

High Availability vSphere for SQL Server: 5 Things You Need to Know

Like

Tweet

Share

g+1

2



A Contributed Article by Jerry Melnick, COO, SIOS Technology Corp.

SQL Server administrators have many options for implementing high availability (HA) in a VMware environment. VMware offers vSphere HA. Microsoft offers Windows Server Failover Clustering (WSFC). And SQL Server in WSFC has its own HA options with AlwaysOn Availability Groups and AlwaysOn Failover Clusters.

Third party vendors also provide solutions purpose-built for HA and disaster recovery, and these often integrate with other solutions to create even more options. For example, some solutions leverage the AlwaysOn Failover Cluster feature included with SQL Server to deliver robust HA and data protection for less than the cost of AlwaysOn Availability Groups that require the more expensive Enterprise Edition.

This article highlights five things every SQL Server administrator should know before formulating a high availability strategy for mission-critical applications in a vSphere environment. Such a strategy is likely to resemble the multi-site configuration shown in Figure 1, which is not possible with some HA options.

1. High-Availability Clusters for vSphere require Raw Disk Mapping

The layers of abstraction used in virtualized servers afford substantial flexibility, but such abstractions can cause problems when a virtual machine (VM) must interface with a physical device. This is the case for vSphere with Storage Area Networks (SANs).

To enable compatibility with certain SAN and other shared-storage features, such as I/O fencing and SCSI reservations, vSphere utilizes a technology called Raw Device Mapping (RDM) to create a direct link through the hypervisor between the VM and the external storage system. The requirement for using RDM with shared storage exists for any cluster, including a SQL Server Failover Cluster.

In a traditional cluster created with WSFC in vSphere, RDM must be used to provide virtual machines (VMs) direct access to the underlying storage (SAN). RDM is able to maintain 100 percent compatibility with all SAN commands, making virtualized storage access seamless to the operating system and applications which is an essential requirement of WSFC.

RDM can be made to work effectively, but achieving the desired result is not always easy, and may not even be possible. For example, RDM does not support disk partitions, so it is necessary to use "raw" or whole LUNs (logical unit numbers), and mapping is not available for direct-attached block storage and certain RAID devices.

2. Use of Raw Disk Mapping means Sacrificing Popular VMware Features

Another important aspect of being fully informed about RDM involves understanding the hurdles it can create for using other VMware features, many of which are popular with SQL Server administrators. When these hurdles are deemed unacceptable, as they often are, they eliminate Raw Device Mapping as an option for implementing high availability.

The underlying problem is how RDM interferes with VMware features that employ virtual machine disk (VMDK) files. For example, RDM prevents the use of VMware snapshots, and this in turn prevents the use of any feature that requires snapshots, such as Virtual Consolidated Backups (VCBs).

Raw Disk Mapping also complicates data mobility, which creates impediments to using the features that make server virtualization so beneficial, including converting VMs into templates to simplify deployment, and using vMotion to migrate VMs dynamically among hosts.

Another potential problem for transaction-intensive applications like SQL Server is the inability to utilize Flash Read Cache when RDM is configured.

3. Shared Storage can create a Single Point of Failure

The traditional need for clustered servers to have direct access to shared storage can create limitations for high availability and disaster recovery provisions, and these limitations can, in turn, create a barrier to migrating business-critical applications to vSphere.

In a traditional failover cluster, two or more physical servers (cluster nodes) are connected to a shared storage system. The application runs on one server, and in the event of a failure, clustering software, such as Windows Server Failover Clustering, moves the application to a standby node. Similar clustering is also possible with virtualized servers in a vSphere environment, but this requires a technology like Raw Disk Mapping so that the VMs can access the shared storage directly.

Whether the servers are physical or virtual, the use of shared storage can create a single point of failure. A SAN can have a high availability configuration, of course, but that increases its complexity and cost, and can adversely affect performance, especially for transaction-intensive applications like SQL Server.

4. HA vSphere Clusters can be built without Sacrificing VMware Functionality

Some third-party solutions are purpose-built to overcome the limitations associated with shared storage and the requirement to use RDM with SQL Server's AlwaysOn Failover Clusters and Windows Server Failover Clusters.

The best of these solutions provide complete configuration flexibility, making it possible to create a SANLess cluster to meet a wide range of needs - from a two-node cluster in a single site, to a multinode cluster, to a cluster with nodes in different geographic locations for disaster protection as shown in Figure 1. Some of these solutions also make it possible to implement LAN/WAN-optimized, real-time block-level replication in either a synchronous or asynchronous manner. In effect, these solutions are capable of creating a RAID 1 mirror across the network, automatically changing the direction of the data replication (source and target) as needed after failover and failback.

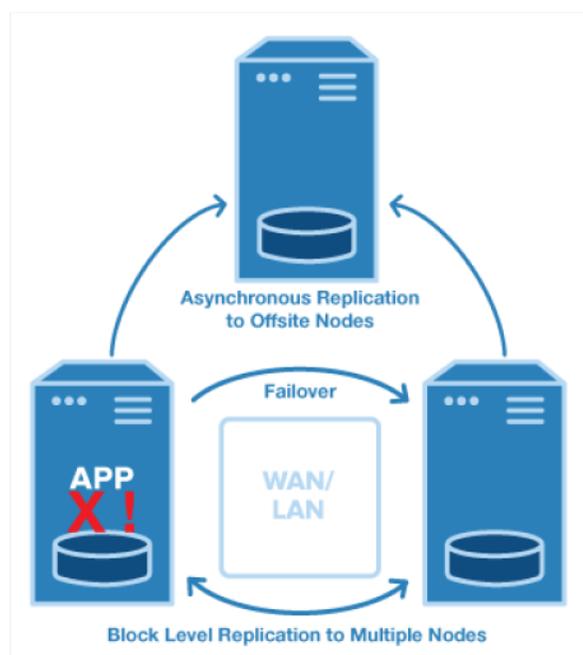


Figure 1 - A multi-site high-availability configuration protects applications from outages that affect an entire data center.

Just as importantly, a SANLess cluster is often easier to implement and operate with both physical and virtual servers. For example, for solutions that are integrated with WSFC, administrators are able to configure high-availability clusters using a familiar feature in a way that avoids the use of shared storage as a potential single point of failure. Once configured, most solutions then automatically synchronize the local storage in two or more servers (in one or more data centers), making them appear to WSFC as if it was a shared storage device.

5. HA SANLess Clusters deliver Superior Capabilities and Performance

In addition to creating a single point of failure, replicating data on a SAN can significantly reduce throughput performance in VMware environments. Highly transactional applications like SQL Server are particularly vulnerable to these performance-related factors.

Figure 2 summarizes test results that show the 60-70 percent performance penalty associated with using SQL Server AlwaysOn Availability Groups to replicate data. These test results also show how a purpose-built high-availability SANLess cluster, which utilizes local storage, is able to perform nearly as well as configurations not protected with any data replication or mirroring.

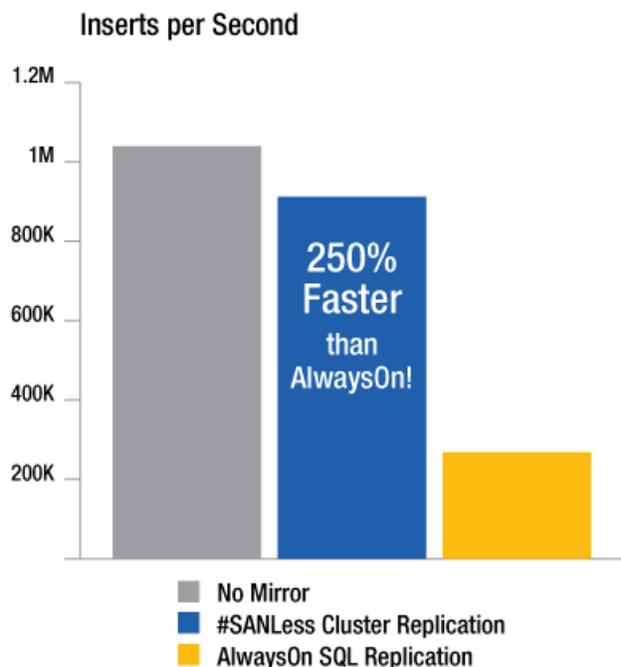


Figure 2 - Testing of SQL Server's AlwaysOn Availability Groups and SIOS #SANLess clusters shows the throughput advantage possible with replication techniques purpose built for high availability and high performance.

The #SANLess cluster tested is able to achieve this impressive performance because its driver sits immediately below NTFS. As writes occur on the primary server, the driver writes one copy of the block to the local VMDK and another copy simultaneously across the network to the secondary server which has its own independent VMDK.

SANLess clusters have many other advantages, as well. For example, those that use block-level replication technology that is fully integrated with WSFC are able to protect the entire SQL Server instance, including the database, logons and agent jobs—all in an integrated fashion. Contrast this approach with AlwaysOn Availability Groups, which failover only user-defined databases, and require IT staff to manage other data objects for every cluster node separately and manually.

##

About the Author

Jerry Melnick, COO, SIOS Technology Corp.

Jerry Melnick (jmelnick@us.sios.com) is responsible for defining corporate strategy and operations at SIOS Technology Corp. (www.us.sios.com), maker of SIOS SAN and #SANLess cluster software (www.clustersyourway.com). He more than 25 years of experience in the enterprise and high availability software industries. He holds a Bachelor of Science degree from Beloit College with graduate work in Computer Engineering and Computer Science at Boston University.

Published Wednesday, February 11, 2015 6:46 AM by [David Marshall](#)

Filed under: [VMBlog Info](#), [Contributed](#)