VERITAS

Veritas InfoScale in a Distributed Cloud

Create availability and resiliency for critical applications.

White Paper | February 2022

Contents

Executive Summary

Cloud services have transformed the IT landscape and several cloud platforms have become mainstream alternatives to hosting IT infrastructure in traditional data centers. With numerous public cloud options available, many business technology deployment architectures have evolved to include multiple cloud services as part of an overall cloud strategy. This approach has resulted in a distributed cloud paradigm that introduces challenges around making IT services highly available and resilient.

Veritas provides enterprise data services focused on application availability, protection and insights. As a market leader in application availability and storage management, Veritas has a long-standing history of providing innovative solutions for managing applications on any platform, including public clouds.

This white paper will discuss the concept of distributed cloud, how it's typically architected and implemented and the Veritas InfoScale™ solution that provides a software-defined approach for managing high availability and storage for business applications in distributed cloud environments.

What is a Distributed Cloud?

Organizations are looking to achieve the following goals[1]:

> *By the end of 2021, based on lessons learned, 80 percent of enterprises will put a mechanism in place to shift to cloud-centric digital infrastructure twice as fast as before the COVID-19 pandemic.*
>
> *By 2024, 80 percent of enterprises will overhaul relationships with infrastructure providers to better execute their digital strategy for ubiquitous deployment of resources and for more autonomous IT operations.*

Distributed cloud is a lexicon for connecting services and infrastructure between discrete public cloud platforms such as Amazon Web Services (AWS), Microsoft Azure and Google Cloud Platform (GCP). An organization interested in moving business-critical applications into the cloud to increase efficiency and reduce operational expenses may consider extending its reach into more than one public cloud.

Costs for public cloud platforms will evolve as new features and improved infrastructure and systems are brought to market. Organizations will benefit from these improvements by distributing their applications and workloads across multiple cloud platforms based on their specific requirements. InfoScale provides organizations with the ability to distribute or move their applications and workloads across public cloud platforms.

What Problems can you Solve with a Distributed Cloud?

It's fairly common to think of a public cloud as a way to solve a myriad IT infrastructure problems at a reduced cost with less operational complexity. In practice, however, using the cloud to reduce costs and complexity always depends on several factors and considerations that vary for every business. InfoScale helps businesses optimize cloud services by providing the tools to move IT services into the cloud, out of the cloud and into another cloud, as desired.

When designing your IT infrastructure to operate in a distributed cloud topology, it is essential to consider the following:

- High availability and disaster recovery (HADR)
- Platform failure
- Innovation

High Availability and Disaster Recovery

Digital transformation, or the integration of technology into all traditional business areas, is in constant flux for any organization. The financial and operational benefits are generally realized as greater efficiency, increased revenue or broader market proliferation, to name a few.

However, a cloud-based infrastructure outage can weaken a business (as it could on-premises), both financially and in reputation, if appropriate HADR safeguards aren't implemented. Public cloud infrastructure is often built for redundancy and may also offer DR services. Although these services are designed to provide infrastructure-level redundancy, they aren't integrated with business applications and services. This design leaves an operational gap around effectively managing application and service-level recovery point (RPO) and recovery time objectives (RTOs).

A distributed cloud model that includes two cloud services enables an organization to distribute the risk of disruption between the two different clouds, reducing the risk of application downtime in the case of a (single) cloud provider service disruption.

Platform Failure

Cloud infrastructure is complex because there is a trade-off with usability. The public cloud is graphically driven and generally easy to use. Unfortunately, public clouds use a complex layer of abstraction to simplify infrastructure operations.

A complex infrastructure can lead to unforeseen problems, which lead to downtime, despite a public cloud's best efforts.

A distributed cloud can help mitigate public cloud failure by isolation. Outages that simultaneously affect multiple public clouds are significantly less likely to occur than individual cloud service outages.

Hosting infrastructure across two or more cloud vendors can mitigate this problem, ensuring applications remain available and operational should one cloud fail.

Innovation

Different cloud providers are innovating to offer customers better, faster and cheaper technology to run their applications. Not all cloud platform providers offer the same level of innovation, however, because they may have different priorities driving feature and functionality development. This disparity can result in a functionality gap that may not satisfy an organization's requirements.

For example, one public cloud provider may be focused on providing user or customer front-end services, and another cloud provider may be focused on enterprise infrastructure.

Tools such as automation, observability and security also vary between public cloud providers. Each provider has similar implementations of the required tools, but the differences can be significant, depending on an organization's requirements. It may be advantageous for an organization to use multiple public clouds to take advantage of features best suited to its particular requirements. A distributed cloud architecture can help an organization take advantage of these differences by integrating features across clouds.

How can InfoScale Help you Build a Distributed Cloud?

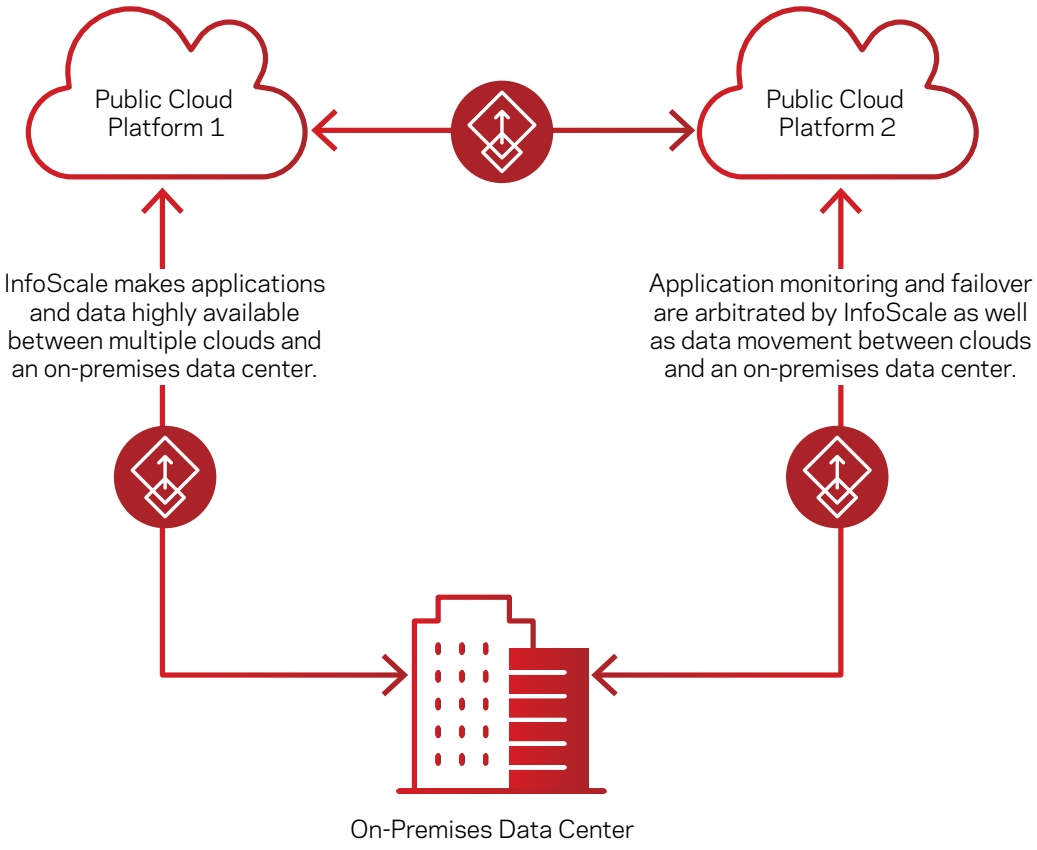

Figure 1. An example of a distributed cloud architecture with Veritas InfoScale.

Public clouds provide redundancy and disaster recovery at the infrastructure level. Figure 1 shows how this architecture can be expanded across clouds. Doing so not only mitigates cloud failures but also provides flexibility by preventing organizations from being locked into a single cloud platform. InfoScale achieves this goal by providing the following capabilities for hybrid cloud and distributed cloud architectures:

- Application-level HADR
- Integration with software-defined storage with scale-out shared storage capability
- Native replication
- HADR automation
- HADR rehearsals ("fire drill") to ensure your HADR plan is compliant

You can deploy InfoScale on compute instances in the cloud, so you can use these features to ensure your business-critical applications are running with minimal interruption on a single or multiple cloud platforms at the same time.

Application-Level High Availability and Disaster Recovery

InfoScale provides enterprise software-defined storage management and HADR for any data center, including on-premises, hybrid and multi-cloud. InfoScale is a comprehensive, industry-proven solution that helps users manage enterprise readiness for modern, mission-critical applications.

InfoScale integrates directly with application components to ensure that both the underlying infrastructure and the application itself are managed to provide the highest possible performance and uptime. InfoScale also provides instantaneous automated recovery that minimizes the impact of service disruptions and outages.

InfoScale can respond to application failures almost instantly when they occur and can take action to maximize the application uptime or, if required, failover to another site. This design also significantly reduces the compute resources required for monitoring and can help prevent data corruption by reducing the time to act in the event of a failure.

Integrated Software-Defined Storage with Scale-Out Shared Storage Capability

InfoScale storage consists of six major components that provide the basis for building a highly available clustered storage infrastructure:

1. **Veritas File System (VxFS)**—A POSIX-compliant enterprise file system designed to maximize application performance.
2. **Veritas Volume Manager (VxVM)**—The storage management subsystem that allows you to create, organize and manage logical data volumes and disk groups used by applications and databases from physical disks and logical unit numbers (LUNs).
3. **Cluster Volume Manager (CVM)**—The storage virtualization layer that enables storage devices to be accessible by multiple nodes in the cluster simultaneously.
4. **Cluster File System (CFS)**—Creates a file system that multiple nodes in the cluster can access concurrently.
5. **Flexible Storage Sharing (FSS)**—Enables individual nodes to share direct-attached storage with other cluster nodes at the physical disk level. FSS is a feature of CVM and CFS that allows a file system to be built on top of a volume shared with FSS.
6. **Veritas Volume Replicator (VVR)**—Enables optimized replication of data between InfoScale-managed data volumes.

These components form the basis for a software-defined, clustered storage infrastructure that enables an organization to scale out storage as requirements change. It also provides fast application failover capabilities to minimize business interruptions. Figure 2 is a high-level block diagram that shows how these components work with each other.

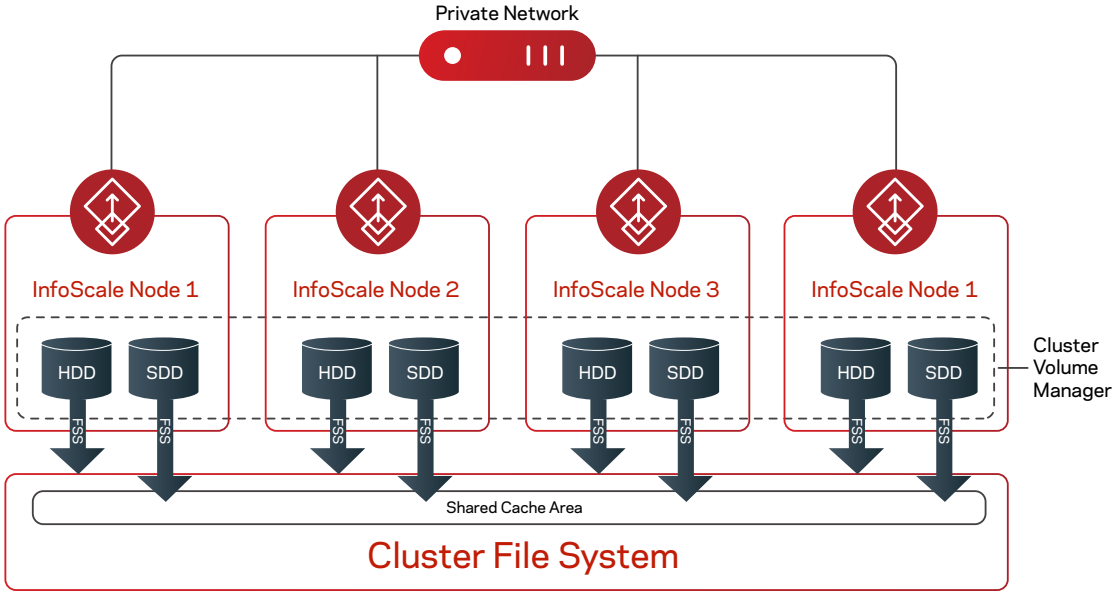

Figure 2. An overview of InfoScale storage management.

With InfoScale availability providing HADR capability for cloud environments, InfoScale storage can significantly improve the underlying cloud-native storage service performance and efficiency with features like FSS and SmartIO. InfoScale storage provides enterprise functionality for cloud environments beyond what's available with native cloud tools, offering some key benefits:

- **Performance**—Although public cloud infrastructure offers higher-performance storage options, there are limitations at the system level that minimize overall performance (IOPS). With InfoScale SmartIO intelligent caching, application reads can be served from faster volumes using SSD storage while writes can be served from a cheaper storage tier. This approach significantly improves application performance with minimal additional cost.

- **Scalability**—With FSS, you can create the resilient shared storage volumes needed to horizontally scale enterprise applications using public cloud infrastructure. InfoScale also enables granular resource scaling. When an application needs additional compute or storage resources, it can be scaled dynamically and independently, reducing operating costs and providing infrastructure flexibility for your applications.

Native Replication

InfoScale provides the flexibility to integrate with either third-party storage replication technology or Veritas Volume Replicator (VVR). Both options can provide a near-zero RTO/RPO for mission-critical data and scale to support the largest workloads. VVR also provides some additional benefits: maintaining write-order fidelity, multi-target support, heterogeneous system configuration (including public cloud) and zero data loss.

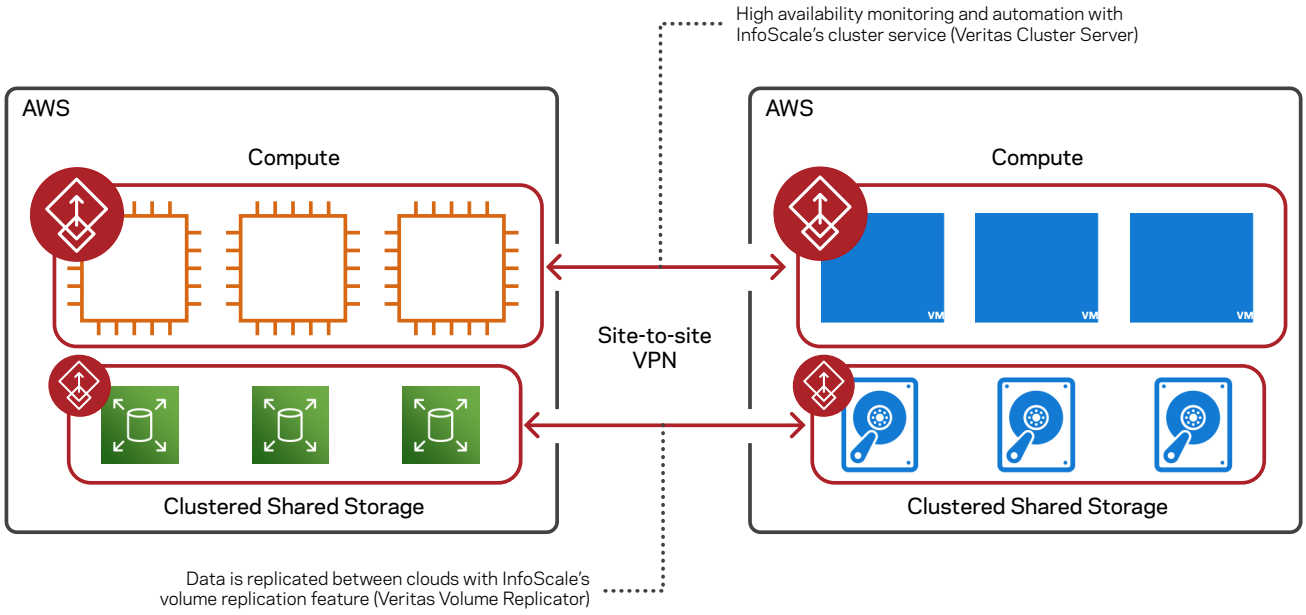

Figure 3. The process of ensuring cloud-to-cloud high availability with InfoScale.

VVR works in concert with InfoScale's Veritas Cluster Server (VCS) to ensure the application state is preserved when an application is failed over between clouds. When a failover event occurs, VCS initiates the application failover by unmounting the data volumes (if possible), switching over the application services to the target cluster and restarting the application with the replicated data. Figure 3 is a high-level diagram of this operation.

Automation for High Availability and Disaster Recovery

InfoScale can further facilitate an automated, enterprise HADR solution with an integrated feature that groups application tiers to represent an entire, multitiered business service. A Virtual Business Service (VBS) represents a multitier application as a single, consolidated entity that augments the HADR provided for the individual application tiers. Using VBS, you can completely automate the recovery or migration of a complex multitier application, making it easy to provide HADR for an entire business service.

High Availability and Disaster Recovery Rehearsals (Fire Drills)

InfoScale can manage and run a simulated test on an isolated, non-production network segment to ensure systems at the secondary site are working correctly before a failover event. To do so, InfoScale uses snapshots of production data attached to temporarily provisioned systems used for testing purposes. InfoScale also manages the clean-up of the fire drill environment when it's no longer needed, removing the temporary resources.

The Value of InfoScale

A distributed cloud architecture provides a wide range of options and features to optimize your IT infrastructure and services. InfoScale abstracts cloud platform architectures to enable organizations to run applications using the cloud services that best meet their requirements. InfoScale eliminates cloud platform lock-in by enabling organizations to move their data and applications to different clouds or repatriate them back to on-premises data centers, if required. InfoScale also supports hybrid environments and ensures data is synchronized in either scenario: between clouds or in cloud-to-on-premises topologies.

Although point-based solutions that provide similar functionality exist, none integrate HADR management and automation with software-defined storage. These solutions may also depend on cloud providers to offer monolithic storage arrays or vendor-specific storage software outside an organization's requirements. InfoScale has no such dependencies because it is software-based and platform-agnostic (see Table 1).

Storage Array–Based HADR	Application-Based HADR
Isilon SyncIQ	Clustered MSSQL
NetApp SnapMirror/SnapVault	Oracle RAC
Pure Asynchronous Replication	Enterprise Applications (SAP, Informatica, etc.)

Table 1. Point-Based Solutions for HADR

What else can InfoScale do?

InfoScale includes a wide variety of features to enhance HADR and application performance and optimize resource utilization. All these features and more are included with an Enterprise License. Most important, these features are software-based and are available on any platform or operating system supported by InfoScale.

SmartIO

The SmartIO feature of InfoScale storage enables data efficiency on your higher-performance storage such as solid-state drives (SSDs) through intelligent I/O caching. Using SmartIO to improve efficiency, you can optimize the storage cost per IOPS. SmartIO does not require in-depth knowledge of the underlying hardware technologies. SmartIO uses advanced heuristics to determine what data to cache and how it is removed from the cache. The heuristics take advantage of InfoScale's visibility into the characteristics of the workload.

SmartIO uses a cache area on the target device or devices. The cache area is the storage space SmartIO uses to store the cached data along with the associated metadata. SmartIO supports different types of read and write caching. The type of cache area used determines whether it supports file system caching or volume caching. To start using SmartIO, you can create a cache area with a single command while the application is online. A tool called SmartAssist is available to help estimate the optimal SmartIO cache size based on analysis of system components and data.

When the application issues an I/O request, SmartIO checks to see if the I/O can be serviced from the cache. As applications access data from the underlying volumes or file systems, certain data is moved to the cache based on internal heuristics. Subsequent I/Os are processed from the cache.

SmartTier

InfoScale helps organizations manage storage costs and growing data footprints by dynamically and automatically moving data from primary storage devices to lower-cost storage with SmartTier. Although many storage solutions include some tiering capability, they lack a heterogeneous enterprise-wide view, resulting in IT having to manage multiple different tiering solutions individually. SmartTier enables the tiering of data to almost any storage solution, minimizing the complexity of managing multiple point solutions.

SmartTier allows custom and flexible configuration of file placement through policies. The policies can be applied to volumes managed in the Veritas File System, then moving files to low-cost secondary storage if policy criteria are met, such as I/O activity.

In practice, organizations typically generate a mix of frequently and less frequently accessed data. To optimize performance while minimizing the expense of moving entire volumes to costlier SSD storage, SmartTier places the frequently accessed files on the faster SSD storage and leaves the infrequently accessed files on spinning disk.

Snapshots

InfoScale can provide volume snapshots that enable the fast and simple creation of backup copies of volumes while the application is online with minimal interruption to users. The backup copies can then be used to restore data that has been lost due to disk failure, software errors or human errors or to create replica volumes for report generation, application development or testing.

With Snapshots, you can manage system backups, upgrades, maintenance and development using point-in-time copies without affecting production data. You can process snapshot copies on the same host as the active data or on a different host. If required, you can offload the processing of snapshots to another host to avoid contention for system resources on your production servers. This method is called off-host processing. When implemented correctly, off-host processing solutions have almost no impact on the performance of the primary production system.

Kubernetes Support

InfoScale has been designed to integrate natively with Kubernetes with a containerized architecture that can be easily deployed within Kubernetes environments. InfoScale provides the storage management functionality needed by stateful applications running in containers that is not available natively in a Kubernetes environment. InfoScale delivers the following functionality for Kubernetes to enable a container management platform suitable for running stateful applications that require:

- **Advanced storage management**—InfoScale's CSI plug-in allows Kubernetes to provide InfoScale persistent storage volumes to containerized applications being managed within a Kubernetes namespace. InfoScale's FSS can provide high-performance storage using the disks directly attached to Kubernetes cluster nodes.
- **Application and system availability**—InfoScale's advanced IO Fencing capabilities enable Kubernetes to bring failed nodes and/or application pods back online quickly and automatically in the event of a failure within the cluster. InfoScale prevents the corruption of data and ensures applications remain online until the failure is resolved.

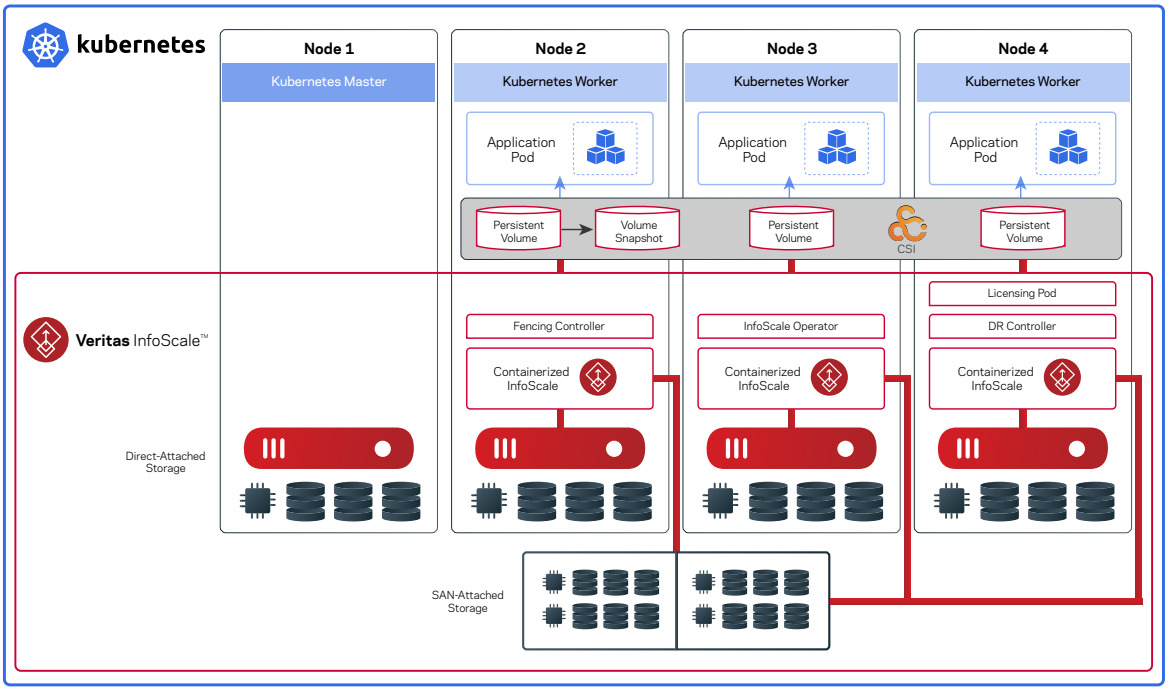

Figure 4. How Kubernetes and InfoScale integration provides high availability and persistent storage for containerized applications.

InfoScale is deployed as containers in the Kubernetes cluster and the InfoScale CSI plug-in provides the interface between Kubernetes and InfoScale. Figure 4 provides an overview of how InfoScale integrates with Kubernetes and containers to provide high availability and persistent storage for containerized applications. In addition to standard Kubernetes environments, InfoScale is also available as a storage provider for the Red Hat OpenShift Container Platform.

Summary

Designing and implementing a robust distributed cloud strategy can be complex. There are many different variables to consider when architecting, deploying and managing distributed cloud environments while also providing high availability and resiliency for your IT applications. InfoScale reduces the complexity and challenges associated with making your applications highly available in distributed cloud environments. InfoScale helps you manage a distributed cloud strategy by providing:

- **Flexibility**—Seamlessly move applications and data between cloud providers as needed.
- **Optimization**—Reduce operating costs by running applications and data on the most cost-effective cloud services while maintaining high availability and resiliency.
- **Innovation**—Avoid being limited to specific cloud provider features; run your IT services in the cloud that's best tailored to provide your ideal user experience.
- **Availability**—Protect your IT services against cloud provider outages and disruptions by using multiple cloud services as part of a cohesive HADR strategy.

With distributed cloud becoming an increasingly common way to host IT infrastructure, InfoScale is the ideal software-defined solution to help you reduce costs and complexity for distributed cloud operations while providing maximum availability, resiliency and performance for your IT services.

Further Reading

Veritas InfoScale Enterprise: Managing Mission-Critical Applications in a Software-Defined Data Center

(Technical Overview): https://www.veritas.com/content/dam/Veritas/docs/white-papers/V0995-veritas-infoscale-managing-mission-critical-applications.pdf

Infrastructure Autonomy Video: https://bcove.video/30VcILT

InfoScale Technical Library: https://www.veritas.com/availability/infoscale/resources

Availability and Resiliency for the Modern Enterprise Solution Overview: https://bit.ly/3d4cq6X

1. Future of Digital Infrastructure: Ever Faster Delivery of Reliable Digital Services and Experiences, Sep 2020 - IDC Perspective - Doc # US46807920, https://www.idc.com/getdoc.jsp?containerId=US46807920

VERITAS

2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
veritas.com

For global contact information visit:
veritas.com/company/contact

V1337 01/22