "On" as shown in **Figure 44**.

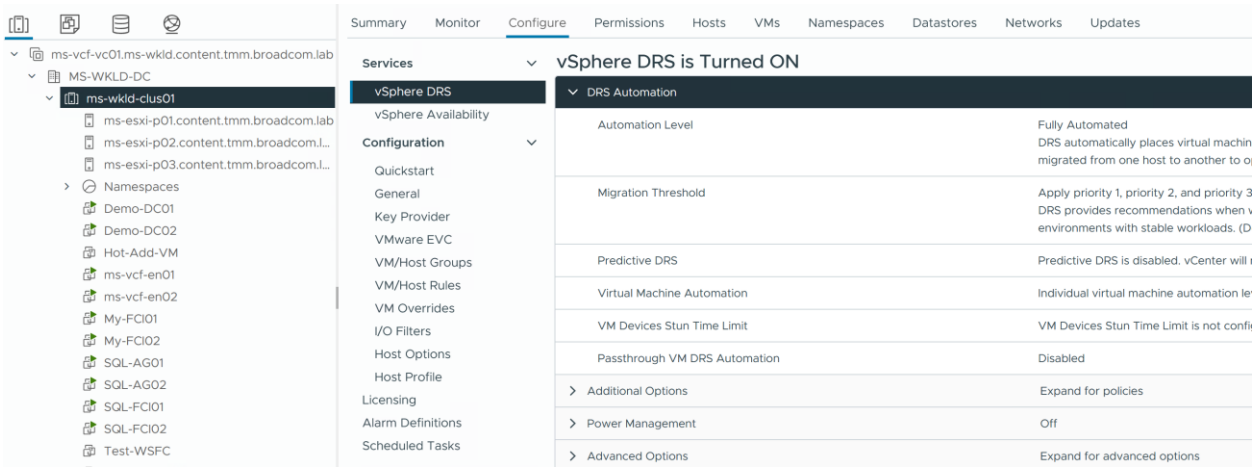


Figure 44 - Verify that DRS is Enabled

4. Navigate to the **VM/Host Rules** section and click **Add** as shown in **Figure 45**.

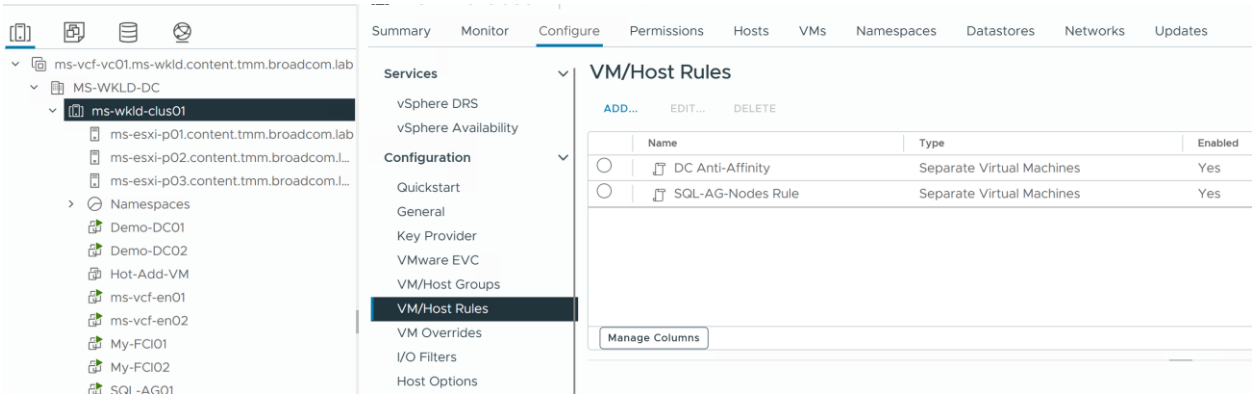


Figure 45 - Creating a DRS VM Rule

5. Give the Rule a name, then choose **Separate Virtual Machines** as the rule type, as shown in **Figure 46**.

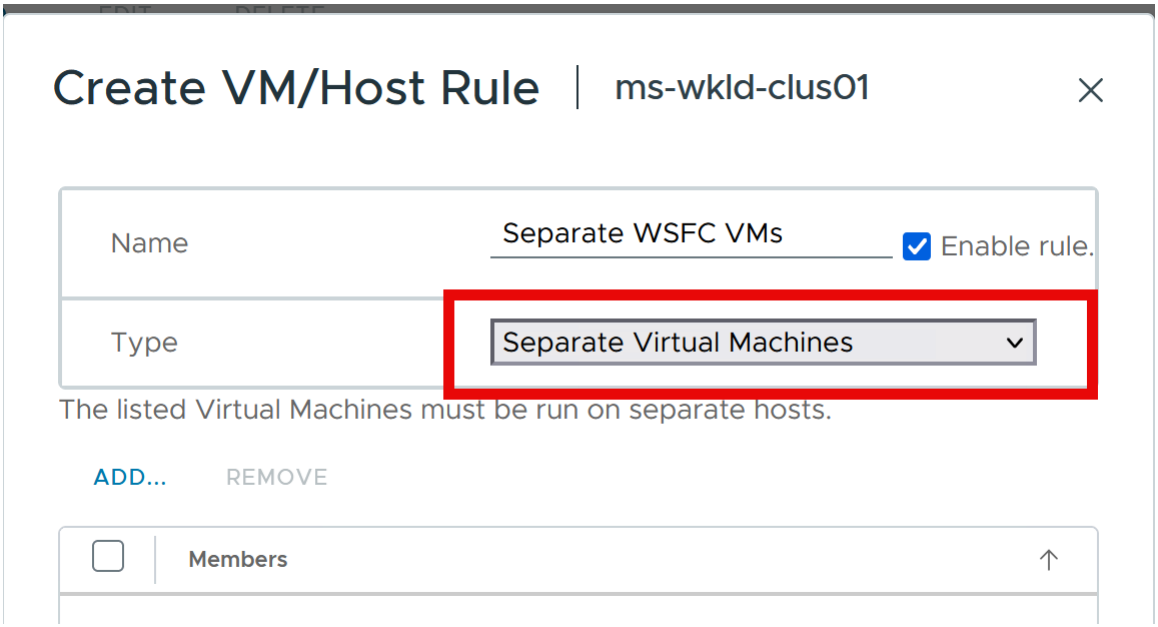


Figure 46 - Creating a VM-VM Anti-Affinity Rule

6. Click **Add** within the rule window and select the VMs that need to be separated. See **Figure 47** below. Click **OK** when done.

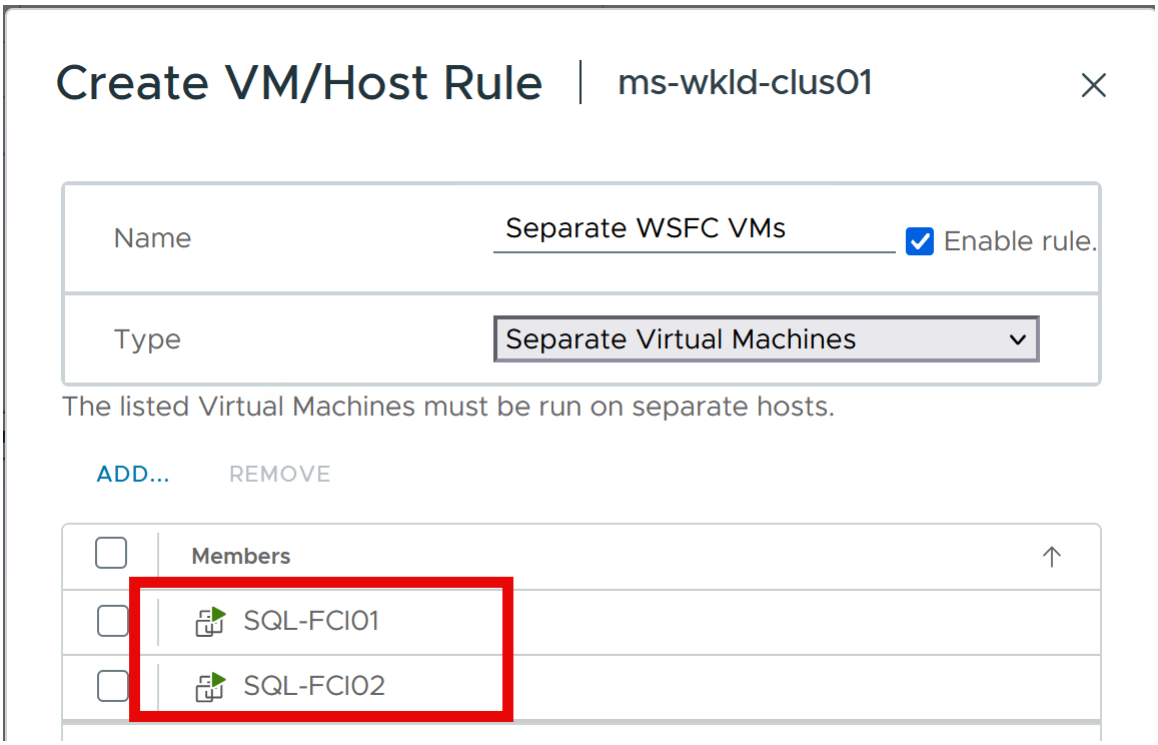


Figure 47 - Applying the Rule to Selected VMs

7. Confirm that the rule is listed and enabled, as shown in **Figure 48**.

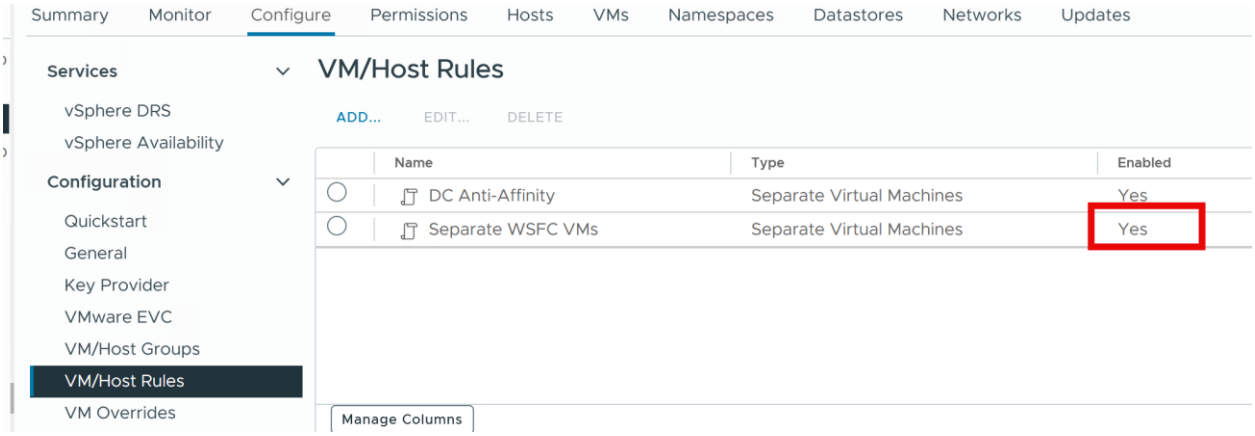


Figure 48 - Verify that Rule is Created

From this point forward, vSphere DRS will ensure that, to the extent possible, the two VMs are not co-located on the same ESXi Host.

Automation and VM Overrides

As previously mentioned, the default automation behavior for DRS is "Manual." This means DRS will not re-balance VMs even if it detects unbalanced resource utilization. For efficiency and optimal performance, Administrators should consider changing this setting. DRS uses vMotion to migrate a VM from one host to another if it determines that doing so will benefit the VM and its workload. Substantial improvements in vMotion logic and operations over the years have made these operations very suitable for even the most resource-intensive and sensitive applications in a vSphere environment.

However, if the policies and rules configured for DRS seem to cause a noticeable impact on SQL Server's performance due to unwanted migration, DBAs and virtualization administrators have the ability to selectively disable DRS for specific VMs. This can be done without disabling DRS for the entire cluster.

Instead of completely disabling DRS for the VM, you can use an option called **VM Overrides** to fine-tune vSphere's automation levels for specific VMs. Here is a high-level description of how to accomplish this:

1. From the DRS configuration screen used in previous sections, click on **VM Overrides**, as shown in **Figure 49**. Then click **Add**.

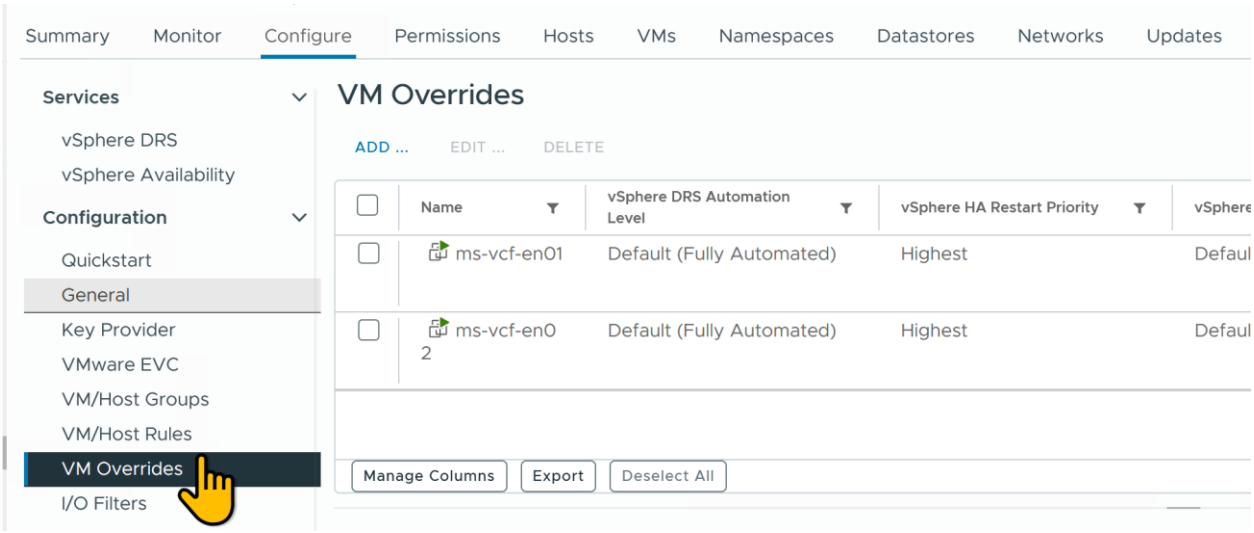


Figure 49 - Create DRS Rules Overrides on for a VM

2. Select the VM that will have overrides configured and click **Next**. An example is shown in **Figure 50**.

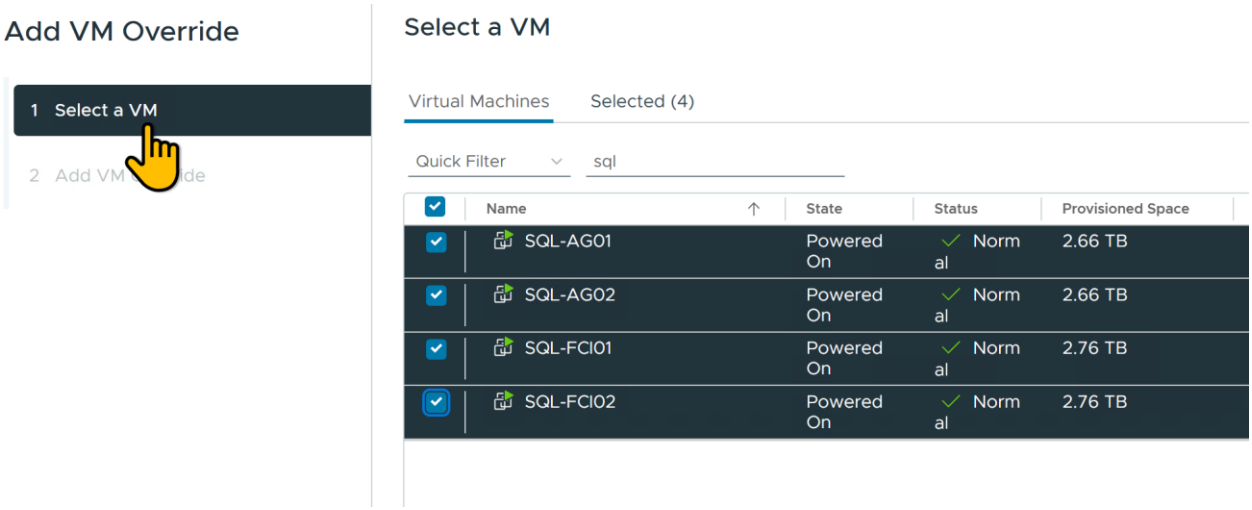


Figure 50 - Apply the Overrides to Desired VM

3. Select the desired automation levels for the VM (e.g., Partially Automated or Disabled) as shown in **Figure 51**, then click **Finish**.

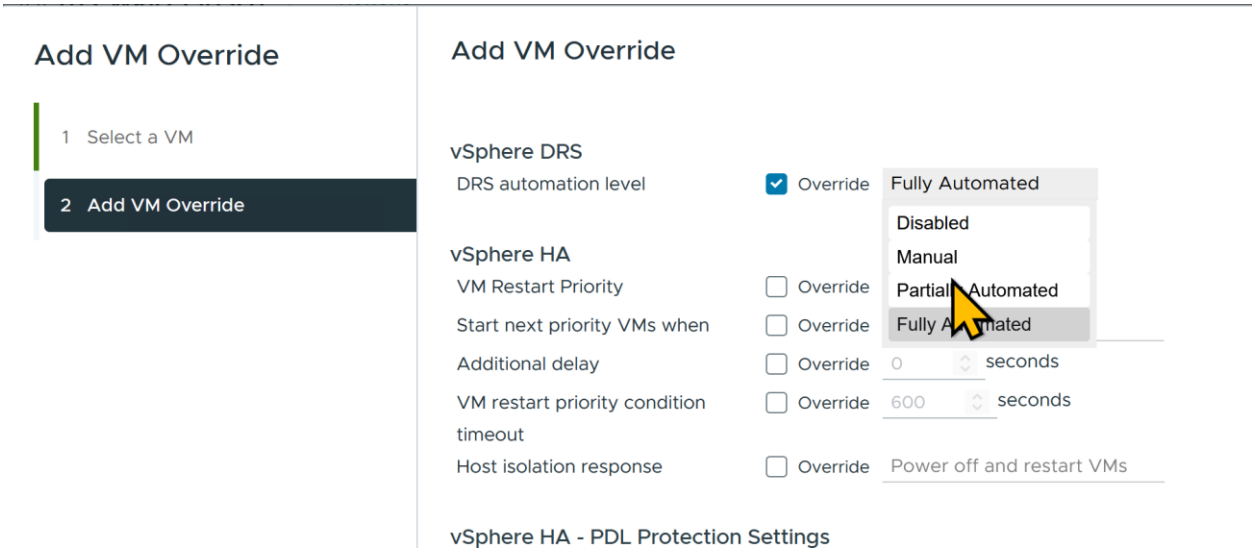


Figure 51 - Configure Desired VM-Specific Response Rule

If vSphere High Availability is enabled on the cluster, the applicable automation settings for HA can also be adjusted through this same override menu.

When enabled, DRS also plays a role in determining the initial placement of a VM in a vSphere cluster. It detects the most suitable ESXi Host for the VM and, depending on the automation level, either places the VM on that host automatically or makes a recommendation for placement.

DRS continuously evaluates the health of the cluster and re-evaluates its guidance every 5 minutes in response to changing conditions. However, other conditions (for example, a vSphere HA event caused by the sudden non-availability of an ESXi Host) may trigger a more immediate response and application of rules.

Deprecated Behavior

In earlier vSphere versions (6.x → early 7.x), vSphere HA could *choose* to violate anti-affinity rules during an HA failover unless you explicitly enable the option for this to not happen. The **"vSphere HA must respect VM anti-affinity rules during failover"** and **"vSphere HA should**

respect VM to Host affinity rules during failover" settings no longer exist because **HA now *always* respects DRS rules by design**, and there's no toggle to disable or enable these settings anymore. Now:

- **HA always honors VM-VM anti-affinity rules**
- **HA always honors VM-Host affinity rules**
- **DRS and HA jointly enforce placement** during failover

This change took effect when VMware unified DRS/HA rule handling in later 7.x releases and is carried forward into 8.x and VCF9.

vSphere vMotion

vSphere vMotion gives a VM the ability to move from host to host, network to network, storage to storage, and datacenter to datacenter while the VM remains online and continues to provide services. This means that if SQL Server is running in the guest, there is no downtime - applications and end users can still use it during the migration with minimal impact on performance.

For mission-critical workloads that need to be rehosted, this is one of the biggest benefits of virtualization. vMotion enables seamless hardware migrations for vSphere architectures with no downtime.

For recommended configuration information and considerations regarding SQL Server clustering, see vMotion support for WSFC. Additional information for architecting your VMware infrastructure to optimally support vMotion for enterprise-class SQL Server instances is available in Architecting Microsoft SQL Server on VMware vSphere.

vMotion is a standard feature in the vSphere infrastructure, so there is no VM-specific configuration required for the VM to be vMotion-eligible. Improvements in vMotion logic have made it suitable and non-disruptive for even the most demanding SQL Server instances.

The "Stun" and Heartbeat Sensitivity

One common challenge when performing a vMotion operation on a VM participating in a clustered SQL Server configuration is the possibility of a temporary service outage. It is useful to describe the conditions that trigger this.

When a vMotion operation is invoked (manually by an Administrator or automatically by DRS), vSphere copies the VM's state from its current host to a target host. This copy operation is iterative and incremental. The duration is dictated by factors such as the size of compute resources, the frequency of change in the guest OS (memory dirty rate), and network bandwidth.

When vSphere has copied enough of the VM's state and determines that the remainder can be transferred in one final transaction, vMotion invokes the **Windows Volume Shadow Service (VSS)** and requests that it freeze the VM so vMotion can perform the last copy (often called the "stun").

When Windows freezes the VM, all operations inside Windows are suspended. During this suspended state, the VM is unable to communicate with its WSFC peer - specifically, it is unable to send or receive cluster heartbeat probes.

- **Legacy Windows:** In earlier versions of Windows, the threshold for missed heartbeats before WSFC declared a node unavailable was 5 seconds. If the VM remained suspended for longer than 5 seconds, WSFC would trigger a resource failover.[12]
- **Modern Windows:** In later versions, this threshold has been increased to **10 seconds** (same subnet) and **20 seconds** (cross-subnet).

This interaction is why unintended resource failover incidents sometimes occur during vMotion. This behavior is neither peculiar nor unique to virtualized SQL Server; the trigger is the invocation of VSS, which happens similarly on physical servers during snapshot operations.

In a properly configured vSphere infrastructure (e.g., adequate compute resources, 10Gb+ network cards, multi-NIC vMotion Portgroups, and Jumbo Frames enabled), vMotion's last copy operation should complete well before reaching the point of triggering WSFC's resource failover behavior.

Impact on Anti-Affinity Rules

When an affinity/anti-affinity rule is employed with DRS, it can technically be violated by a manual vMotion. However, because DRS rules run periodically, this issue will correct itself. DRS will detect the violation and move the VMs that should be separated to appropriate hosts automatically.

Performance Benefits

From a SQL Server perspective, one of the biggest benefits of vMotion is that everything in memory - such as execution plans and recently used data in cache - remains intact. Therefore, the performance of SQL Server instances and databases remains consistent, rather than hitting the "reset button" as occurs in a traditional failover.

vSphere High Availability

vSphere High Availability (HA) provides high availability for virtual machines by pooling the VMs and the hosts they reside on into a protected cluster. Hosts in the cluster are monitored continuously; in the event of a failure, the virtual machines on a failed host are restarted on alternate hosts.

When vSphere HA is enabled, one host in the cluster is elected as the **Primary Host**. All other hosts are considered **Secondary Hosts** and can become the primary host if the original primary becomes unavailable. The primary host is responsible for monitoring the state of all VMs and hosts in the cluster. It communicates this status to vCenter and determines the appropriate response in the event of a host failure.

vSphere HA is a matured technology which has become a staple feature of the vSphere/VCF platform. When enabled, it is responsible for providing continuity for a protected VM if the VM's host becomes unavailable. Specifically, vSphere HA restarts the impacted VMs on a surviving host.

Infrastructure HA vs. Application HA

It is important to note that, although vSphere HA has introspective insights into the health of a VM and its operating system, vSphere HA is not an application-level HA solution. It is not a replacement for WSFC, FCI, or AG.

vSphere HA protects the **VM**, not the application within it. As long as a VM's heartbeat is detectable and determined to be operational, vSphere HA does not intervene - even if the application inside the VM (e.g., SQL Server) has crashed or become inaccessible.

How They Work Together

vSphere HA complements - and does not interfere with - WSFC, AG, and FCI. By ensuring that a failed VM (which has failed because its physical host failed) is promptly restarted on another host, vSphere HA dramatically reduces the duration of the outage and helps improve the application's native high availability.

For illustration, consider a scenario in a 2-node WSFC-AG configuration:

- **Steady State:** "VM-A" is on Host A. Everything is operational. VM-A is the primary node for a critical database.
- **The Failure:** In the middle of the night, Host A fails physically, taking VM-A down with it.
- **WSFC Response:** WSFC detects the loss, moves the clustered resources to the second Node, and the SQL Server AG makes the database accessible within an acceptable window.
- **The Problem:** The SQL Server infrastructure is now operating in a degraded, non-highly available, single-node state because VM-A is down. Without vSphere HA, VM-A will remain down until an Administrator intervenes.
- **vSphere HA Response:** With vSphere HA, VM-A is immediately powered on on a different Host while Host A is still down. Within a minute or two, VM-A has rejoined the WSFC cluster and HA stability is restored.

Of course, VM-A is no longer the primary node at that point (unless auto-failback has been configured in WSFC), but this is immaterial from a high availability perspective. The critical factor is that redundancy has been restored automatically.

Can vSphere HA Replace WSFC?

Some customers inquire about the feasibility of using vSphere HA to provide "good enough" high availability for their SQL Server workloads. This inquiry is primarily driven by cost considerations (e.g., avoiding the cost of extra licenses for the extra node, WSFC and SQL clustering complexities, additional storage for AGs, etc.).

VMware's position is that this choice should be informed primarily by the Customer's and Business Owner's SLA and resource availability expectations. If a short interruption of service is acceptable, then vSphere HA provides enough protection for SQL Server.

Important Caveats:

1. **Manual Verification:** If relying solely on vSphere HA, someone responsible for SQL Server must check that the databases are healthy after the VM reboots.
2. **App Awareness:** If using vSphere HA without native Windows/SQL Server HA options, applications and end-users *will* notice that SQL Server was unavailable while the VM was being brought online on the new host.
3. **Resiliency:** Unlike WSFC/AG, where applications can be coded to be cluster-aware and resilient, there is no equivalent API for vSphere HA. vSphere HA is essentially a "stop and start" event, identical to a physical host power cycle, so SQL Server will go through its standard crash recovery mechanisms upon boot.

Key design considerations include:

- Memory reservations: All SQL Server VMs participating in WSFC should have full memory reservations to prevent ballooning and swapping.
- HA isolation response: Misconfigured isolation responses may cause WSFC nodes to power off during transient network issues, triggering unnecessary cluster failovers.
- Restart priority: FCIs and synchronous AG replicas should be assigned higher restart priority to reduce recovery time after host failures.

vSphere Fault Tolerance

vSphere Fault Tolerance (FT) creates an identical clone of a VM and keeps it continuously up-to-date. This means that if the ESXi host of the original VM encounters a problem, the FT copy can replace it instantly. However, this failover only helps if the infrastructure fails; it does not protect against application-level issues, such as a corrupted installation of SQL Server or a Windows Server OS crash (BSOD).

FT maintains a 1:1 ratio from the primary to the secondary VM, meaning you can only have a single shadow copy of a source VM.

If a failover occurs and the secondary VM becomes the new Primary VM, a **new** secondary VM will be automatically deployed for its replacement on another ESXi Host to restore protection. The primary and secondary FT VMs cannot reside on the same host; therefore, a minimum of three ESXi Hosts is required in the vSphere cluster for full protection. Only one copy of an FT VM is accessible (readable/writable) at any given time.

Limitations

VMware Fault Tolerance has strict limitations on the VM's compute resource capacity, which often makes it unsuitable for large enterprise SQL Server instances. Table 3 below shows the maximum supported compute resource allocation for VMs participating in Fault Tolerance:

Config Maximum	vSphere 9.x, 8.x, 7.x
Max CPU Per VM	8
Max RAM Per VM	128 GB
Max Virtual Disk Per VM	16
Max Disk Size Per VM	2 TB
Max FT VMs Per Host	4
Max FT CPU Per Host	8

Table 2 - vSphere Fault Tolerance Maximums

It is also important to note that because FT protects at the VM layer, the secondary VM is an exact execution mirror of the primary VM. This means that any OS or application defect on the primary VM is replicated intact to the Secondary VM. VMware FT is not an exact replacement for application-level HA.

Note: VMware FT is **not supported** with clustered configurations of SQL Server (WSFC).

vMotion Application Notification

vMotion Application Notification is a significant enhancement introduced in vSphere 8 and fully integrated within vSphere and VMware Cloud Foundation (VCF) 9. This feature enables applications running inside a virtual machine - including SQL Server 2025 - to receive advance notice of an impending live migration event. This capability allows SQL Server to prepare for the brief operational pause associated with vMotion, reducing the likelihood of unintentional failovers or service disruptions in Windows Server Failover Clustering (WSFC), Always On Failover Cluster Instances (FCIs), and Always On Availability Groups (AGs).

vMotion and vMotion Application Notification

Live migration of SQL Server VMs can momentarily pause execution during the switchover phase. Historically, this pause - known as stun time - could trigger WSFC heartbeat loss or AG replica degradation. Modern vSphere dramatically reduces stun times, especially for large-memory VMs, and VCF 9 incorporates further optimizations that improve reliability of SQL workloads during maintenance windows.

The **vMotion Application Notification** framework introduces a new interaction path between the hypervisor and guest OS:

- Before migration begins, VMware Tools sends a pre-migration event.
- The guest OS and applications can prepare for short interruptions.
- SQL Server may quiesce critical operations, reduce the risk of replica demotion, or adjust log synchronization.
- After the guest acknowledges readiness or a timeout is reached, vMotion proceeds.

This feature should be enabled on all SQL Server VMs participating in FCIs or AGs, particularly synchronous commit replicas and primary roles.

vMotion and Microsoft Cluster Stability

Clustered Microsoft SQL Server replicas may be impacted by transient delays during live migrations. VMware's "vMotion Application Notification" framework - fully integrated in vSphere - significantly reduces the risk of replica demotion or failover during vMotion.

When enabled:

- The guest OS receives an event indicating that a migration is about to begin.
- SQL Server or supporting agents can quiesce sensitive operations.
- Replica log send/receive threads can be momentarily buffered.
- The risk of failover due to heartbeat timeouts is reduced.

This feature is strongly recommended for AG nodes, particularly those running synchronous commit replicas or hosting the AG primary. Combined with proper DRS anti-affinity rules, redundant networking designs, and storage best practices, SQL Server 2025 AGs offer robust and predictable performance under VMware vSphere and VCF 9.

The following is a high-level description of how vMotion Application Notification works to help minimize unplanned outage or failover during vMotion or DRS operations.

Pre-Migration Signaling

When vMotion Application Notification is enabled for a VM, the following sequence occurs:

1. vMotion is initiated by DRS, maintenance mode, or administrator action.
2. VMware Tools inside the guest OS receives a pre-migration notification event.
3. The guest OS or application-level agents may take preparatory actions, such as:
    - Temporarily pausing latency-sensitive operations.
    - Flushing buffers or stabilizing in-flight transactions.
    - Preparing for WSFC heartbeat sensitivity.
    - Managing AG replica synchronization pipelines.

SQL Server 2025, working together with WSFC and AG health logic, can use this notification window to ensure stability during the subsequent migration process.

SQL Server Behavior During Migration

Historically, the "stun time" at the end of a vMotion operation could disrupt WSFC communication or AG replica synchronization, especially for synchronous commit configurations. These disruptions sometimes triggered:

- Unexpected AG failovers
- Replica demotions
- FCI resource group movement
- WSFC heartbeat warnings or failures

With Application Notification:

- SQL Server can anticipate the stun period and reduce the likelihood of false failovers.
- Replica threads can buffer incoming log blocks.
- Heartbeat sensitivity can be accommodated.
- Highly active workloads can avoid mid-transaction sensitivity.

This is especially important for synchronous replicas, where even small pauses can cause significant cluster reactions.

Timeouts and Fallback Behavior

If SQL Server or other processes inside the VM do not acknowledge the pre-migration event within a defined timeout period:

- vMotion continues regardless, ensuring migration operations do not stall.
- Cluster sensitivity remains higher, and administrators should investigate whether relevant agents or scripts are misconfigured.

Integration with DRS and vSphere HA

vMotion Application Notification complements DRS and vSphere HA in the following ways:

- During **DRS-initiated migrations**, workload movement becomes more predictable and cluster-safe.
- During **Host maintenance mode evacuations**, SQL Server VMs can migrate without destabilizing WSFC or AG topologies.
- In **vSphere HA failover scenarios**, restarted VMs may migrate automatically after recovery; Application Notification reduces post-restart instability.

Implementation Requirements

To use vMotion Application Notification:

- VMware Tools must be installed and up to date.
- Set a Timeout value for vMotion Application Notification on all Hosts that the VM can potentially run on ***(VmOpNotificationToApp.Timeout = "90")***

Note: Because vMotion will automatically start once this timeout period is reached (regardless of whether the VM is ready or not), we strongly recommend that you set this value as high as reasonably tolerable.

- The advanced VM setting **"vmOpNotificationToAppEnabled"** must be enabled ***(vmx. vmOpNotificationToApp.Enabled=True)***.
- Set a Timeout value for how long vMotion Application Notification should wait for the Guest-side application preparation to complete ***(vmx.vmOpNotificationToApp.timeout = "60")***
- From within the VM's Guest Operating System, create and register an "Application" to periodically listen and poll for vMotion Application Notification events.

Note: These events are communicated to the VM through the in-Guest VMware Tools. Because a vMotion operation can be triggered at any given time, we highly recommend that the polling interval be as low as possible

(in our sample script, we poll every second). See How to Register an Application for vSphere vMotion Notifications for more details

Important Notes:

- Application registration does NOT persist across VM reboots. This means that a registration must be performed every time the VM is powered off and on, or when rebooted. In Windows, this is easily achieved by adding the registration process to Windows Task Scheduler as a Machine Startup script.
- "Application" (as used in this registration context) does not refer to application(s) installed on the VM. It is called "Application" only within the vMotion Application Notification naming constructs.
- The effective Timeout period is whichever is lower between the Host's and the VM's Timeout values. For example, if you set it to 60 seconds on the Host and 120 seconds on the Guest, vMotion will wait for 60 seconds and begin immediately thereafter.
- Regardless of what it's set to on either side, vMotion will automatically start once it receives acknowledgment from the VM's registered "Application". For example, if you set it to 60 on the VM and 90 on the Host, but the VM sends a notification to the Host after 10 seconds, vMotion will commence immediately. It will not wait for the 60 seconds (the lower value) specified on the VM
- Microsoft does not ship a built-in SQL Server or WSFC listener for vMotion events, but third-party solutions and custom scripts can be used to automate quiescing and preparation steps.

Recommended Usage

vMotion Application Notification is strongly recommended for:

- SQL Server 2025 FCIs
- Synchronous commit AG replicas
- Mission-critical primary AG replicas
- Latency-sensitive workloads
- Large memory SQL Server VMs with potentially longer stun windows

When properly enabled and integrated, this feature significantly improves operational stability during both planned maintenance and automated VM migrations, ensuring SQL Server 2025 high availability configurations function predictably in modern VMware vSphere and VCF environments.

Site Recovery Manager

vSphere Site Recovery Manager (SRM) is an orchestration solution that enables disaster recovery (DR), avoidance, and site migration. DR allows you to come online elsewhere if your primary data center experiences a major downtime ev[6]ent. SRM requires the use of either block-level replication (storage array-based) or vSphere Replication to synchronize VMs between sites.

Non-Disruptive Testing

One of the core value-adds of SRM as a Business Continuity and Disaster Recovery (BCDR) solution is that it enables Administrators, Business Owners, and Auditors to perform simulated BCDR scenarios without any negative impact on the Production environment. Customers can failover an entire SQL Server farm and all its upstream dependencies into an SRM "Bubble Network" (an isolated test network) without interrupting continuous access to the live environment. This allows customers to continuously verify, validate, and refine their protection and recovery plans on-demand.

Automated Recovery Workflows

Programmable and configurable recovery automation is essential for an optimal BCDR Plan. Manual administrative tasks are prone to human error, especially during the stress of a catastrophic disaster. SRM minimizes these manual tasks through specific features:

1. IP Reconfiguration:

A DR site is usually geographically separate from the production site. Without expensive network stretching technologies, the IP addresses at the DR site are usually different. If you recover a protected VM with an IP of a.b.c.d to a DR site where the subnet is e.f.g.x, the recovered VM will not be able to communicate.

SRM solves this by providing a mechanism to preconfigure the target IP address that a recovered VM should adopt immediately upon recovery at the DR site.

2. Recovery Dependencies:

   SQL Server instances are usually part of multi-tiered application constructs. They provide services leveraged by other servers (e.g., Web Servers) and depend on infrastructure services like Active Directory and DNS. In a DR event, the order of recovery is critical. If you recover an AD-joined SQL Server VM before the Domain Controller is available, services will fail.

   SRM solves this by enabling an Administrator to preconfigure the precise order in which all protected VMs are recovered.

3. Script Execution:

   Restoring service continuity often requires more than just powering on servers. Specifically for WSFC, AGs, and FCIs, the cluster has additional IP addresses (Cluster IP, Listener IP) that are distinct from the OS-level TCP/IP address.

   While SRM can automatically change the Windows OS IP address (as described in point 1), it does not natively change the SQL Cluster or Listener IP addresses inside the application. SRM solves this by allowing Administrators to invoke post-recovery scripts. These scripts can automatically update the cluster resources to match the new DR site configuration.

An exhaustive and comprehensive documentation of how to protect and recover a production Microsoft SQL Server infrastructure is available in Protecting and Recovering Mission-Critical Applications in a VMware Hybrid Cloud with Site Recovery Manager.

Of importance in this document is the fact that SRM supports WSFC, AG, FCI, and standalone SQL Server instances without any limitations.

VCF Features and Microsoft Cluster Alignment

A well-engineered SQL Server high availability environment requires alignment between hypervisor-level and cluster-level policies:

- **Anti-affinity rules** prevent co-location of critical replicas.
- **Restart priorities** ensure essential nodes recover first.
- **Quorum design** aligns with potential host failures.
- **Network redundancy** in the VDS or NSX-T ensures stable cluster heartbeats.
- **Storage availability** across hosts ensures rapid VM restart after host failures.

When designed correctly, SQL Server 2025, WSFC, and vSphere operate collaboratively: WSFC provides database-level failover logic, vSphere HA ensures rapid VM recovery, DRS maintains safe placement, and vMotion Application Notification enhances workload stability during maintenance operations.

Storage Planning for SQL Server Under vSphere

Storage planning remains one of the most important design considerations when deploying SQL Server 2025 on VMware Cloud Foundation (VCF) 9. SQL Server performance is fundamentally tied to predictable storage latency, stable throughput, and properly isolated I/O patterns. While virtualization abstracts physical storage, the underlying design still determines the behavior, stability, and performance of SQL Server workloads - particularly those involving Always On Failover Cluster Instances (FCIs) or Always On Availability Groups (AGs).

This Section is devoted to ensuring that SQL Server storage under vSphere/VCF is deployed using reliable, supportable, and high-performing architectures aligned with both Microsoft's and VMware's 2025 guidance. Let's first take cognizance of the fact that recent improvements in both VCF and Microsoft SQL Server have altered some of the storage-related guidance in previous versions in the following ways:

- vSAN ESA (Express Storage Architecture) has become the preferred vSAN deployment model for SQL Server, offering significantly improved latency and throughput.
- "Clustered VMDK" datastores have matured into the primary shared-disk technologies for FCIs, reducing reliance on Raw Device Mappings (RDMs).
- NVMe-based storage systems - whether vSAN ESA, VMFS on NVMe arrays, or vVols backed by NVMe fabrics - are now standard for performance-sensitive SQL Server deployments.
- SQL Server 2025 includes improved I/O scheduling, TempDB enhancements, and more predictable write patterns that impact storage selection and sizing.
- VCF 9 storage stack enhancements (improved I/O scheduler behavior and optimized virtual NVMe controllers) further influence recommended design.

Although virtualization abstracts the hardware layer, storage misconfiguration remains one of the most common causes of SQL Server performance issues. SQL Server's sensitivity to storage latency - especially for write-heavy OLTP workloads and synchronous AG replicas - makes proper sizing and design essential. VMware recommends following the principles in this section when planning storage for SQL Server under vSphere or VCF.

Regardless of the selected storage platform - VMFS, vVols, vSAN ESA, or external array - storage for SQL Server 2025 should be designed to meet the following core objectives:

- Low and predictable latency
- High IOPS availability during peak conditions
- Isolation of data and log workloads when beneficial
- Full redundancy across hosts and failure domains
- SCSI controller optimization and distribution across paravirtual adapters
- Consistent configuration across all SQL Server nodes

This section also outlines how modern VMware storage technologies interact with SQL Server high availability features. FCIs require SCSI-3 Persistent Reservation support, which is now fully provided by Clustered VMDKs, vVols, and vSAN ESA. AGs, by contrast, have no shared storage requirements but depend heavily on network throughput and replica database storage sizing.

Finally, this storage guidance aligns with VMware's most recent support statements for SQL Server on vSphere, including storage controller recommendations, disk provisioning formats, and datastore design. Because storage misconfiguration can manifest as SQL Server performance stalls, AG replica lag, or FCI failover delays, careful planning remains essential.

The following subsections provide updated guidance for datastore selection, storage controller configuration, disk layout, vSAN ESA and OSA considerations, vVol splanning, FCI shared-disk architecture, and performance tuning aligned with VMware's 2024–2025 best practices.

Datastore Options and Considerations

Selecting the appropriate datastore type is a foundational decision when deploying SQL Server 2025 on VMware vSphere or VMware Cloud Foundation (VCF) 9. Although SQL Server is highly flexible in terms of where its data and logs can reside, not all datastores behave the same under load, and not all are equally suitable for mission-critical, latency-sensitive workloads. This subsection provides updated guidance for VMFS, vSAN (ESA and OSA), and Virtual Volumes (vVols), and describes how these datastore technologies affect SQL Server performance, recoverability, and support boundaries.

VMware no longer treats the datastore layer as simply a storage endpoint. Modern vSphere and VCF storage platforms provide integrated intelligence for I/O scheduling, multipathing, storage policy enforcement, and metadata operations. SQL Server 2025 introduces its own

enhancements - such as improved parallel file processing in TempDB initialization, better I/O queuing behavior, and enhanced logging throughput - that interact closely with these underlying datastore technologies. Choosing the correct datastore design is therefore essential for achieving predictable performance at scale.

VMFS Datastores

VMFS (Virtual Machine File System) continues to be widely deployed and remains fully supported for SQL Server 2025 workloads. VMFS is simple, resilient, and highly compatible with the full vSphere feature set. When backed by modern NVMe arrays or all-flash SAN appliances, VMFS offers excellent latency and throughput characteristics suitable for OLTP and analytics workloads.

For SQL Server deployments on VMFS:

- Use thick provision eager-zeroed VMDKs for SQL data and log files.
- Place VMDKs across multiple Paravirtual SCSI (PVSCSI) controllers to distribute I/O queues.
- Ensure underlying storage arrays offer redundant fabric paths with multipathing enhancements such as NMP or vendor-specific PSPs.
- Avoid mixing high-throughput SQL workloads with unrelated VM workloads on the same datastore during peak periods.
- Prefer arrays with NVMe-TCP, NVMe-FC, or high-speed SAS backends, as these deliver the lowest write latency.

VMFS remains an excellent choice when properly sized and isolated, especially for AG-based architectures where each replica uses independent storage.

vSAN Express Storage Architecture (ESA)

vSAN ESA represents VMware's modern distributed storage architecture and provides substantial performance benefits for SQL Server 2025 workloads:

- Predictable low latency for both reads and writes
- Improved resiliency via log-structured writes
- Enhanced compression, checksum, and metadata handling
- High throughput with NVMe-first data paths

For SQL Server clusters:

- ESA supports Clustered VMDKs, enabling WSFC-based FCIs without RDMs.
- Storage policies allow administrators to enforce IOPS limits, availability rules, and resilience settings per VMDK.
- Write-intensive log volumes benefit significantly from ESA's log-optimized storage engine.

When deploying AGs, ESA distributes each replica's storage independently, which aligns with the AG architecture's no-shared-disk model.

ESA is now VMware's recommended architecture for new vSAN deployments hosting SQL Server.

vSAN Original Storage Architecture (OSA)

Although still supported, vSAN OSA should be considered a legacy platform for SQL Server 2025 deployments. OSA relies on caching tiers and disk groups that were designed for earlier generations of storage hardware. SQL Server workloads will still run correctly, but:

- Latency may be higher compared to ESA.
- Write-intensive workloads may place significant pressure on the caching tier.
- Future VMware optimizations will target ESA rather than OSA.

Customers deploying new environments should strongly consider ESA, while existing OSA clusters should follow VMware's SQL Server design guidelines carefully.

vVols (Virtual Volumes – Deprecated in VCF 9)

vVols provide the most flexible storage integration for SQL Server under vSphere 9. Because each VMDK becomes an object managed directly by the storage array:

- Storage policies can enforce performance, redundancy, or encryption per disk.
- Arrays can directly offload snapshots, clones, replication, and backup operations.
- vVols fully support SCSI-3 Persistent Reservations, making them an ideal option for FCIs.
- Each SQL Server database component (data, log, TempDB) can be assigned its own performance class.

vVols are especially beneficial for enterprise SQL Server deployments where granular SLAs are required at the disk or database level.

Key Considerations Across All Datastore Types

Regardless of datastore selection, SQL Server 2025 deployments under vSphere must follow several universal rules:

- Ensure sufficient IOPS headroom: SQL Server spikes can exceed baseline patterns dramatically.
- Monitor storage contention: SQL workloads often reveal datastore bottlenecks before other workloads do.
- Use PVSCSI controllers: They remain best practice for all SQL data, log, TempDB, and backup disks.
- Distribute VMDKs across controllers: SQL Server benefits from multiple queue depths.
- Avoid oversubscription: Particularly on shared arrays or vSAN OSA clusters.
- Validate latency: Synchronous AG replicas require low, consistent latency - typically under 5–10 ms.

As SQL Server continues to push the boundaries of storage throughput, the datastore layer must be engineered to support both peak I/O demand and long-term operational stability.

Storage Controller and Virtual Hardware Configuration

Storage controller configuration is a critical element of SQL Server performance under VMware vSphere and VMware Cloud Foundation (VCF) 9. While storage platforms have evolved significantly, SQL Server's fundamental dependency on predictable disk I/O patterns has not. Incorrect controller configuration, insufficient queue depth, or suboptimal virtual hardware choices can lead to latency spikes, throughput collapse, and degraded performance for SQL Server 2025 workloads - particularly those involving Always On Availability Groups (AGs) or Failover Cluster Instances (FCIs).

This subsection provides updated 2025 guidance on VMware storage controllers, multi-controller layouts, NVMe virtualization, queue-depth distribution, and virtual hardware configuration to ensure optimal performance and predictable failover behavior.

Paravirtual SCSI (PVSCSI) Recommendation

VMware Paravirtual SCSI (PVSCSI) remains the recommended controller for all SQL Server 2025 data, log, TempDB, and backup disks. PVSCSI provides the following benefits:

- High queue depths suitable for SQL OLTP workloads
- Low CPU overhead
- Predictable latency under heavy load
- Multi-queue parallelism for storage-intensive workloads

All SQL Server VMDKs should be attached to PVSCSI controllers unless using virtual NVMe devices.

Deprecated controllers such as LSI Logic SAS or LSI Parallel should not be used for SQL Server deployments in 2025.

Multiple PVSCSI Controllers

To maximize parallelism and reduce contention, VMware recommends distributing SQL Server VMDKs across multiple PVSCSI controllers. SQL Server 2025 can issue a high volume of I/O operations, and a single PVSCSI controller may become a bottleneck during periods of peak write activity or heavy read operations.

Best practices include:

- Assign separate controllers for data, logs, and TempDB
- Limit each controller to 1–4 VMDKs for high-throughput workloads
- Ensure controller assignments are consistent across AG replicas or FCI nodes

Queue depth improves when workloads are spread across controllers, helping prevent latency spikes and improving AG synchronization throughput.

Virtual NVMe Controllers

VMware's virtual NVMe controller has matured significantly in vSphere 8 and 9 and is now a strong option for SQL Server 2025 deployments, especially on NVMe-backed storage arrays or vSAN ESA. NVMe controllers offer:

- Lower latency than PVSCSI
- Higher I/O parallelism
- Reduced CPU overhead
- Improved efficiency for log-intensive workloads

However, NVMe virtual controllers do not support Clustered VMDKs, which makes them unsuitable for FCIs. They are recommended for:

- Standalone SQL Server instances
- AG replicas (primary or secondary)
- Read-scale replicas

Administrators should ensure:

- Compatibility between storage platform and VMware NVMe controller
- Consistent controller type across AG nodes

Queue Depth and Performance Considerations

Queue depth remains a critical tuning area. SQL Server generates bursty I/O patterns, especially during:

- Checkpoint operations
- TempDB spills
- Log flush cycles
- AG synchronization under synchronous commit

PVSCSI controllers allow high queue depths by default, but the underlying storage must be able to sustain the volume. On VMFS or vVols, ensure the array supports parallel I/O scaling. On vSAN ESA, queue depths dynamically adjust based on cluster conditions but benefit from:

- Adequate host-side NVMe capacity
- Proper storage policy sizing
- Avoiding unnecessary contention with non-database workloads

Queue-depth bottlenecks may manifest as:

- Write latency spikes
- AG replica lag
- FCI failover delays

Virtual Hardware Version

SQL Server 2025 deployments should use the latest supported virtual hardware version on vSphere 9, which provides:

- Enhanced vNUMA topology
- Optimized interrupt handling
- Virtual NVMe support improvements
- More efficient PVSCSI mappings
- Better CPU scheduling behavior

Virtual hardware upgrades should be planned during maintenance windows and validated before being applied to production SQL workloads.

vNUMA Alignment

SQL Server 2025 is NUMA-aware, and improper virtual NUMA alignment can adversely affect memory and I/O performance. Best practices:

- Ensure vNUMA topology matches physical NUMA boundaries
- Avoid spanning vCPUs across NUMA nodes unnecessarily
- Align storage-intensive threads (such as log writers) with consistent NUMA scheduling

Misalignment can impact I/O operations significantly, particularly under heavy OLTP workloads and AG synchronous commit configurations.

Cluster-Specific Considerations

For FCIs:

- Only PVSCSI controllers should be used
- SCSI bus sharing must align with VMware's FCI documentation
- Controller configuration must be identical between nodes

For AGs:

- NVMe controllers may be used for performance gains
- VMDKs do not need to be shared, simplifying controller layout

Consistency across replicas is still critical for predictable performance

Summary

Correct storage controller configuration is fundamental to achieving predictable and stable SQL Server 2025 performance under vSphere 9/VCF 9. Multiple PVSCSI controllers, optional NVMe virtualization, queue-depth optimization, and appropriate virtual hardware versions ensure SQL Server performs reliably across all availability models.

Disk Layout and File Placement

Proper disk layout and file placement remain central to achieving predictable and stable SQL Server performance under VMware vSphere and VMware Cloud Foundation (VCF) 9. While virtualization abstracts physical storage and provides flexible provisioning, SQL Server 2025 continues to rely heavily on well-structured, isolated I/O patterns to minimize latency, prevent contention, and ensure availability across shared-nothing architectures like Always On Availability Groups (AGs) and shared-disk architectures like Failover Cluster Instances (FCIs).

This subsection provides updated best practices for VMDK layout, separation of database components, TempDB optimization, backup storage placement, and VMware-specific considerations such as Storage vMotion, snapshots, and vVols. Where relevant, SQL Server 2025 enhancements and VMware's 2024–2025 guidance are incorporated.

General Principles

Regardless of underlying datastore type (VMFS, vSAN ESA, vVols), disk layout for SQL Server should follow these core principles:

- Predictability over consolidation: Separate components that generate distinct I/O patterns.
- Isolation of write-heavy workloads: Log files and TempDB should not share disks with data files.
- Multiple controllers: Distribute VMDKs across multiple PVSCSI or NVMe controllers to improve queue parallelism.
- Consistent layout across nodes: AG replicas and FCI nodes must follow identical disk structures.
- Avoid over-provisioning: SQL Server 2025 can generate rapid I/O spikes; sustained headroom is required.

These principles apply universally across standalone, AG, and FCI deployments.

Data File Placement

Database data files (.mdf, .ndf) generate mixed read-write workloads, with random access patterns common in OLTP environments. For SQL Server 2025:

- Place database data files on dedicated VMDKs attached to a data-only PVSCSI or NVMe controller.
- Large databases may benefit from multiple data VMDKs to distribute I/O.
- On vSAN ESA, use storage policies with appropriate durability and performance characteristics.
- On vVols, assign per-disk policies aligned with the database's SLA requirements.

For AG replicas, each replica should have the same data layout, ensuring predictable replica synchronization performance.

Transaction Log File Placement

SQL Server log files (.ldf) generate sequential write-heavy workloads. Logs are uniquely sensitive to latency, especially under synchronous AG commit mode.

Best practices:

- Use dedicated VMDKs for transaction log files.
- Place log VMDKs on a separate PVSCSI/NVMe controller from data and TempDB.
- Avoid mixing multiple database log files in a single VMDK unless they have identical write patterns.
- Provide adequate IOPs and low write latency; log write stalls directly impact application throughput.
- On vSAN ESA, consider storage policies with performance/throughput enhancements tailored for write-optimized workloads.

For FCIs, shared volume (data and log) disks must be located in Clustered VMDK datastores or vVols capable of supporting SCSI-3 PR.

TempDB Placement

TempDB is one of the most critical components of SQL Server I/O, especially with TempDB-intensive workloads such as sorts, spills, hash operations, row versioning, and memory pressure events.

SQL Server 2025 introduces improvements in TempDB initialization and concurrent allocation, but optimal placement still matters:

- Place TempDB files on their own dedicated VMDKs.
- Attach TempDB disks to their own PVSCSI/NVMe controller to prevent interference with data/log operations.
- Use multiple TempDB data files, typically one per logical processor up to eight, then tune based on contention.
- Consider increasing TempDB data file sizes upfront, reducing autogrowth events.
- Avoid placing TempDB on shared disks or disks with unrelated workloads.

On vSAN ESA, SQL TempDB performs exceptionally well when policies are configured for high performance.

Backup Target Placement

Backup operations generate large sequential read and write patterns. Recommendations include:

- Place backup target VMDKs on a separate PVSCSI controller.
- Use thick-provisioned disks for high-throughput backup operations.
- For environments using backup appliances or file-based backup repositories, ensure adequate network throughput.
- Avoid performing backups to the same disk group or datastore hosting production data files.

Backup placement is especially important in AG environments where backups may be offloaded to secondary replicas.

Multiple VMDK Strategy

Using multiple VMDKs for SQL Server 2025 deployments is recommended for:

- Parallelizing I/O across controllers
- Preventing hotspots during peak loads
- Supporting multiple TempDB data files
- Allowing storage policy granularity in vVols or vSAN ESA

A typical modern layout might include:

- 1–4 data VMDKs
- 1–2 log VMDKs
- 1–4 TempDB VMDKs
- 1 backup VMDK
- Optional: file stream, full-text, or DQS-specific disks

This structure provides predictable performance and is compatible with all vSphere features, including Storage vMotion and snapshots (when permitted by SQL operational policies).

VMware Snapshot and Storage vMotion Considerations

SQL Server–based VMs can technically use snapshots and Storage vMotion, but administrators should follow best practices:

- Avoid application-consistent snapshots during periods of high write load.
- Never use snapshots for long-term retention; they negatively impact disk performance.
- Storage vMotion is supported for SQL Server VMs and AG/FCI nodes.
- For FCIs on clustered VMDKs or vVols, follow VMware's guidance regarding shared-disk movement limitations.
- Use vMotion Application Notification to minimize the risk of cluster instability during migrations.

AG and FCI-Specific Placement Notes

Always On Availability Groups

AG replicas operate independently on separate storage. Ensure:

- Identical layout across replicas
- Consistent disk sizes
- Matching controller architecture
- Fast underlying storage for synchronous replicas
- Additional headroom for log send/receive operations

Failover Cluster Instances

FCIs require:

- Shared disks implemented with Clustered VMDKs or vVols
- Identical SCSI controller numbering and bus sharing

- Placement consistency across nodes

Disk layout impacts failover time and recovery stability.

Summary

Proper disk and file placement remains a foundational factor for SQL Server 2025 performance and stability under vSphere 9/VCF 9. By isolating workloads across multiple VMDKs, distributing disks across multiple controllers, optimizing TempDB, and aligning layouts across AG and FCI nodes, organizations can achieve predictable performance and high availability for mission-critical SQL workloads.

vSAN Considerations for SQL Server

VMware vSAN has become one of the most widely deployed storage platforms for SQL Server on vSphere. With the introduction of vSAN Express Storage Architecture (ESA), VMware has significantly modernized its data path, resiliency model, and performance characteristics. SQL Server 2025 benefits from these advancements, particularly with improvements in write handling, TempDB concurrency, and predictable I/O behavior.

This subsection provides updated guidance for deploying SQL Server on vSAN ESA and legacy vSAN Original Storage Architecture (OSA). It also outlines considerations for AGs, FCIs, and standalone SQL Server workloads operating on distributed storage in VMware Cloud Foundation (VCF) 9 environments.

vSAN Express Storage Architecture (ESA)

vSAN ESA is the recommended architecture for all new SQL Server deployments. It replaces the traditional caching-tier model with a log-structured filesystem optimized for NVMe-based devices. ESA dramatically improves latency, throughput, and consistency under mixed SQL Server workloads.

Key benefits for SQL Server 2025 include:

- Low write latency due to log-structured data placement and NVMe-first pipeline
- High throughput under concurrent mixed read/write conditions
- Improved resiliency with modern RAID-5/6 implementations
- Optimized checksum, compression, and metadata handling
- No dependence on cache/buffer device tiers

ESA inherently aligns with SQL Server's pattern of:

- Log writes
- Checkpoint surges
- TempDB bursts
- AG synchronization cycles

As a result, SQL Server workloads generally exhibit more stable and predictable performance on ESA compared to OSA.

ESA Storage Policy Recommendations

Key SPBM attributes relevant to Microsoft SQL Server workloads include:

- **Failures To Tolerate (FTT)**: RAID 5/6 is recommended for ESA. If Auto-Policy management is enabled, use the optimal storage policy defined by the number of hosts in the vSAN cluster.
- **Object space reservation**: Thin provisioning is supported; thick provisioning is optional for predictable capacity management

- **Space Efficiency**: Compression is recommended for SQL Server workloads on ESA. Global deduplication (first introduced in VCF 9.0 P01) is recommended for SQL Server Always On Availability Group workloads running on the same vSAN cluster.
- IO Size and Performance Class: Automatically determined by ESA.
- **Checksum**: Always enabled and recommended

For TempDB and log volumes, consider a performance-focused policy with enhanced stripe width if required.

vSAN Data Protection

vSAN Data Protection in VMware Cloud Foundation 9 helps protect customer workloads against ransomware attacks and disaster recovery locally and remotely using vSAN ESA snapshot solutions that are easy to manage at scale.

- vSAN Data Protection enables the following use cases:
- Reverting existing VMs
- Restoring deleted VMs
- Clone VMs from snapshots
- Remote replication with VMware Live Recovery (VLR)

vSAN data protection allows up to 200 snapshots per VM, which extended the 32 snapshots per VM limit through traditional VADP UI or APIs. For mission-critical production SQL Server workloads:

- Use vSAN data protection to augment your existing backup and recovery strategy by offering new levels of protection and flexibility through native snapshots.
- Always maintain a short snapshot chain and avoid taking snapshot during periods of peak workload activity.
- vSAN does not support snapshot for shared disks consumed by SQL FCI.

vSAN Original Storage Architecture (OSA)

Although still supported, OSA should be considered legacy for SQL Server 2025 deployments. OSA's architecture relies on separate caching and capacity tiers, which are no longer aligned with modern NVMe-based designs.

SQL Server will run correctly on OSA, but administrators must be aware of:

- Increased write amplification
- Potential caching-tier saturation under heavy TempDB loads
- Higher tail latencies
- Reduced fault-domain flexibility compared to ESA

Only choose OSA if constrained by existing infrastructure or upgrade timelines.

Clustered VMDKs and FCIs on vSAN

vSAN ESA and OSA both support Clustered VMDKs, enabling SQL Server Failover Cluster Instances (FCIs) without RDMs.

Key considerations:

- SCSI-3 Persistent Reservations are fully supported
- Storage objects must be accessible from all FCI nodes
- SCSI controller numbering must be identical between nodes
- vSAN storage policies apply uniformly to all nodes accessing the shared disk

Due to ESA's improved write path, FCI failovers and recovery operations typically complete more quickly.

Availability Groups on vSAN

Availability Groups do not require shared disks, making them naturally aligned with vSAN's distributed architecture.

Best practices:

- Each replica's database disks should follow identical storage policies
- Synchronous replicas require sub–5 ms latency for best performance
- vSAN ESA's consistent write performance benefits synchronous commit AGs
- Replica placement should align with fault domains or availability zones

AG performance often improves noticeably when migrated from SAN or OSA-based systems to ESA due to lower write latencies and more predictable log throughput.

vSAN Networking Considerations

vSAN is sensitive to network performance. SQL Server adds additional pressure through AG synchronization, client connections, and cluster heartbeat traffic.

Ensure:

- Dedicated or well-segmented 25/40/100 GbE connectivity for vSAN
- Redundant physical uplinks
- Network consistent across hosts
- Low jitter and minimal packet loss

For AGs running synchronous commit, ensure that vSAN and AG traffic do not compete aggressively for bandwidth.

DRS, HA, and vMotion Behavior on vSAN

vSphere HA

When a host fails, SQL Server VMs restart on surviving hosts. Because vSAN is distributed, storage remains immediately accessible.

DRS

DRS host balancing is fully compatible with SQL Server on vSAN. However:

- AG replicas and FCI nodes require anti-affinity rules
- VM evacuation during maintenance mode should leverage vMotion Application Notification
- Ensure adequate capacity across hosts to avoid forced co-location of replicas

vMotion

vSAN accelerates vMotion through distributed metadata handling. When combined with vMotion Application Notification, SQL Server failover risks during migrations drop significantly.

Sizing Considerations for SQL Server on vSAN ESA

For production SQL workloads:

- Use NVMe TLC devices with high endurance ratings
- Ensure adequate capacity for log-structured writes
- Allocate per-database storage policies when necessary
- Avoid extreme oversubscription of shared capacity
- Maintain headroom to allow vSAN metadata operations to run unhindered

ESA minimizes the need for manual tuning, but SQL Server performance still relies on appropriate host sizing and network throughput.

Summary

vSAN ESA is the optimal storage platform for SQL Server 2025 on vSphere and VCF 9. Its modern architecture aligns naturally with SQL Server's I/O behaviors, offering consistent performance for OLTP workloads, TempDB-intensive operations, and synchronous AG replicas. Legacy vSAN OSA remains supported but is no longer recommended for new SQL Server deployments.

Whether using AGs or FCIs, SQL Server on vSAN benefits from predictable performance, simplified storage management, and seamless integration with HA, DRS, and vMotion Application Notification.

vSphere Storage Policy-Based Management (SPBM)

vSphere Storage Policy-Based Management (SPBM) is an essential component of modern VMware storage design and has become the recommended mechanism for ensuring SQL Server workloads receive the correct performance, availability, and data services across vSphere and VMware Cloud Foundation (VCF) 9 environments. SPBM provides a consistent way to define and enforce storage capabilities across VMFS, vSAN ESA, and vVols, ensuring that SQL Server data, log, TempDB, and backup volumes are provisioned with the correct characteristics for their workload profiles.

SQL Server 2025 benefits significantly from the granularity and predictability provided by SPBM. With increasingly diverse I/O patterns - ranging from sequential log writes to TempDB bursts to read-heavy analytical workloads - policy-driven storage allocation ensures each disk is provisioned according to its operational requirements.

This subsection provides modern guidance for using SPBM in SQL Server deployments across all supported VMware storage platforms.

SPBM Fundamentals in vSphere 9

SPBM creates a logical abstraction layer between virtual machines and the underlying storage. Instead of manually selecting datastores, administrators associate each VMDK (or entire VM) with a storage policy that defines:

- Availability characteristics
- Performance requirements
- Data services (compression, encryption, checksums)
- Replication and protection rules

When storage conditions change or underlying resources become noncompliant, SPBM:

- Detects noncompliance
- Provides remediation workflows
- Ensures policies are reapplied during vMotion, Storage vMotion, or failover events

This guarantees that SQL Server disks retain their intended performance behaviors regardless of cluster changes.

SPBM for vSAN ESA

vSAN ESA integrates deeply with SPBM. Unlike OSA, ESA does not rely on separate disk tiers; instead, ESA enforces capabilities directly through the SPBM policy.

Each VMDK (data, logs, TempDB) can be assigned its own policy, ensuring that high-write workloads like TempDB and log files receive enhanced availability or performance settings.

ESA's modern architecture ensures that SQL Server performance characteristics are mostly governed by storage policy rather than hardware tiering.

vSAN ESA introduced auto-policy management for SPBM which offers an optimized, cluster-specific default storage policy that will help administrators run SQL Server workloads on an ESA cluster using the optimal level of resilience and efficiency.

SPBM for vVols

Virtual Volumes (vVols) offer the most granular policy control, with each SQL Server VMDK mapped to an independent storage object on the array.

SPBM enables SQL Server disks to use:

- Array-provided performance tiers
- Hardware-based snapshots and clones
- Per-disk replication policies
- Encryption at the array level
- QoS controls (if supported by the vendor)

With vVols:

- Data, log, and TempDB VMDKs can each receive dedicated performance policies
- SQL Server backup disks can be placed on separate policy-driven storage classes
- FCIs can use vVols with SCSI-3 Persistent Reservations and fully policy-managed shared disks

SPBM for VMFS

Although VMFS does not support native SPBM features beyond tagging and datastore groups, SPBM still provides:

- Logical grouping of datastores
- Placement control based on administrator-defined tags
- Matching SQL workload classes to approved VMFS volumes

VMFS lacks the dynamic, object-based capabilities of ESA or vVols, but SPBM still offers consistency in mapping SQL disks to appropriate backend storage tiers.

Policy Recommendations for SQL Server 2025

Data Files

- ESA: RAID-5 or RAID-1 depending on latency requirements
- vVols: Array-tier associated with sustained mixed I/O
- VMFS: Tag-based mapping to high-performance NVMe-backed storage

Transaction Log Files

- Highest performance and availability class
- Prefer RAID-1 on ESA
- Disable compression in ESA if latency-sensitive (optional)
- vVols: Map to latency-optimized storage class
- Always place logs on a dedicated VMDK with a dedicated controller

TempDB

- Performance-oriented policy
- High stripe width optional on ESA
- Compression optional; often unnecessary for TempDB
- vVols: Associate with the array's fastest performance tier

Backups

- Lower-tier or capacity-optimized policy acceptable
- SPBM can direct backup VMDKs to lower-cost storage
- Ensure read bandwidth is sufficient for restore testing

Policy Compliance and Operations

SPBM continuously evaluates compliance and flags noncompliant disks when:

- A VM migrates to a host with incompatible storage
- A vSAN or vVol policy changes
- A datastore's capabilities change
- Administrators modify storage profiles

Administrators can remediate compliance with one click, ensuring that SQL Server workloads always adhere to the intended performance parameters. This is especially important for AG replicas and FCI shared disks, where policy drift may cause inconsistent performance between nodes.

Interaction with vMotion, DRS, and Storage vMotion

vMotion

- SPBM enforces storage compatibility during host migration
- vMotion Application Notification ensures SQL Server stability during movement

DRS

- DRS respects policy placement rules
- Prevents placement on incompatible hosts or datastores

Storage vMotion

- Safe for SQL Server VMs
- Automatically reapplies policies to destination storage
- Required for certain FCI shared-disk operations in maintenance events

SPBM ensures that SQL Server workloads maintain consistent storage performance regardless of cluster activity.

Summary

SPBM is a critical component of modern SQL Server deployments on vSphere and VCF 9. By applying per-disk storage policies to data, log, TempDB, and backup VMDKs, organizations gain predictable performance, simplified management, and guaranteed compliance across AGs, FCIs, and standalone SQL Server workloads. vSAN ESA and vVols offer the deepest integrations, while VMFS benefits from policy tagging and placement frameworks.

4.6 Storage Performance Monitoring

Monitoring storage performance is essential for maintaining stable and predictable SQL Server behavior under VMware vSphere and VMware Cloud Foundation (VCF) 9. SQL Server remains highly sensitive to fluctuations in disk latency, throughput, and I/O consistency, especially under mission-critical workloads involving Always On Availability Groups (AGs) and Failover Cluster Instances (FCIs). VMware's storage stack - whether vSAN ESA, vSAN OSA, vVols, or VMFS-based SAN/NVMe arrays - provides extensive telemetry that can be used to correlate virtualization-layer metrics with SQL Server's internal performance indicators.

This subsection provides updated 2025 guidance on monitoring SQL Server storage performance from both the SQL layer and the VMware layer, and explains how to interpret telemetry during normal operations, AG synchronization cycles, failover scenarios, and DRS/vMotion events.

SQL Server 2025 Storage Performance Indicators

SQL Server 2025 offers improved instrumentation for tracking I/O patterns, wait states, and replica synchronization health. The following metrics should be monitored consistently.

Key SQL Server wait types

- WRITELOG - Write latency to transaction log
- PAGEIOLATCH_* - Data file read latency
- LOGBUFFER - Log flush pressure
- HADR_SYNC_COMMIT - Synchronous AG commit latency
- HADR_DATABASE_FLOW_CONTROL - AG sustained send/receive queue pressure
- HADR_WORK_QUEUE / HADR_TRANSPORT_DRAIN - AG backlog or network constraints
- PREEMPTIVE_OS_FLUSHFILEBUFFERS - Storage flushing pressure

Sustained increases in these wait types typically indicate underlying storage or network pressure.

SQL Server I/O performance counters

Modern SQL Server 2025 counters include:

- Avg. Log Write Latency (ms)
- Avg. Data File Read/Write Latency (ms)
- Log Bytes Flushed/sec
- Write Transactions/sec
- TempDB PFS/GAM contention metrics (system views)
- AG replica log send rate and redo queue metrics

Synchronous AG replicas require:

- Log write latency consistently < 5 ms
- AG send/redo queue near zero under steady-state OLTP

Windows Performance Monitor (PerfMon) Counters

Only a subset of PerfMon counters remain relevant in 2025 due to SQL Server's improved DMVs and VMware's more advanced telemetry.

Recommended counters include:

LogicalDisk & PhysicalDisk

- Avg. Disk sec/Read
- Avg. Disk sec/Write
- Disk Reads/sec
- Disk Writes/sec
- Disk Queue Length (supplemental only)

SQLServer:Databases

- Log Flush Wait Time
- Log Flushes/sec

SQLServer:Buffer Manager

- Page life expectancy

SQLServer:HADR Replica

- Log Send Queue
- Redo Queue
- Flow Control Time

VMware vSphere Storage Monitoring

VMware provides several storage monitoring surfaces that complement SQL Server telemetry and help identify cluster-level issues.

Key vSphere metrics for SQL workloads

- Datastore Latency (Read/Write)
- Virtual Disk Latency (per-VMDK)
- Kernel Latency
- Device (Array) Latency
- Outstanding I/O Count
- Throughput (MBps)
- I/O rate (IOPS)

SQL Server–critical thresholds:

- < 5 ms ideal for log volumes
- < 10 ms ideal for data volumes
- Consistent I/O more important than transient spikes

Spikes in kernel latency often correlate with CPU scheduling or queue-depth exhaustion rather than raw storage issues.

vSAN ESA Monitoring

vSAN ESA provides modern, distributed telemetry that aligns closely with SQL Server I/O behaviors.

Critical ESA metrics

- DOM Latency (Read/Write)
- Resync I/O
- Write Buffer Utilization
- Congestion Levels
- Checksum/Compression Overheads

ESA simplifies monitoring because cache-vs-capacity architecture is removed.

Latency expectations

- Log write workloads: usually 1–3 ms
- Data file workloads: 2–5 ms for OLTP
- TempDB bursts: may momentarily increase ESA write buffer usage

Any sustained increase in ESA DOM Write Latency may correlate with:

- TempDB spills
- AG synchronous commit queues
- Host CPU saturation
- Insufficient NVMe endurance or oversubscription

vSAN OSA Monitoring Notes

For customers still running OSA:

- Monitor cache tier usage, write buffer fullness, and disk group health
- Elevated read or write latency often indicates cache exhaustion
- Rebuild/resync events have higher impact on I/O compared to ESA

vVols Storage Monitoring

When SQL Server uses vVols:

- Monitor per-disk latency directly from the storage array
- Integrate with SPBM compliance monitoring
- Monitor array-level QoS (if available)
- Validate snapshot/cloning offload performance

vVols offer granular visibility, especially for AG replicas with independent performance tiers.

Correlating SQL Server and VMware Telemetry

A critical 2025 practice is correlating SQL wait types with VMware latency metrics:

SQL Wait / Symptom	Likely VMware Indicator	Meaning
WRITELOG spikes	Virtual Disk Write Latency ↑	Log volume stress
PAGEIOLATCH_*	Read Latency ↑	Data volume read bottleneck
HADR_SYNC_COMMIT	Log Write or Network Latency ↑	Sync AG commit delay
FLOW_CONTROL	vSAN DOM Congestion ↑	Replica throttling
Backup slowness	High array throughput usage	Backup target contention

Table 3 - Sample SQL Server Symptoms Correlation

NOTE: Cross-platform correlation is essential for root-cause analysis. Windows Performance Monitor (and the applicable SQL Server's Dynamic Management View query results) should be used primarily when correlating SQL internals with virtualization-layer telemetry.

Monitoring DRS, HA, and vMotion Events

DRS

During automatic placement or balancing:

- Monitor latency spikes on log and TempDB volumes
- Ensure anti-affinity rules prevent co-location of synchronous replicas

vSphere HA

After HA-triggered VM restarts:

- Check AG/FCI failover behavior
- Monitor log send/redo queues after node restarts
- Validate storage access and policy compliance

vMotion

During vMotion:

- Expect a short pause (stun event)
- AG synchronous commit may show brief HADR_SYNC_COMMIT spikes
- vSAN ESA minimizes stun duration
- vMotion Application Notification reduces failover risk

AG and FCI-Specific Monitoring Notes

Availability Groups

- Monitor log send/redo queues, latency, and replica synchronization
- Synchronous commit replicas should show near-zero queues

- Log write latency must remain low

Failover Cluster Instances

- Monitor shared disk latency
- Watch for rare SCSI-3 PR delays
- Validate access path consistency during failovers

Summary

Effective storage performance monitoring for SQL Server 2025 on vSphere and VCF 9 requires correlating SQL Server wait types, Windows counters, and VMware storage telemetry. With modern capabilities - especially on vSAN ESA and vVols - administrators can quickly identify storage, network, or cluster issues affecting SQL performance. Monitoring AG and FCI health alongside virtualization metrics ensures predictable performance and minimizes synchronization issues.

Storage design remains one of the most critical components of deploying SQL Server 2025 on VMware vSphere and VMware Cloud Foundation (VCF) 9. Although virtualization abstracts the underlying hardware, SQL Server continues to require predictable, low-latency storage performance - particularly for write-heavy transaction log workloads, TempDB operations, and synchronous Availability Group (AG) replicas.

This section has provided updated guidance for the full spectrum of storage considerations, beginning with datastore selection (VMFS, vVols, and especially vSAN ESA), and progressing through optimal controller configurations, disk layout and file placement, vSAN-specific design principles, policy-based management using SPBM, and modern storage monitoring methodologies. Taken together, these recommendations form a comprehensive storage strategy aligned with current VMware and Microsoft engineering guidance.

vSAN ESA has become the preferred storage architecture for SQL Server due to its NVMe-optimized log-structured design, predictable latency characteristics, and deep integration with SPBM. Meanwhile, vVols (which is still supported on non-VCF Platforms) continues to offer unmatched granularity for enterprises requiring per-disk policy control or array-based data services. VMFS remains a valid option when backed by modern NVMe storage systems and properly isolated for SQL Server workloads.

SQL Server 2025 also benefits from advances in virtualization tooling such as vMotion Application Notification, enhanced DRS placement logic, and improved vSphere HA behavior. These features reduce the risk of cluster instability during maintenance operations and VM migrations while maintaining predictable performance for both AG and FCI configurations.

Finally, modern telemetry from SQL Server (DMVs), Windows' PerfMon, vSphere, and vSAN ESA allows administrators to correlate performance signals more accurately than ever before. This enables faster detection of storage bottlenecks, AG synchronization issues, or virtualization-layer contention, resulting in more stable high-availability deployments.

A well-designed storage architecture - built on modern VMware platforms, managed by SPBM, aligned with SQL Server I/O characteristics, and validated through continuous monitoring - is essential for achieving a resilient, high-performance SQL Server environment in 2025 and beyond.

Acknowledgments

Authors:

Dèjì Akọ́mọláfẹ́ – Staff Solutions Architect, Microsoft Applications Practice Lead
Mark Xu – Product Marketing Engineer, VCF
Christian Rauber – Product Marketing Engineer, VCF

Item No: vmw-bc-wp-tech-a4-word-2025 Feb-26