VERITAS

Veritas InfoScale

Managing mission-critical applications in a software-defined data center.

Technical Overview | May 2023

Contents

Introduction

Veritas Technologies is an industry leader in developing data resiliency and availability solutions that focus on the protection and management of digital assets critical for a company's success and business continuity. One of our flagship products, Veritas InfoScale™, enables enterprise-class, software-defined storage management along with high availability and disaster recovery (HADR) for all data centers, including on-premises, hybrid, and multi-cloud. InfoScale works with applications running on physical, virtual, and cloud platforms, with support extended in version 8.0 to include applications running in a Kubernetes infrastructure (or environment). InfoScale is a comprehensive, industry-proven solution that helps organizations manage enterprise readiness for modern, mission-critical applications, focused on three key principles:

1. Application availability—Integrate directly with native application components to ensure the underlying infrastructure and the application itself (which includes protecting and providing access to data necessary for that application) are managed to provide the highest possible performance and uptime. InfoScale can also provide near-instantaneous automated recovery that minimizes the impact of service disruptions and outages.
2. Application performance—Eliminate overhead and complexity by using InfoScale to build a high-performance, software-defined storage environment using commodity hardware that intelligently manages data so it's always available on the fastest storage tier. This approach improves application performance and maximizes resource utilization while significantly reducing infrastructure overhead, costs, and complexity.
3. Application agility—Enable applications to be managed for the highest performance and uptime across different operating systems and platforms that may span multiple geographic regions. InfoScale helps avoid vendor lock-in by enabling applications to run in a highly available configuration on any operating system (OS) and any platform, including public cloud and hybrid-cloud configurations. InfoScale also provides automated cloud migration from on-premises systems and between different cloud providers.

This technical overview will explain the Veritas InfoScale Enterprise solution, which includes InfoScale Availability and InfoScale Storage. InfoScale Enterprise enables organizations to combine the HADR requirements of IT applications with highly performant and scalable software-defined storage to achieve maximum application uptime and performance using any OS or platform.

InfoScale Solution Value

InfoScale has several unique features that offer significant value to organizations looking to improve application performance and reduce costs by maximizing architectural flexibility without being locked into a specific technology. Creating an environment that's capable of supporting your most mission-critical applications presents some challenges that cannot be completely resolved with native tools. InfoScale is designed to integrate with nearly any IT infrastructure to provide resiliency and high performance for the world's most demanding 'always-on' applications (see Figure 1).

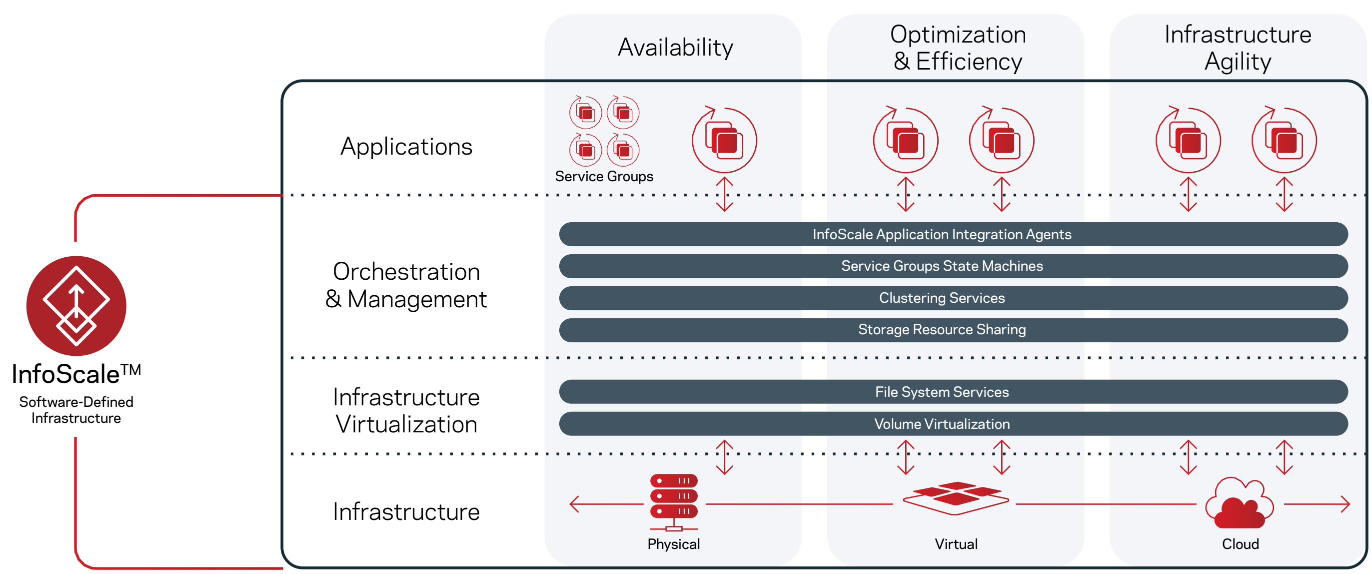


Figure 1. An overview of InfoScale's solution value.

InfoScale has customized agents designed for mission-critical applications that understand and manage the application components and resources not normally monitored by native system tools. This functionality ensures the application has the highest possible uptime and either meets or exceeds performance expectations. InfoScale can also intelligently manage the HADR process and nearly eliminate the need for manual user intervention for applications to be successfully brought online or recovered in a DR scenario. There are several other benefits of using InfoScale to manage both application storage and the overall HADR solution:

- Near-instant fault detection—By introducing processes that run at the kernel level, InfoScale Availability can respond to application failures almost instantly when they occur and can take action to maximize application uptime or failover to another site, if required. This design also significantly reduces compute resources required for monitoring and can help prevent data corruption by reducing the time to action in the event of a failure.
- Data availability and management—Dynamic Multi-Pathing (DMP) along with industry-leading snapshot and checkpoint features ensure applications can access vital data. Additionally, I/O fencing and I/O shipping technologies protect that data from hardware and software failures in the underlying infrastructure.
- Maximized performance—Regardless of the underlying infrastructure, InfoScale can intelligently cache application data so it's always available on the highest performance storage tier. Going one step further, InfoScale can manage the intelligent data caching process on highly available storage clusters that can be built using direct-attached storage on commodity hardware while still leveraging storage area network (SAN) assets. This software-defined approach provides enterprise storage features and better application performance at a significantly reduced cost, with no vendor lock-in.
- Replication management—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 recovery time objective/recovery point objective (RTO/RPO) for mission-critical data and can scale to support the largest workloads. With VVR, there are some additional benefits such as maintaining write order fidelity, multi-target support, heterogeneous system configuration (including public cloud), and zero data loss.
- Cloud integration—InfoScale can manage HADR functionality for both hybrid-cloud and native cloud environments. InfoScale Availability can also manage data transfer between on-premises systems and public cloud environments, making it easy to move applications between platforms. Organizations can also move applications between public cloud platforms such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform to support a multi-cloud HADR strategy.
- Kubernetes integration—With the 8.0 release, InfoScale now provides cloud-native storage and data services in a Kubernetes environment. By deploying containerized InfoScale alongside a Container Storage Interface (CSI) plugin, organizations can leverage InfoScale's industry-leading data services with containerized applications.
- HADR Firedrill—InfoScale Availability can use the Firedrill feature to manage and run a simulated test on an isolated, non-production network segment to ensure systems at the secondary site are working properly prior to a full failover event. It does so by using snapshots of production data that are then attached to temporarily provisioned systems used for testing purposes. InfoScale also manages the clean- up of the Firedrill environment when it's no longer needed. This feature validates that DR is not only set up correctly, but also is operating correctly in your environment.
- Management simplicity—Veritas InfoScale Operations Manager (VIOM) is a single-pane-of-glass management console that lets you manage any InfoScale component from one place. In addition to management of InfoScale solutions, VIOM also provides visibility into non-InfoScale infrastructure, so you can use it to identify potential issues within an environment that could lead to unexpected downtime.
- In-depth security—InfoScale can protect and keep systems and data in a known-good state with a comprehensive and multi-platform security stance. Use an external key management service to keep data encrypted at rest whether on Windows or Linux, restrict nodes to known-good code using SecureBoot kernel and module signing, commit data to immutable storage or schedule read-only snapshots to protect it from ransomware, whether in the context of the cloud or on-premises.

InfoScale can further facilitate an enterprise HADR solution with an integrated feature where application tiers can be grouped together in a way that represents the entire business service the application provides. This group is known as a Virtual Business Service (VBS). A VBS represents a multi-tier application as a single, consolidated entity that augments the HADR provided for the individual application tiers. Using a VBS, you can completely automate the recovery or migration of a complex, multi-tier application, making it easy to provide HADR for an entire business service.

InfoScale Architecture

InfoScale is a fully software-defined solution that has an extensive compatibility matrix and is OS and platform agnostic. You can deploy it on industry-standard hardware, and it can support a wide range of enterprise applications, regardless of their underlying infrastructure. InfoScale is composed of several products, mostly delineated by licensing scheme. At its core, InfoScale has two distinct parts: InfoScale Availability and InfoScale Storage. (See Figure 2.)

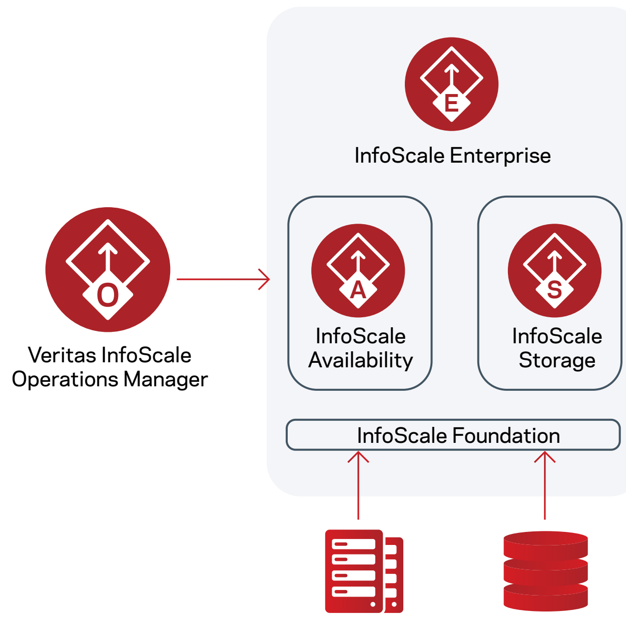


Figure 2. The components of the InfoScale product suite.

InfoScale Availability consists of the components shown in Table 1. It encompasses all the services and features designed to ensure applications are highly available. There are several application-specific agents available with InfoScale Availability that ensure optimal application performance and resiliency.

Component	Description
Service Groups	A collection of hardware and software resources and their dependencies that are required for an application to run.
High-Availability Daemon	A process that collects information about resource states from the application agents on the local system and forwards it to all cluster members.
Application Agents	A software package for systems managed by InfoScale that enables the management of hardware and software resources.
Intelligent Monitoring Framework (IMF)	An event-based notification system integrated with certain application agents that provides near-instantaneous notification of resource state changes.
Veritas InfoScale Operations Manager (VIOM)	A web-based centralized management console for InfoScale Enterprise that provides monitoring, visualization, and management of resources.

Table 1. InfoScale Availability Component

InfoScale Storage is the software-defined storage group of services and features. It enables you to build an enterprise-class storage solution to manage and protect any type of enterprise data and can be scaled up or down to accommodate any environment or budget. InfoScale Storage consists of the components shown in Table 2.

Component	Description
Veritas File System (VxFS)	A POSIX-compliant extent-based file system that runs on most Unix and Linux variant operating systems.
Veritas Volume Manager (VxVM)	A storage management utility that manages physical disks as logical data volumes that are presented to an operating system as a storage device on which you can create file systems.
Veritas Volume Replicator (VVR)	A software-based data replication utility that enables consistent block-level data replication between a VxVM-managed source data volume and one or more remote data volumes.
Cluster File System (CFS)	A VxFS file system running in parallel access mode that permits multiple systems (cluster nodes) to access the same file system data simultaneously.
Veritas File Replicator	A periodic file-level replication utility that tracks and replicates changed files in a VxFS file system transparently to any concurrent application usage, leveraging file system checkpoints to read and replicate changed files between a source checkpoint and one or more target checkpoints.

Table 2. InfoScale Storage Components

InfoScale Availability

InfoScale Availability consists of five main components that work together to create a highly available clustered system infrastructure:

1. Service Groups
2. High Availability Daemon (had[1])
3. Application Agents
4. Monitoring – Intelligent Monitoring Framework (IMF)
5. Veritas InfoScale Operations Manager (VIOM)

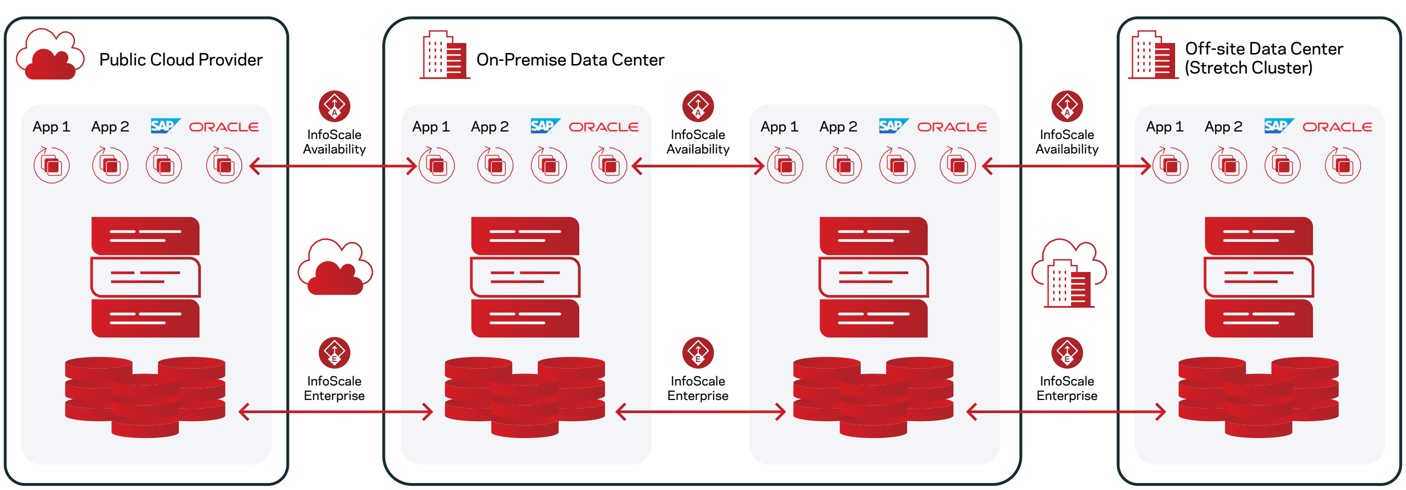


Figure 3. An example of different HADR configuration scenarios.

Figure 3 shows the different HADR configurations that are possible with InfoScale Availability. These range from HA within a local cluster or even a single-node, single-site cluster to a multi-site clustered configuration that spans geographic regions. Replication is managed by InfoScale Availability using either VVR or direct integration with supported storage subsystems that provide native replication functionality.

Service Groups

A service group is a virtual container of resources managed by InfoScale that are required for an application to run. The service group defines the individual resources as well as the dependencies between them and the order in which they need to be managed for an application to be online. Service groups allow InfoScale to manage all the hardware and software resources of a managed application as a single unit. Resources include components such as Network Interface Cards (NICs), IP addresses, disk groups, volumes, and mount points as well as application-specific processes. InfoScale manages both the resources and their dependencies to ensure the application is online and can immediately act in the event of a resource failure.

Service groups have multiple attributes that define how it will manage resources for an application. Attributes each have a definition and a value. The definition describes the scope of the attribute and the value contains the input.

For example, the SystemList attribute is a user-defined attribute that contains the list of systems on which the service group is configured to run as well as their priorities.

There are three types of service groups:

1. Failover service groups—Service groups that run on one system in the cluster at a time. Failover service groups are used for most applications that do not support multiple systems simultaneously accessing the application's data.
2. Parallel service groups—Service groups that run simultaneously on more than one system in the cluster. A parallel service group is more complex than a failover group and is appropriate for applications that manage multiple application instances running simultaneously without data corruption.
3. Hybrid service groups—Service groups used for replicated data clusters that combine both failover and parallel service groups. The hybrid service group behaves like a failover service group within a system zone and a parallel service group across system zones.

Service groups can be dependent on each other. For example, a managed application might be a web application that is dependent on a database application. Because the managed application consists of all the components required to provide the service, service group dependencies create more complex managed applications. When you use service group dependencies, the managed application is the entire dependency tree.

High Availability Daemon

The Veritas Cluster Server (VCS) High Availability Daemon (*had*) is the cluster "engine" and is also known as the Veritas High Availability Engine. An instance of had runs on each cluster node and dynamically maintains a replicated state machine that provides all cluster nodes with a consistent representation of the cluster's state and membership.

The *had* manages the following functions:

- Building the cluster configuration using the configuration files
- Managing information distribution when new nodes are added to the cluster
- Responding to user input and propagating configuration changes
- Responding to resource failure or unavailability

The *had* uses agents to monitor and manage resources. It collects information about resource states from the agents on the local system and forwards it to all cluster members. The engine that runs on each node has a completely synchronized view of the resource status on each node.

Application Agents

In general terms, an agent is a software package that runs on the systems being managed as part of an InfoScale HADR solution. Agents are multi-threaded processes that provide the logic to manage resources. Agents manage the system components (or resources) required for an application to be online. Data from the agents is also fed back to had for analysis and actionability and to VIOM for reporting.

InfoScale agents provide HA for specific resources and applications. Each agent manages resources of a certain type. For example, the SAP agent manages SAP components such as NetWeaver and HANA databases. Typically, agents start, stop, and monitor resources and report state changes.

InfoScale Availability has agents for most tier 1 applications such as SAP, Oracle, Tibco, and Microsoft applications. In situations where an application agent is not available, a generic resource group option is available that enables functionality like that provided by the InfoScale Availability agents.

There are three types of agents available with InfoScale Availability:

1. Application agents for managing enterprise applications such as SAP, Tibco, and WebSphere as well as custom applications.
2. Database agents for managing database applications such as Oracle, MySQL, SAP HANA, and Microsoft SQL Server.
3. Replication agents that monitor and manage hardware and software data replication technologies such as EMC SRDF, NetApp SnapMirror, and Hitachi TrueCopy.

InfoScale Availability also supports third-party custom agents for use in situations where a specific agent is not available. These third-party custom agents are developed and supported by the third parties.

Intelligent Monitoring Framework

The IMF is an extension of the application agent. It provides near-instant fault detection and is essentially a kernel-level process that is installed as part of the InfoScale Availability agent software.

IMF is an integrated part of the InfoScale Availability agent framework. IMF allows the agents to register the resources to be monitored with a notification module that enables immediate (event-based) notification of resource state changes without having to periodically poll the resources to find the resource current state. This process enables InfoScale Availability agents to act immediately in the event of a system fault.

The IMF is a key component of InfoScale Availability that helps eliminate downtime for enterprise applications with near-zero RPO/RTO requirements. Not all InfoScale Availability agents are IMF-aware. You can find a list of agents that support IMF here.

InfoScale Operations Manager

VIOM is a platform- and vendor-agnostic centralized management console for Veritas InfoScale Availability, InfoScale Storage, and other third-party infrastructure. VIOM is used for monitoring, visualization, and management of system and storage resources (see Figure 4). VIOM is also a reporting engine and can generate multiple reports, including a risk analysis report that can summarize issues that may arise within an environment that could reduce HADR readiness.

A typical VIOM deployment consists of two main components: a management server and managed hosts. The management server hosts the web-based user interface (default URL: https://<hostname>:14161/vom). Depending on the usage scenario, VIOM may also discover virtualization environments and SAN/NAS infrastructure as well as SAN fabrics.

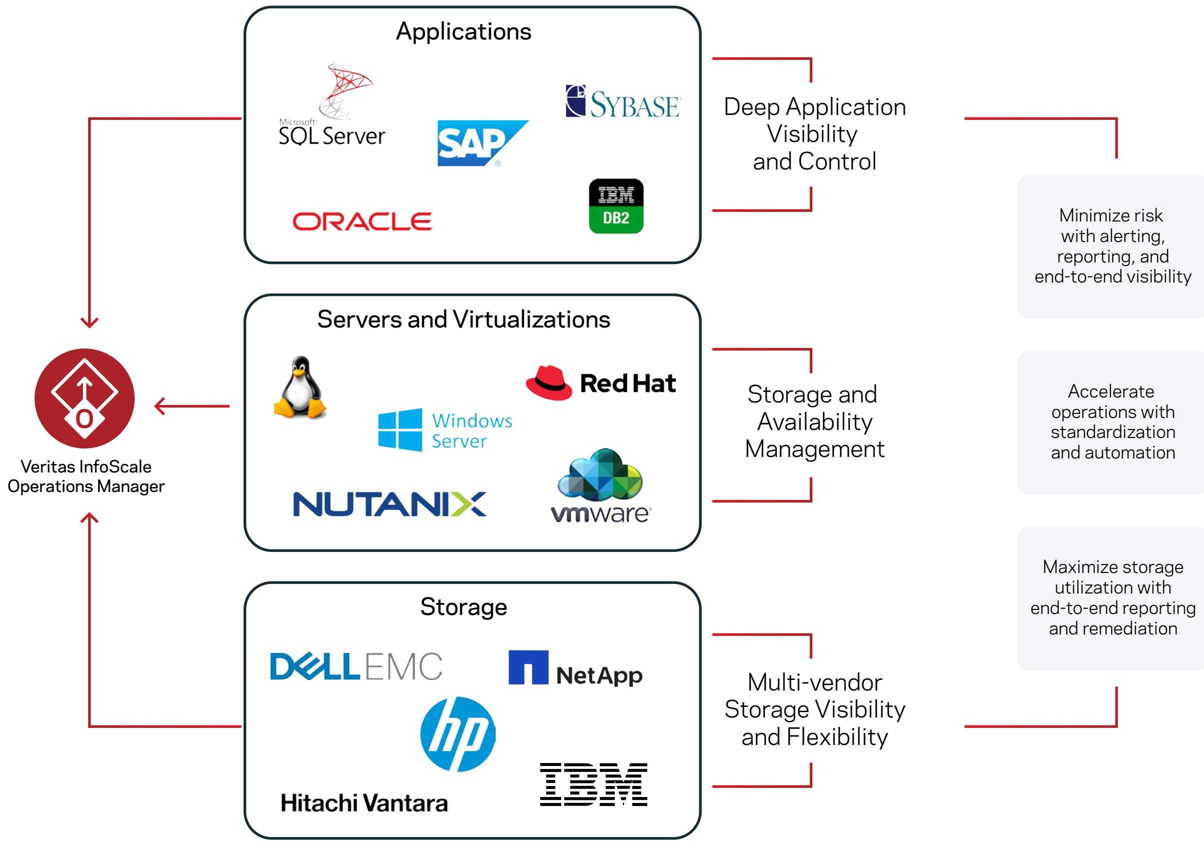


Figure 4. An overview of VIOM management integration points.

You can add managed hosts into VIOM either using agents or as agentless hosts in situations where there is no InfoScale software installed on the target hosts. VIOM manages agentless hosts using SSH or WMI. The level of visibility within VIOM for agentless hosts is the infrastructure only; there is no application visibility for agentless hosts.

Virtual Business Services

Enterprise applications typically consist of multiple systems deployed in tiers that work together to provide an overall business service. A VBS is designed to manage complex, multi-tier applications, even those that span multiple HA (VCS) clusters, as a single entity that represents the overall business service the application tiers are providing.

With the VBS, you can work across a heterogeneous environment, which enables IT to ensure the availability of multi-tiered applications across almost any platform or infrastructure. It doesn't matter if the web server sits in a VMware virtual machine (VM), the application server in KVM on Linux, or the database on physical Big Iron. If the platform falls within the InfoScale support matrix, it can be included as a tier in a VBS.

A VBS manages dependencies between the service tiers by allowing you to configure the order in which the service groups in the VBS are brought online in a start operation and taken offline in a stop operation. The VBS does not alter the dependencies that are configured for clusters included in the VBS tiers.

VIOM provides the management framework and is required to create and manage a VBS with InfoScale Availability (see Figure 5).

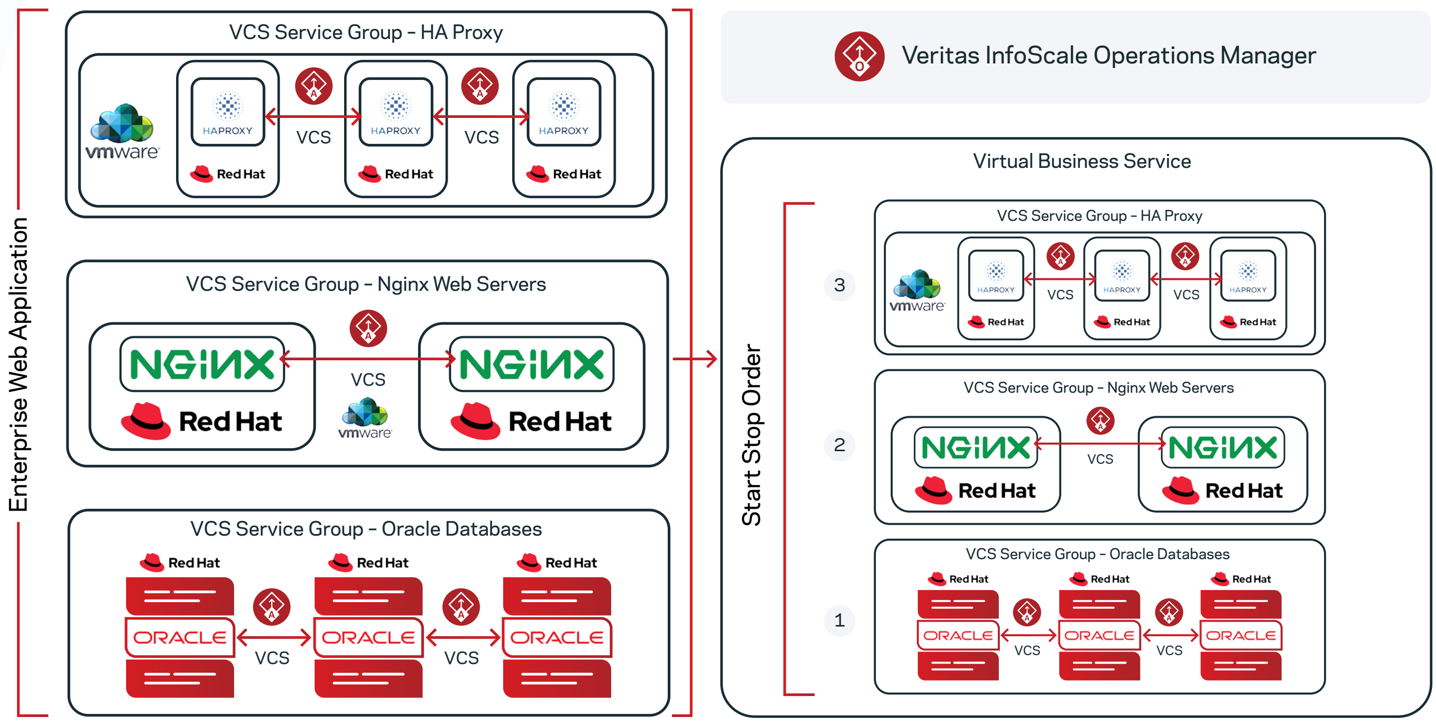


Figure 5. A sample VBS configuration managed by VIOM.

InfoScale Storage

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

1. Veritas File System (VxFS) is a POSIX-compliant journaling enterprise file system designed to maximize application performance.
2. Veritas Volume Manager (VxVM) is the storage management subsystem that allows you to create, organize, and manage logical data volumes and disk groups used by applications and databases and the underlying physical disks and logical unit numbers (LUNs).
3. Cluster Volume Manager (CVM) is the add-on to VxVM's storage virtualization layer that enables the storage devices to be accessed by multiple nodes in the cluster simultaneously.
4. Cluster File System (CFS) is the extension to VxFS that creates file systems that can be shared by multiple nodes in the cluster concurrently.
5. Flexible Storage Sharing (FSS) enables individual nodes to share direct-attached storage with other nodes in the cluster 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.

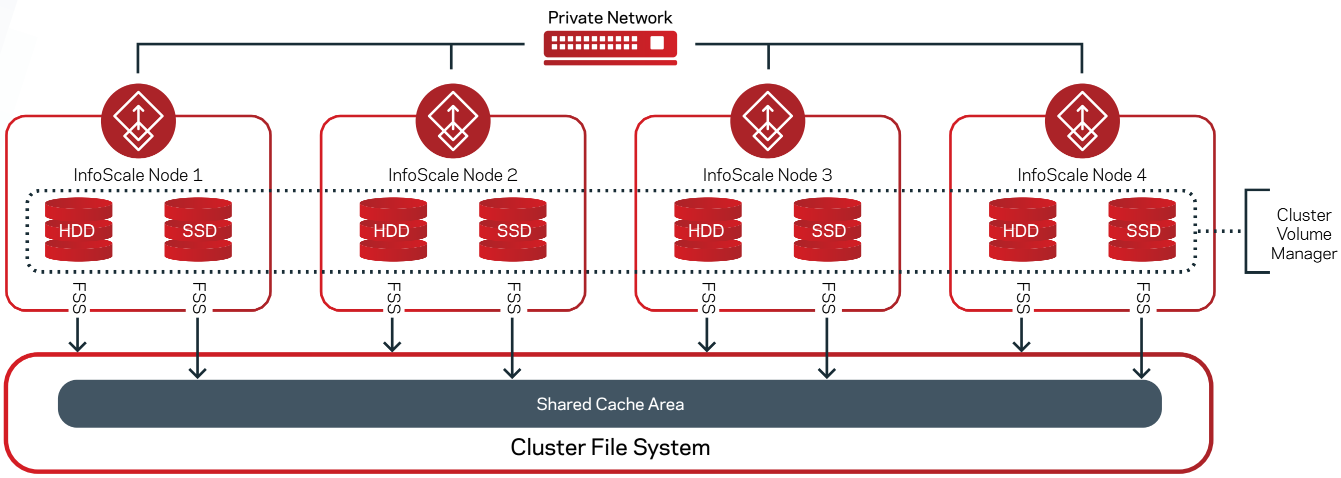


Figure 6. A depiction of how the Cluster File System (CFS) and Flexible Storage Sharing (FSS) enable node sharing.

Figure 6 shows a four-node InfoScale cluster mounting a CFS consisting of storage exported from local disks with FSS. The CFS allows all nodes in the cluster to read and write data to the shared namespace. Data stored on the CFS can be modified by any node on the cluster.

Veritas File System

The VxFS is an extent-based, POSIX-compliant journaling file system capable of managing large volumes of data that is designed to provide high performance and availability for applications. The VxFS has several advanced features that maximize application performance while optimizing the data footprint. The VxFS supports online growing and shrinking of the file system, compression, encryption, and thin reclamation, which allows you to release free data blocks of a VxFS file system to the free storage pool of a thin-provisioned storage LUN. The VxFS also enables you to create storage checkpoints that can be easily restored. You can designate checkpoints as read-only for compliance purposes and as a method of ensuring immutability for your data from threats such as ransomware .

The VxFS allocates disk space to files in groups of one or more adjacent blocks called extents. The VxFS defines an application interface that allows programs to control various aspects of the extent allocation for a given file.

The VxFS supports large file systems of up to 16 petabytes.

Veritas Volume Manager

The VxVM is a software-based data management utility that provides online disk storage management for compute environments and SAN environments. With support for the Redundant Array of Independent Disks (RAID) model as well as erasure coding, the VxVM can be configured to protect against disk and hardware failure and to increase I/O throughput.

The VxVM overcomes restrictions imposed by hardware disk devices and LUNs by providing a logical, software-defined volume management layer. After first creating disk groups, which consist of all the physical disks or LUNs that will be used for a single purpose, you can create volumes that span multiple disks and LUNs. The disk group is the atomic unit in which VxVM instances import (gain access to), deport (relinquish access to), activate (present to VxFS), and deactivate (withdraw accessibility to) disks. The VxVM maintains a redundant, persistent record of each disk group’s membership, volumes, and other underlying structures in dedicated private regions of storage on the disks in a disk group. The VxVM provides the tools to improve performance and ensure data availability and integrity. You can also use the VxVM to dynamically configure storage while the system is active.

The VxVM has several other advanced features for data management, such as FlashSnap for optimized point-in-time data copies and Portable Data Containers that enable the migration of data between different OS platforms.

Cluster File System

The CFS enables you to concurrently mount the same file system on multiple nodes and is an extension of the industry-standard VxFS. Unlike other file systems that send data through another node to the underlying storage, the CFS is a true SAN file system. All data traffic happens over the SAN, and only the metadata traverses the cluster interconnect.

The CFS uses a distributed locking mechanism called Global Lock Manager (GLM) to ensure all nodes have a consistent view of the file system. The GLM provides metadata and cache coherency across multiple nodes by coordinating access to file system metadata such as inodes and free lists. The role of the GLM is set on a per-file system basis to enable load balancing.

Cluster Volume Manager

In general terms, a volume is a unit of storage carved out of a physical disk device. The CVM presents a consistent volume state across an InfoScale cluster as nodes import and access volumes concurrently. It also enables all nodes in a cluster to access their underlying storage devices concurrently. The CVM transforms the read and write requests that CFS addresses to volume blocks into I/O commands that it issues to the underlying disks.

The primary difference between CVM and VxVM is that CVM allows disk groups to be imported on all the systems in the cluster concurrently, whereas VxVM only allows a disk group to be imported on a single node at a time.

All CVM instances in a cluster must always present the same view of disk group and volume configuration, even in the event of:

- Storage device failure—For example, if a disk is added to or removed from a mirrored volume, all CVM instances must effect the change and adjust their I/O algorithms at the same logical instant.
- Cluster node failure—If a cluster node fails while it is updating one or more mirrored volumes, CVM instances on the surviving nodes must become aware of the failure promptly, so they can cooperate to restore volume integrity.

The CVM always guarantees that all instances in a cluster have the same view of shared volumes, including their names, capacities, access paths, and "geometries." Most important, the CVM also manages volume states, including whether the volume is online, the number of operational mirrors, whether mirror resynchronization is in progress, and so forth. A volume's state may change if a device fails or a node fails or an administrative command is issued.

Kubernetes Container Storage Interface Plug-In

InfoScale can also provide software-defined storage for Kubernetes (K8s) environments. The InfoScale Container Storage Interface (CSI) plug-in allows you to use InfoScale volumes created and managed via CVM and CFS as persistent storage for your stateful containerized applications running in Kubernetes.

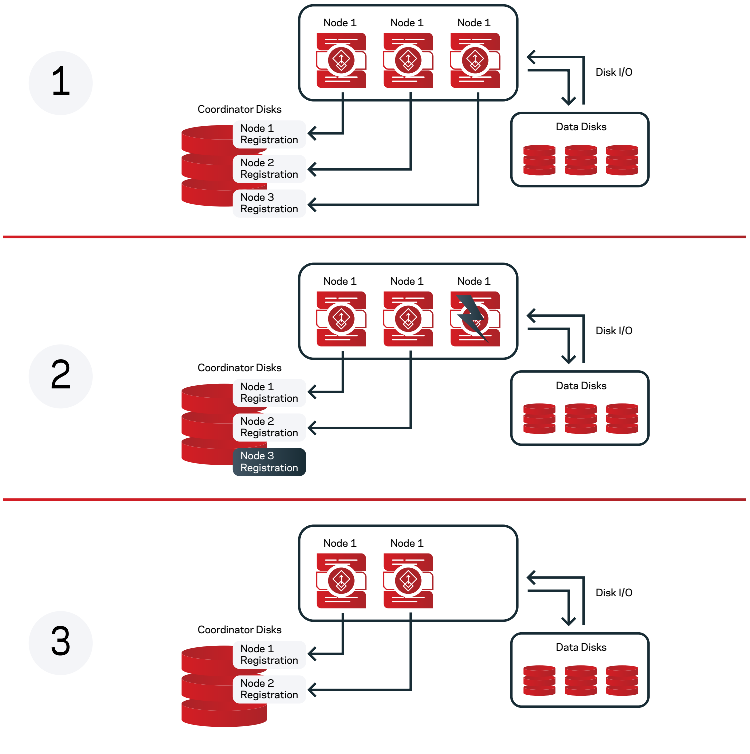
I/O Fencing

A condition known as "split brain" occurs when there is communication disruption between cluster nodes. This disruption can result in data corruption due to the fact that InfoScale (and other cluster software) cannot always distinguish between a system failure and an interconnect failure. The split-brain condition can also occur if a node within the cluster is so busy that it appears to be hung and pauses communication with the other cluster nodes. The split-brain condition can occur in all clustered storage implementations, including K8s. To mitigate and resolve the split-brain condition, InfoScale implements an I/O fencing system that guarantees data integrity by determining which nodes in the cluster should remain in the event of a communication disruption. When a disruption occurs, the node (or nodes) that has failed is ejected from the cluster and prevented from accessing the data disks.

The key to protecting data in a shared-storage cluster environment is guaranteeing there is always just a single consistent view of cluster membership. In other words, when one or more systems stop sending heartbeats, the InfoScale software must determine which nodes can continue to participate in the cluster membership and how to handle the other nodes.

There are three I/O fencing options available in InfoScale:

1. Disk-based I/O fencing—Members of a cluster notify other nodes in the cluster that they are still present by registering themselves to special "coordinator disks." If there is a failure, the coordinator disks ensure the surviving nodes permit write operations to data disks. Disk-based I/O fencing requires SCSI-3 Persistent Reservation–compatible disk devices. (See Figure 7.)
2. Server-based I/O fencing—A special coordination point server maintains a registry of node membership in a cluster instead of disks, as in disk-based I/O fencing.
3. Majority-based I/O fencing—This option is used when coordinator disks or coordination point servers are not available. When a network partition happens, one node in each sub-cluster is elected as the racer node, and the other nodes are designated as spectator nodes. The sub-cluster with the majority number of nodes survives and nodes in the rest of the sub-clusters are taken offline.



1 The cluster is operating nominally
2 Node 3 loses communication with the rest of the cluster. it loses its registration on the coordinator list.
3 Node 3 is ejected from the cluster. Writes from Node 3 are no longer permitted.

Figure 7. A depiction of SCSI-3 I/O fencing in an InfoScale cluster.

For more information on I/O fencing behavior, see "How I/O fencing works in different event scenarios."

The CFS involves a primary/secondary architecture. One of the nodes in the cluster is the primary node for a file system. Although any node can initiate an operation to create, delete, or resize data, the GLM master node actually carries out the operation. After creating a file, the GLM master node grants locks for data coherency across cluster nodes. For example, if a node tries to modify a block in a file, it must obtain an exclusive lock on that block to ensure other nodes that may have the same file cached invalidate this cached block copy.

Flexible Storage Sharing

As an advanced feature of CVM and CFS, FSS enables the network sharing of local storage across a cluster. The local storage can be in the form of either Direct-Attached Storage (DAS) or internal disk drives and can be either spinning disk or solid-state disk. Network sharing of storage is enabled by using a network interconnect between the nodes of a cluster and leveraging CVM and CFS to allow multiple systems to access the disks simultaneously.

FSS allows network-shared storage to co-exist with physically shared storage, and CVM can create logical volumes using both types of storage, enabling a common storage namespace. Logical volumes using network-shared storage provide data redundancy, HA, and DR capabilities without requiring physically shared storage. This process is transparent to file systems and applications.

You can also use FSS with SmartIO technology for remote caching to service nodes that may not have locally attached solid-state drives (SSDs).

SmartIO

The SmartIO feature of InfoScale Storage enables data efficiency on higher-performance storage (such as 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. Rather, SmartIO uses advanced heuristics to determine what data to cache and how that data 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 metadata for the cached data. 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 the internal heuristics. Subsequent I/Os can then be processed from the cache.

SmartIO and FSS

SmartIO supports the use of SSDs exported by FSS to provide caching services for applications running on Veritas Volume Manager (VxVM) and Veritas File System (VxFS). In this scenario, FSS exports SSDs from nodes that have a local SSD. FSS then creates a pool of the exported SSDs in the cluster. From this shared pool, FSS creates a cache area for any or all nodes in the cluster. Each cache area is accessible only to the node for which it is created. (See Figure 8.)

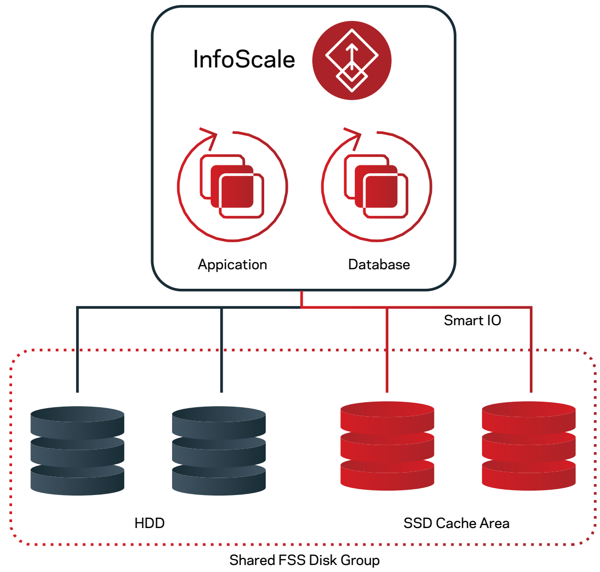


Figure 8. An overview of how InfoScale uses FSS and SmartIO to provide caching services.

Veritas Volume Replicator

VVR is a comprehensive solution for platform-independent replication of data volumes managed by VxVM. VVR enables cost-effective replication of data over IP networks. This capability provides organizations with an extremely flexible, storage hardware–independent alternative to traditional array-based replication solutions as well as a robust mechanism for moving data into public cloud environments. VVR provides the flexibility of block-based continuous replication as well as file-based periodic replication with VFR.

VVR includes the following components, as shown in Figure 9:

- Replicated Volume Group (RVG)—A group of related volumes within a given VxVM disk group configured for replication. A related volume is a set of volumes where writes must be replicated in order on a secondary site.
- Storage Replicator Log (SRL)—A buffer of writes for an RVG. Each RVG contains one SRL and writes to data volumes in the RVG are first queued in the SRL on the primary host before they are sent to the secondary.
- Replication Link (RLINK)—The link between a primary and secondary RVG. A primary RVG can have up to 32 associated RLINKs
- Data Change Map (DCM)—Tracks writes when the SRL overflows. The DCM is also used for the initial synchronization of volumes with a secondary site when starting replication. The DCM becomes active only when the SRL no longer has space to hold accumulated updates
- Replicated Data Set (RDS)—A grouping of a primary RVG and one or more secondary RVGs.

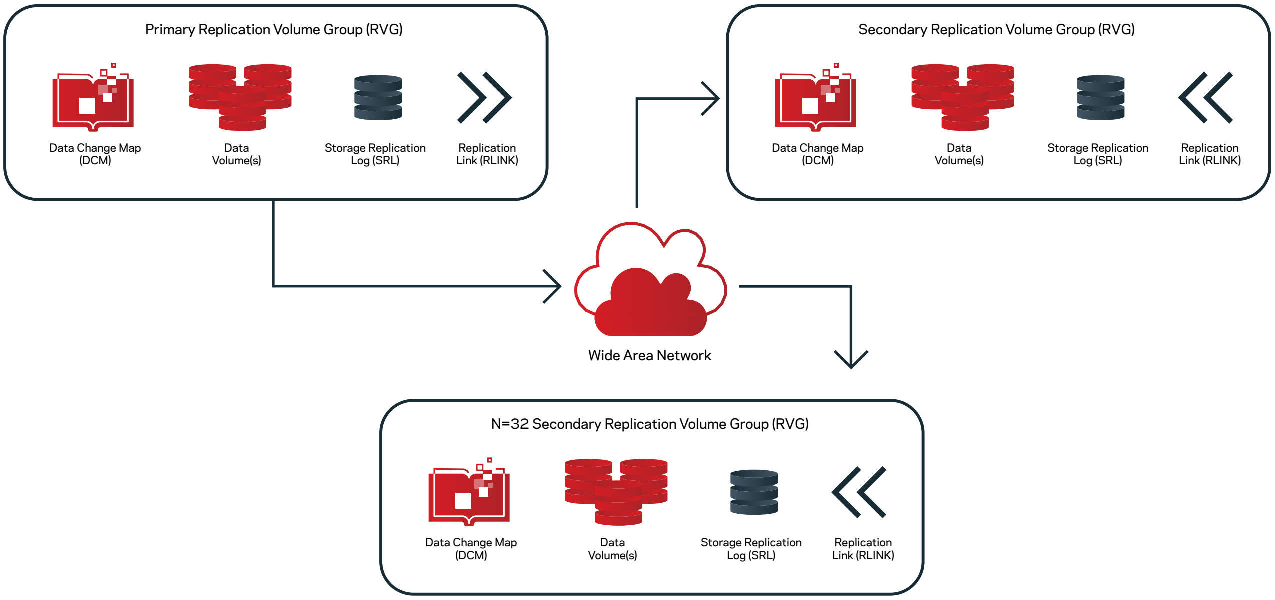


Figure 9. A summary of VVR components.

VVR also manages and maintains write-order fidelity. This is an important feature that ensures data consistency, which is not typically managed by native storage replication solutions. Write-order fidelity means that VVR tracks writes on the primary volume in the order in which they are received and then applies them on the secondary volume in the same order. Maintaining write-order fidelity ensures the data on the secondary volume is consistent with the data on the primary volume. Data on the secondary volume can be behind in time (async replication), but VVR ensures it is a consistent image of the primary RVG at a point in the past.

VVR is a component of InfoScale Storage and is included with InfoScale Enterprise. Its primary use cases include:

- Replication between like or heterogeneous on-premises systems
- Cloud data mobility (data replication from on-prem to cloud, cloud to on-prem, and cloud to cloud)

VVR can work independently or as part of an InfoScale Storage or InfoScale Enterprise cluster. In scenarios where InfoScale is managing storage and HADR for a cloud environment, VVR manages the data replication. VVR supports replication of data stored in cloud volumes, regardless of the storage type, to other public cloud environments or to a customer-managed on-premises storage volume.

VVR has an advanced feature called Adaptive Sync that improves sustained throughput for latency-sensitive applications by automatically switching from synchronous to asynchronous mode and vice versa based on cross-site latency. Adaptive Sync enables the configuration of time-outs for I/O so that if any I/O duration exceeds the time-out, an acknowledgment that the write operation has completed is returned immediately regardless of whether the write operation is completed at the remote site. When VVR is in asynchronous mode, it detects when latency returns to normal and automatically switches replication back to synchronous mode. The Adaptive Sync functionality allows your applications to withstand higher I/O latencies while preserving the ability to maintain remote copies of your application data with a low RPO.

InfoScale Networking

InfoScale clusters are formed from multiple discrete compute nodes connected by private networks, enabling inter-cluster communication. This communication is required for all the nodes to perform functions such as acquiring exclusive "locks" to storage and file system resources, updating cluster node membership (preventing split-brain errors), and transferring data from one node to another to fulfill I/O requests for data not stored on the node local to the request.

Figure 10 shows two private networks configured for the InfoScale cluster. This approach is required to provide redundancy in case one of the private networks fails.

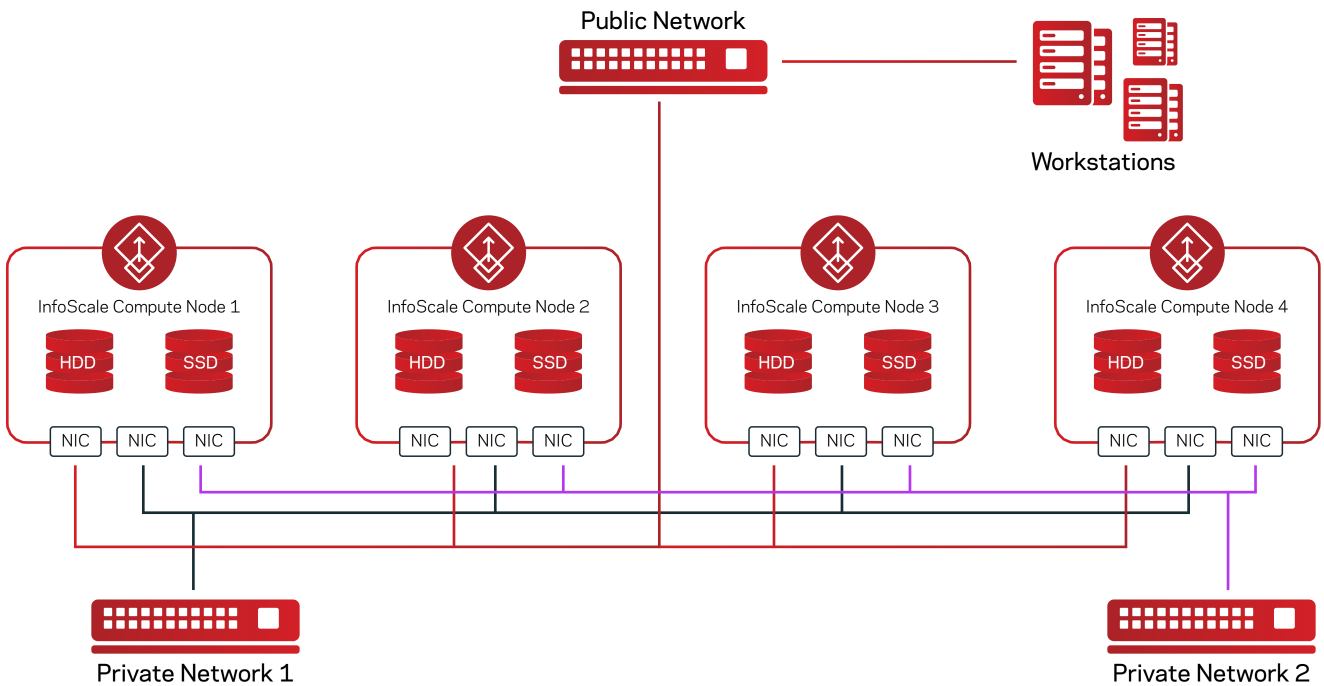


Figure 10. An overview of InfoScale Nnetworking between two private networks.

Cluster communication is conducted over a special high-performance, Low Latency Transport (LLT) protocol. LLT enables each node to discover other nodes in the cluster by means of a broadcast. Once the nodes have been discovered, heartbeats are sent between all the nodes so each node is aware of any changes in the status of other nodes in the cluster.

InfoScale Enterprise

Figure 11 shows an example of an InfoScale Enterprise solution that combines both the HADR functionality of InfoScale Availability as well as the software-defined storage features and functionality offered by InfoScale Storage. Together, the two work to create a highly available, highly scalable, and high-performance environment that maximizes application performance and availability.

Each InfoScale Availability had instance monitors the resources connected to its node and takes appropriate action (for example, initiating application failover) if it detects a critical resource failure.

All *had* instances obtain cluster configuration and policy information from a single cluster configuration file normally called main.cf. The configuration file specifies cluster resources and their organization into service groups.

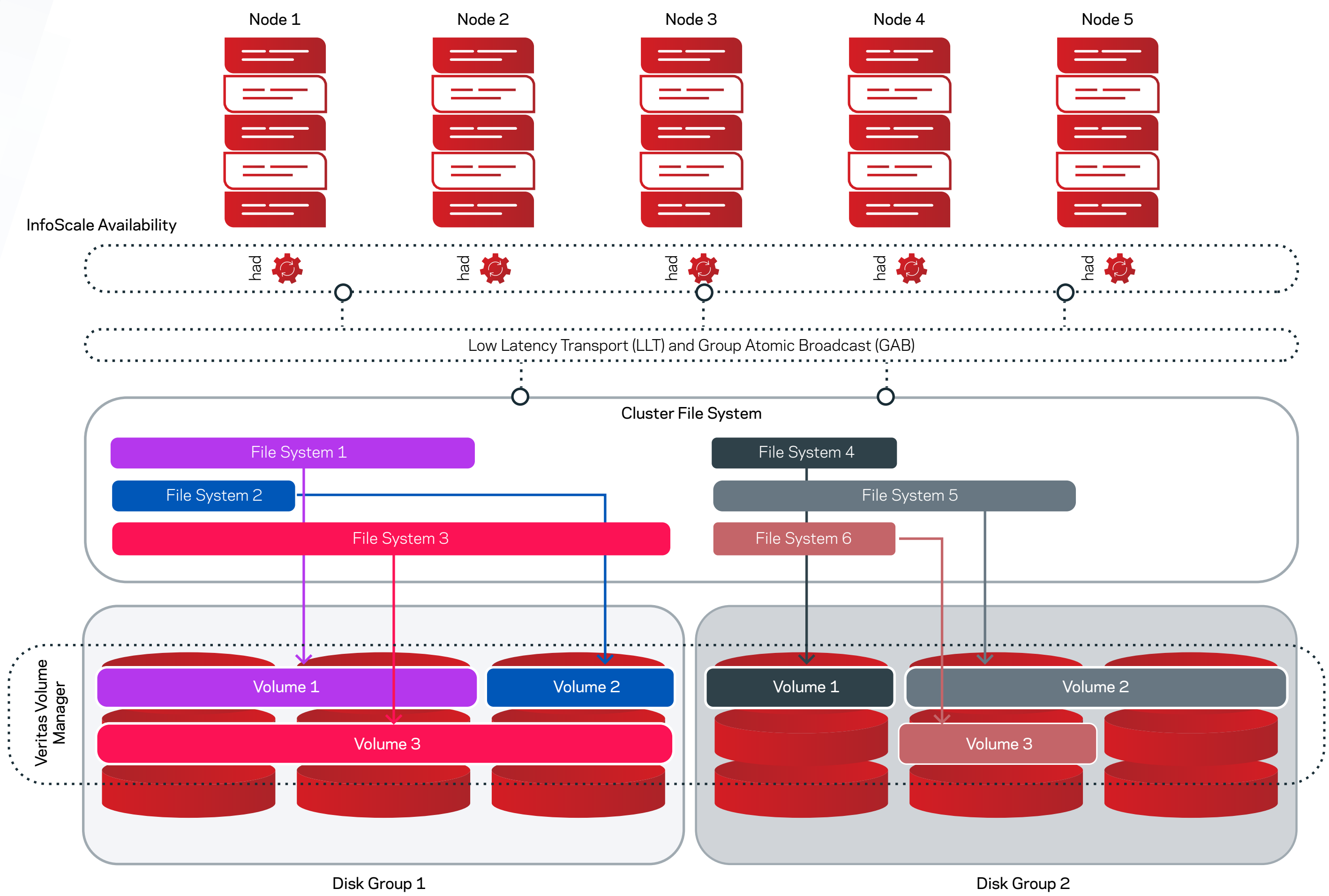


Figure 11. An example of an InfoScale Enterprise solution that includes VCS, CFS, and VxVM.

InfoScale Availability encapsulates applications and the resources they require to run as service groups. Examples of components typically included in service groups are disk groups, CVM volumes, CFS file system mounts, network interfaces, and IP addresses. The InfoScale Availability framework manages service groups by monitoring their resources while they are operating and by starting and stopping them in response to changes in the cluster state and/or administrative commands.

CVM instances are structured as VCS parallel service groups, with an instance running on each cluster node. All instances access storage devices directly through a storage network using Fibre Channel, iSCSI, SAS technology, or FSS. Each CVM instance transforms CFS read and write requests into I/O operations on the underlying devices and executes the corresponding commands.

Cloud Integration

HADR in the Cloud

InfoScale supports several private and public cloud providers and can be used in these cloud environments to provide the HADR capability that is not commonly available with cloud-native tools. Using InfoScale in the cloud provides several benefits:

- Application focus—Although public cloud infrastructure is designed to provide excellent availability and durability for compute and storage systems, InfoScale is focused on the applications that run on top of this infrastructure. Providing all the same benefits as an on-premises deployment, InfoScale Availability cloud agents are developed specifically for cloud services and can manage the cloud compute, network, and storage resources required for both your infrastructure and application to be online in the cloud.
- Quality of Service—With InfoScale, your applications can be deployed in a highly available configuration that spans across cloud zones, regions, or even cloud service providers to ensure you're protected against cloud service disruptions and outages.

- Architectural flexibility—InfoScale supports multiple HADR configurations that are not limited to specific cloud providers. InfoScale can manage applications regardless of the underlying infrastructure, including public cloud infrastructure, so you can easily implement a multi-cloud strategy with a single solution.
- Performance—Like on-premises InfoScale environments, instant fault detection for applications running in cloud environments ensures action is taken instantly in the event of an application failure. InfoScale cloud agents integrate with native cloud monitoring and reporting tools to provide increased operational visibility into the application.

With InfoScale, applications can be replicated to supported public and private cloud environments. InfoScale supports operations such as migrating from physical UNIX systems on-premises to Linux systems in the cloud with minimal configuration. VVR also provides the functionality to extract your data from the cloud to either bring data back on-premises or move data to another cloud provider. This capability gives you the architectural flexibility required to implement a multi-cloud strategy without needing multiple tools or professional services to manage data transfer between different cloud providers.

Cloud-Based Software-Defined Storage

With InfoScale Availability providing HADR capability for cloud environments, InfoScale Storage can significantly improve the performance and efficiency of the underlying cloud-native storage services 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 and 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, they can be scaled dynamically and independently, which reduces operating costs and provides infrastructure flexibility for your applications.

Enterprise Storage and Availability in the Cloud

InfoScale Enterprise combines the HADR features of InfoScale Availability and the storage management and performance benefits of InfoScale Storage to provide the enterprise functionality needed to confidently run a tier 1 application in a public cloud environment. With InfoScale Enterprise, you have the tools needed to manage your high-priority applications in the cloud:

- Application-specific integration for near-instant fault detection and failover
- Highly performant block-level shared storage using cloud-native infrastructure
- Support and tooling for sharing AWS or Azure high-performance block storage devices between nodes
- Simplified multi-tier application management from a single console
- Flexibility to architect your applications to run on any or across multiple cloud platforms

InfoScale currently has solution templates available in public cloud marketplaces to simplify the purchasing and deployment experience. Solution templates are available in the AWS Marketplace (AWS CloudFormation Template), the Azure Marketplace (Azure ARM Template), and the Google Cloud Platform Marketplace (Deployment Manager Template).

Hybrid Cloud

Managing storage and high availability for environments that consist of both on-premises and public cloud infrastructure can be challenging and may require multiple point tools. InfoScale can fully support a hybrid approach to public cloud consumption, offering bi-directional HADR and storage management between on-premises environments and supported public cloud providers. InfoScale manages the application components both on-premises and in the cloud environment, and VVR manages the data replication between on-premises and cloud data volumes. (See Figure 12.)

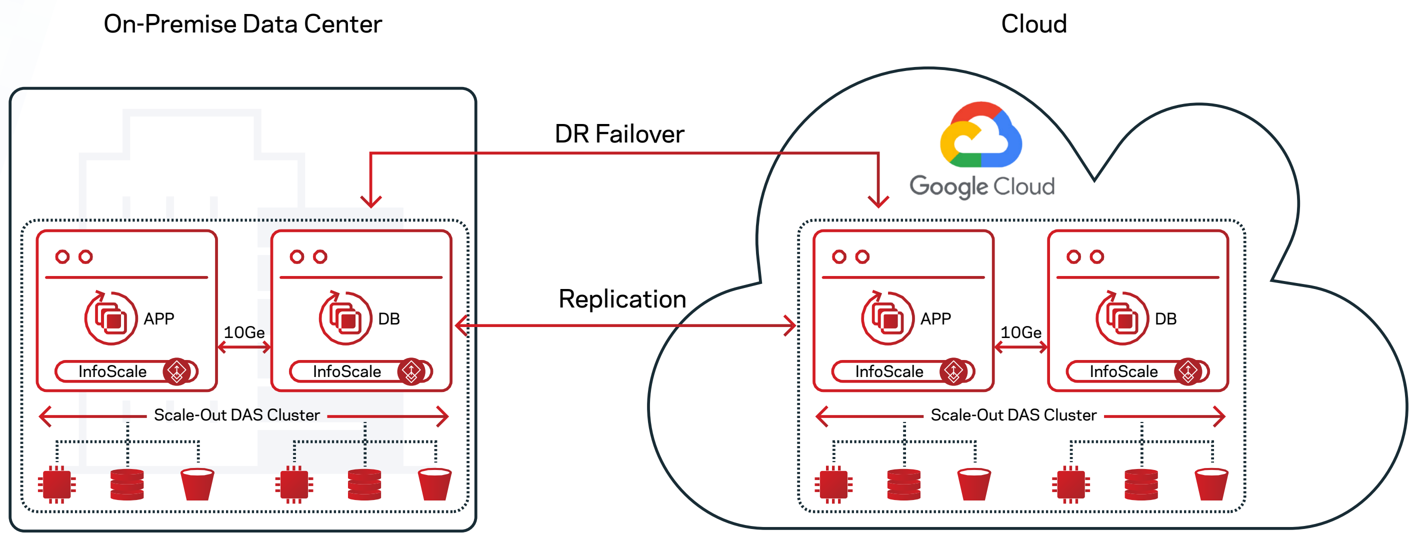


Figure 12. A sample InfoScale hybrid-cloud design concept.

With an agnostic approach to operating systems and platforms, InfoScale is well suited for deployment in hybrid-cloud HADR configurations and can be tailored to support varying RPO and RTO requirements. InfoScale can also support hybrid-cloud HADR configurations with multiple cloud providers, enabling a resilient and performant multi-cloud strategy.

Public Cloud

Public cloud services have limited options for providing HADR and enterprise storage functionality for cloud-native applications deployed in an infrastructure as a service (IaaS) model. In most cases, cloud-native applications have the same RPO and RTO requirements as on-premises applications running in a clustered or other HA configuration. InfoScale offers the same benefits, functionality, and configuration options for the HADR and storage management of cloud-native applications with the added benefit of providing data mobility between cloud providers. This capability helps eliminate being locked into any specific cloud service provider and enables an HA configuration that can protect against cloud provider outages. (See Figure 13.)

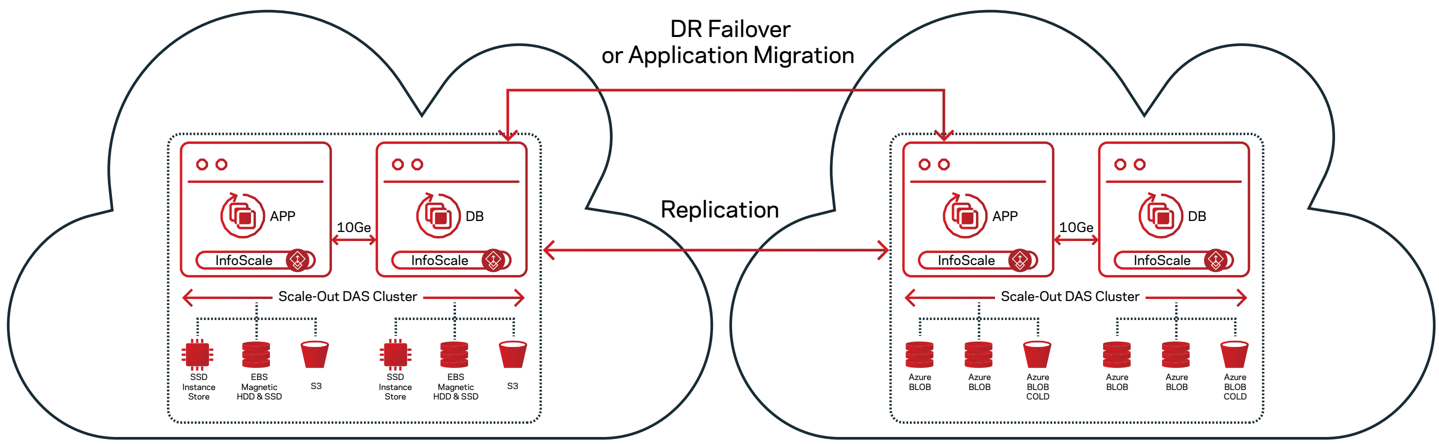


Figure 13. A sample InfoScale public cloud design concept.

InfoScale provides flexible configuration options and is a certified solution for tier 1 applications such as SAP and Oracle that run in public cloud environments. For more information on using InfoScale in public cloud environments, see the following documents:

- SAP NetWeaver in AWS
- SAP HANA in AWS
- Oracle in AWS
- InfoScale in Microsoft Azure

Cloud Migration

There are several tools available for cloud migrations, including those offered by most public cloud service providers. These tools are designed for migration purposes only, however, and don't offer any additional benefit beyond a one-time migration.

InfoScale can support the migration of nearly any application to the cloud while also offering some added benefits over cloud provider migration tools:

- Rehearsal—Using the Firedrill feature, InfoScale manages the testing of an application on a non-production network segment in the cloud using temporarily provisioned cloud compute instances and snapshots of the production data volumes. It can do so on demand using cloud resources, minimizing cost and operational overhead.
- Enterprise readiness—In addition to managing application migration to the cloud, InfoScale also provides HA for your applications once they're migrated to the cloud environment.
- Failback—With full bi-directional operations support for public cloud services, InfoScale can also move applications back on-premises for any reason once they're migrated and online in the cloud environment.

InfoScale supports cloud migrations for on-premises environments being migrated to the cloud and for cloud-native environments being migrated within the cloud or to a different cloud service provider.

As enterprise IT infrastructures become increasingly software-defined, a key driver of efficiency has become the interoperability and addressability of that infrastructure via API, or Application Programming Interface. Exposing the configurability and power of InfoScale's components with an API allows you to programmatically take control them and automate or script common procedures, and allows other infrastructure components you use to query its operating state, provision and update resources, and track metrics.

InfoScale 8.0.2 and later present a RESTful API interface, allowing clients to use common HTTP methods such as GET to (for instance) query a resource, PUT to change it, and POST to create a new one. In short, a RESTful API allows the state of InfoScale to be represented elsewhere, and for changes to that state to be implemented by InfoScale, without the client needing to keep track of that stateful representation itself. The API interface also allows granular control over permissions, permitting or denying access to InfoScale's commands and resources based on an assigned role.

InfoScale for Containers/Kubernetes

Containers have become a popular solution for efficiently developing and running applications. Containers provide excellent application portability for moving applications between environments and they also help improve efficiencies by making it easy to standardize the resources your applications require. However, containers and container orchestration engines like Kubernetes (K8s) do not natively provide all the functionality needed to manage stateful and mission-critical applications.

InfoScale's Container Storage Interface (CSI) plug-in works with Kubernetes to provide advanced storage management for containerized applications. InfoScale Storage services provide the functionality needed by stateful applications running in containers that isn't natively available in a K8s environment, using the RESTful API endpoint provided to make requests of InfoScale's storage and availability resources. InfoScale's enterprise functionality integrates with K8s to provide a container management platform suitable for running stateful and mission-critical applications that require the following:

- Advanced storage management—InfoScale's CSI plug-in allows K8s to provide InfoScale persistent storage volumes to containerized applications being managed within a Kubernetes namespace. You can use InfoScale's FSS to provide high-performance storage using the disks directly attached to the K8s cluster nodes.
- Application and system availability—InfoScale's advanced I/O fencing capabilities let you bring failed nodes and/or application pods back online quickly without worrying that data has become corrupted. By quickly preventing access to shared data from failed nodes or pods, InfoScale prevents the corruption of data that may be in use by other nodes or pods. This approach enables the applications to remain online and servicing requests while administrators work to repair the failures.

InfoScale is deployed directly as a containerized application on K8s cluster nodes and the InfoScale CSI plug-in provides the interface between Kubernetes and InfoScale. Figure 14 provides an overview of how InfoScale integrates with K8s and containers to provide persistent storage for containerized applications.

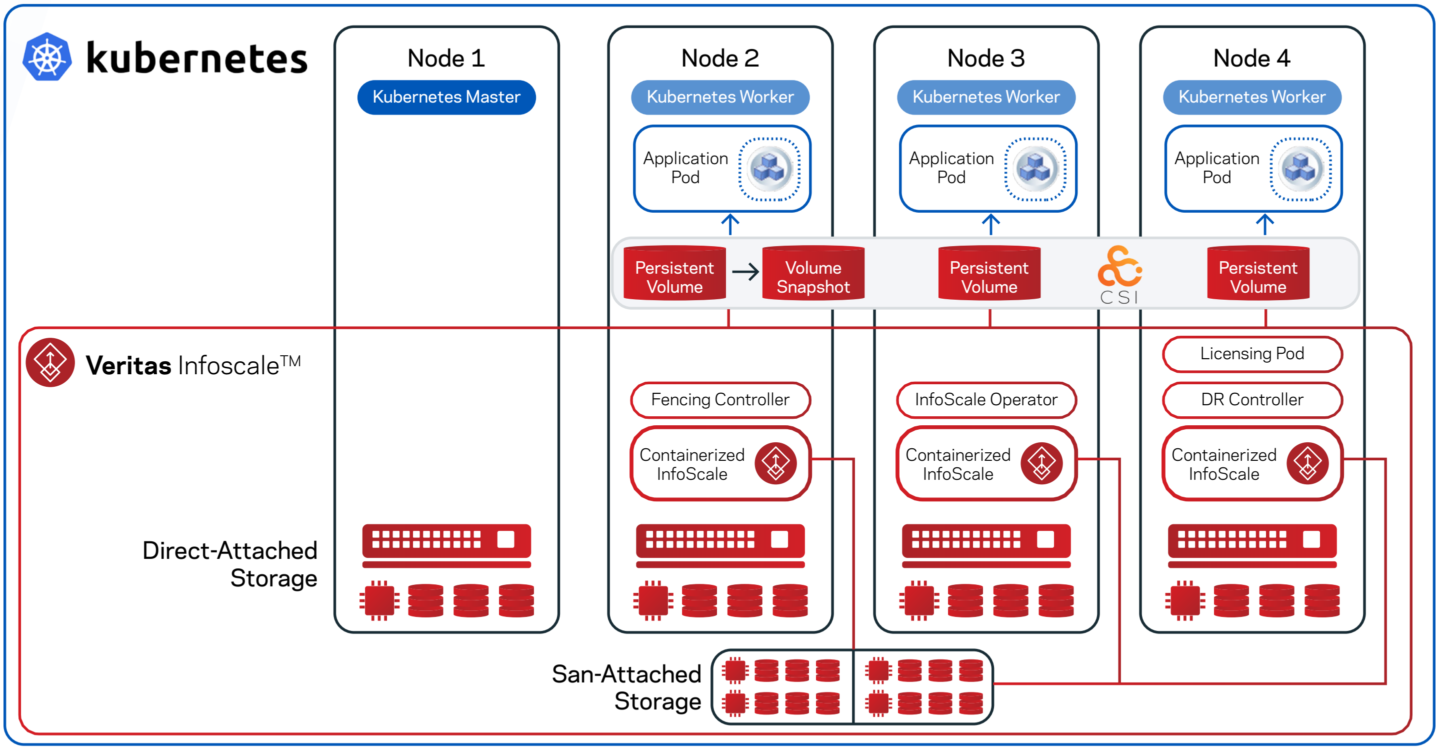


Figure 14. An overview of InfoScale running in a Kubernetes (K8s) environment.

Storage Management

The InfoScale CSI plug-in allows K8s to mount InfoScale Storage volumes formatted with VxFS inside the containers being managed, which stateful applications can then use without the risk of data loss if a container is powered off or removed from the pod. Kubernetes uses the InfoScale CSI plug-in to interface with the InfoScale Storage volumes that are created on the K8s cluster nodes. A storage request is made by K8s using a persistent volume claim (PVC) described in a yaml file. Once a persistent volume is created, it is bound to the PVC and made available to the application in the container.

InfoScale Persistent Volumes

Kubernetes Storage Classes are used to manage the attributes of InfoScale storage volumes that are mounted by K8s inside application pods using the InfoScale CSI plug-in. InfoScale provides several different storage class configuration options that can be used to create persistent storage volumes. Storage classes are defined for performance, resiliency, and security and can be customized to meet application requirements.

InfoScale persistent storage volumes are provisioned by K8s using the InfoScale CSI plug-in either dynamically or statically.

- Dynamic provisioning—Volumes are created at the same time as the container and application pod using K8s with the InfoScale CSI plug-in and InfoScale container installed. An InfoScale persistent volume claim binds the storage accessible to the application pods to the InfoScale persistent volume that is available to K8s cluster nodes.
- Static provisioning—Volumes are created within the K8s cluster by directly accessing the InfoScale containers, creating an InfoScale volume and exposing this volume to containerized applications within the cluster. You can also use InfoScale statically provisioned volumes to simplify the process of migrating traditional applications into containers. Application data volumes outside a K8s cluster can be migrated to volumes within the cluster and then managed by InfoScale within the cluster using the InfoScale CSI plugin. Contact Veritas Support to help validate requirements and environmental variables involved in the migration process.

Figure 15 shows the InfoScale components installed on a Kubernetes system.

The InfoScale CSI plug-in enables K8s to use InfoScale to create persistent volume snapshots that can be used by containerized applications for multiple purposes such as data protection, resiliency, and analytics. Snapshots are used as static persistent volumes and the snapshot creation process has no impact on production data. In addition to persistent volume snapshots, you can also create persistent volume clones that help protect containerized applications against hardware failures.

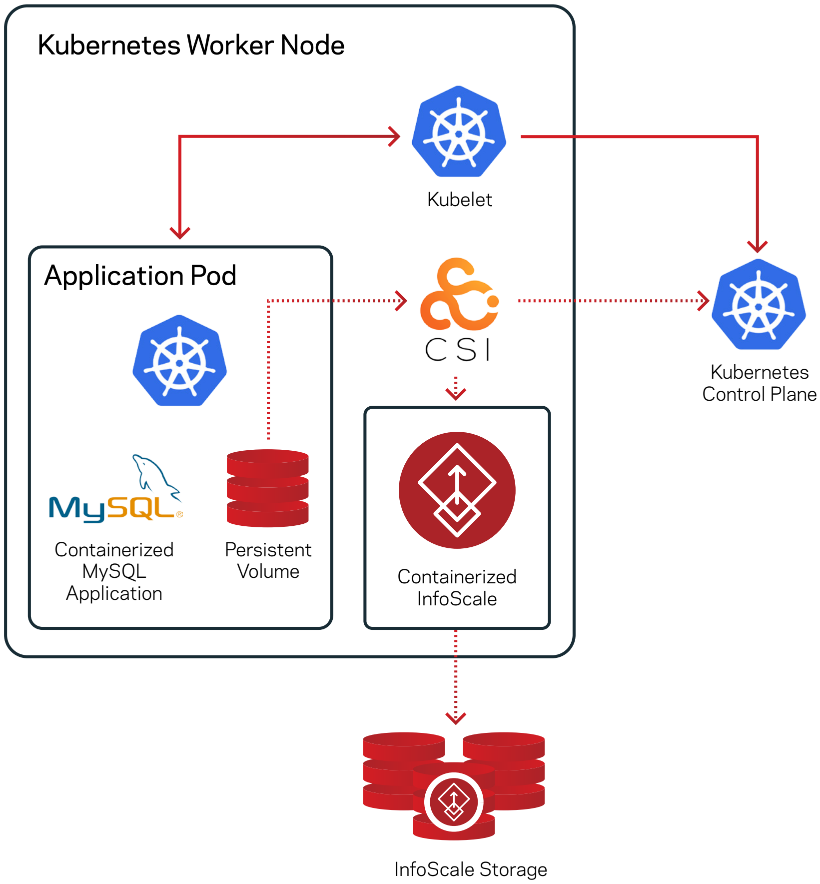


Figure 15. An example of an InfoScale installation on a Kubernetes system.

I/O Fencing

Containerized applications with shared storage provided by InfoScale are automatically protected against data corruption due to a split-brain scenario, which can occur in any clustered environment in the event of a node/hardware failure that disrupts cluster communications and membership. InfoScale provides advanced I/O fencing by preventing data from being written by nodes within the K8s cluster that have failed due to hardware or network communication failures. If a node failure is detected by K8s, the InfoScale fencing driver can ensure the persistent volumes being used by application pods on the failed node are no longer accessible by fencing this node out of the cluster. This process prevents data corruption by allowing only the working nodes to continue normal operations. In the event of a communication loss between cluster nodes (or "worker nodes"), InfoScale's fencing driver relays this information to the K8s master, which can then mark the node as failed and move pods to another node. (See Figure 16.)

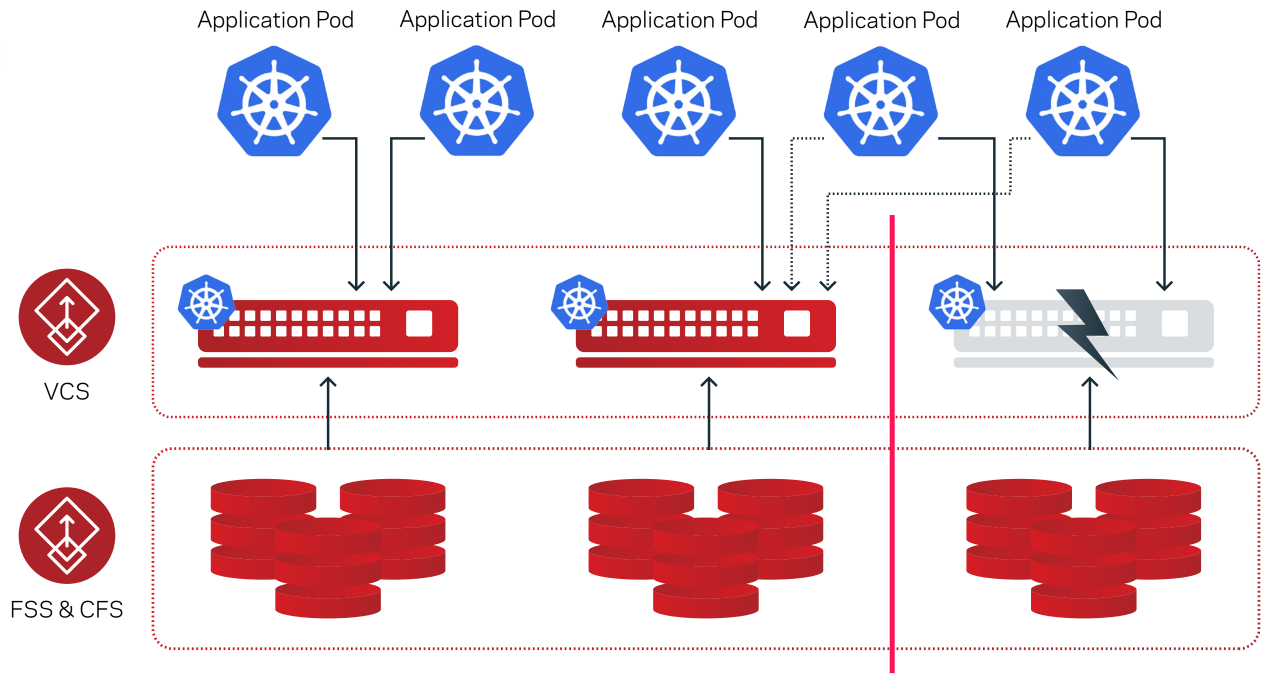


Figure 16. The I/O fencing process in Kubernetes using InfoScale.

Red Hat OpenShift Container Platform Integration

Starting with version 8.0, InfoScale can be used within the Red Hat OpenShift Container Platform (OCP) to provide persistent software-defined storage using local disk or SAN. Leveraging the OpenShift platform lets you take advantage of a fully integrated, enterprise-ready Kubernetes deployment architecture backed by Red Hat's customer support and experience.

InfoScale is a certified Red Hat operator and available as an Operator bundle on the Red Hat Catalog and OperatorHub.io. This availability will enable developers to deploy InfoScale with a single-click native deployment and provision static or dynamic storage based on application needs from the Red Hat OpenShift console. (See Figure 17.)

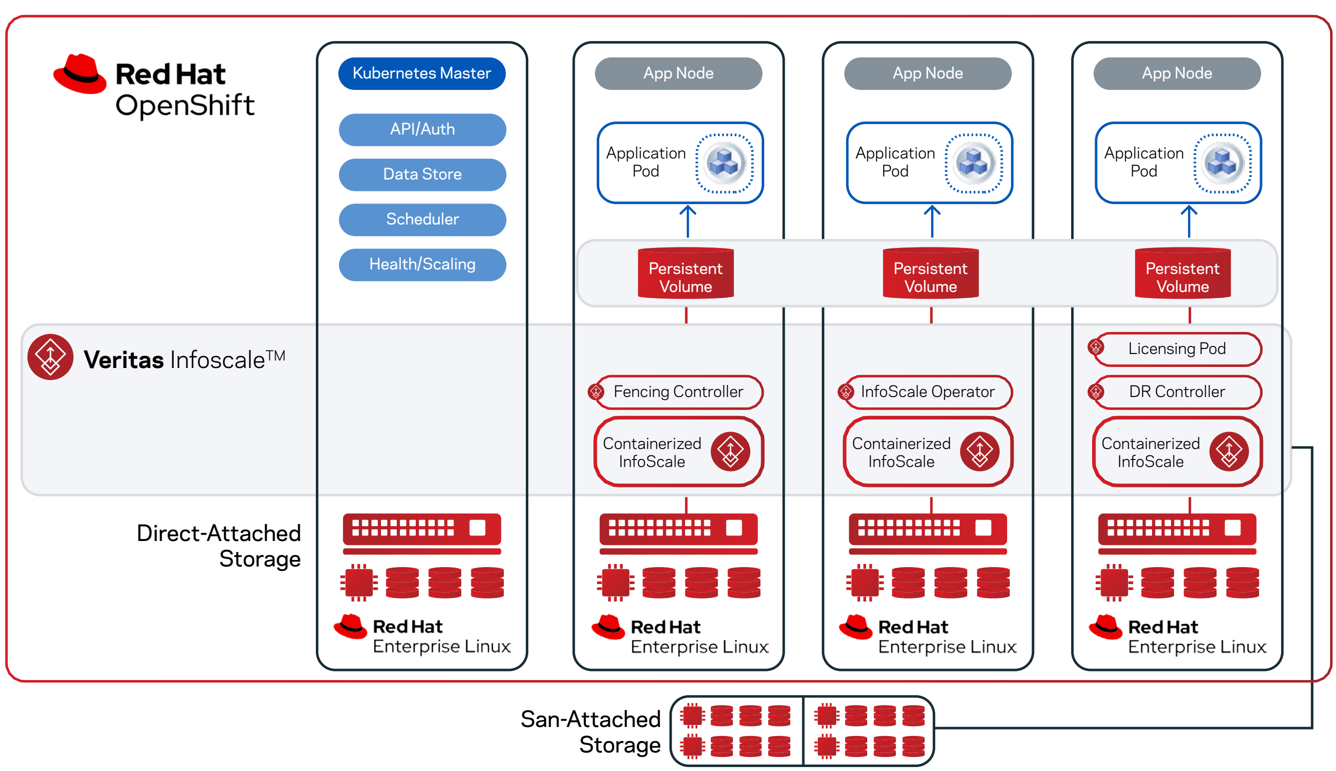


Figure 17. An example of InfoScale in a Red Hat OpenShift environment.

Conclusion

The InfoScale Enterprise architecture provides a highly available, highly scalable, high-performance storage and application management suite that enables organizations to deliver their business services with the confidence they'll always be available and running in an optimal state. InfoScale is a proven solution that offers several key benefits for enterprises with demanding service-level agreements (SLAs):

- Flexibility—Enterprise-level HADR and storage management for any application, any platform, anywhere. InfoScale's agnostic approach to HA and storage management lets you architect the best solution for your applications without being locked into a specific technology or service provider.
- Simplified management—Use a single console and reduce the need for manual intervention in the HADR process, which can be time-consuming and error prone. Virtual Business Services (VBS) further simplifies managing system availability by logically representing complex, multi-tier applications as a single entity you can manage with a single click.
- Enterprise-level availability—With instant application fault detection and the ability to manage several HADR configurations across operating systems, platforms, and even cloud services, InfoScale provides best-in-class options for managing HADR for your most important applications.
- Increased confidence—InfoScale's integrated, non-disruptive DR Firedrill feature lets you maximize production uptime and increase your confidence in the HADR and cloud migration processes.
- Best-in-class performance—As a software-defined solution, InfoScale can maximize resource utilization and significantly improve application performance with an intelligent data management capability that works with any application, on any type of infrastructure. This approach significantly reduces operational costs, eliminates vendor lock-in, and provides complete freedom of architecture to build and run your applications on the infrastructure that best fits your business requirements.

With the ability to provide exceptional application performance and high availability as well as architectural flexibility, InfoScale enables businesses to improve their application SLAs while reducing infrastructure footprints by integrating cloud into their IT strategy. InfoScale can help ensure maximum application performance and uptime in heterogeneous environments. Whether running on-premises, in a hybrid-cloud configuration, or entirely within a cloud environment, InfoScale provides an enterprise software-defined storage and application availability solution for any platform.

References

Cluster File System in Veritas InfoScale Data Sheet

Storage Foundation Cluster File System High Availability 8.0 Administrator's Guide – Linux

Veritas Flexible Storage Sharing Data Sheet

Veritas InfoScale™ 8.0 SmartIO for Solid-State Drives Solutions Guide – Linux

Veritas InfoScale™ 8.0 Getting Started Guide - Linux

Veritas InfoScale™ Enterprise Data Sheet

Veritas InfoScale™ Intelligent Monitoring Framework

Veritas InfoScale™ Main Site

Veritas InfoScale™ Solutions for Cloud Environments

Veritas InfoScale™ Storage 8.0 Administrator's Guide – Linux

Veritas Volume Replicator Administrator's Guide – Windows

Veritas Volume Replicator Administrator's Guide – Linux

Virtual Business Service – Availability User Guide

Table of Figures

Table of Tables

About Veritas

Veritas Technologies is a leader in multi-cloud data management. Over 80,000 customers—including 95 percent of the Fortune 100—rely on Veritas to help ensure the protection, recoverability, and compliance of their data. Veritas has a reputation for reliability at scale, which delivers the resilience its customers need against the disruptions threatened by cyberattacks, like ransomware. No other vendor is able to match the ability of Veritas to execute, with support for 800+ data sources, 100+ operating systems, 1,400+ storage targets, and 60+ clouds through a single, unified approach. Powered by Cloud Scale Technology, Veritas is delivering today on its strategy for Autonomous Data Management that reduces operational overhead while delivering greater value. Learn more at www.veritas.com. Follow us on Twitter at @veritastechllc.

VERITAS

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

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

V0995 05/23