TECHNOLOGY TRENDS

VERITAS's Diverse Server-Based Replication Technologies Meet Multiple Business Continuity Needs

Harvey Hindin ◆ VP and Lead Analyst, Business Continuity ◆ hhindin@dhbrown.com ◆ August 24, 2004

A full disaster recovery or business continuity solution requires both the data and application to be available at a secondary site. The process of replicating the data ensures it will be available if and when data recovery is required. Data replication, for disaster recovery or business continuity, can be server or storage array based. Often, storage vendors, the traditional suppliers of data replication solutions, will handle data directly and provide replication services through the user's storage resources. In a large number of cases however, it is more advantageous to employ server-based data replication. This *Technology Trends* discusses the operation of server-based data replication and the VERITAS software offerings related to it.

The advantage of server-based replication is that there is no lock-in to a specific storage vendor – any suitable storage will do and the suppliers can be different at both the primary and secondary sites. Server-based replication allows replication to occur between the same storage from a single vendor, between dissimilar storage from the same vendor, or between dissimilar storage from different vendors. Sidebar 1, "Server-Based Replication Advantages," summarizes some of the additional benefits associated with this method. VERITAS has played in the data replication markets from the very beginning and offers a number of software products for server-based replication, which are analyzed in this *Technology Trends*. These products include: the VERITAS Volume Manager (VVM), which can be used for synchronous replication; the VERITAS Volume Replicator (VVR), which provides synchronous and asynchronous replication; VERITAS FlashSnap, which creates "point-in-time" copies; and the VERITAS Storage Replicator, which enables remote office data protection. These offerings provide data replication capabilities for different Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO), enabling planners to engineer systems with a range of recovery options – from zero data loss to fractions of a second to minutes – depending on their environment's needs and budget.

The Role of Volume Manager Software

Software-based volume managers have existed for a long time and there are currently many volume manager products on the market (often associated with operating systems). This software provides an interface to multiple physical disks so that they appear to be one logical and therefore "physical" disk. While there are many advantages to utilizing volume manager software (such as the ability to automatically grow a volume) the ability to construct multiple synchronous copies (mirrors) of the same data is the one of concern here.

In the days of SCSI connection to disks, volume-manager-based synchronous mirrors were limited by SCSI cable lengths, and therefore could only achieve minimal distances. But now in a SAN environment using Fibre Channel, it is possible to achieve these volume-manager-

D.H. Brown Associates, Inc.
222 Grace Church Street
Port Chester, NY 10573

Phone: (914) 937-4302
Fax: (914) 937-2485
www.dhbrown.com

An ideas International Company

Sidebar 1: Server-Based Replication Advantages

- With server-based replication, not only can the storage vendors differ at the primary and secondary sites, it is also not necessary to have the same physical storage layout or configuration. All that is required is that there be enough storage at the secondary site to accept replication.
- With server-based replication, distances between primary and secondary sites can be almost unlimited (depending on the available bandwidth) since IP networks are used for data transmission. The use of IP networks over long distances for storage array-based replication requires Fibre Channel (the native connectivity in storage arrays) to IP conversion gear (either externally or built into the array).
- Storage array-based asynchronous replication can have data integrity issues under certain conditions as there is no knowledge of the "write" activity at the application level – consequently "writes" cannot be readily logged and tracked. While there are workarounds for this (e.g., periodic synchronization, clever firmware in the storage that can properly reorder the data at the secondary site), data integrity questions may persist. Additionally, periodic synchronization has its own consequences (e.g., stale data, the need for a lot of extra storage at the secondary site, data consistency) while server-based replication does not suffer from the same data integrity concerns.
- Server-based synchronous replication may be faster than storage array-based replication in many instances because of shorter data paths.
- Finally, storage-based asynchronous replication can require more storage at the secondary site due to the need to stage/store accumulated data and guarantee its consistency by having to create multiple copies of data and use up valuable storage space).

based mirrors over much greater distances. The distance between the separate mirrors is limited only by the length of the fibre infrastructure (often tens of kilometers, and even this distance can be extended by special Fibre Channel SAN hardware).

As a result of SAN advances, it is possible to use volume manager-based mirroring as the replication component of a business continuity or disaster recovery solution for environments that need to achieve zero data loss and have Fibre Channel connectivity between geographically dispersed environments.

Inside VERITAS Volume Manager (VVM)

The VERITAS Volume Manager, a core component of the VERITAS Storage Foundations package, is useful for synchronous mirrors over a Fibre Channel SAN with storage spanning less than 100 kilometers (Metropolitan Area Network). VVM can provide data availability for this environment (cluster technologies should be employed for application availability). Note that it is the SAN infrastructure "radius" – not the VVM software – that limits the achievable distance. It takes a finite amount of time for data to get from one point in the SAN to another. The application's performance and VVM's establishment of a remote synchronous mirror cannot operate any faster than is permitted by this distance constraint.

The net here is that within a Fibre Channel SAN radius, it may be far less expensive, and equally effective, to employ a simple server-based solution, using a product like VERITAS' VVM, to achieve a synchronous mirror at a secondary site rather than to use a storage array-based solution. Infrastructure managers must investigate whether or not this is the case when an infrastructure design or redesign is required.

Inside VERITAS Volume Replicator (VVR)

Simple SANs are generally not able (or available) to replicate data over distances more than 100 kilometers, and even extended SANs have distance limitations. When it is necessary to replicate data over such distances, WAN connectivity must be used between the primary and secondary sites over the IP protocol. VERITAS' Volume Replicator is suitable for replication over an IP network in these extended environments and can provide either synchronous or asynchronous replication between a total of thirty-two locations (for one-to-one, many-to-one, and one-to-many replication topologies). VVR natively works over IP and can be used even for short distances where no SAN is available.

VVR can be thought of as an IP extension to VVM that provides the ability to expand VVM distances. VVR is an option to VERITAS' VVM and is turned on by a license key. Another VVR benefit is that it requires little more additional technical expertise beyond understanding VERITAS' VVM to manage the replication environment.[1] VVR offers the ability to replicate data in synchronous or asynchronous mode. All system volume management capabilities that pertain to VVM also apply to the secondary system's replicated volumes. No management capabilities are lost.

Synchronous Replication with Volume Replicator

In synchronous replication, all of the changes incurred at the primary site must be propagated to and made at the secondary site in order for the application to acknowledge that changes have occurred. This acknowledgement ensures that there is no data loss between the primary and the secondary sites. The acknowledgment also helps ensure that there is no difference between the primary and secondary data.

The latter statement also means that synchronous replication may be thought of as supplying the ultimate Recovery Point Objective – zero data loss. However, there is a potential price associated with server-based synchronous replication: The application operation can incur a penalty due to the time required for the trip from the primary to the secondary site and back. Depending on the application and its surrounding environment, this penalty may, or may not, be of significance. It should be noted that this penalty will be incurred with any replication technology attempting to replicate data in synchronous mode.

Determining the suitability of synchronous replication (i.e., its application impact) for a given application and environment is a non-trivial task. A variety of both simple and complex tools are available to help in this regard from VERITAS, consulting firms, outsourcers, and third parties. The tools available from VERITAS include the free VRAdvisor, which collects data change rate information to both help determine the

required network bandwidth size and to configure VVR's Storage Replicator Log (SRL). After the VRAdvisor has accumulated sufficient typical network "write" data, the system administrator can model the network to see how much data would be lost in an outage time period based on a number of parameters set by the client. Typically, for write-intensive, or very long distance replication scenarios, synchronous replication is not optimal as there is a direct impact on the application.

Asynchronous Replication with VVR

Asynchronous replication is designed for customers who can tolerate some possible data loss but cannot allow application performance to be slowed by the synchronous replication technology's network round-trip time. When using VERITAS's Volume Replicator in asynchronous mode there is no application performance penalty.

When VVR is used in asynchronous replication mode, the first step in the operation requires the application to issue a "write request" to VVR. This write request is recorded on the VVR log (SRL) and committed to disk. VVR then acknowledges the write request to the application.

A write request is then sent to the secondary site server (the network may or may not transmit this immediately). The secondary server acknowledges the write request to the primary server (and uses its SRL for record keeping). VVR then issues a command to generate a write request to the secondary storage. The "write" is then committed at the secondary server. When VERITAS Volume Replicator is used in its asynchronous mode, there is no application impact. In other words, the "application write" performance is not affected by the network latency and data is immediately available at the secondary site. Network outages have a minimal effect since the writes can be batched at the primary site should a failure occur.

What's the CPU Cost?

According to VERITAS, VVM "typically" consumes 3% to 5% of the server CPU cycles. Moreover, VERITAS says that VVR only "typically" adds an incremental 2% over VVM for a total of 5% to 7% .

VERITAS also states that its tests show that VVM and VVR can exceed the performance of storage array-based replication (such as EMC's SRDF) by as much as 30% to 72%[2] This is because VVM and VVR "live" in the host and can perform "writes" at the same time to both primary and secondary locations without the bottleneck of going through storage arrays and additional data paths.

Three Components

VERITAS' VVR adds[3] three key components to the VVM code base. The first of these, Rlinks (Replication Links), can establish up to thirty-two native IP connections; at least one Rlink is required at each connected host. Each Rlink is configured for either synchronous or asynchronous replication as required by the environment in question. When more than one

Sidebar 2: Secondary Site Tape Initialization

When a secondary site is being set up, process replication and initialization has to occur with an existing and functioning primary site. The data that is already at the primary site must be sent to the secondary site before replication can begin. There are several options that are available to initialize the secondary site. Clearly this data can be sent over the network. However, this method can be very expensive depending on how much data there is. Another solution is to mirror the data onto a new storage array and physically ship the storage array to the second site. Again, this can be a costly endeavor due to shipping prices.

VERITAS tape-based initialization offers a third option for initializing a secondary site. This feature is part of VERITAS Volume Replicator. With tape-based initialization a tape backup is done at the primary site with a checkpoint inserted in the primary SRL. The tape is then sent to the secondary location. After the tape is loaded at the secondary site, only the changes that occurred after the time the primary checkpoint was inserted in the primary SRL are sent over the network.

Rlink is set up it is possible, for example, to replicate from site A to site B synchronously and, at the same time, replicate from site A to site C asynchronously.

The second VVR component, Replicated Volume Group (RVG), allows system administrators to choose the data they want to replicate (based on the volumes spanned by the data). RVG also groups volumes together so the correct write order is preserved.[4] The utility of being able to choose the data to be replicated (for example, it may be desirable to replicate some data, and just backup other data on tape) is critical.

The final Volume Replicator component, Storage Replicator Log (SRL), helps ensure data consistency between the primary and the secondary sites during replication (for both synchronous and asynchronous modes), by tracking data changes and guaranteeing that the correct write order is preserved. The SRL software, (which lives in the administrator-selected RVG), creates the logs that provide a record of the data write requests that have taken place at the primary and secondary sites.

The SRL serves other purposes as well. For example, it assists in the process of initializing a secondary site. Sidebar 2 provides an overview of this technique. It also can assist in the timing of the use of the VERITAS FlashSnap product (discussed later in this document) by scheduling when point-in-time copies of data are taken.

As stated earlier, the storage at the primary and secondary sites do not need to be from the same manufacturer nor do they need to be the same model (as is the case for storage array-based replication). But, the servers at the primary and secondary sites must be using the same operating system even though they can be different types of servers.

Protection against Network Outages

VERITAS' VVR can also protect the primary site from a network outage. This capability is useful should the network between the primary and secondary sites go down. The writes to

the secondary site are "batched" and held at the primary site until the network comes back online.

If replication is occurring in synchronous mode and the network fails, VVR can switch into asynchronous mode and begin tracking the writes at the primary site. As soon as the network comes back online the log is drained and then the replication is automatically switched back into synchronous mode. If replication is occurring in asynchronous mode, then the secondary site is caught up as much as possible.

This procedure can be set up so that no user intervention is required. Since WANs can fail, this could be a critical feature to planners. With this capability a redundant network link may not be necessary, depending on the applications and the environment.

Utilizing the Secondary Site for Site Maintenance Purposes

There is a further advantage to VERITAS's VVR, particularly when compared to asynchronous storage array-based solutions. One function of a secondary site is to use it as a backup to the primary site when the primary undergoes site maintenance. Operations are failed over to the secondary site and the primary maintenance is accomplished. When the maintenance is complete, failback to the primary only requires that "changes" be sent over the network. A complete initialization is not needed.

Using Data at the Secondary Site with VERITAS FlashSnap

The VERITAS FlashSnap[5] offering found in the VERITAS Storage Foundations allows secondary site data to be used for a variety of real-world purposes. Fundamentally, FlashSnap allows point-in-time snapshots of the secondary site's data. These snapshots can be used for any purpose that requires data access at the secondary site where the data does not have to be totally up to date (i.e, backup or the off-host processing used in decision support systems/data mining).

A recent modification to VERITAS FlashSnap, available in VERITAS VVR 4.0, enables "space-optimized" snapshots. With this technology VERITAS stores only data changes from the last snapshot in some of the secondary site's storage, creating far less of a storage-provisioning burden. Typically, this optimized technique only consumes an additional 10% of storage space versus traditional alternatives, which consume 100% .

Remote Office Data Protection with VERITAS Storage Replicator

A final technology offered by VERITAS is the VERITAS Storage Replicator. Storage Replicator is a server-based replication technology available on the Windows platform and is typically used for remote office data protection as an extension to the backup processes. Storage Replicator eliminates the need for remote office hardware and administrators by replicating the remote office data to a central location and enabling all backups to be performed at the central site. This functionality allows organizations to ensure that the

remote office data is protected, enabling backup administrators to remain at a central location.

Put It All Together

There are many user installations that employ the VERITAS software discussed in this *Technology Trends*. One VERITAS client is providing application availability locally using VERITAS Cluster Server (VCS). VERITAS Volume Replicator then replicates the data to the secondary site and disaster recovery management is provided by the Global Cluster option through a T-3 WAN over a distance of 888 miles.

The applications, which fail over locally, include Siebel on DB2 and SAP on Oracle (both running on IBM hardware), as well as imaging and mail applications on DB2 running on Sun hardware. All of these can also fail over to secondary sites running IBM and Sun hardware respectively through the Global Cluster option. Data is sent asynchronously using the VERITAS Volume Replicator with the secondary site keeping copies of all the primary site volumes.

The solution described replaces three different clustering environments and an outsourced secondary datacenter. It achieves an RPO of milliseconds and an RTO of less than forty-five minutes for the applications and databases indicated.

The DHBA Bottom Line

When synchronous replication is of concern, particular effort must be expended to carefully determine just what performance impact the particular application in question will bear as a result of the latency of network connections. Should this latency be acceptable, server-based synchronous replication needs to be on the infrastructure manager's short list of techniques to consider for disaster recovery since it can provide a fast-performing solution compared to storage-array replication. Server-based replication also provides asynchronous replication without affecting the primary applications. The traditional thinking that storage arrays are the only solutions is no longer valid.

In many applications and environments, server-based replication, either synchronous or asynchronous, can compare favorably with storage array-based replication. Moreover, server-based replication can offer many distinct advantages, as outlined in this *Technology Trends*. VERITAS offers a complete product suite in this area. Additionally, VERITAS' solutions are operating system-independent,[6] use clustering technology for application failover and a comprehensive disaster recovery solution that ensures that data and applications remain highly available.

[1] For example, just as the VERITAS Enterprise Architecture (VEA) GUI manages the VVM, the same GUI also manages VVR. VVR also has a CLI and a web-based GUI.

[2] http://eval.veritas.com/mktginfo/products/White_Papers/High_Availability/giantloop_replication_performance.pdf.

[3] VVM is available on AIX, HP-UX, Linux, Solaris, and Windows as is VVR (with the exception of Linux, which will enjoy VVR capability in Q3 2004.)

[4] Oracle 9i RAC and Oracle 10g data usually spans volumes. RVG is used here to group Oracle requests together so that they are consistent at the secondary site. VVR is Oracle-certified in both its synchronous and asynchronous modes. DB2 and Sybase are also VVR-supported, however neither IBM nor Sybase have VVR certification programs.

[5] Analogous to EMC TimeFinder, but server-based.

[6] VERITAS cluster failover software can accommodate AIX, HP-UX, Linux, Solaris, and Windows.