



BakBone NetVault:Replicator with RedHat Enterprise Linux and RedHat Global File System

Abstract

This application brief addresses how to use BakBone's NetVault:Replicator to add continuous real-time byte-level data replication to RedHat Enterprise Linux and optionally RedHat Global Cluster File System for scalable long distance protection.

May 2006

Data Protection Objectives

The growing reliance on IT solutions and the constantly increasing storage capacities makes protecting an organization's data resources both more important and more difficult at the same time.

Because of the overwhelming reliance on technology, organizations simply cannot afford to experience a failure of critical applications. Even worse, they cannot afford to lose the only copy of recently changed or created data.

History continually demonstrates the importance of taking proactive measures to protect an organization. Preparing for natural or man-made disasters is something that every organization, large and small, needs to do. Organizations that rely on technology to drive their business cannot afford to experience downtime, but if they do, they need to recover as quickly as possible to the point right before the interruption.

What is Continuous Data Protection?

What is true continuous data protection and how does an organization effectively protect their data? The high availability of services and data is critical for all organizations, regardless of size. This continuous dependence on an organization's changing data requires a sufficient level of data protection to meet the needs of the organization. The constant growth of an organization's data results in increased challenges in providing high levels of data availability and reliability.

Continuous Data Protection is the level of availability of IT services to provide application availability for an organization's critical services, applications and corresponding critical data. Critical environments require the addition of continuous data protection solutions to control their stringent recovery time objectives (RTO) and recovery point objectives (RPO) , rather than only relying on slow and out-of-date tape based strategies.

IT infrastructure services like email, office file servers, database servers, accounting and inventory systems and other applications are critical to the operations of any company. Protecting those services and underlying data assets while complying with new government regulations requires that companies plan to effectively manage the availability of their mission-critical applications and data.

Data Availability Solutions Need to Scale

Most organizations have already implemented disaster recovery solutions through the use of traditional tape backups. Many of these organizations also store computer backup tapes safely off-site. Although tape backups provide a basic level of data protection to recover small numbers of lost files, they fail to continuously protect rapidly changing data stores on critical servers and are ineffective at recovering from a large-scale disaster.

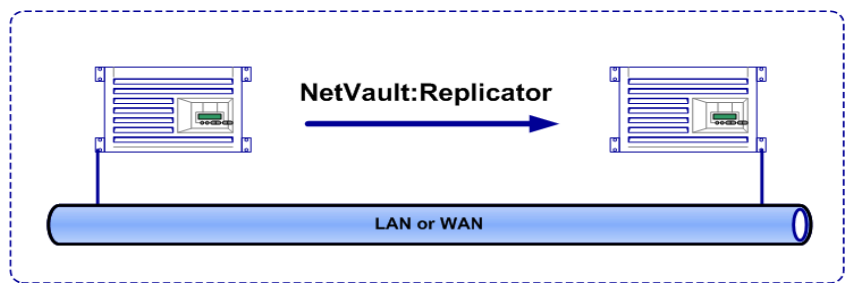
The challenges with data availability solutions that only rely on backups is that they fail to sufficiently protect the data that is continuously changing. The more data that needs to be protected, the longer it takes and the more expensive it is to perform tape backups. Conversely, the more data that is being backed up, the

longer that data will take to recover and the less effective the data protection solution.

The Replication Solution: RedHat Enterprise Linux and BakBone NetVault:Replicator

Recent global events have stressed the importance of continuous data availability solutions for business continuity. Data protection through off-site real-time data replication for data availability has become a higher priority and a growing requirement for enterprises, large and small. BakBone and RedHat helps companies prepare for and minimize interruptions.

NetVault:Replicator installs on servers to continuously maintain a usable on-line mirror backup of changing file system data on another remote server for increased protection. Since online replicas of data remain on a hot-standby server readily available and verifiable, recovery times from unplanned outages are minimized.



Supports RedHat Enterprise Linux

NetVault:Replicator provides support for all 2.4 and 2.6 Linux kernels and is certified on all versions of RedHat Enterprise Linux. The software can be installed on production Linux servers without downtime.

Efficient Replication of Byte-Level Changes over Long Distances

Once installed and configured, NetVault:Replicator continuously replicates changes to filesystem data between servers over any TCP/IP network regardless of the distance between the servers. This storage-independent architecture allows the software to be installed on critical servers in the data center or on servers in other remote offices or off-site data centers. NetVault:Replicator provides unattended data protection with operational simplicity.

Supports Popular File Systems

NetVault:Replicator replicates changes to files on a mounted file system. Customers can configure replication via a list of regular expressions that define files to include and optionally exclude from replication. The software supports all popular file systems.

Application Independence for Greater Flexibility

NetVault:Replicator replicates file changes made to the file system regardless of the application modifying data.

Unattended Failover Support

NetVault:Replicator is a cluster-aware solution that maximizes application and data availability when combined with a third party cluster management software

like RedHat Cluster Suite.

RedHat Enterprise Linux and the RedHat Global File System Provides Scalable High Availability on a Local SAN

Many organizations have also taken steps to implement clustered solutions for high performance load balancing and high availability of lower-cost and high-performance server hardware with the increased demands for solving complex and compute-intensive problems. Increased levels of application performance often requires increased numbers of servers clustered together with access to the same storage.

Organizations are scaling out to take advantage of cost-effective high performance blade servers or clusters of servers. These servers have access to common local storage usually through Storage Area Networks (SANs).

RedHat's Global File System (GFS) allows a cluster of RedHat Enterprise Linux servers to share data in a common pool of storage, allowing you to:

- Greatly simplify your data infrastructure:
- Install and patch applications once, for the entire cluster
- Reduce the need for redundant copies of data
- Simplify back-up and disaster recovery tasks
- Maximize use of storage resources and minimize your storage costs:
- Manage your storage capacity as a whole vs. by partition
- Decrease your overall storage needs by reducing data duplication
- Scale clusters seamlessly, adding storage or servers on the fly:
- No more partitioning storage with complicated techniques
- Add servers simply by mounting them to the common file system
- Achieve maximum application uptime with optional Cluster Suite for unattended application failover

The Long Distance Problem caused by Network Latency

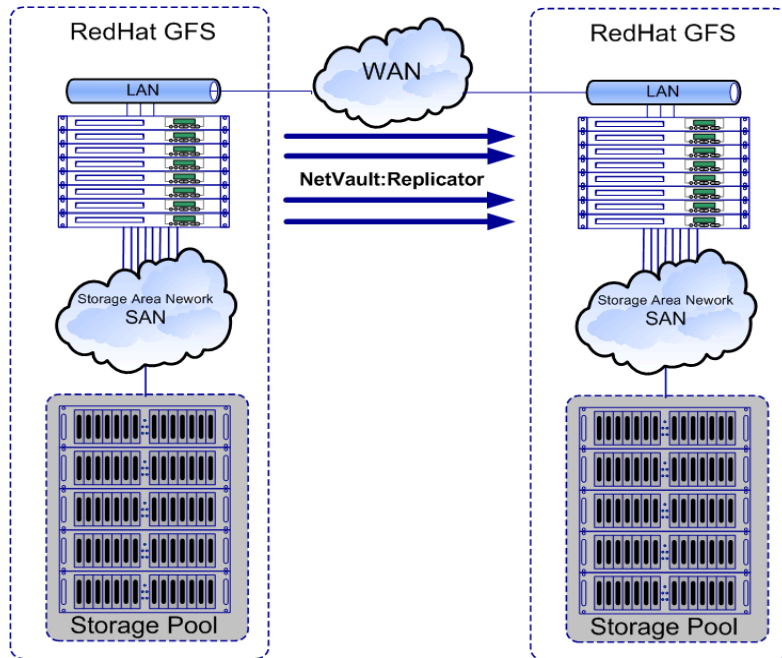
To meet the demands of increased levels of corporate-wide data availability many organizations require storage clusters to share file data with other servers or RedHat GFS storage clusters located at remote offices or data centers. This long-distance requirement is not a capability provided by storage clusters because of the use of common storage usually in local SAN in the same data center.

The RedHat GFS File System is installed on the individual nodes in the storage cluster to create common storage. Each node in the local RedHat GFS cluster has extremely fast access to common storage. However, due to the physical limitations of network performance and network latency, the RedHat Global Clustered File System is unable to provide clustering over long distances.

The Long Distance Linux Enterprise Solution: RedHat Enterprise Linux, RedHat GFS and BakBone NetVault:Replicator

The RedHat GFS shared storage clusters can be extended over long distances by installing BakBone's NetVault:Replicator on each node in the GFS storage cluster. The combined solution will replicate data changes that are made on each node in the cluster to a corresponding destination server or node in a remote RedHat GFS cluster.

Customers seeking higher levels of data availability for fault isolation and increased levels of application performance benefit from adding BakBone NetVault:Replicator to existing RedHat storage clusters on RedHat Enterprise Linux to interconnect existing geographically distributed clusters providing universal data availability.

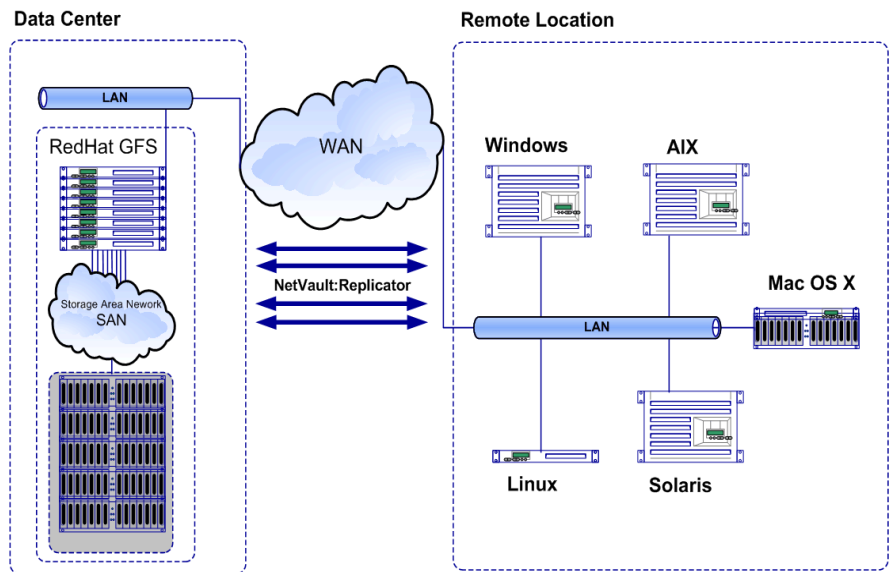


BakBone's NetVault:Replicator is a low-impact solution that is installed on every node in a RedHat GFS shared storage clusters and is independent of the clustered file system and existing clustered applications.

NetVault:Replicator is unaffected by network latency so replication is not affected by the distance between the replicated servers. With no distance limitation, data changes on the existing nodes in a cluster are asynchronously mirrored between different nodes in distributed remote RedHat GFS clusters.

With the combined RedHat and BakBone solution, Linux enterprise customers in clustered or grid computing environments can ensure that large amounts of rapidly changing critical data will remain universally available across distributed clusters even in the event of multiple node failures, no matter where those failures occur, or how far apart those clusters might be located.

Additionally, customers may wish to replicate data between a RedHat GFS Cluster and another remote Linux, Unix or Windows server:



Continuous Byte-level Replication Scales with Growing Storage

NetVault: Replicator scales along with your growth regardless of the size of the underlying storage or the distance between the nodes in the replicated mirror. The application-independent replication software extends the Storage Area Network over IP networks between remote data centers beyond the typical distance limitation of Fibrechannel SANs. Data is kept safely up-to-date off-site over existing IP network infrastructures.

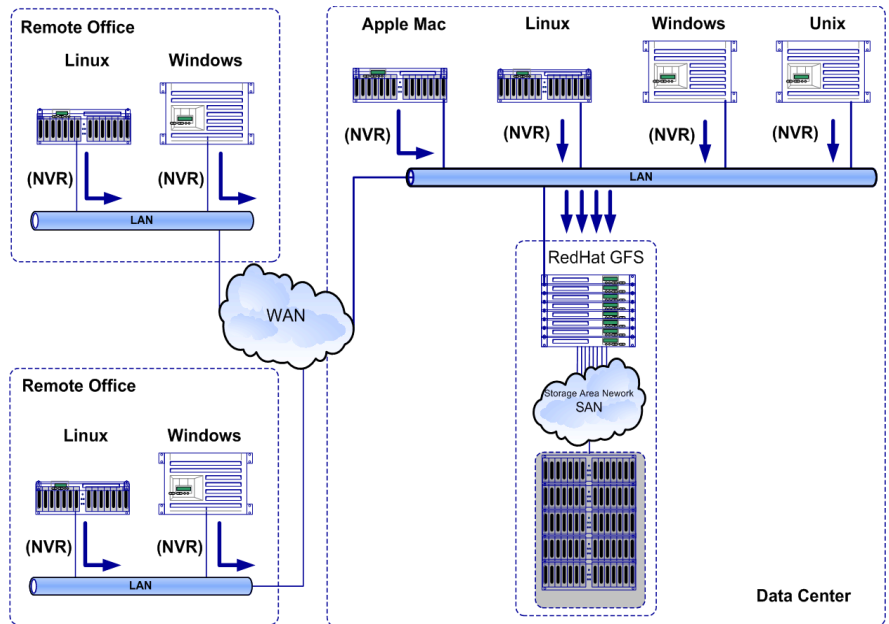
Data Consolidation from Remote Sites

Many customers may also wish to consolidate content from servers located in remote offices, remote suppliers, data centers or remote research centers as is often the case in Grid computing environments.

Continuous real-time data consolidation offers organizations with the opportunity to reduce the cost of existing server backups while simplifying workflow challenges by consolidating data to a centralized location.

NetVault:Replicator provides the instant benefits of better collaboration, improved business processes, and increased compliance while also providing the long-term benefits of allowing an organization to survive in the event of a disaster.

In this configuration, the RedHat GFS clusters would have NetVault:Replicator installed on individual nodes in the cluster and be configured as a destination of replication:



NetVault:Replicator for Heterogeneous IT Environments

Most IT environments have a combination of servers installed from three main operating systems: Windows, Unix and Linux. Traditionally, Linux was installed into predominantly Unix or Windows IT computing environments for edge computing services like web servers running apache or file/print servers in remote offices. As Linux continues its widespread adoption and moves upstream from dedicated edge computing services to more critical core data center applications like hosting database servers, the requirement for Linux-centric enterprise data protection solutions becomes more of a priority.

One of the unique advantages of BakBone's NetVault solutions for data protection is the heterogeneous cross-platform capabilities of its backup and recovery products as well as its continuous real-time data replication solution. The heterogeneous capabilities of NetVault:Backup and NetVault:Replicator allow Linux to play a significant role in enterprise data protection.

Unique Cross-Platform ACL Metadata Support

NetVault:Replicator replicates changing files at the byte-level not only between servers of similar operating systems, but also cross-platform between servers of different operating systems. Identical copies of files are maintained across multiple servers regardless of their operating system. In addition to preserving the underlying file data, a unique feature of the replication technology translates and preserves the file metadata between servers of different operating systems. This capability allows NetVault:Replicator for Linux to be installed in heterogeneous computer environments with traditional Unix or Windows servers. The only limitations to this feature is the requirement for the file system on the destination server to support extended attributes.

For example, customers can install NetVault:Replicator on Sun Solaris servers and continuously replicate files to remote Linux servers or visa versa. The Sun Solaris file ACLs are preserved on the Linux server, assuming the file system on the Linux server supports extended attributes. This same capability applies to replication between Windows servers and Linux servers. Customers can use NetVault:Replicator to continuously replicate file changes between Windows servers and Linux servers. The Windows metadata attributes are stored within

the extended attributes on the remote destination file system.

Summary

BakBone's NetVault:Replicator can be used to add continuous real-time byte-level data replication to RedHat Enterprise Linux and optionally RedHat Global Cluster File System for scalable long distance protection.

www.redhat.com

www.bakbone.com/products/replication/