Western Group (Thailand).,Ltd.
31/92 Rangsit-Klong 7 Rd., Lam luk ka,
Patum-tani, 12150, THAILAND
Tel: +662 909-3691, Mobile: +66(0)8-1908-1052
Fax: +662 909-3691

VERITAS Cluster Server support

VERITAS Cluster Server can be used to eliminate both planned and unplanned downtime. It can facilitate server consolidation and effectively manage a wide range of applications in heterogeneous environments. VERITAS Cluster Server supports up to 32 node clusters in both storage area network (SAN) and traditional client/server environments, VERITAS Cluster Server can protect everything from a single critical database instance, to very large multi-application clusters in networked storage environments. This section provides a brief summary of the features of VERITAS Cluster Server.

Hardware Requirements

Following is a list of hardware currently supported by VERITAS Cluster Server:

- For server nodes:
  - Any SPARC/Solaris server from Sun(TM) Microsystems running Solaris(TM) 2.6 or later with a minimum of 128MB RAM.
- For disk storage:
  - EMC Symmetrix, IBM(R) Enterprise Storage Server(R), HDS 7700 and 9xxx, Sun T3, Sun A5000, Sun A1000, Sun D1000 and any other disk storage supported by VCS 2.0 or later; your VERITAS representative can confirm which disk subsystems are supported or you can refer to VCS documentation.
  - Typical environments will require mirrored private disks (in each cluster node) for the DB2(R) UDB binaries and shared disks between nodes for the DB2 UDB data.
- For network interconnects:
  - For the public network connections, any network connection supporting IP-based addressing.
  - For the heartbeat connections (internal to the cluster), redundant heartbeat connections are required; this requirement can be met through the use of two additional Ethernet controllers per server or one additional Ethernet controller per server and the use of one shared GABdisk per cluster

Software Requirements

The following VERITAS software components are qualified configurations:

- VERITAS Volume Manager 3.2 or later, VERITAS File System 3.4 or later, VERITAS Cluster Server 2.0 or later.
- DB Edition for DB2 for Solaris 1.0 or later.

While VERITAS Cluster Server does not require a volume manager, the use of VERITAS Volume Manager is strongly recommended for ease of installation, configuration and management.

Failover

VERITAS Cluster Server is an availability clustering solution that manages the availability of application services, such as DB2 UDB, by enabling application failover. The state of each individual cluster node and its associated software services are regularly monitored. When a failure occurs that disrupts the application service (in this case, the DB2 UDB service), VERITAS Cluster Server and/or the VCS HA-DB2 Agent detect the failure and automatically take steps to restore the service. This can include restarting DB2 UDB on the same node or moving DB2 UDB to another node in the cluster and restarting it on that node. If an application needs to be migrated to a new node, VERITAS Cluster Server moves everything associated with the application (i.e., network IP addresses, ownership of underlying storage) to the new node so that users will not be aware that the service is actually running on another node. They will still access the service using the same IP addresses, but those addresses will now point to a different cluster node.

When a failover occurs with VERITAS Cluster Server, users may or may not see a disruption in service. This will be based on the type of connection (stateful or stateless) that the client has with the application service. In application environments with stateful connections (like DB2 UDB), users may see a brief interruption in service and may need to reconnect after the failover has completed. In application environments with stateless connections (like NFS), users may see a brief delay in service but generally will not see a disruption and will not need to log back on.

By supporting an application as a service that can be automatically migrated between cluster nodes, VERITAS Cluster Server can not only reduce unplanned downtime, but can also shorten the duration of outages associated with planned

Western Group (Thailand).,Ltd.
31/92 Rangsit-Klong 7 Rd., Lam luk ka,
Patum-tani, 12150, THAILAND
Tel: +662 909-3691, Mobile: +66(0)8-1908-1052
Fax: +662 909-3691

downtime (i.e., for maintenance and upgrades). Failovers can also be initiated manually. If a hardware or operating system upgrade must be performed on a particular node, DB2 UDB can be migrated to another node in the cluster, the upgrade can be performed, and then DB2 UDB can be migrated back to the original node.

Applications recommended for use in these types of clustering environments should be crash tolerant. A crash tolerant application can recover from an unexpected crash while still maintaining the integrity of committed data. Crash tolerant applications are sometimes referred to as *cluster friendly* applications. DB2 UDB is a crash tolerant application.

7 For information on how to decrease the amount of time it takes 7 to perform a failover using a VERITAS CFS, CVM, and VCS solution, see the white paper 7 entitled "DB2 UDB Version 8 and VERITAS Database Edition: Accelerating Failover Times in DB2 UDB Database Environments", which is available from the "DB2 UDB and DB2 Connect(TM) Online Support" web site 7 (http://www.ibm.com/software/data/pubs/papers/).

Shared Storage

When used with the VCS HA-DB2 Agent, Veritas Cluster Server requires shared storage. Shared storage is storage that has a physical connection to multiple nodes in the cluster. Disk devices resident on shared storage can tolerate node failures since a physical path to the disk devices still exists through one or more alternate cluster nodes.

Through the control of VERITAS Cluster Server, cluster nodes can access shared storage through a logical construct called "disk groups". Disk groups represent a collection of logically defined storage devices whose ownership can be atomically migrated between nodes in a cluster. A disk group can only be imported to a single node at any given time. For example, if Disk Group A is imported to Node 1 and Node 1 fails, Disk Group A can be exported from the failed node and imported to a new node in the cluster. VERITAS Cluster Server can simultaneously control multiple disk groups within a single cluster.

In addition to allowing disk group definition, a volume manager can provide for redundant data configurations, using mirroring or RAID 5, on shared storage. VERITAS Cluster Server supports VERITAS Volume Manager and Solstice DiskSuite as logical volume managers. Combining shared storage with disk mirroring and striping can protect against both node failure and individual disk or controller failure.

VERITAS Cluster Server Global Atomic Broadcast(GAB) and Low Latency Transport (LLT)

An internode communication mechanism is required in cluster configurations so that nodes can exchange information concerning hardware and software status, keep track of cluster membership, and keep this information synchronized across all cluster nodes. The Global Atomic Broadcast (GAB) facility, running across a low latency transport (LLT), provides the high speed, low latency mechanism used by VERITAS Cluster Server to do this. GAB is loaded as a kernel module on each cluster node and provides an atomic broadcast mechanism that ensures that all nodes get status update information at the same time.

By leveraging kernel-to-kernel communication capabilities, LLT provides high speed, low latency transport for all information that needs to be exchanged and synchronized between cluster nodes. GAB runs on top of LLT. VERITAS Cluster Server does not use IP as a heartbeat mechanism, but offers two other more reliable options. GAB with LLT, can be configured to act as a heartbeat mechanism, or a GABdisk can be configured as a disk-based heartbeat. The heartbeat must run over redundant connections. These connections can either be two private Ethernet connections between cluster nodes, or one private Ethernet connection and one GABdisk connection. The use of two GABdisks is not a supported configuration since the exchange of cluster status between nodes requires a private Ethernet connection.

For more information about GAB or LLT, or how to configure them in VERITAS Cluster Server configurations, please consult the VERITAS Cluster Server 2.0 User's Guide for Solaris.

Bundled and Enterprise Agents

An agent is a program that is designed to manage the availability of a particular resource or application. When an agent is started, it obtains the necessary configuration information from VCS and then periodically monitors the resource or application and updates VCS with the status. In general, agents are used to bring resources online, take resources offline, or monitor resources and provide four types of services: start, stop, monitor and clean. Start and stop are used to bring resources online or offline, monitor is used to test a particular resource or application for its status, and clean is used in the recovery process.

Western Group (Thailand).,Ltd.
31/92 Rangsit-Klong 7 Rd., Lam luk ka,
Patum-tani, 12150, THAILAND
Tel: +662 909-3691, Mobile: +66(0)8-1908-1052
Fax: +662 909-3691

A variety of bundled agents are included as part of VERITAS Cluster Server and are installed when VERITAS Cluster Server is installed. The bundled agents are VCS processes that manage predefined resource types commonly found in cluster configurations (i.e., IP, mount, process and share), and they help to simplify cluster installation and configuration considerably. There are over 20 bundled agents with VERITAS Cluster Server.

Enterprise agents tend to focus on specific applications such as DB2 UDB. The VCS HA-DB2 Agent can be considered an Enterprise Agent, and it interfaces with VCS through the VCS Agent framework.

VCS Resources, Resource Types and Resource Groups

A resource type is an object definition used to define resources within a VCS cluster that will be monitored. A resource type includes the resource type name and a set of properties associated with that resource that are salient from a high availability point of view. A resource inherits the properties and values of its resource type, and resource names must be unique on a cluster-wide basis.

There are two types of resources: persistent and standard (non-persistent). Persistent resources are resources such as network interface controllers (NICs) that are monitored but are not brought online or taken offline by VCS. Standard resources are those whose online and offline status is controlled by VCS.

The lowest level object that is monitored is a resource, and there are various resource types (i.e., share, mount). Each resource must be configured into a resource group, and VCS will bring all resources in a particular resource group online and offline together. To bring a resource group online or offline, VCS will invoke the start or stop methods for each of the resources in the group. There are two types of resource groups: failover and parallel. A highly available DB2 UDB configuration, regardless of whether it is partitioned or not, will use failover resource groups.

A "primary" or "master" node is a node that can potentially host a resource. A resource group attribute called `systemlist` is used to specify which nodes within a cluster can be primaries for a particular resource group. In a two node cluster, usually both nodes are included in the `systemlist`, but in larger, multi-node clusters that may be hosting several highly available applications there may be a requirement to ensure that certain application services (defined by their resources at the lowest level) can never fail over to certain nodes.

Dependencies can be defined between resource groups, and VERITAS Cluster Server depends on this resource group dependency hierarchy in assessing the impact of various resource failures and in managing recovery. For example, if the resource group ClientApp1 can not be brought online unless the resource group DB2 has already been successfully started, resource group ClientApp1 is considered dependent on resource group DB2.

For detailed information on the implementation and design of highly available IBM DB2 Universal Database environments with the VERITAS Cluster Server see the technote entitled "DB2 UDB and High Availability with VERITAS Cluster Server" which you can view by going to the following web site: http://www.ibm.com/support, and searching for the keyword "1045033".