Technical White Paper: **January 2023**

Cloud Flex Storage and VMware & Cloud Disaster Recovery

Availability and Resilience

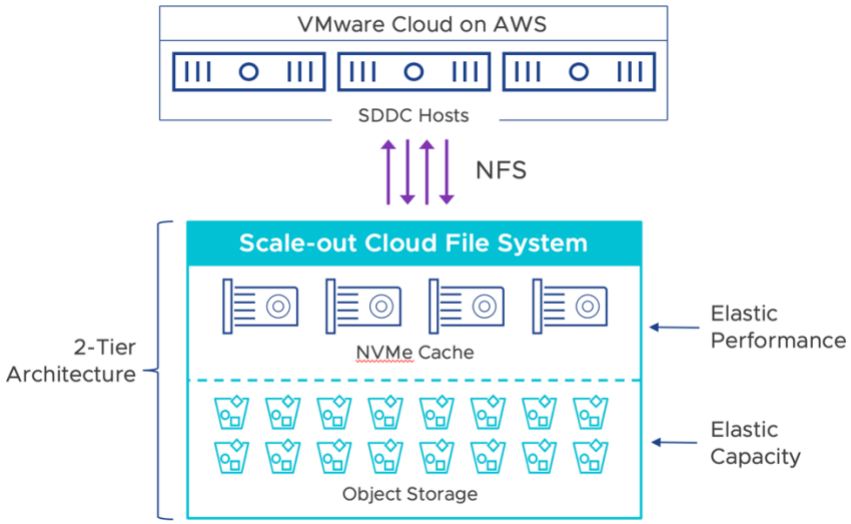
vmware®

Table of contents

Scale-out Cloud Filesystem (SCFS)

Cloud Flex storage uses the same multi-purpose Scale-out Cloud Filesystem (SCFS) used by VMware Cloud Disaster Recovery with some slight modifications to better align it for continuous usage.

NVMe and Object Storage Design



All incoming data is first acknowledged from within the NVMe device's cache layer. Large NVMe devices are deployed for the file systems to accelerate read performance.

We employ the LFS techniques to store data in S3. All incoming data is converted to large ~10MB sequential segments, and these large segments are stored as S3 objects, and S3 is excellent at large sequential IOs. This allows data to be stored in S3 at high-speed. The S3 buckets are Designed to provide 99.999999999% durability and 99.99% availability of objects over a given year. S3 buckets are durable to an entire region and can survive the total persistent loss of an AZ. All new incoming data "always" goes to new locations (because it is a log). By design, there is never any danger of overwriting blocks containing old data. The combination of LFS and the 2-tier designs are what makes SCFS a multi-purpose filesystem.

Failure Handling

The most common question asked when discussing Cloud Flex Storage is how it protects against a storage device or server failure. The failure of any server used by Scale-out Cloud Filesystem (SCFS) will result in a replacement being rapidly provisioned to take its place. Failure of components within the solution is automatic, and transparent to the virtual machine (Sub 30 seconds failover time), as the environment continuously monitors for failure. Failure of a host running virtual machines connected to the NFS Datastore, will trigger VMware vSphere High Availability (HA) and automatically failover the virtual machines to a different host within the SDDC cluster. While write cache is mirrored, the read cache is not.

Mirrored Write Cache

A common question for Cloud Flex Storage is what happens to the data being collected before it is sent to the object storage buckets. To reduce latency and maintain durability and resiliency this data is protected using write buffers. All writes committed to the NFS front end server are mirrored to two different servers local NVMe devices, forming a distributed mirrored write cache. Writes are not acknowledged to vSphere or the

guest OS in virtual machines until data has been persistently written to these two locations. This guarantees data durability, while allowing for the NFS front end to keep data in memory for perpetration to send to the back-end object storage. The vSphere NFS client used always sends data using the "sync" option and confirms a full acknowledgment before notifying the guest has been acknowledged as written. At no point is data acknowledged as written that is solely stored in a volatile RAM cache.

Cloud Flex Storage

Partition Placement Groups

VMware Cloud on AWS leverages AWS Partition Placement Groups (PPG). With this feature, EC2 instances are spread across logical partitions that do not share underlying hardware, including the physical rack, to minimize the impact of host and rack failures.

VMware Cloud on AWS automatically allocates Partition Placement Groups within an SDDC cluster. Instances in a cluster are placed on a best-effort basis in separate logical partitions that do not share underlying hardware. Placement happens automatically for every new SDDC, cluster or host add operation. Customers simply benefit from the increased availability of Partition Placement Groups; there is no configuration or effort required.

VMware Cloud Flex Storage leverages Partition Placement Groups to make sure that both copies of data currently in the write cache and not yet de-staged to S3 are protected from drive, system, rack or switch failure.

Partition Placement Groups is an AWS feature that provides customers with added protection against single points of failure. With the collaboration between VMware and AWS, the VMware Cloud on AWS service brings this capability automatically and transparently to customers.

See AWS documentation for more information on *Partition Placement Groups*.

VMware Cloud Disaster Recovery

VMware Cloud Disaster Recovery offers easy to use, on demand disaster protection and recovery, delivered as a SaaS service, with cloud economics. It combines cost-efficient cloud storage with simple SaaS-based management for IT resiliency at scale, and customers benefit from a 'pay when you need' failover capacity model for DR resources.

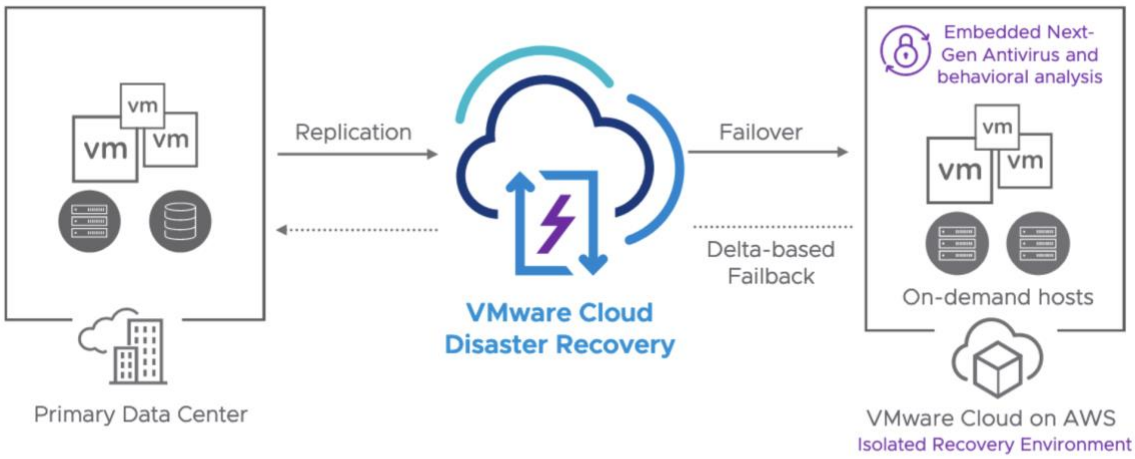


Figure 1: VMware Ransomware Recovery.

Leveraging the elasticity of cloud computing, VMware Cloud Disaster Recovery spins up VMware Cloud on AWS infrastructure only during a DR testing or failover event. It utilizes a highly efficient cloud storage layer for storing backups, lowering DR costs. It delivers fast recovery with zero copy and no rehydration of data from cloud storage to VMware Cloud on AWS hosts where the recovered VMs can be immediately powered-on. For longer term use, and achieving production-level performance, workloads can be migrated (using Storage vMotion) to vSAN storage in the SDDC. Using optional pilot light clusters makes the recovery time even faster. VMs are maintained in their native VMware vSphere format, eliminating the need for brittle and time-consuming VM disk format conversions. Instant power-on of VMs is very powerful for rapid identification of the best recovery point when recovering from a ransomware attack. VMware Cloud Disaster Recovery can protect a very broad set of IT services in a cost-efficient manner, with fast recovery capabilities (On-demand DRaaS). Learn more about VMware Cloud Disaster Recovery *here*. Importantly it can perform virtual machine, folder level, and file level recovery.

Today, VMware Cloud Disaster recovery has several unique features that enhance further the durability of the backups it provides beyond the capabilities of Cloud Flex Storage and when using it to protect vSAN and Cloud Flex Storage based virtual machines.

Immutability

A filesystem also needs metadata (pointers) to remember where all the data blocks are stored. We use content-based crypto-hashes as pointers for data blocks. Content-based crypto-hashes are immutable. Each backup is represented by a tree of crypto-hashes, and the root of the tree is also a crypto-hash (see Merkle-Trees usage in blockchain). All of this makes each backup immutable, and these backups are hidden and not accessible directly to the outside world. Even to recover, a backup is never directly used. We clone the needed backup into a new object for recovery purposes, which leaves the original backup copy untouched. The clone operation is instantaneous no matter how many backups are being cloned or how big they are.

Additionally, SCFS checks the data integrity of each backup copy every single day. The goal is that we want to ensure that your backup copies are ready for use when you need them in a ransomware-attack emergency.

Immutable backups

There are a few things SCFS does to protect backups from damage:

- Each backup is represented by a set of content-based crypto-hash trees (Merkle-Trees), and this makes each backup immutable.
- All backups are hidden from the normal access modes (like NFS, etc.). Even viewing backups indirectly via UI is done with multi-factor authentication.
- A backup's data cannot be changed. A backup must first be cloned if there is a need to use that data or modify it.
- VMware Cloud DR also supports multi-factor authentication, role-based access controls, and different administrator domains to protect against ransomware gaining DR administrator privileges.

Storage of last resort

SCFS uses LFS to eliminate the danger of new backups accidentally overwriting an old backup copy (e.g., due to software bug). Additionally, all data is verified every day to ensure that your backup copies are ready for use when you need them in a ransomware-attack emergency.

VMware Ransomware Recovery

VMware Ransomware Recovery for VMware Cloud DR is a purpose-built ransomware recovery as-a-service solution that delivers safe, controlled recovery using an on-demand Isolated Recovery Environment (IRE) in the cloud. For more information about how this server can help provide you with an immutable, durable recovery environment see the solution brief.

Conclusion

VMware Cloud Flex Storage and VMware Cloud Disaster Recovery delivers a seamlessly integrated Storage and disaster recovery capabilities. It provides a highly resilient and available platform with capabilities that enable customers to focus on their applications rather than the infrastructure. In summary, key highlights Include:

- Unique Two-Tier log structured file system.
- Combines NVMe and object storage for the best benefits of cost, performance and durability of data.
- Partition Placement Groups to minimize the impact of rack level failures
- VMware Cloud Disaster Recovery offers Immutable backups
- VMware Ransomware Recovery offers a turnkey Isolated Recovery Environment (IRE) to rapidly respond to advanced ransomware attacks.

Combined, these capabilities make VMware Cloud on AWS the easiest way for customers to migrate to the cloud, with built-in availability and resiliency. For more information, please contact your VMware or AWS representative or visit our website at *https:// cloud.vmware.com/vmc-aws/*.

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