HITACHI
DATA SYSTEMS

Enterprise Storage

Disaster Recovery Issues and Solutions

A White Paper

By Roselinda R. Schulman

September 2004

Executive Summary

Well-planned business continuity and disaster recovery solutions are critical to organizations operating in 24/7 environments. Minimal to zero disruption is the goal of today's enterprises when faced with planned or unplanned outages. In addition to these business needs, recent government recommendations are driving businesses to look at out-of-region replication options, to enable them to recover from larger-scale events.

This paper considers technological approaches to meeting disaster recovery needs, and emerging configuration alternatives and technologies that provide cost-effective solutions for disaster recovery and business resilience.

This paper does not address traditional tape backup technology or disk-based backup options, but focuses on remote replication approaches. It familiarizes readers with the vocabulary of software copy alternatives and defines currently available remote copy techniques, including network options for replication. It reviews a variety of remote copy technologies that are widely used for in-region and out-of-region replication, offered by Hitachi and other major storage suppliers. And it outlines cost-effective replication approaches—including two data center and three data center configurations—that meet a wide range of business needs.

In each area of replication technology, this paper shows that Hitachi Data Systems provides customers with a wide range of flexibility and choice in technologies and solutions for both open systems and mainframe environments.

Contents

Disaster Recovery Issues and Solutions

A White Paper

By Roselinda R. Schulman

Introduction

The world has changed significantly in the past few years. Devastating terrorist acts and threats, the seemingly increased frequency of widespread power-grid disruptions, and the emergence of regulatory requirements for infrastructure protection are all placing stringent, yet necessary, data protection requirements on many organizations. Regardless of the industry, as more and more businesses operate in a 24/7 environment—especially large enterprises where global operations are the norm—they need an increasingly competitive edge to maintain profitability and stay in business.

In the complex and challenging global environment, well-planned business continuity or proven disaster recovery practices for nonstop data availability have become critical to organizations if they are to survive any type of outage.

Most information technology-related disasters are actually logical disasters, such as data corruption, viruses, and human error, as opposed to physical disasters like fire, earthquakes, hurricanes, etc. Logical disasters occur all the time and pose a bigger threat to businesses. However, because they're less visible to the general public, logical disasters tend to be taken less seriously.

The real challenge lies in your organization's ability to think proactively and deploy best practices and technologies that can be leveraged to maximize business operations instead of adopting a reactive "fix-it" posture. The true test lies in the ability to prevent outages from occurring in the first place, and minimizing the effects of those incidents when they do occur. Companies today must follow the continuous business paradigm, which combines high-availability solutions with advanced disaster recovery techniques. The ultimate goal is to be able to manage both planned and unplanned situations with minimal or zero disruption.

When an unplanned event does occur, the ideal scenario is:

- Recovery will happen almost automatically with no loss of data
- Costs of the solution and resources are minimal
- Impact to the production environment is zero

While technology is moving forward at a rapid pace to reach this ideal scenario, many other business and technology concerns exist, including some significant trade-offs dictated by technology, budgets, and personnel resources.

Recovery Objectives: All Data Is Not Created Equal

Recovery time objective (RTO) and recovery point objective (RPO), along with their associated costs, are important criteria when evaluating the right solution.

- **RTO** describes the time frame in which business functions or applications must be restored (includes time before disaster is declared and time to perform tasks).
- **RPO** describes the point in time to which data must be restored to successfully resume processing (often thought of as time between last backup and when outage occurred).

Many solutions are available depending on your organization's recovery objectives. For example, when looking at your RPO, you may be concerned with the cost of some data loss (typically less than five minutes, depending on the replication methodology). You may prefer having the ability to quickly perform a database restart instead of a no-data-loss option. Or, you may prioritize limiting possible impacts to the production environment and ensuring easy recovery (minimizing rolling disaster and database corruption) at the secondary site.

Consider the different tiers of recovery available to you (see Figure 1). While Hitachi Data Systems focuses on the higher recovery tiers in terms of its software and hardware replication offerings, we understand that not all data is created equal when it comes to disaster recovery protection (see Figure 2).

Remember that these trade-offs are business- and application-driven. A thorough business impact analysis is a good starting point to determine the best course of action for data protection and business continuity.

Figure 1. Available Tiers of Recovery

Technology Tier	RPO Range[1]	RTO Range[1]	Minimum # Disk Copies	Distance	Regional Disaster Support
Tier 1—Tape Backup	24–168 hours	48–168 hours	N/A	Any	Maybe[2]
Tier 2—Disk Point-in-time copies	4–36 hours	4–24 hours	3[3]	Any	Maybe[2]
Tier 3a—Sync	0–2 minutes	1–8 hours	2[3]	Limited	No
Tier 3b—Sync w/failover	0–2 minutes	5–60 minutes	2[3]	Limited	No
Tier 4a—Async	0–5 minutes[4]	1–8 hours	2[3]	Any	Yes
Tier 4b—Async w/failover	0–5 minutes[4]	30–90 minutes	2[3]	Any	Yes
Tier 5—Three data center	0–2 minutes	1–8 hours	3-7 [3,5]	Any	Yes

Note 1 The RPO and RTO ranges in this table are Hitachi Data Systems estimates based on customer experience in real-world enterprise application environments.

Note 2 If the tapes or disk backups are relocated to an out-of-region recovery site

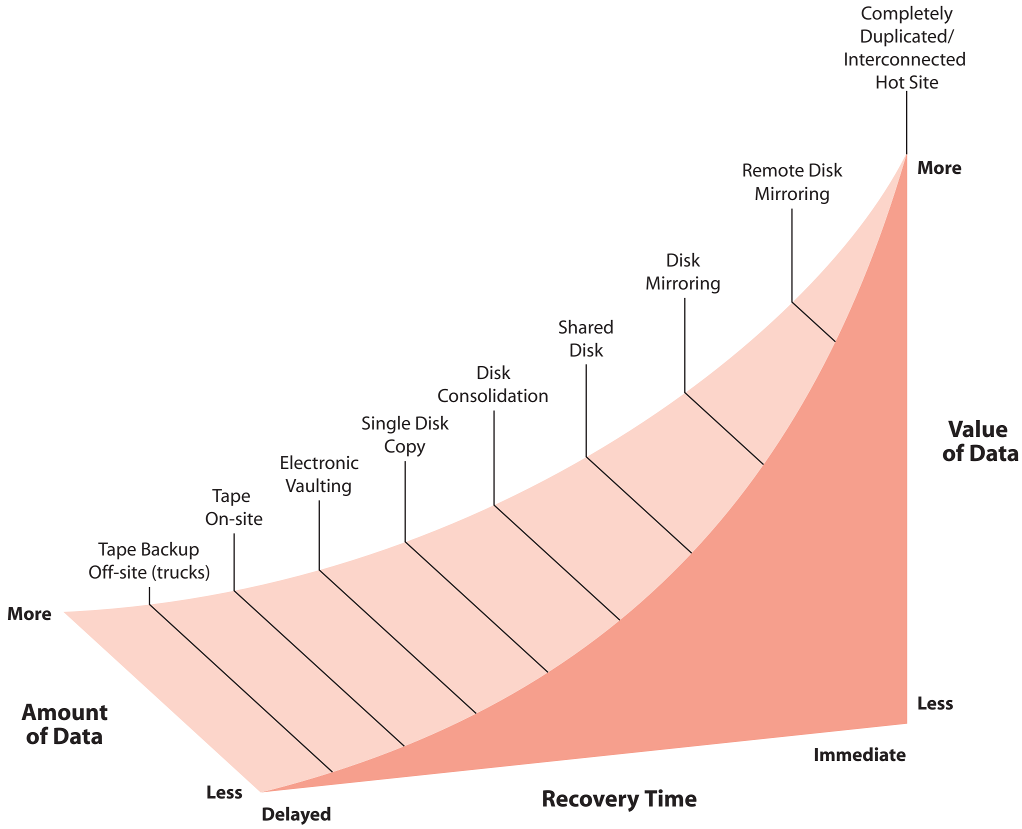
Note 3 Best practice is one additional copy for doing DR testing without impacting the ongoing replication session

Note 4 Network problems will extend the RPO

Note 5 Depends on vendor and method deployed

Many organizations still fall into Tier 1, although some of them will also use techniques such as remote logging of data. Typically, the higher the tier, the greater the cost. You can, however, achieve significant improvements in data currency and recovery time with the higher tiers.

Figure 2. Data Types and Disaster Recovery



Different types of data require different levels of protection. A data audit is required to assess business criticality and cost to recover.

First, we should consider your environment. As mentioned earlier, all data is not created equal. It is likely that only a portion of a corporation's data is critical to its basic operation and that a variety of techniques could be used to secure that data, depending on the criticality of that particular business function.

Depending on how intertwined the data and applications are and the degree to which they are segregated, this can become a big undertaking. Many organizations choose not to take this approach, but rather choose to copy everything.

This is a trade-off in terms of the cost of potentially having to re-engineer your environment compared to the cost of using a higher tier for all data.

For companies that have a fairly local site for replication and significant bandwidth available, the option to copy everything is very attractive. However, this can change as protection from both local and regional disasters becomes necessary. While evolving network technologies help lower the cost of replicating data over significant distances, the price can still be extremely high. Therefore, it may be necessary to prioritize data or use a less current copy of data (for example, a copy that is four hours old) for the second copy. However, newer approaches such as Hitachi Universal Replicator for TagmaStore™ Universal Storage Platform can help by enabling enterprises to support replication needs without over-provisioning bandwidth based on peak-load traffic.

Before we start focusing on the technology, it is important to understand some of the basic terms we use when discussing remote-copy alternatives. Advances in technology have brought new words and phrases, such as "real time," "point in time" (PiT), and "snapshot" into the language of enterprise-class storage. Copy products are designed to allow an enterprise to replicate, protect, and share data in dynamic new ways. Some of the terms used in copy technology are:

- **Remote copy—**This refers to the mirroring of data, typically in real time, to provide an I/O-consistent remote copy of that data. The purpose of remote copy is to protect the data in the event of a business interruption at a production location.
- **Point-in-time (PiT) copy—**PiT copy refers to a copy of data that is taken at a specific point in time. Ideally, this copy should be I/O-consistent. PiT copies are used in many ways, including backups and checkpoints. More recently, PiT copies have been used in architected disaster recovery solutions.
- **Data duplication—**This software duplicates data, as in remote copy or PiT snapshots. Data duplication differs from data migration in that with data duplication there are two copies of data at the end of the process, while with data migration there is only one.
- **Data migration—**This software migrates data from one storage device to another. Data migration differs from data duplication in that at the end of the process there is only one copy of data. One purpose of data migration is to reduce operational complexity and costs for storage system upgrades or equipment refurbishment.
- **Synchronous replication—**Requires the application to wait for the remote site to confirm each write operation before sending the next write operation. As the replication distance increases, the time lag between synchronous write operations gets longer, and these delays become intolerable for high-volume, write-intensive transaction processing applications.
- **Asynchronous replication—**Allows the production application to continue, and keeps track of the pending writes until they are complete. A well-designed asynchronous replication solution maintains an I/O-consistent copy at the remote site. It may deliver a somewhat lengthened RPO and RTO, but it does protect the enterprise from major data loss in case of a regional disaster. For further discussion, see the resources in Appendix A.

Over the last few years, many significant new technologies in both the software and hardware arenas have come to market. These technologies can reduce time-to-business resumption from days to hours and shorten the downtime required for backups to near zero.

When evaluating alternative copy technologies, there are some important points to consider. One is the consistency or integrity of the copy. While replicating data may sound simple in practice, the ability to recover from that copy of the data can be extremely complex. This depends not only on the technology but also on the processes that were employed. We will discuss this in further detail as we look at available technologies with a focus on Hitachi Data Systems offerings. But first, we look at the impact of rolling disasters on disaster recovery strategies and technologies.

Rolling Disasters

Real-time copy products are designed to maintain a duplicate image of data at a remote location so that if the primary location is lost due to a disaster, processing can continue at the second site. Although the concept of replicating updates sounds simple, surviving a disaster is actually extraordinarily difficult and challenging. To address this, three basic disaster recovery requirements should be satisfied by any disaster recovery solution:

- Surviving a rolling disaster
- Preserving write sequencing
- Emergency restart capability following a disaster

Two discrete points in time define any catastrophic disaster: when the disaster first strikes (the beginning), and when the disaster finally completes (the end). Many seconds or even minutes may follow the beginning of a disaster. The period of time between these two events is the "rolling disaster

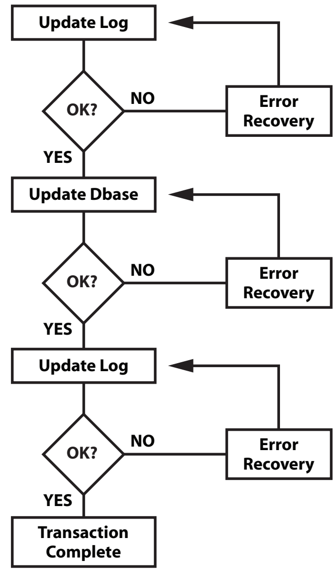
Figure 3. Rolling Disaster Window

Surviving a rolling disaster is the true test of any disaster recovery solution, because data corruption might occur within this rolling disaster window.

The real objective of any disaster recovery solution is to provide the capability to produce an image or I/O-consistent copy of data at the secondary location, as it existed at a point in time prior to the beginning of the disaster. This can be likened to the state in which data exists after a server or system crash. If update activity during the rolling disaster is also shadowed to the backup site, the backup copy may also be corrupted as the write order cannot always be preserved during this time. We know that the image of the shadowed data is usable at any point prior to the disaster occurring, but the image may not be immediately usable if the potentially corrupted updates that occur during the rolling disaster are copied. In a rolling disaster, the data image may be corrupted due to write sequencing and write dependency. Write sequencing is the notion that the order or sequence of updates to the primary data structure must be maintained to insure the integrity of the data (see Figure 4).

Figure 4. Crash Recovery Example

- :: Many applications control update sequence
- :: Updates applied out of sequence violate data integrity
- :: One transaction may involve updates to multiple volumes
- :: Databases have been carefully designed to protect integrity
- :: A DBMS has no knowledge of remote copy
- :: A DBMS does not control remote copy updates



The sequence in which a database and a log are updated allows the database management system (DBMS) to instantly recover the database, with data integrity, following any sudden outage.

That means a remote copy solution must be able to replicate the original sequence of updates; failure to do so will result in corrupted data at the backup site. Write dependency implies that there is a logical relationship among a series of updates, and if there is a particular update failure, the sequence of subsequent updates might change. The application controls this write sequence/dependency, but the application has no knowledge of the remote copy.

There are different ways to preserve write dependency, and vendors have chosen various approaches in their remote copy products. Hitachi Data Systems believes that customers should look to proven technologies. These include true synchronous remote copy products with appropriate controls, such as freeze functions. A number of asynchronous replication approaches also satisfactorily solve the write-sequencing problem, including Hitachi Universal Replicator, Hitachi TrueCopy™ Asynchronous Remote Replication, Compatible Replication for IBM® XRC (formerly known as Hitachi Extended Remote Copy, or HXRC), IBM XRC, IBM GDPS, IBM PPRC Global Mirror, EMC SRDF Asynchronous, and Hitachi NanoCopy™ technology.

For applications that span multiple volumes, including many production databases, the remote copy technology must also maintain consistency across all related volumes. This can be a challenge in a rolling disaster, in which different storage systems or communication links can cease transmitting remote copy updates at different points in time. Remote copy vendors employ a number of different approaches to define consistency groups and to maintain consistent remote copy status across all the volumes in a consistency group.

Network Options for Replication

Before we take a look at the various remote copy technologies, we will consider some of the network options available.

ESCON-based Options

Hardware-based replication techniques had their origins in the S/390® world, and therefore used ESCON channels as the transport mechanism for links between the systems. Today, most customers employ Fibre Channel to transport data between systems. This is available for both open systems and mainframe environments and provides some additional benefits over ESCON-based copy.

Fibre Channel-based Options

Fibre Channel can be used in a manner similar to ESCON or it can be directly connected using dark fibre and dense wavelength division multiplexing (DWDM) technology. In this environment, we may see a lower total cost of ownership (TCO) due to improved performance and the need for fewer links. Other benefits accrue if Fibre Channel storage area network (SAN) management expertise is already in place: the customer doesn't have to manage yet another type of interconnection, and this network option can fit into the existing SAN infrastructure environment.

When looking at Fibre Channel for extended distance copy, there are several options including replication over IP networks connecting into traditional networks, such as point-to-point OC3s. Newer extender technologies also allow you to connect directly through synchronous optical network technologies (SONET) networks with potentially lower overhead. When comparing the different options, you have to consider cost, ease of procurement, and the ability to control the available pipe and quality of service (QoS), so as not to compromise the real-time application by not having enough bandwidth. However, IP replication does seem to offer the ability to cover extreme distances at a lower cost.

Hitachi Data Systems and other leading IT and storage vendors were invited to participate in a transcontinental IP storage demonstration named the Promontory Project. Hitachi storage systems were connected to Fibre Channel SANs in California and New Jersey that were bridged (with gigabit IP "pipes") with IP storage switches from Nishan Systems (now part of McDATA). This demonstration was undertaken to show that distance and speed are no longer barriers to mass storage architectures and solutions. Storage locations can now be implemented far enough away from primary centers so as not to share common disaster areas. Hitachi TrueCopy™ Remote Replication software over Fibre Channel was tested over the gigabit IP links; basic features of the software were run without modifications over the IP backbone.

While the Promontory Project demonstration involved Nishan Systems, other vendors such as Ciena and CNT are currently qualified for replication over IP networks and over more traditional point-to-point protocols with Fibre Channel. Many other vendors are looking to get into this market as well as other emerging network replication protocols, such as iSCSI.

Currently Available Remote Copy Techniques

Organizations have implemented a variety of replication approaches based on server, storage, and network capabilities. Each has its strengths and weaknesses, but storage-based replication is generally considered the best approach for replication of critical data, particularly in heterogeneous server and application environments. This white paper focuses primarily on storage-based approaches, and considers a number of approaches that can be used alone or in combination to meet enterprise disaster recovery requirements:

Synchronous remote copy

- Hitachi TrueCopy Remote Replication and similar approaches from EMC and IBM

Asynchronous remote copy

- IBM XRC and Hitachi Compatible Replication for IBM® XRC (for z/OS® and OS/390® environments only)
- Hitachi TrueCopy Asynchronous
- Hitachi Universal Replicator for TagmaStore Universal Storage Platform
- EMC SRDF Asynchronous

PiT-mediated remote copy (using point-in-time copies for consistency)

- Hitachi NanoCopy technology
- IBM PPRC Global Mirror
- Cascade copy—typically EMC and IBM using SRDF Adaptive or PPRC-XD

All of these offerings—except IBM XRC and Hitachi Compatible Replication for IBM XRC—are available for UNIX, Microsoft Windows, and z/OS and OS/390 operating systems.

Implementations differ widely based on customer requirements, and no single alternative will satisfy every customer's objectives. In many circumstances there is little choice; for example, if a customer's backup location is hundreds of kilometers away, then only asynchronous or PiT copy technologies may be practical. It is also worth noting that not all options deserve equal merit. The concepts of write dependency and write sequencing are not addressed in all remote copy techniques, notably PPRC-XD (Peer-to-Peer Remote Copy) and adaptive copy. Other techniques, such as multi-hop, cascade, IBM PPRC Global Mirror, and Hitachi NanoCopy, address the issue by creating consistent PiT copies on a fairly regular basis, although they all have different characteristics and cost structures.

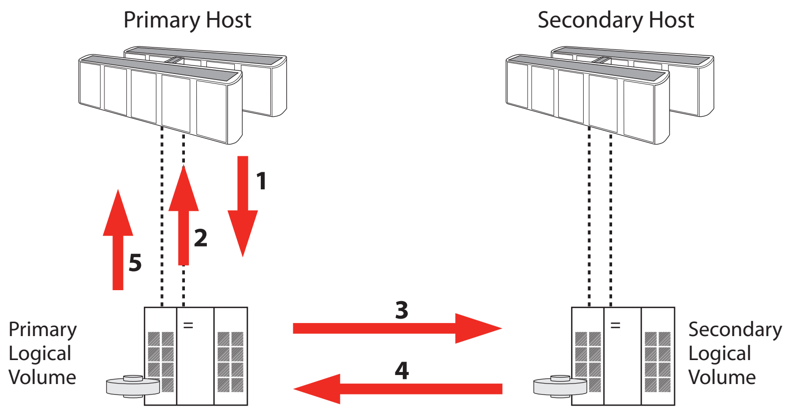
Hitachi Data Systems has the widest range of offerings of any storage vendor, and, our competitors' protests to the contrary, this does not make things confusing for the customer. Hitachi has taken the initiative of providing both IBM-compatible and proprietary technologies to give customers the most flexibility in choosing offerings. We use the concept of building blocks for all our software to architect the right solution, based on the customer's goals and objectives, rather than trying to fit a square peg solution into a round hole. For example, our flexibility is evident in long-distance replication. Depending on the requirements, we can offer a solution with which a customer could replicate directly to a site halfway around the world, or, for the ultimate protection, we can offer advanced three data center solutions. Other vendors can only offer cascade-style solutions, as they cannot support direct copy for write-dependent applications over distance.

Synchronous Remote Copy

Hitachi TrueCopy Remote Replication software is one of three hardware-based, synchronous remote copy solutions from major storage vendors. The other synchronous solutions include PPRC from IBM and SRDF (Symmetrix Remote Data Facility) Synchronous from EMC.

In these solutions, an update to the primary data is not allowed to complete to the application until it has also been successfully secured at the secondary location. It is widely but erroneously believed that with synchronous copy you are protected against a rolling disaster. In fact, if the update does not complete successfully at both locations, the action taken is dependent on how the environment has been set up and what controls are in place. Without the correct procedures, the production environment can be impacted and/or the data at the secondary location can be corrupted (see Figure 5).

Figure 5. Synchronous Replication



1. Write to primary logical volume
2. Disconnect from path—path free
3. Write to secondary logical volume
4. Write complete on secondary logical volume
5. I/O completion—application posted

The remote link is storage controller to controller. Remote copy activity is serverless and remote copy is at the LUN/volume level. Issues with this solution include performance, distance, and multiple controller coordination.

Performance is another consideration with synchronous copy, and we often get asked what the distance limitation is. The answer is that it depends on the performance sensitivity of your applications. While newer technologies, such as Fibre Channel and DWDM, provide improvements, you still cannot exceed the speed of light. It takes about 1ms for light to travel 199.64km (124 miles); both ESCON and Fibre Channel protocols support multiple round trips when used for data replication. This means that even at short distances there will be overhead when using synchronous remote copy. Having said this, many satisfied customers are using our synchronous technology very successfully. One other benefit of using Hitachi TrueCopy Remote Replication software is our ability to fully support IBM GDPS (an IBM service offering) for system failover, workload balancing, and data mirroring on systems spread across two or more sites up to 40km (25 miles) apart. While this may not be of interest to you today, it is an evolving technology that could be an option at a future time. Other vendors also support GDPS, but nearly all of the production GDPS sites around the world are currently using Hitachi storage systems.

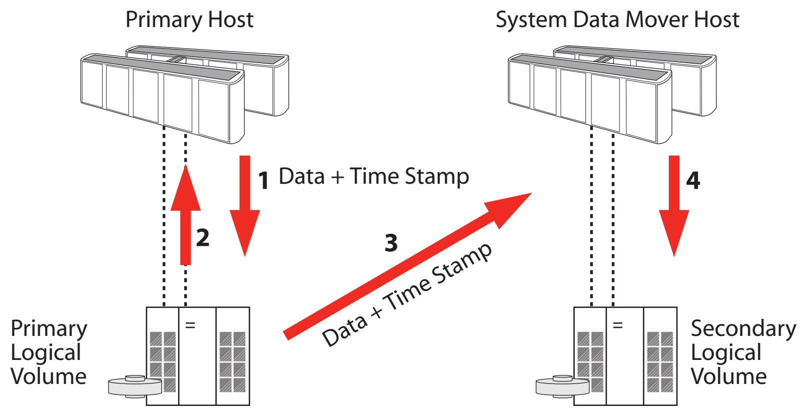
Asynchronous Remote Copy

Examples in this category include several remote copy technologies: Hitachi Compatible Replication for IBM XRC, Hitachi TrueCopy Asynchronous, Hitachi Universal Replicator for TagmaStore Universal Storage Platform, and EMC SRDF Asynchronous.

Hitachi Compatible Replication for IBM XRC

This offering from Hitachi Data Systems is based on the same architecture as IBM's XRC. Its strength is in its ability to survive a rolling disaster. This asynchronous remote copy technique is designed to provide complete data integrity for primary and remote systems. This is accomplished through the use of time stamps that allow the asynchronous updates to be applied consistently by the system data mover (SDM) host software. Time stamping and SDM algorithms are the heart of XRC technology. While it is often considered a software offering due to the requirement of host software, the hardware has to be able to interface with the software and is therefore not supported by some vendors (see Figure 6).

Figure 6. Hitachi Compatible Replication for IBM® XRC



This option provides great data integrity, long distance capability, and time stamps on all updates. Issues with this option include the requirement of host MIPs and a server at the secondary site, and cost (software license at secondary site).

This solution does require a secondary site server and software, but in order to recover quickly from any disaster, that is a requirement anyway. Recent enhancements have improved performance, reliability, and scalability, as well as adding unplanned outage support. EMC has recently added support for XRC, although it is believed the initial version is a very early one and may not incorporate the recent enhancements.

Hitachi TrueCopy Asynchronous Remote Replication

This technology is unique to Hitachi Data Systems and was developed based on a requirement to achieve consistent disaster recovery solutions in both open and mainframe environments—without host software and the need for intermediate systems. Using this technology, our customers can be confident that they can recover mission-critical business functions at facilities across the world, without the encumbrance and overhead that is inherent in other vendors' solutions.

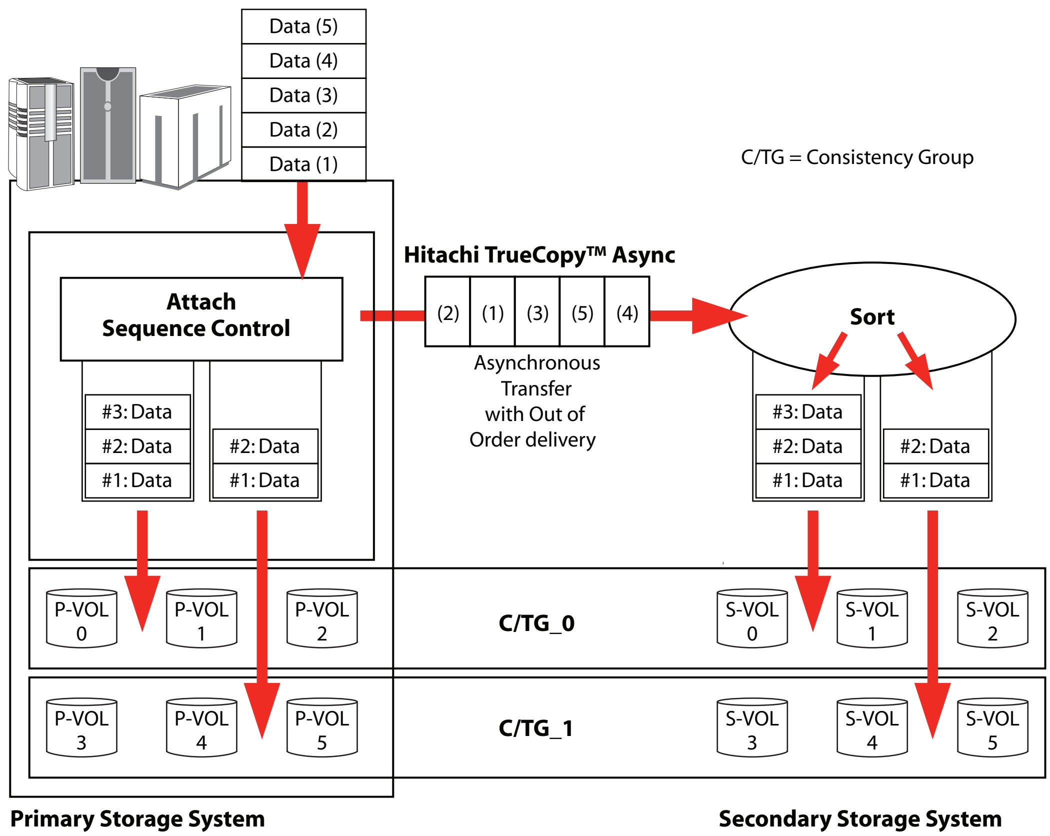
In a mainframe environment, Hitachi TrueCopy Asynchronous software utilizes reliable time stamps, along with other information created by the primary system(s), to enable Hitachi storage systems to transmit updates directly to the secondary system without any host intervention. These writes are then buffered in the secondary system's cache, queued, sorted by the time stamp, and written to the correspondent volumes in the same sequence. These rewrites are issued by the primary system over the remote link, checking sequence numbers embedded in records to ensure no records are missing. In this way, I/O consistency is maintained.

In open systems, TrueCopy Asynchronous uses sequence information to achieve the same level of consistency, though currently an application must reside on a single system. In contrast, an application in the S/390 or z/OS environment can be spread across up to four primary controllers. Hitachi Data Systems is working on eliminating some of these restrictions in the future.

This approach allows the enterprise to achieve data integrity in a significantly simpler and less costly hardware-based solution, with no impact on server or application performance.

TrueCopy Asynchronous also supports consistency groups, which enable you to logically group applications together for consistency purposes (see Figure 7).

Figure 7. Hitachi TrueCopy™ Asynchronous Remote Replication: Write Sequencing and Consistency Groups



TrueCopy Asynchronous uses write sequencing and consistency groups to ensure data integrity and allow users to perform operations on single applications.

This capability allows you to execute operations on single applications, for example, during disaster recovery testing.

In areas where customers require long-distance replication to ensure business continuity or for highly performance-sensitive environments, Hitachi Data Systems has demonstrated clear leadership with TrueCopy Asynchronous software:

- Data integrity is guaranteed for dependent write applications
- Excellent performance is achieved for both long- and short-distance requirements, due to its asynchronous nature
- Future enhancements will provide additional configuration flexibility

Let's compare TrueCopy Asynchronous with other asynchronous implementations:

- Other implementations are based on sending changed tracks to the secondary system instead of time-stamped I/Os
- Application of changed tracks by other implementations (and not individually time-stamped I/Os) cannot preserve the original sequence of writes, and therefore should not be used for real-time disaster recovery unless used as part of a properly architected PiT solution.

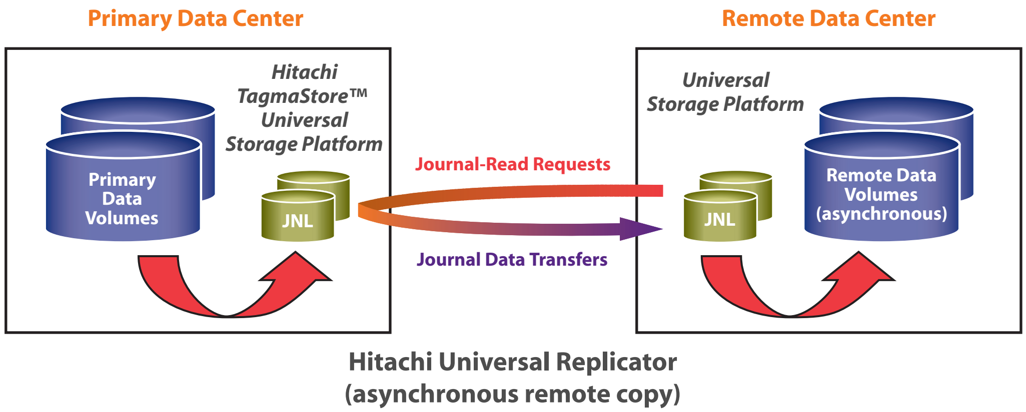
Hitachi Universal Replicator for TagmaStore™ Universal Storage Platform

Hitachi Universal Replicator for TagmaStore Universal Storage Platform is an advanced technology for asynchronously replicating data hosted on Universal Storage Platform or on externally attached storage systems from Hitachi Data Systems and other vendors.[1] Universal Replicator builds on the proven algorithms for data consistency of TrueCopy Asynchronous, but differs from other solutions in that it *pulls* the data to the remote site, instead of *pushing* it from the primary site.

Figure 8 illustrates Universal Replicator's application of performance-optimized, disk-based journaling. In this illustration, primary data volumes at the primary data center are being replicated to a second Universal Storage Platform at the remote data center.

1 For a discussion of the TagmaStore Universal Storage Platform and business continuity applications of its capabilities, including Universal Replicator, see *Business Continuity and the Hitachi TagmaStore™ Universal Storage Platform*, Hitachi Data Systems white paper WHP-163. Customers should check with Hitachi Data Systems on supported platforms and configurations.

Figure 8. Hitachi Universal Replicator: Asynchronous Remote Copy



Universal Replicator uses disk-based journal volumes and a pull-style replication engine to move data from the primary site to the remote site.

When collecting the data to be replicated, the primary Universal Storage Platform writes the designated records to a special set of journal volumes. The remote system then reads the records from the journal volumes, pulling them across the communication link as described in the next section.

Universal Replicator's disk-based journaling and pull-style replication engine help reduce resource consumption and costs, while increasing performance and operational resilience. In effect, Universal Replicator restores primary site storage to its intended role as a transaction processing resource, not a replication engine.

Additionally, by using a disk-based journaling technique, Universal Replicator prevents network issues or spikes in workload from causing the replication process to suspend: the overflow is buffered in the journal disks until the replication process can catch up. This is a significant benefit, as recovery from a suspended state (in cache-based replication methods) requires a destructive resynchronization process that causes RPOs to be significantly elongated.[2]

EMC SRDF Asynchronous

This is a new asynchronous replication technology from EMC, available only on the EMC Symmetrix DMX storage series. Its technique for preserving write dependency is essentially just honoring writes, without the use of other proven techniques such as sequencing or time stamping.

SRDF Asynchronous uses the concept of timed cycles known as delta sets, typically every 30 seconds, and captures all host I/O during that period in cache. If the same record is updated more than once, only the most recent update is kept. Any dependent I/O will in theory be in that delta set or in a subsequent one. Once the time is up, SRDF Asynchronous starts another delta set cycle at the primary site and begins to transmit the previous delta set to the remote side cache. Once all the data is received at the secondary site, it then promotes it to an active cycle and the data can be destaged to the back-end disk at the secondary site. If there is a problem during transmission or a disaster occurs, then the copy of

2 For a more detailed discussion of Universal Replicator and its application to advanced disaster-recovery strategies, see *Universal Replicator Advanced Technology*, Hitachi Data Systems white paper WHP-165.

consistent data at the remote site should be the previous completely transmitted delta set. However, with multiple links and network retries, that may be of concern.

This technology also appears to use large amounts of cache, as the data is held there much longer than with other asynchronous techniques. EMC claims significant improvement in required bandwidth, but this will be highly data and network dependent. Additionally, since the time RPO of the data will be at least the 30-second cycle and the time required to transmit it, using too little bandwidth will result in elongated RPOs and the possible dropping of the SRDF Asynchronous environment due to cache overload. SRDF Asynchronous has a number of other restrictions, at least in the initial release, such as limitation to a single consistency group.

PiT-mediated Remote Copy

Examples in this category include Hitachi NanoCopy technology, IBM PPRC Global Mirror, and various cascade copy approaches such as those offered by EMC and IBM.

Hitachi NanoCopy Technology

Hitachi NanoCopy technology is a PiT-mediated remote copy solution that allows customers to take a point-in-time copy (snapshot) of data without having to quiesce or otherwise interrupt the application accessing the data. It is, in fact, the only storage-based solution from any vendor that can do this. This version of NanoCopy technology is available for OS/390 and z/OS environments on Hitachi TrueCopy Remote Replication software for any number of logical volumes and storage systems. This capability has also been significantly enhanced through scripting to provide an alternative disaster recovery solution by taking periodic, non-disruptive PiT copies. These copies may be split off every 15 minutes or every few hours, according to requirements. NanoCopy technology is an architected solution that is based on customer-specific goals and objectives. For more information on Hitachi NanoCopy configurations, see *The Hitachi NanoCopy Advantage*, Hitachi Data Systems white paper WHP-134.

IBM PPRC Global Mirror

This is essentially IBM's implementation of NanoCopy, but it uses an adaptive copy type technology PPRC-XD along with FlashCopy. IBM claims it can cycle rapidly, updating the remote PiT copy every 30 seconds. However, as always, it depends on the type and rate of change of the data in addition to available bandwidth.

Cascade Copy Techniques

Cascade copies refer to a method of creating remote PiT copies of data, using PiT copy methods such as EMC TimeFinder or IBM FlashCopy together with remote copy technologies such as SRDF Adaptive or PPRC-XD. The approach is similar to NanoCopy or PPRC Global Mirror, but there are a few major differences. First, most solutions in this area require you to quiesce the application in order to spin off the local PiT copy. As this will affect the application, you may not want to do it very often; so the age of the PiT copy may be affected. Typically these approaches are designed using four copies of data (i.e., two local copies and two remote copies) instead of just two or three. This technology may seem appropriate for an environment where a customer wants to create a PIT copy every four hours; there is some expectation of lower bandwidth requirements in such an environment, since not every update is replicated to the remote site. We will revisit the bandwidth trade-offs in a later section of this paper.

Three Data Center Copy

Up to this point, we have been describing remote replication solutions involving two data centers: either synchronous replication to an in-region recovery site or asynchronous replication to an out-of-region recovery site. To obtain the benefits of both approaches—combining in-region and out-of-region replication—enterprises may consider techniques involving three data center (3DC) replication. The following sections describe traditional and advanced 3DC approaches.

Multi-hop and Traditional Three Data Center Copy

These offerings revolve around creating multiple PiT copies of data and require anything from two to six copies of data and two or three storage systems. EMC's multi-hop capability was primarily devised for the express purpose of achieving consistency of data over distance, as they did not support any true data-consistent asynchronous technology. The obvious downside of these solutions is the requirement of many additional copies of data and extra storage systems just to achieve consistency over distance, as is still the case for EMC customers who do not have Symmetrix DMX storage systems or who require consistency across more than a single storage system in a mainframe environment (see the discussion of SRDF Asynchronous, above). This is not the case with Hitachi Data Systems replication offerings.

Zero data loss over distance is a typical reason to adopt these solutions. By running synchronous and then asynchronous technology, you can essentially get a synchronous-style replication over distance. It's important to remember that in order to protect your enterprise from a rolling disaster, it is unlikely that you will get zero data loss, but you may get close to it. Additionally, you need to review your RTO, as there is a cyclical catch-up phase to the recovery site before you can use the data. How long this takes will depend on the locality of reference[3] of the data and the available bandwidth. This catch-up lag is only true of the traditional 3DC approaches; advanced 3DC solutions like those from Hitachi Data Systems (below) do not use PiT copies and therefore do not need a catch-up cycle, which makes them ideal for those customers whose business requirements demand a very quick recovery.

Another issue with the multi-hop and traditional three data center copy approaches is the requirement to run synchronous for the first hop, as you need to be close enough to avoid any synchronous performance issues, but far enough away to avoid the impact of a region-wide disaster as this could significantly affect the RPO of the recovery copy. To mitigate a potential negative dependence on the middle regional site, the best option is to replicate to two or more sites at the same time, including out of region.

This approach can work, for example, in an environment where a customer wants to have two copies of data at different locations. In this scenario where you need two copies of data, one fairly local and one over an extreme distance, multi-hop and traditional three data center copy may seem to be a good choice. The first hop provides the local backup, with an additional copy at an alternate location. However, if a problem arises at the local backup site, you have no disaster recovery at all unless you replicate directly from your production center.

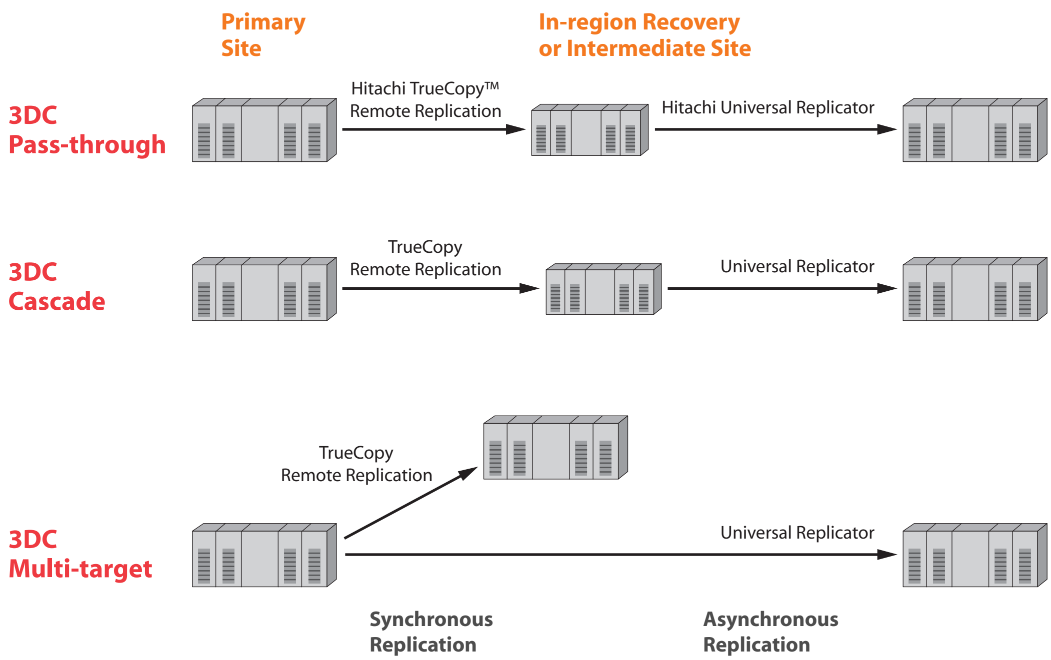
Multi-hop and traditional three data center copy are only really applicable in very specific circumstances and for a limited number of customers, and rely heavily on complex scripts and many copies of data. EMC offered multi-hop prior to the availability of SRDF Asynchronous, and may still offer it to customers who have not moved to the DMX product line.

3 "Locality of reference" refers to the observation that references to data tend to cluster. A particular piece of data, once referenced, is often referenced again in the near future. Similarly, once a particular piece of data is referenced, nearby data is often referenced in the near future.

Advanced Three Data Center Copy with Universal Replicator

With its introduction of the TagmaStore Universal Storage Platform, Hitachi Data Systems offers additional replication choices that reduce the cost of a 3DC configuration, making this approach affordable for a broader range of enterprises and applications. Unlike traditional methodologies that were complex and expensive to implement due to large numbers of copies, the Universal Storage Platform reduces the number of copies required and uses simple combinations of TrueCopy Remote Replication and Universal Replicator to enable three different 3DC configurations: cascade, pass-through, and multi-target (see Figure 9).

Figure 9. Advanced Three Data Center Strategies



With configuration options available, three data center copy is now an affordable, realistic alternative with TrueCopy and Universal Replicator.

Figure 10 summarizes the RTO and RPO impact of several different disaster recovery approaches based on two data center (2DC) and 3DC replication.

Figure 10. Benefits Comparison for Selected Disaster Recovery Configurations

Data Center Strategy	1DC*	2DC		3DC Traditional	3DC Advanced		
Replication Configuration	On-site	Sync Near	Async Far		Pass-through	Cascade	Multi-target
Primary site failure/failover							
Speed of recovery (RTO)	bad	better	good	depends	good	better	better
Data currency (RPO)	bad	better	good	depends	better	better	better
Regional disaster (RTO)	bad	bad	good	depends	good	good	good
Protection after failure outside primary site	n/a	bad	bad	bad	bad	depends	better

* 1DC strategy represents a traditional tape-based disaster recovery approach.

Alternative replication strategies provide different levels of protection.

The chart shows the relative benefits of these configurations in terms of recovery speed and data currency after a primary site failure—and in terms of recovery speed after a regional disaster. The chart also illustrates the impact of a failure outside the primary site, which could affect the ongoing level of protection and recovery capability in case of an additional site failure. For a more detailed discussion of these 2DC and 3DC configurations and their pros and cons, see *Hitachi Universal Replicator Advanced Technology*, Hitachi Data Systems white paper WHP-165.

Bandwidth Considerations

One of the biggest considerations for your organization is the cost of the solution, and network bandwidth is one of the biggest contributors to that cost. In real-time remote copy, every update is sent to the remote location. If your application executes 100 writes of 10K blocks/sec you need bandwidth that accommodates the writes, plus any control information that is also sent. In PiT copy solutions for disaster recovery, data is only replicated at pre-set intervals, such as every 15 minutes, every hour, etc. During the period of time that data is not being replicated, for data that is updated at the primary site, tracks are marked as changed by the storage system. If the same record is updated 100 times, then only the last change to the track will be shipped when the data is sent to the secondary location. This in theory means that the bandwidth requirement may be lower when using PiT technologies; however, there are many factors to consider before coming to that conclusion. If the PiT copy is frequent—under six hours, for example—as is often required, then the chances are that there will not be significant reduction in bandwidth. In certain circumstances, bandwidth requirements may be the same or not significantly different. The bandwidth requirement will be very dependent on locality of reference of the data and how current you want the PiT to be at the remote location.

Careful consideration should be used when architecting remote copy solutions with a view to saving network bandwidth. This includes technologies such as cascade, NanoCopy, multi-hop, three data center copy, and even SRDF Asynchronous. It is important to understand your data patterns, to try and determine if there are any benefits, and to consider your RPO and RTO. Using too little bandwidth can elongate your RPO and cause other problems in the environment. Where catch-up is required, the available bandwidth will affect the time you can begin your recovery (RTO).

Universal Replicator is unique in that customers can improve bandwidth utilization and lower their communication costs by sizing bandwidth for average bandwidth needs—not for peak usage. This simplifies bandwidth planning, and empowers users to better control their RPO in relation to infrastructure and communication costs.

In environments in which RPO of the data can be looked at in hours versus minutes, a PiT solution may be a viable alternative to a real-time solution.

Summary

Enterprises today must meet business applications' service levels—including increased resilience and protection from local and regional disruptions—while dealing with complex infrastructures and tight budgets. Hitachi Application Optimized Storage™ solutions enable organizations to precisely match business application requirements to Hitachi storage system attributes (performance, availability, data value, and cost), leveraging the application-centric management capabilities of the HiCommand® Storage Area Management (SAM) Suite.

To meet these requirements, Hitachi Data Systems provides customers with a wide range of flexibility and choice in technologies and solutions for both open systems and mainframe environments.

- The focus of Hitachi Data Systems product offerings is on standards and interoperability
- TagmaStore Universal Storage Platform and Lightning 9900™ V Series systems provide fully compatible, high-performance S/390 (PPRC, GDPS, and XRC) solutions; this positions customers to take advantage of future software enhancements
- Hitachi TrueCopy Asynchronous software, Hitachi Universal Replicator, and advanced 3DC configurations provide unique, enterprise-wide, simple, elegant disaster recovery solutions

Many copy technologies today can be considered when implementing business continuity solutions, especially when you include traditional backup methods of copying data sets. This is why it is important to choose only from best-of-breed solutions when addressing disaster recovery objectives. Hitachi Data Systems provides not only a superior range of offerings in this area, but also the expertise necessary to help you achieve your disaster recovery goals..

Appendix A: Bibliography and Resources

White Papers

Business Continuity and the Hitachi TagmaStore™ Universal Storage Platform
Hitachi Data Systems white paper WHP-163.

Hitachi Universal Replicator Advanced Technology
Hitachi Data Systems white paper WHP-165.

Addressing Federal Government Disaster Recovery Requirements with Hitachi Freedom Storage
Hitachi Data Systems white paper PERI-128.

The Hitachi NanoCopy™ Advantage
Hitachi Data Systems white paper WHP-134.

Books

Disaster Recovery Planning (3rd Edition)
by Jon William Toigo
ISBN: 0130462829

Business Continuity: Best Practices (2nd Edition)
by Andrew Hiles
ISBN: 1931332223

Web Sites

Contingency Planning and Management
http://www.contingencyplanning.com/

Continuity Central
http://www.continuitycentral.com/

Disaster Recovery Journal
http://www.drj.com/

DRI International (Disaster Recovery Institute)
http://www.drii.org/

Hitachi Data Systems Corporation

Corporate Headquarters
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
www.hds.com
info@hds.com

Asia Pacific and Americas
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
info@hds.com

Europe Headquarters
Sefton Park
Stoke Poges
Buckinghamshire SL2 4HD
United Kingdom
Phone: + 44 (0)1753 618000
info.eu@hds.com

Hitachi Data Systems is registered with the U.S. Patent and Trademark Office as a trademark and service mark of Hitachi, Ltd. The Hitachi Data Systems logotype is a trademark and service mark of Hitachi, Ltd. HiCommand is a registered trademark of Hitachi, Ltd.

TagmaStore, Lightning 9900V, NanoCopy, Application Optimized Storage, and TrueCopy are trademarks of Hitachi Data Systems Corporation.

All other trade names, trademarks, and service marks used herein are the rightful property of their respective owners.

All other company names are, or may be, trademarks or service marks of their respective owners.

Notice: This document is for informational purposes only, and does not set forth any warranty, express or implied, concerning any equipment or service offered or to be offered by Hitachi Data Systems. This document describes some capabilities that are conditioned on a maintenance contract with Hitachi Data Systems being in effect, and that may be configuration-dependent, and features that may not be currently available. Contact your local Hitachi Data Systems sales office for information on feature and product availability.

Hitachi Data Systems sells and licenses its products subject to certain terms and conditions, including limited warranties. To see a copy of these terms and conditions prior to purchase or license, please go to http://www.hds.com/products_services/support/warranty.html or call your local sales representative to obtain a printed copy. If you purchase or license the product, you are deemed to have accepted these terms and conditions.

©2004, Hitachi Data Systems Corporation. All Rights Reserved.

WHP-117-02 September 2004