VERITAS Global Cluster Manager™
Heterogeneous Cluster Management and Disaster Recovery Protection
# Table of Contents

- Executive Overview ................................................. 1
- VERITAS Global Cluster Manager ................................. 1
- VERITAS Global Cluster Manager Architecture .............. 3
- Disaster Recovery Design and Management .................. 6
- Replication Components and Monitoring ..................... 7
- Replication Agent Example ........................................ 8
- Putting the Pieces Together ....................................... 8
- Summary ............................................................ 10
Executive Overview

Enterprise implementation of clustering technologies traditionally was centered around protecting the availability of back end databases. Providing a means to ensure availability to the critical heart of a company's data store was essential for corporate viability in a data-driven digital economy. As business enterprises grew and their data availability needs similarly expanded, the use of clustering evolved to include a wider range of application service protection beyond just the availability of back-end data sets. Availability of the entire IT infrastructure was required from the transaction tier through the application layer and down into the database level, no matter what the cause of failure or even planned downtime. The type of applications requiring high availability grew to encompass a wider array of business requirements. Additionally, corporations began structuring their high-availability needs around the business units of their organizations. Clusters of applications and their associated resources were defined and controlled with greater granularity, imposing strict guidelines of where and how resource groups could be failed over and similarly how their resources could (and could not) be used as failover targets for other business units' needs. In order to provide this type of protection, more complex clustering or groups of systems are configured to satisfy a company's high-availability needs.

The next level of high availability surfaced with the need for disaster recovery protection; protecting and providing failover not only application groups from node to node in a local cluster but from entire clusters to clusters. A natural disaster or even a more benign rolling blackout may cause the complete shutdown of a company's data center. Data availability needs in today's information driven economy require all associated resource groups and accompanying applications to cleanly fail over to another site geographically in order to provide uninterrupted access to the information managed at the affected site. These geographically disperse take-over sites add yet another dimension to the enterprise's availability infrastructure.

As a result of these complex and sometimes competing requirements, a company's high-availability landscape can become populated with dozens, sometimes hundreds and even larger numbers, of individual clusters. Each cluster itself may be composed of perhaps 20+ systems managing a similar number of resource groups. Efficient management of this environment, which may span from just a few feet to thousands of miles between cluster groups and across multiple operating systems, is essential for the smooth and proper operation of a well-planned highly available data center.

The purpose of this white paper is to provide an overview of VERITAS Global Cluster Manager™— wide area application availability management software that functions as a critical part of an overall high-availability plan providing wide area network cluster and application group visibility, data replication management and disaster recovery protection.

VERITAS Global Cluster Manager™

VERITAS Global Cluster Manager is a complete Web-based cluster management solution that allows a single consolidated view of your VERITAS Cluster Server™ environment. The Web-based GUI allows administrators to view and manage multiple, discrete and geographically dispersed Cluster Server clusters. Events and actions can be coordinated among the separate clusters, creating a virtual global cluster that provides the framework for wide area controlled administrative migration and disaster recovery protection.

The foundation of the Global Cluster Manager framework is Cluster Server-controlled clusters. VERITAS Cluster Server is a highly scalable clustering solution that monitors and controls the availability of applications running in a cluster and restarts applications in response to a variety of hardware and/or software faults. A highly scalable solution, allowing up to 32 nodes within a single cluster, Cluster Server is currently available for Solaris, HP-UX and Windows NT platforms (Windows 2000 and IBM AIX platforms available soon).

VERITAS Global Cluster Manager extends the Cluster Server environment to a wide area configuration by managing Cluster Server application and cluster state information and broadcasting status updates, event handling and replication commands over standard TCP/IP connections to remote sites. The interface to Global Cluster Manager is completely browser based, providing administrators with sufficient access privileges and a location-independent command post from which to monitor, manage and control their global enterprise clustered environments.
Development motivation:

There are several instances where managing multiple clusters may be preferable to managing one:

- **Consolidated Cluster Management**—Customers may have multiple Cluster Server clusters but want to see the configuration of all applications from one explorer tree.

- **Functional Disparity**—Separate clusters are running applications that are only marginally related. These clusters perform independent tasks, and there is no advantage, yet possible disadvantages, for creating a single cluster to control the unrelated services.

- **Business Unit Divisions**—Separate business units within an organization may require complete division of system resources and application usage requiring unique cluster failover and disaster recovery requirements and failover policies.

- **Architectural Disparity**—Clusters consist of machines of different architectures that cannot be easily unified into a single Cluster Server cluster.

- **SAN scalability**—As more and more systems are connected to SAN infrastructures, the need for multiple clusters becomes apparent.

- **Geographic Diversification**—Clusters are too far apart physically to be part of the same GAB (Global Atomic Broadcast; the interconnect protocol of nodes in a single Cluster Server cluster) membership.

- **Disaster Tolerance**—Customers want a second cluster or site to take over applications in case an application cannot be run on any system in a primary cluster.

Distributed Cluster Management

Global Cluster Manager can be used to monitor and manage multiple clusters of different architectures and different operating systems from a single Web-based console. Events may be coordinated, and automatic responses may be provided. Managing multiple clusters from a unified console has four principle purposes:

- **Cluster Scalability**—Clusters may be added dynamically to an existing configuration without offlineing the existing Global Cluster Manager configuration. Once the new cluster is configured, it is automatically registered in the Global Cluster Manager framework.

- **Monitoring and Management**—Clusters of various operating systems can coexist within the same site (a virtual collection of clusters), viewed logically from a single explorer tree and controlled by common command syntax.

- **Administrative Migration**—Planned downtime for upgrades and maintenance can be virtually avoided by migrating site resources to an alternate site. Global enterprises can implement “follow-the-sun” policies and ensure that their support and data infrastructures are operating in proximity to their active business requests and queries.

- **Disaster Recovery**—Administrators with clusters in physically disparate areas can use Global Cluster Manager to set policy for migrating applications from one location to another in the event that clusters in one geographic area become unavailable because of an unforeseen disaster. In cases where it is not feasible to link all the systems as a single cluster, Global Cluster Manager enables them to be configured as a cluster of clusters to keep applications available in the event of failure of an entire location.
VERITAS Global Cluster Manager Architecture

Within the Global Cluster Manager management framework lies an extremely scalable architecture with remote node management of 1 to 16,000 separate systems and virtually an unlimited number of resource groups from a single console!

Object Hierarchy

- **Site**—virtual collection of one or more VERITAS Cluster Server clusters (up to 8 can be managed simultaneously)
- **Cluster**—one or more systems that are part of the same GAB membership (max: 64 clusters per site)
- **System**—a computer that is part of the cluster (max: 32 systems/cluster)
- **Group**—a collection of resources that are required to be online for a service to be available

Process Architecture

**Global Cluster Manager Master**

The Site Master runs on one system in one cluster per site. The Site Master is itself a highly available VCS service group that can be locally failed over to another node in case of server or service fault. A site is a virtual grouping of one or more Cluster Server clusters. The Site Master receives connection and state information from every slave in a site, providing a consolidating view of all clusters in a site.

Site Master consolidates system and application information for each cluster in a site. The site master connects to other site masters that exchange information about the site they manage. As a result, each master stores information about the entire enterprise. State changes are therefore forwarded from Cluster Server to local slave to site master to all other site masters.

Site masters are robust in that they can “infer” failed or state changes even when cluster slaves may be down during the event change. Global Cluster Manager is tolerant of WAN/network connection issues and disconnects.

**Global Cluster Manager Slave**—Runs on one system in each Cluster Server cluster; receives configuration, state information and updates of Cluster Server-controlled application groups. The slave runs as a Cluster Server application itself, and the replicated state machine control of Cluster Server allows for only one slave instance per cluster.
Component Architecture

Global Cluster Manager components themselves run as individual Cluster Server applications for manageability and availability. All application components can fail over, with local slaves and remote site masters seamlessly reconnecting. Upgrade process is the same as for any Cluster Server application (offline, remove package, reinstall, online).

Cluster Service Group components:

- Master: GCMmaster
- Slave: GCMslave
- DNS update: GCMdns
- SNMP/SMTP notifier: GCMnotify
- IP alias: GCMalias
- GUI Web Server: VRTSweb

Slave clusters contain only the GCMslave and VERITAS Web components.
Command Process

- Clients (commands, GUI) can connect to Global Cluster Manager master and issue commands.
- Master forwards commands to appropriate Cluster Server cluster (via another master if necessary and then to appropriate slave/Cluster Server cluster).
- Result is returned to the client.
- Corruption is impossible because Cluster Server treats slaves like any other client (rejects competing online commands, etc.).
- VERITAS Cluster Server is required for Global Cluster Manager; no other product is required.

Interface

The Web browser GUI provides an intuitive, dynamic and omnipresent interface. The browser-based interface provides summary information, hierarchy structure, including state and global information, policy definition, logs, real-time state graph and online help.

Command line includes same features and command set available from GUI. Under the Global Cluster Manager framework, group online/offline commands extend to support any group anywhere. Logs are replicated and written to disk for auditing purposes. Logs are searchable and sortable directly from the GUI.
Global Cluster Manager now provides an integrated interface with Cluster Server through links to the Cluster Server Web GUI. This enhancement offers additional monitoring of cluster, group and system-level information and management options.

Future direction will link the cluster management framework to the SAN infrastructure layer, allowing complete visibility and control from heterogeneous SAN environments to individual storage resources.

Disaster Recovery Design and Management

Building a robust disaster recovery infrastructure
Always start with a strong foundation of local availability before proceeding to wide area disaster recovery implementation:

- VERITAS Cluster Server for application fail over
- VERITAS Volume Manager™ for disk mirroring
- VERITAS File System™ for fast file recovery
- Dual dedicated paths between nodes and from host to disk
- Adequate cooling, bracing, power and data backup protection

Stretching the definition of “local”

- Placing all equipment in the same lab exposes equipment to local outage.
- Volume Manager Mirroring is limited only by fibre-channel distance ~10km.
- Local cluster (via VCS) interconnect is restricted only by private network cabling limitations.
- Such a configuration could survive a rolling blackout that usually affects one or a local group of buildings at a time.

Extending to Wide Area Management and Failover
Once a resilient and robust local high-availability environment is provided, extension to a wide area framework should be considered in order to provide capabilities for:

- Global application and cluster management
- Disaster Recovery
- Follow the Sun Administration Migration and Call Center Support
- Exceeding geographic limitations for mirroring imposed by fibre and scsi architectures
- Providing dual dedicated LLT links are not feasible
- Business Continuity
Replication Components and Monitoring

Global Cluster Manager provides the Replication Framework, Replication Agent, and Global Application for monitoring the status of replication processes. These enable you to determine outstanding writes, link status and other replication information useful in planning and timing failover responses to brownouts or other marginal conditions as well as site disasters.

Global Cluster Manager works with host-based replication applications (e.g., VERITAS Volume Replicator™) and array-based replication solutions to monitor and manage data replication to remote sites to support disaster recovery. Global Cluster Manager provides complete replication configuration, control, management and status updates for these replication solutions completely from within the Global Cluster Manager Web console.

Replication Framework

VERITAS Global Cluster Manager provides a replication framework that provides an extensible platform for plugging in popular third-party replication solutions.

- Replication management is flexible, providing mechanisms for both graceful migration and controlled failover.
- Delivery of actions and logging of results are performed locally—replication commands are performed at the location where replication is occurring, not through remote commands such as rsh/rcp.
- Global Cluster Manager replication framework is similar to Cluster Server agent framework.
- Agents linking to Cluster Server, Volume Replicator, and third-party solutions allow for flexible policy-based failover.
- Agent binaries are installed on every node where replication is configured.
- Failover and replication management is available for both synchronous and asynchronous replication environments.
- Trigger automatic action based on events at multiple levels of granularity: resource group, cluster and site.
- Agent periodically calls getinfo entry point to gather real-time information about replication and deliver to Global Cluster Manager interface. This allows the administrator to notice any replication problems in addition to application status from the same console. This is also useful for failover decision—for example, in determining how much data can be lost. Action is useful to modularize various actions and guarantee they will run in the appropriate location. Agents issue commands locally where the replication group is running; result is returned to the Global Cluster Manager master.

Global Application

The Global Application enables monitoring at a higher level than the Cluster Server or group object because it holds information for application groups that span clusters and/or sites.

Replication is configured through the Global Cluster Manager global application module. The global application virtually links one or more applications in any cluster. By specifying a replication relationship, Global Cluster Manager will automatically configure the specified replication agent and add it to the specified application group, providing control of where it will start, stop and fail over the group.

The global application intelligently handles provided agent parameters and passes them in proper order to their respective application handler (e.g., Cluster Server commands, Volume Replicator commands). All results are consolidated and displayed in the GUI and command line.
Replication Agent Example

VERITAS Volume Replicator™ agent
Getinfo performs vxrlink status periodically and displays how far out of synch Primary and Secondary are.

Action can perform any of the following:
- Promote
- Demote
- Convert
- Resynch

The global application command gagrp will deliver the parameters to the agent for the specified group wherever it is running.

If primary site is down:
  Determine if data loss is acceptable.
  If so, take over.
When old primary comes back up:
  Convert to secondary.
  Resynch with new primary.
Once data synch is complete, perform administrative migration.
  Demote “new” primary.
  Promote secondary.

Putting the Pieces Together

Once resilient local availability architecture has been provided and a suitable replication technology implemented, Global Cluster Manager acts as the command center for strong policy-based and dynamic WAN HA configurations.

Use Global Cluster Manager to:
- Monitor health of sites—remote master knows that the primary cluster is completely down because it cannot ping it across the WAN (WAN heartbeating). Global Cluster Manager uses by default an ICMP ping on the remote site’s master alias to determine site failure. Additional heartbeats (via other physical mechanisms such as wireless or satellite) can be configured to report to Global Cluster Manager in order to increase confidence in remote site status detection.
- Multiple site links can be configured to each other site, providing a redundant wide area detection schema. A site is considered faulted only if all sites lose all heartbeats to the respective site. Additionally, within a site, a resource (application) group is considered faulted only if it is faulted on every running node in a cluster or that entire cluster and/or site faults.
- Application failover can be configured with either manual confirmation (default) or automatic.
- Control application migration between clusters—remote master can start application in local cluster by issuing commands to Cluster Server on that cluster.
- Replication Management—interfaces to popular replication software and can dynamically change parameters to reflect takeover by new cluster and/or site.
Normal Conditions

Global Cluster Manager monitors and manages applications and cluster status across all sites. Data is replicated to the remote site while Global Cluster Manager tracks and displays replication status and reports any lag or issues with replication activity. Integration with replication technologies (host and array based) allow for pro-active management of replication activity.

Local Application Fault and Group Failover

Global Cluster Manager monitors local application faults and reports and tracks local application failover within the cluster via VERITAS Cluster Server. If necessary, modification of VERITAS Cluster Server events and triggers can be controlled directly from the Global Cluster Manager interface.

Data replication continues and client access to the application is uninterrupted.
Complete Site Fault and Takeover

Complete site takeover (whether planned or unexpected) is efficiently handled by Global Cluster Manager configured policies. Name Server updates trigger, providing seamless redirecting of client access to the remote/disaster recovery site.

Data consistency has been maintained prior to the site failure via ongoing replication allowing client access with virtually no perception of moved resources. The remote site is promoted as the primary site and replication roles are reversed in order to prepare the original site for eventual failback when site issues are resolved and replication consistency provided.

- Update name service (supports Bind v.8)—interface to DNS to cause original address to resolve to IP address of takeover cluster.
- Support both host-based and storage device-based replication solutions.
- Support both synchronous and asynchronous replication control.
- Synchronous mode—both sides must receive data before I/O OK is returned to writing application. Data integrity is maintained, but performance is affected if network is slow.
- Asynchronous—Data writes are queued—first returns I/O OK and sends to remote site at specified later time. Data may be lost if other side takes over while out of synch; but application is performance is same as mirroring only.

Summary

As the high-availability needs of enterprise storage systems continue to expand, efficient management of distributed cluster architectures becomes increasingly important. Global Cluster Manager can serve as the foundation of such an enterprise’s data center. Whether the needs are simple wide area monitoring, dual site takeover configurations or multiple, redundant failover locations handling hundreds of cluster groups and located thousands of miles apart, Global Cluster Manager can provide the framework for protecting business-critical data and ensuring Business Without Interruption™.