Technical White Paper: April 2023

VMware Cloud on AWS Storage Resilience

VMware vSAN and VMware Cloud Flex Storage

vmware®

Table of contents

Introduction

Running business-critical workloads in the cloud has become an increasingly popular option for organizations due to its convenience, scalability, and cost-effectiveness. However, with the rise of cyberattacks and natural disasters, the resilience of cloud storage has become a crucial factor to consider. Cloud providers have implemented various strategies to ensure data availability, durability, and integrity. In this context, the resilience of cloud storage refers to the ability of the system to withstand and recover from disruptions while maintaining the continuity of operations and data access. This document explores the key factors contributing to the resilience of VMware storage offerings in VMware Cloud on AWS™.

VMware Cloud on AWS provides a seamlessly integrated hybrid cloud platform for running enterprise workloads of today and tomorrow. With VMware Cloud on AWS, you can start your application and infrastructure modernization journey with minimal disruption to your business. You can rapidly extend and migrate their applications to the cloud in an AWS Region of their choice, live and without the need to change the applications. VMware and AWS jointly engineer this cloud service and combines the benefits of VMware's enterprise-class Software-Defined Data Center (SDDC) solution and the AWS global cloud infrastructure.

VMware vSAN™ and VMware Cloud Flex Storage™ are two highly resilient VMware Cloud on AWS storage solutions that provide organizations with the flexibility and scalability needed to meet their growing storage demands. These solutions work seamlessly with VMware's cloud platform, providing a complete software-defined data center solution that enables organizations to easily manage and optimize their cloud storage infrastructure. vSAN and VMware Cloud Flex Storage deliver cost-effective and highly available storage solutions that can scale to meet the needs of nearly any organization.

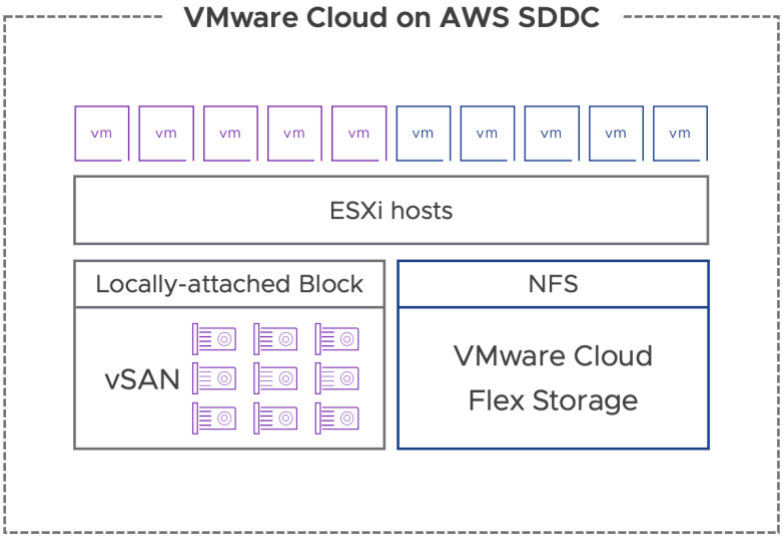

Figure 1. VMware storage solutions in VMware Cloud on AWS.

VMware vSAN

VMware vSAN is a software-defined storage solution that delivers high performance, availability, and scalability for hyper-converged infrastructure (HCI). It is tightly integrated with the VMware ESXi hypervisor and consolidates computing, storage, and networking resources into a single platform. VMware vSAN is a foundational component that enables the seamless extension of your existing on-premises vSAN infrastructure to the cloud. With vSAN in VMware Cloud on AWS, you can experience the same high-performance, resilient storage solution as your on-premises deployment but with the added benefits of on-demand cloud flexibility and scale.

vSAN provides the core storage platform for VMware Cloud on AWS, utilizing storage-optimized Amazon EC2 bare-metal instances. Powered by locally attached Non-Volatile Memory Express (NVMe) flash storage, vSAN aggregates storage from each EC2 instance connected to the cluster into a distributed vSphere datastore. The vSAN datastore consists of two logical entities: vsanDatastore and WorkloadDatastore. VMware developed this specifically for this service to restrict permissions between the vsanDatastore, which hosts the management components, and WorkloadDatastore, which stores your virtual machine disks, configuration files, and swap files.

Figure 2. vSAN datastores.

Storage Policies

vSAN's storage policy-based management allows administrators to set policies for individual virtual machines or groups of virtual machines based on their resilience and capacity requirements. The policies define the availability level and storage space consumption for each VM. vSAN automatically ensures the enforcement of storage policies. Storage policy-based management eliminates manual storage provisioning and enables the dynamic allocation of virtual machine resources as needed. Storage policies are easily modified to adapt to changing workload requirements, and vSAN provides real-time monitoring and alerting if policy violations occur. This approach to storage management is highly flexible and efficient and enables organizations to achieve high levels of performance and availability while optimizing their storage resources.

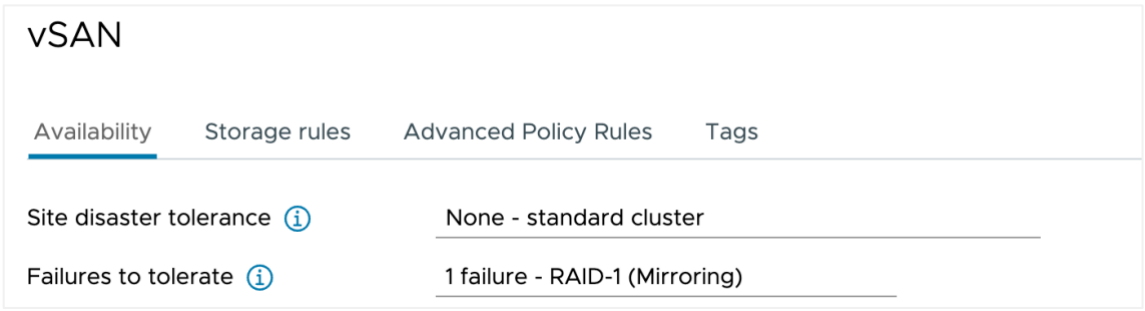

Figure 3. vSAN storage policy.

One of the key features of vSAN is its ability to protect data using different levels of redundancy, such as RAID-1 mirroring, RAID-5 erasure coding, and RAID-6 erasure coding.

RAID-1 mirroring creates an exact copy of data across two or more disks in the vSAN cluster. This method provides resilience against up to three drive or host failures. Mirroring requires more storage capacity than erasure coding, as it creates complete copies of the data.

RAID-5 erasure coding uses mathematical algorithms to create parity data distributed across multiple disks in the vSAN cluster. This method provides better storage capacity efficiency than mirroring. However, it has computational overhead and might have a slight performance impact on workloads. RAID-5 erasure coding tolerates the loss of one drive or host.

RAID-6 erasure coding is like RAID-5 erasure coding but provides additional protection against data loss. It uses two sets of parity data instead of one, which allows the vSAN cluster to recover data even if two drives or hosts fail simultaneously. However, it requires more storage resources than RAID-5 erasure coding and has a higher computational overhead.

When deciding which data protection method to use in vSAN, it's essential to consider the requirements of the application and the resources available in the vSAN cluster. RAID-1 mirroring provides the highest levels of protection and performance but requires more storage capacity. RAID-5 erasure coding offers better storage efficiency but might introduce a slight performance impact on specific latency-sensitive workloads. RAID-6 erasure coding provides higher resilience but increases storage consumption and computational overhead. VMware recommends starting with an erasure coding policy to minimize storage capacity consumption while providing the necessary level of redundancy. It is easy and non-disruptive to switch to a mirroring policy if needed.

Integration with vSphere High Availability

VMware vSphere High Availability (HA) monitors ESXi hosts for any signs of failure. In the event of a host failure, the affected virtual machines are restarted on other healthy hosts in the cluster. The downtime is equivalent to the time required to restart the virtual machines on other hosts. It is important to remember that this includes the time it takes for the operating system to boot up and applications to start on the virtual machines. vSphere HA is a solution that operates at the virtual machine level and can be utilized even if applications lack inherent high availability capabilities. vSphere HA does not rely on the guest operating system or software installed in the virtual machine.

A datastore shared across all hosts in the cluster is required for vSphere HA to work correctly. All hosts have access to all copies of the virtual machine data on the vSAN datastore, which enables vSphere HA to recover workloads when a host goes offline. For example, a virtual machine with a RAID-1 mirroring policy assigned has a full copy of a virtual machine's data on two hosts. If one of those hosts fails, vSphere HA uses the mirrored data on the other host to restart the virtual machine.

Stretched Clusters

AWS Availability Zones (AZs) are fault domains designed to be independently resilient. A Stretched Cluster SDDC is a highly available deployment across two AZs where the hosts in the cluster are divided across the AZs, along with a witness host in a third AZ. The SDDC provides two data sites and one witness host per cluster. Stretched clusters are supported in AWS regions that provide at least three AZs. Stretched clusters offer an additional layer of resiliency to protect workloads against host failures and AZ failures within the region. Still, it's essential to remember that they do not protect against all failure scenarios.

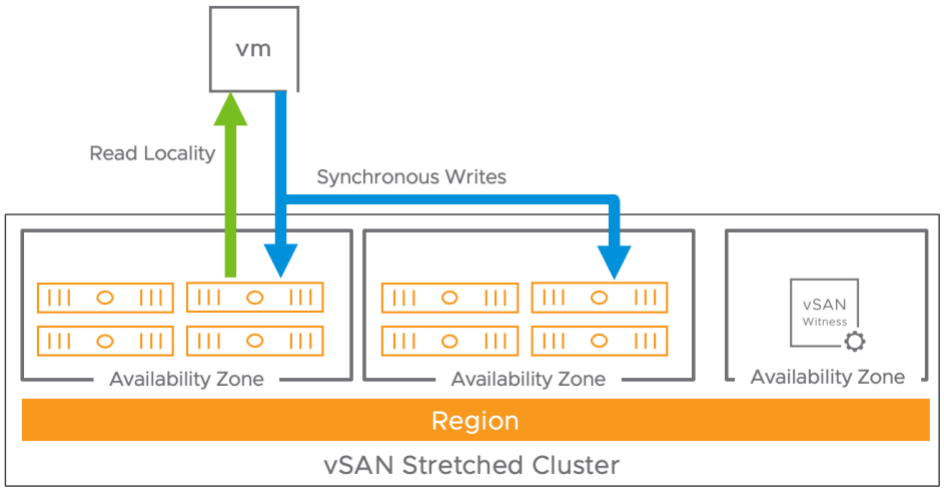

Figure 4. vSAN Stretched Cluster across Availability Zones.

Understanding workloads' read/write behavior that utilizes a dual-site mirroring vSAN policy is also essential. Data is read from the local copy, written locally, and writes are synchronously replicated to the hosts across AZs. This method helps ensure no data loss if a host or an entire AZ is offline. Adding and removing hosts in a stretched cluster must be done in pairs, and the maximum host per cluster for a stretched cluster is identical to that of a standard SDDC. Stretched clusters are deployed with as few as two hosts and covered by a 99.9% SLA.

VMware Cloud Flex Storage

VMware Cloud Flex Storage is built on a mature, enterprise-class filesystem developed and production-hardened over many years, dating back to Datrium's DHCI storage product, which VMware acquired in 2020. The same filesystem has supported the VMware Cloud Disaster Recovery and Ransomware Recovery services. With VMware Cloud Flex Storage, this proven technology extends to primary storage by making it available in the public cloud. It delivers storage performance, scale, and cost efficiency for traditional and modern virtual machine workloads. VMware Cloud Flex Storage has built-in enterprise-class storage features, including at-rest encryption, deduplication, compression, and data integrity checks. You pay only for the storage you consume under a simple pricing model based on the storage used.

Use cases include but are not limited to virtual machines running file services and media storage, which require a fixed amount of CPU usage and increasing amounts of storage. New analytics initiatives may require significant amounts of storage, and changing regulations may also require unplanned storage retention. VMware Cloud Flex Storage datastores presented to multiple clusters within an SDDC are particularly useful for specific workflows, such as virtualized desktops.

VMware Cloud Flex Storage offers independent scaling of storage capacity without purchasing additional hosts. It provides on-demand and elastic scaling of storage, enabling you to adjust storage capacity up or down as needed. The storage service is natively integrated with VMware Cloud on AWS, and you can easily attach data stores with a few clicks from the VMware Cloud Console. The storage is presented to the cluster using NFS.

Scale-out Cloud File System

The Scale-out Cloud File System (SCFS) is a multi-purpose filesystem. It has a 2-tier design, using EC2 instances with local NVMe for caching and S3 for capacity storage. The SCFS employs the Log-Structured Filesystem (LFS) technique to store data. The method involves appending incoming data into a sequential log and doing garbage collection later. All incoming data are converted to large sequential segments and stored as S3 objects, which allows data storage at high-speed. Initially, all incoming data is acknowledged from within the cache layer of the NVMe device. Large NVMe devices are utilized to accelerate read performance. LFS techniques are employed for data in S3, whereby incoming data is converted to large sequential segments of approximately 10MB. These segments are then stored as S3 objects, which enables high-speed storage due to S3's exceptional handling of large sequential IOs. Amazon S3 is designed to provide 99.999999999% durability and 99.99% availability of objects over a given year.

Additionally, S3 is resilient to the loss of an AZ. As new incoming data is always considered a log, it is always saved in new locations, thereby eliminating the risk of overwriting blocks containing old data. The combination of LFS and the 2-tier designs make SCFS an all-purpose filesystem.

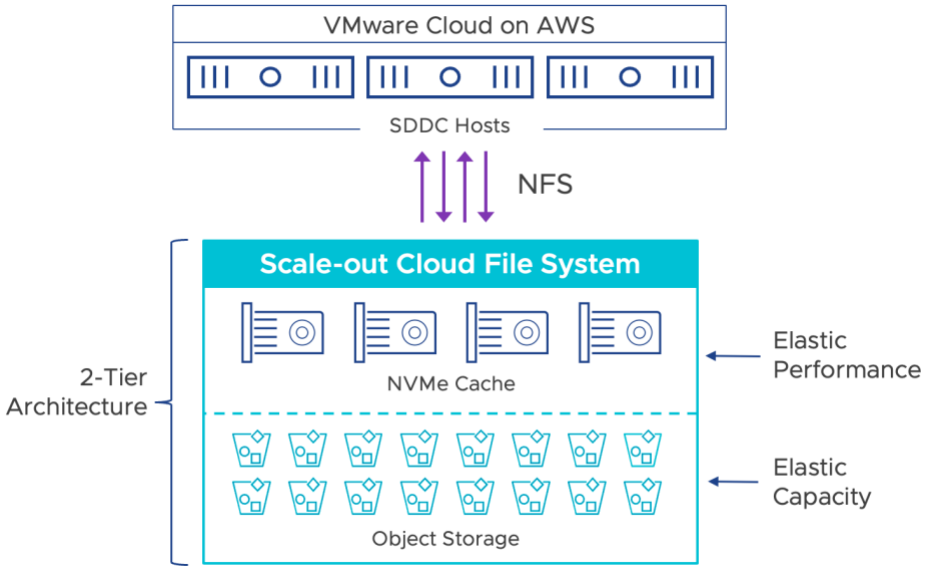

Figure 5. VMware Cloud Flex Storage architecture.

Failure Handling

When discussing VMware Cloud Flex Storage, the most frequently asked question is how it protects against the failure of a storage device or server. In the event of an SCFS server failure, a replacement server is quickly provisioned to take its place. If any components within the system fail, this is automatically detected and resolved without any impact on the virtual machines. Failover time is commonly less than 30 seconds.

vSphere HA works with VMware Cloud Flex Storage the same as vSAN. VMware Cloud Flex Storage provides a shared datastore that is accessible by all hosts in the cluster. Suppose a host running virtual machines connected to the VMware Cloud Flex Storage datastore fails. In that case, VMware vSphere HA is activated, and the impacted virtual machines are automatically restarted on other hosts within the cluster.

Mirrored Write Cache

Every NFS write is duplicated onto two different servers' local NVMe devices, resulting in a distributed mirrored write cache. Only when data has been consistently written to these two locations will writes be acknowledged to the guest OS in virtual machines. This approach guarantees data durability while enabling the NFS front-end to hold data in memory for later transfer to the back-end object storage. The vSphere NFS client always employs the "sync" option to send data and confirms that it has been fully acknowledged before notifying the guest that it has been recorded.

Partition Placement Groups

AWS Partition Placement Groups are a strategic instance placement approach that reduces the probability of co-related host failures caused by hardware issues. vSAN and VMware Cloud Flex Storage utilize this solution. The availability of applications is increased by deploying hosts in different logical partitions that do not share the same underlying hardware. The algorithm used in Partition Placement Groups is "best effort," which automatically deploys hosts across as many different partitions as available within an Availability Zone. Each partition has its set of racks with a distinct network and power source. In addition, no two partitions share the same racks, effectively isolating host failures within an SDDC cluster. This approach significantly enhances the availability of applications while minimizing the impact of hardware failures. More information on AWS partition placement groups: [Placement Groups].

Conclusion

VMware vSAN and VMware Cloud Flex Storage provide highly resilient and flexible options to manage growing storage demands in VMware Cloud on AWS. These cloud storage solutions work seamlessly with VMware's cloud platform, delivering highly available storage that can scale. While considering the application's requirements, you can achieve high levels of performance and availability while minimizing costs. VMware's partnership with AWS offers seamless integration of a hybrid cloud platform, allowing you to modernize applications and infrastructure with minimal disruption to your business.

Item No: vmw-wp-tech-temp-uslet-word-2021 8/21