HITACHI
DATA SYSTEMS

Hitachi TagmaStore™ Adaptable Modular Storage

TrueCopy™ Extended Distance Software User's Guide

MK-95DF778-06

Trademarks

Hitachi Data Systems is a registered trademark and service mark of Hitachi, Ltd., and the Hitachi Data Systems design mark is a trademark and service mark of Hitachi, Ltd.

TagmaStore™, TrueCopy™ and ShadowImage are trademarks of Hitachi Data Systems Corporation.

Serviceguard and Tru64 UNIX are trademarks or registered trademarks of Hewlett-Packard Company in the United States and other countries.

IBM, AIX, and HACMP are trademarks or registered trademarks of International Business Machines Corp. in the United States.

Linux is a registered trademark of Linus Torvalds.

Microsoft, Windows, and Windows NT are registered trademarks or trademarks of Microsoft Corporation.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Red Hat is a trademark or registered trademark of Red Hat, Inc.

SGI and IRIX are trademarks or registered trademarks of Silicon Graphics Inc. in the United States.

Sun and Solaris are trademarks of Sun Microsystems, Inc.

VERITAS is a registered trademark and VERITAS Cluster Server is a trademark of VERITAS Software Corporation.

All other brand or product names are or may be trademarks or service marks of and are used to identify products or services of their respective owners.

Notice of Export Controls

Export of technical data contained in this document may require an export license from the United States government and/or the government of Japan. Please contact the Hitachi Data Systems Legal Department for any export compliance questions.

Document Revision Level

Revision	Date	Description
MK-95DF778-00	May 2006	Initial Release
MK-95DF778-01	July 2006	Revision 1, supersedes and replaces MK-95DF778-00
MK-95DF778-02	July 2006	Revision 2, supersedes and replaces MK-95DF778-01
MK-95DF778-03	October 2006	Revision 3, supersedes and replaces MK-95DF778-02
MK-95DF778-04	January 2007	Revision 4, supersedes and replaces MK-95DF778-03
MK-95DF778-05	May 2007	Revision 5, supersedes and replaces MK-95DF778-04
MK-95DF778-06	June 2007	Revision 6, supersedes and replaces MK-95DF778-05

Changes in this Revision

- Added reference to Volume Migration product
- Changed software version to 7.2 or higher
- Added recommendations for setting differential management LUs in Table 3.1 TCE Requirements
- Added supported line band information in Table 3.1 TCE Requirements
- Added support and restrictions for concurrent use of Volume Migration in Table 8.15 TCE restrictions
- Added section 3.2.2 Supported Configurations
- Added section 3.2.3 Unsupported Configurations
- Added section 3.1 TCE Specifications
- Added section 4.2.6.1 Determining Pool Capacity
- Added index entries
- Added a new list of acronyms.

Preface

This document describes TrueCopy™ Extended Distance functionality and provides instructions for performing TrueCopy™ Extended Distance operations on Hitachi TagmaStore™ Adaptable Modular Storage subsystems.

This user guide assumes the following:

- The user has a background in data processing and understands RAID storage subsystems and their basic functions.
- The user is familiar with the Hitachi TagmaStore™ Adaptable Modular Storage array subsystem.
- The user has read and understands the *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).
- The user is familiar with the Windows® 98, Windows NT®, and/or the Windows® 2000 operating system.

Notes:

- The use of TrueCopy™ Extended Distance software and all other Hitachi Data Systems products is governed by the terms of your license agreement(s) with Hitachi Data Systems.
- Throughout this manual, the term **"TCE"** refers to the TrueCopy™ Extended Distance software program.
- Throughout this manual, the term **"TrueCopy"** refers to the TrueCopy™ Synchronous Remote Replication software program.
- Throughout this manual, the term **"Navigator"** refers to the Storage Navigator Modular software.
- Throughout this manual, the term **"ShadowImage"** refers to the ShadowImage™ In-System Replication software.
- Throughout this manual, the term **"SnapShot"** refers to the Copy-on-Write SnapShot software.
- Throughout this manual, the term "**Volume Migration"** refers to the Modular Volume Migration.
- For further information, please contact your Hitachi Data Systems account team, or visit the Hitachi Data Systems worldwide web site at http://www.hds.com.
- The use of TrueCopy™ Synchronous Remote Replication and all other Hitachi Data Systems products is governed by the terms of your agreement(s) with Hitachi Data Systems.

Software Version

This document revision applies to the following Hitachi TagmaStore™ Adaptable Modular Storage products version 7.2 and higher.

Convention for Storage Capacity Values

Storage capacity values for logical units (LUs) on the Adaptable Modular Storage TrueCopy™ Extended Distance are calculated based on the following values:

1 KB (kilobyte) = 1,024 bytes

1 MB (megabyte) = $1{,}024^2$ bytes

1 GB (gigabyte) = $1{,}024^3$ bytes

1 TB (terabyte) = $1{,}024^4$ bytes

Storage capacity values for hard disk drives (HDDs) on the Adaptable Modular Storage TrueCopyTM Extended Distance are calculated based on the following values:

1 KB (kilobyte) = 1,000 bytes

1 MB (megabyte) = $1{,}000^2$ bytes

1 GB (gigabyte) = $1{,}000^3$ bytes

1 TB (terabyte) = $1{,}000^4$ bytes

Referenced Documents

- *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Graphical User Interface (GUI) User's Guide (MK-95DF711)*
- *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Command Line Interface (CLI) User's Guide (MK-95DF712)*
- *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide (MK-95DF701)*
- *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Copy-on-Write SnapShot Software User's Guide (MK-95DF708)*
- *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage*
- *ShadowImage™ In-System Replication Software User's Guide (MK-95DF709)*

Comments

Please send us your comments on this document. Make sure to include the document title, number, and revision. Please refer to specific section(s) and paragraph(s) whenever possible.

E-mail: doc.comments@hds.com

Fax: 858-695-1186

Mail:

Technical Writing, M/S 35-10

Hitachi Data Systems

10277 Scripps Ranch Blvd.

San Diego, CA 92131

Thank you! (All comments become the property of Hitachi Data Systems Corporation.)

Contents

List of Figures

List of Tables

Chapter 1 Overview of Hitachi Remote Replication Solutions

This chapter discusses remote replication concepts and provides an overview of Hitachi's remote replication solutions. In addition to the above, it provides an introduction to Hitachi's TrueCopy™ Extended Distance product. This chapter includes the following sections:

- Replication & Disaster Preparedness Business Objectives (section 1.1)
- Remote Replication Concepts (section 1.2)
- Hitachi's Remote Replication Product Portfolio (section 1.3)
- TrueCopy Extended Distance Product Differentiation (section 1.4)

1.1 Replication & Disaster Preparedness Business Objectives

Today, all areas of business are increasingly dependent on the availability of critical data resources. Most enterprises require access to business data 24/7 and cannot tolerate more than a few hours or even minutes of downtime, without serious impact to the bottom line. Meta Group estimates lost revenue from downtime at an average of US$1 million/hour. Contingency Planning Research says losses rise as high as US$6.45 million/hour for retail brokerages. Beyond the loss of revenue, there are adverse headlines and the potential impact on company valuation to consider, along with lower employee productivity caused by sporadic outages.

Governments have also established mandates, requiring corporations to comply with business data protection regulations. For example, United States regulations on data protection now apply to health care (HIPAA), financial services (SEC 17a4), corporate accountability (Sarbanes-Oxley Act), life sciences (21 CFR Part 11), and government (DoD 5015.2-STD). Beyond government regulators, investors and even insurers insist that businesses should establish feasible disaster recovery and business continuity plans, to protect critical information. In addition, organizations constantly face increasing internal and external threats to the stability of information systems.

In today's extremely competitive business environment, organizations cannot afford to ignore the impact of all the above factors. Since business continuity and data availability have become the top concerns of enterprises globally, it has become important for businesses to address the areas of disaster preparedness and proactively plan for successful disaster recovery. Preparedness measures are often undertaken by analyzing the following business objectives:

1. Disaster scale: What are the assumptions for the scale of a disaster or failure?
2. Recovery Point Objective: Up to what point in time, should data be recovered in the event of a disaster?
3. Recovery Time Objective: How soon business operations should be resumed after the disaster?
4. Cost: How much can be budgeted for preparedness measures?

The overriding goals of any business continuity plan must be to survive a disaster and resume operations as quickly as possible. The best recovery plan to achieve these goals depends on how an organization chooses to prioritize the above business objectives. The following sections discuss each of the above business objectives in more detail.

1.1.1 Disaster Scales

Before determining necessary preparedness measures, assumptions must be made about the scale of the disaster. Disaster scales describe the scope of a disaster, in terms of damage to information systems and the overall business environment. These assumptions often determine how information systems need to be configured.

Table 1.1 shows examples of disaster scales. For example, when a file is accidentally erased, the damage is limited only to the specific file. On the other hand, in the event of a disaster such as an earthquake, the damage can easily extend to an entire city or region.

Table 1.1 Examples of Disaster Scales

Type of Disaster	Example	Disaster Scale
Data destruction	File erased by mistake	File
Drive failure	Dual failure on a drive on which RAID is being reconfigured	Logical Unit
Subsystem failure	Hardware failure in controller, etc.	Subsystem
Site disaster	Fire in building, etc	Site (within some hundreds of meters)
Natural disaster	Earthquake, hurricane, etc.	Extensive (hundreds of kilometers or more)

1.1.2 RPO and RTO

Recovery Point Objective (RPO) and the Recovery Time Objective (RTO) are parameters that define how quickly an organization needs to recover to survive a disaster and how much data loss can be tolerated.

Recovery Point Objective (RPO) is defined by the Storage Networking Industry Association (SNIA) as follows:

> *The maximum desired time period prior to a failure or disaster during which changes to data may be lost as a consequence of recovery. Data changes preceding the failure or disaster by at least this time period are preserved by recovery. Zero is a valid value and is equivalent to a "zero data loss" requirement.*

In other words, RPO is a target value that specifies the point in time starting from which data needs to be recovered, in the event of a failure or disaster. It answers the question: How much data can your business afford to lose? For example, the RPO value for temporary data generated during a calculation is considered low because such data can be generated again if the calculation can be made again. On the other hand, if the transaction data of an Internet shopping site is lost, the image of the business suffers and, consequently, the RPO value is considered high.

Recovery Time Objective (RTO) is defined by the Storage Networking Industry Association (SNIA) as follows:

> *The maximum desired time period required to bring one or more applications and associated data back to a correct operational state.*

In other words, RTO defines the time frame within which specific business operations or data must be restored to avoid any business disruption. It answers the question: how long can a business afford to be down? RTO values for data are determined by how indispensable the data is for continuous site operation. For example, although auditing data is important for business operation and the RPO value is high, the RTO value of this data is considered low if the data is used only for auditing. On the other hand, the RTO value of the transaction data mentioned in the RPO example above is considered high because such data is indispensable to business continuity.

Both RTO and RPO influence the data replication and recovery options an organization needs to adopt. The Recovery Time Objective combined with the Recovery Point Objective of business data, can be mapped to a range of technical approaches, associated costs, and degrees of data protection.

1.1.3 Cost

Cost is the monetary requirement for building and operating a system which achieves its disaster scale, RPO, and RTO targets. Cost and the RPO and RTO values exist in a trade-off relationship. For example, cost increases as the loss of data in preparation for an extensive disaster is minimized and the target for a quick recovery is achieved. Recovery-time objective combined with a recovery point objective based on the value of particular data, map to a range of technical approaches, costs and degrees of data protection. The following figure illustrates the data protection choices based on cost versus value of data.

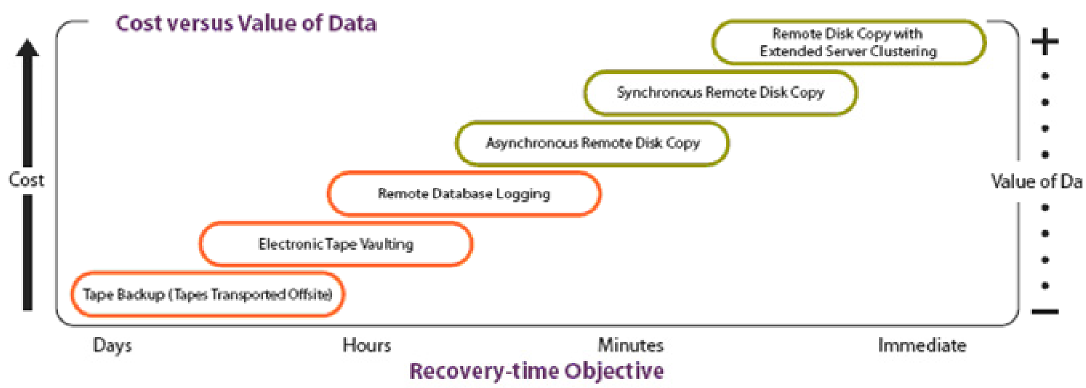


Figure 1.1 Data Protection Options

1.1.4 Disaster Preparedness Measures

Fundamentally, there are two types of disaster preparedness measures:

Data protection measures: These measures are steps taken to guard critical data against physical destruction (such as when a dual failure occurs on a drive), logical destruction (such as when a file is accidentally erased), site failures or natural disasters. One way to protect data is to periodically create data backups on physically independent secondary storage devices. If a primary data storage device is physically destroyed, most of the data can be retrieved from the latest backup data in the independent secondary storage device.

Another way to protect data is to create snapshots of primary data on the local storage device, at specific time intervals. In the case of logical destruction in the storage device, data can be restored from the point-in-time snapshots. Backup data can also be created at a remote location as a protection against site failures or natural disasters. This type of backup is called remote replication.

Disaster recovery measures: These measures are contingency plans that can be implemented in the event of a disaster. These measures help achieve a high RTO and may involve creating a clustering configuration to ensure work can be resumed using the copied data, immediately after a primary site failure is detected. Recovery measures may also involve establishing backup sites far from the location of a possible disaster.

Once the above business objectives are addressed and risks established, businesses can design resilient information systems, to ensure quick recovery time of critical application processing. Assessment based on the discussed business objectives also helps determine which applications require the most protection and the business impact of unexpected downtime.

1.2 Remote Replication Concepts

From the wide range of technical approaches to the business continuity problem, the most effective insurance policy against system downtime is provided by replicating data to a remote secondary site. Data replication over relatively long distances ensures that an organization can resume operations quickly and with a minimum of data loss in the event of a wide area disaster. Moreover, remote data replication can provide a variety of productivity benefits through secondary or parallel access to data that does not impact regular production workloads.

Remote data replication increases data availability in the following ways:

- Automating procedures to reduce the duration of planned failures, such as system maintenance, application testing and development, and data backups
- Allowing non-disruptive backup of current production data with no impact to the production application
- Speeding failover and data restoration in the event of an outage by replacing slow and labor-intensive tape-based restores with continuously available online backups
- Allowing secondary sites to take over primary processing to eliminate scheduled downtime
- Enabling frequent, non-disruptive disaster recovery testing with an online copy of current and accurate production data

1.2.1 Remote Copy

Remote data replication involves backing up data stored in a primary storage device to another geographically remote secondary storage device. The primary data volume (P-VOL) in a local storage system is connected to the secondary data volume (S-VOL) in the remote storage system via a remote path. The connected P-VOL and S-VOL form a pair, which is the fundamental entity for remote copy operations. When data replication is performed data in the primary data volume is copied to the remote secondary data volume. Remote data replication is performed in the background without affecting host application processing.

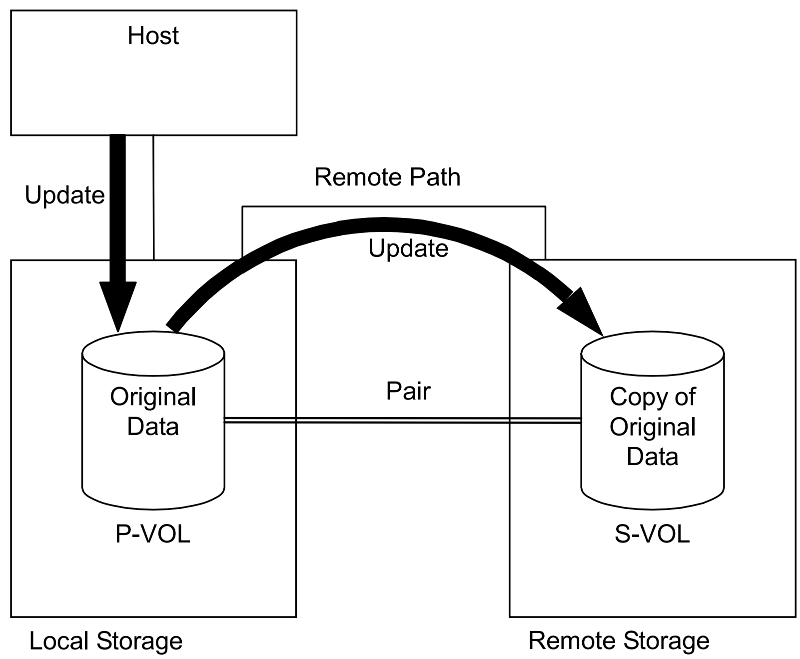


Figure 1.2 Remote Copy Operation

1.2.2 Integrating Local and Remote Replication

Remote data replication functionality can be combined with local replication functions (such as snapshot or image backups) to provide extremely resilient and powerful replication solutions. The ability to integrate these functions differentiates remote data replication technology from traditional backups. Snapshot backups involve creating point-in-time frozen replicas (snapshots) of data in both primary storage device and the secondary storage device. The static images of the primary volume are transferred to the secondary storage device during the remote backup operation. The multiple snapshots created prior to the onset of a disaster, are useful for data recovery, particularly from rolling disasters.

Rolling disasters refer to unplanned outages causing system failures to occur over seconds, minutes, or even hours. For example earthquakes, tornados, and floods which wreak havoc across wide areas. All systems, storage, or network connections may not fail in the same instant, but failure may affect the remote replication connections first. A system in the primary site may still be able to process transactions and issue updates to the local storage devices after the initial failure but may be unable to replicate updates to the recovery site. When the system on the primary site is eventually non-operational, recovery attempts at the remote site may fail due to data inconsistencies. Frozen snapshots allow data restoration from one or more additional copies of critical data.

1.2.3 Remote Data Replication Options

Organizations can choose between two methods of remote data replication:

Synchronous remote replication: Synchronous replication ensures that an identical remote copy is created at the time the primary copy is updated. In this method, an input/output (I/O) update operation is not considered completed until a response to the update operation is received from the mirrored site. If an operation does not complete, it is rolled back at both sites. This ensures that the remote copy is always an exact mirror image of the primary copy. The obvious benefit of synchronous replication is that data can be recovered quickly. Should operations at the primary site be disrupted for any reason, the mirrored site can immediately pick up where the primary site stopped. Only those few operations in progress at the moment of the disruption might be lost. Since neither the primary nor remote copy will have a confirmed record of completion of those last few operations, the database rolls back automatically to the last confirmed state, ensuring database consistency. This replication method is viable for organizations that cannot tolerate any data loss and require a quick recovery time after disasters.

Asynchronous remote replication: Asynchronous remote replication is carried out as a background process, and does not affect the status of host I/O operations. In this method, the remote copy function is independent of I/O operations issued by the host. As a result the I/O operations between the host and P-VOL are faster but there is a trade-off in the concurrency of S-VOL data. Due to the asynchronous nature of the remote copy operations, there is always some differential data in the P-VOL. Asynchronous replication can span virtually any distance since there is no propagation delay involved in confirming transactions at the remote site. This replication method is appropriate for organizations that can tolerate significant down-time but not the performance impact of synchronous propagation delays. Asynchronous remote replication works over any distance and is the appropriate solution if data resiliency is to be ensured even in the face of wide-area outages.

1.3 Hitachi's Remote Replication Product Portfolio

Today, businesses must be able to resume operations fast, almost from the point where they broke off. Hitachi Data Systems provides a comprehensive portfolio of remote replication and disaster recovery solutions for open systems and mainframes. This portfolio includes synchronous solutions for shorter distances or asynchronous solutions for longer distances, depending on business recovery objectives. Figure 1.3 illustrates Hitachi's replication and disaster recovery solutions.

Hitachi's replication solutions enhance operational efficiency and resilience of your information systems. Using the available solutions you can implement customized business continuity solutions, addressing key business problems relating to disaster recovery, data protection and availability. Different solutions are available for the large-size TagmaStore Universal Storage Platform (USP/NSC) subsystems and the mid-range TagmaStore AMS subsystems.

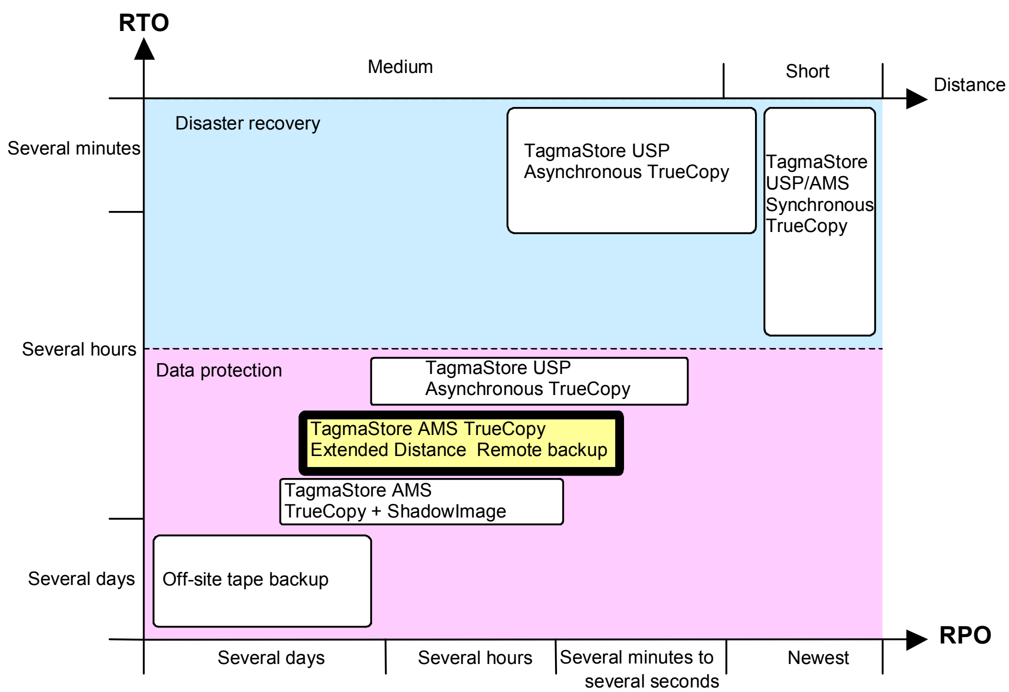


Figure 1.3 Positioning of Hitachi's solutions

Hitachi's remote replication portfolio for TagmaStore AMS subsystems includes the TrueCopy™ Synchronous Remote Replication product (hereafter referred to as TrueCopy) and the TrueCopy Extended Distance product (hereafter referred to as TCE). The following sections discuss the TrueCopy Extended Distance remote replication solution for mid-range TagmaStore AMS subsystems and provide an overview of the other Hitachi copy solutions.

1.3.1 TrueCopy Extended Distance (TCE)

Hitachi Data Systems has integrated its TrueCopy remote replication technology with Copy-on-write Snapshot operations, to provide TrueCopy Extended Distance (TCE), a continuous, non-disruptive, host-independent asynchronous remote replication and disaster recovery solution for midrange Hitachi TagmaStore® Adaptable Modular Storage models AMS500 and AMS1000. TrueCopy Extended Distance enables open-system users to perform remote copy operations between Tagmastore storage sub-systems in geographically distant locations, using fibre-channel interface connections. TCE channel extension options for long-distance remote replication include Internet Protocol (IP), Dense Wavelength Division Multiplexer (DWDM), Synchronous Optical Network (SONET), and Asynchronous Transfer Mode (ATM).

TCE is a scalable solution that addresses the highly visible issues of business continuity, data resilience, disaster preparedness and recovery. In particular TCE is a solution for the following demand drivers:

Longer distance data protection and disaster recovery: TCE is an ideal solution for businesses seeking to build a highly available environment, with data resiliency even in the face of wide-area outages. This product leverages both snapshot local replication and asynchronous data migration technologies. Remote sites can be at virtually any geographically remote distance ensuring that the replicated copy is safely outside any likely disaster zone.

Fewer host performance interventions: TCE is a storage system-based solution that consumes zero host resources and has no impact of host services and performance levels. Remote copy operations using TCE are performed as non-disruptive, asynchronous background processes on the array and are transparent to the host server and other business applications. The remote copy functions are independent of host I/O operations and do not affect the completion of I/O operations or the associated system response time. The primary volume is continuously accessible by the host for I/O operations.

Minimal network latency even in lower speed networks: Hitachi Data Systems has until now supplied a script-based synchronous remote backup solution using TrueCopy remote replication software in combination with ShadowImage in-system replication. This solution is ideal for backup and disaster recovery over short distances but when used for remote replication involving distant recovery sites, carries the overheads associated with network costs and reduced overall performance attributed to slow connections. TCE can be deployed for wide-area disaster protection across virtually any distance.

Quick and accurate data restoration after wide-area disasters: TCE supports accurate data restoration in the event of disasters and the ability to maintain concurrent real-time replicated copy of data (less then 12 hours old) in the remote secondary site. Remote replication using TCE involves determining and transferring differential data from the primary storage subsystem to a remote storage subsystem.

Several point-in-time images of data in the primary storage subsystem are captured at specific time intervals and retained in a data pool. These images are then periodically copied to the remote secondary storage subsystem, at regular time intervals. This replication method reduces backup workload caused when copying large data volumes and enables building of systems optimized for higher RPO performance. The multiple images of critical data allow for data restoration starting at different time frames, in the event of disasters. TCE uses a unique method of sequence numbers and timestamps in each data record to insure proper sequencing and data integrity during transmission and recovery.

Once established, TCE operations continue unattended and provide continuous, real-time data backup. TCE is configured and managed using Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular software and Command Control Interface (CCI) software.

1.3.2 TrueCopy Synchronous Remote Replication Software (TrueCopy)

For distances within the same metropolitan area, TrueCopy Synchronous Remote Replication software (TrueCopy) provides a no data loss, rapid restart solution. TrueCopy solution is available for Hitachi TagmaStore® Adaptable Modular Storage models AMS500 and AMS1000. TrueCopy provides the ability to copy data over short and medium distances, by combining local clones created with ShadowImage software and data movement capabilities of TrueCopy™ technology. Combined with Universal Replicator and the Universal Storage Platform, TrueCopy allows for advanced three data center configurations for optimal data protection.

TrueCopy remote copy operations are interlocked with host I/O operations at the primary storage site. Each I/O update operation at the primary site waits until copy operation is completed and confirmation is received from the remote site. Any incomplete operation is rolled back at both locations; thus, the remote copy is always an exact mirror image of the primary data volume. Synchronous remote copy implementation allows for data consistency at the remote secondary site to be potentially minutes behind primary source data. The probability of losing committed transactions is therefore minimized in the event of an unplanned failover at the secondary site. Rapidly recoverable secondary data is ensured because the remote data volume is always in an I/O-consistent state, requiring only the most recent transaction logs to be reapplied to recover to the point of the outage.

1.3.3 Other Complementary Copy Solutions

The following section provides a brief overview of local data replication technologies that can be used in conjunction with TrueCopy functions, to provide versatile remote replication solutions.

1.3.3.1 ShadowImage

Hitachi's ShadowImage is an in-system replication software that uses a high-speed, undisruptive, local mirroring technology to rapidly create multiple full-volume copies or clones of mission-critical information within storage subsystems, while keeping data RAID-protected and fully recoverable. ShadowImage can be combined with TrueCopy to provide a versatile remote replication solution with maximum flexibility in data backup and restoration.

1.3.3.2 Copy-on-Write Snapshot

Hitachi's Copy-on-Write Snapshot is a local replication solution designed to reduce backup time and provide point-in-time backup copies. Copy-on-Write Snapshot accelerates backup and recovery processes by reducing the amount of disk storage and by providing the capability for quick recovery of production data. When integrated with advanced disaster recovery techniques, it provides the ability to create high-availability solutions for quick restoration of production environments, in the event of a disaster. Copy-on-Write Snapshot accelerates recovery processes and reduces critical application downtime from hours to seconds.

1.4 TrueCopy Extended Distance Product Differentiation

TrueCopy Extended Distance is a unique Hitachi innovation and the best-of-breed controller-based, server-less, long-distance asynchronous remote copy and data protection solution for the mid-range segment.

Competitive Capabilities

TCE offers the following competitive capabilities:

- Heterogeneous server and application support
- Zero host resource consumption and no requirement for remote replication agent
- OS independent and application optimized solution with no negative impact of application performance
- Easily deployed and managed using centralized GUI and supporting CLI
- Efficient bandwidth utilization leveraging existing network bandwidth for transmitting only data updates, not entire data volume
- Deployed over geographically extended distances via DWDM (dark fiber), ATM or IP networks
- Capacity unlimited licensing

Business Benefits

TCE provides the following business benefits to ensure business continuity:

- Facilitates compliance with government mandates and regulations
- Improved service levels by reducing planned and unplanned downtime of customer-facing applications
- Improved business resilience by enabling frequent, undisruptive disaster recovery testing with an online copy of current and accurate production data
- Substantial cost savings over the life of the systems in the form of reduced management costs, floor space savings, reduced power and cooling costs, increased system performance, and increased customer satisfaction
- Fully leverage data warehousing/data mining investments

Operational Benefits

TCE provides the following operational benefits for improved productivity and processes:

- Automated host-independent data mirroring (for UNIX and Microsoft Windows environments) to a remote site over unlimited distances
- Operating system independent array-based replication which is host and application-transparent
- Heterogeneous server and application support (Open Systems)
- Ability to maintain multiple generations of data both locally and remotely by leveraging in-system and remote snapshot technology
- Ability to restore a recent I/O-consistent copy of the production data at the remote site using multiple point-in-time replicas of primary data generated prior to the occurrence of a disaster
- Proven reliable data migration and normal backup operations on a copy of up-to-date production data independent of critical applications
- Bi-directional replication allowing the target system to also be a source, enabling customers to cross protect critical applications and data assets between distant sites
- Scalability of replicated environment to accommodate 10 times the storage capacity with no change in footprint or performance
- Reduced frequency and duration of planned outages, such as system maintenance and data backups
- Optimized resource utilization by off-loading processing and data to alternate systems
- Continuous access to data and increased performance over previous storage solutions

- Significantly reduced recovery time resulting from replacement of slow and expensive tape-based restores with continuously available online copies

1.4.1 Differentiation from Hitachi Copy Solutions

The following sections differentiate TrueCopy Extended Distance from the following Hitachi copy solutions:

- TrueCopy Synchronous Remote Replication software
- Copy-on-Write Snapshot
- TrueCopy Asynchronous (TCA) for USP/NSC

1.4.1.1 Comparison of TCE with TrueCopy

TrueCopy Extended Distance differs from TrueCopy Synchronous Remote Replication product in the following areas:

Distance Limitation: Using TCE overcomes the distance limitation which is the major drawback when using TrueCopy synchronous replication for long distance remote replication. The maximum distance for TrueCopy synchronous theoretically extends to 200 kilometers, depending on channel extender specifications. But latency quickly becomes a problem as propagation delays (the time spent waiting for the update to travel to the remote site and confirmation to come back) lengthen with increased distance.

The practical distance for synchronous replication of a busy transaction system is usually about 30 to 50 kilometers (20 to 30 miles), possibly further, depending on application’s tolerance for delayed response and other factors. These disadvantages make TrueCopy unsuitable for long-distance remote replication.

Performance Impact: The key differentiation between TCE and TrueCopy is with regard to performance impact on host systems. TrueCopy synchronous replication has significantly more performance impact than asynchronous replication, because a write from the host must wait for acknowledgment from the secondary storage system. Due to the asynchronous nature of its copy operations, TCE eliminates the propagation delays involved in confirming transactions at the remote site and improves system I/O performance.

Data Integrity: TrueCopy ensures better data integrity at the remote site. Synchronous replication operations are done in precisely performed in the same write sequence as the primary volume update. Hence data integrity issues caused by out-of-sequence writes do not arise because “dependent writes” are not initiated until prior writes on which they depend complete. TCE uses several mechanisms, such as sequence numbers and time stamps in the transferred data packet, to guarantee write-sequence and ensure data integrity.

Differential Data Loss: In the event of failure in the local TagmaStore AMS subsystem, the probability of differential data loss is higher when using TCE. Data updated in the primary data volume at the time of failure in the local subsystem, cannot be copied to the secondary volume at the recovery site. TCE uses local point-in-time snapshots of primary volume data to eliminate this disadvantage. Multiple snapshots are stored in a data pool and can be used for data restoration, beginning at different time intervals proceeding the time of a disaster. Data loss is a virtually non-existent problem when using TrueCopy, because of the synchronous copy operations. Table 1.2 lists the pros and cons of Hitachi's TCE and TrueCopy software.

Table 1.2 Comparison of TrueCopy and TCE

Areas of Comparison	TrueCopy Synchronous	TrueCopy Extended Distance
Write sequence guarantee when S-VOL is updated	Available	O for each cycle
RPO Performance	O	In the event of failure at the local TagmaStore subsystem, differential data that has not been transferred cannot be copied to the secondary subsystem.
IOPS Performance	O	Pool operations on the local and remote subsystems lower I/O performance
Response Performance	Depends on line delay	O
Long-distance Remote Copy	Line delay time increases and host I/o timeout could occur.	O
Pool	Not Required	Required
Applications	Backup, disaster recovery, data transfer	Backup, disaster recovery, data transfer.

1.4.1.2 Comparison of TCE and Copy-on-Write SnapShot

TCE uses Copy-on-Write SnapShot for local data replication on the primary site and remote site. TCE remote copy operations use snapshot images to determine differential data to be copied to the remote recovery site. TCE and Copy-on-Write SnapShot share a common pool of data snapshots. For more information on TCE and SnapShot product differentiation, see Section 5.3.3.

1.4.1.3 Comparison of TCE and TrueCopy Asynchronous for USP

TCE is an asynchronous remote replication and recovery solution for mid-range Hitachi TagmaStore® Adaptable Modular Storage models AMS500 and AMS1000. TrueCopy Asynchronous (TCA) is a unique disaster recovery solution for the Hitachi TagmaStore® Universal Storage Platform (USP).

Table 1.3 summarizes differences between TCE and TCA.

Table 1.3 Comparison of TCE and TCA

Features	USP TCA	TCE
Architecture	Write order guaranteed data transfer	Snapshot and periodic differential copy
Applications	Mainly disaster recovery	Remote data protection and recovery
Update Cycle	Continuous	5 minutes (default)
Data Buffer	Cache memory (side file)	Disk drives
Consistency Group	Yes	Yes
Initiator and Target on Same Port	No	Yes
Cascading	ShadowImage, 3DC	Copy-on-Write SnapShot
Sync/Async	Both	No-coexistence
Pair operations	All	No swap-takeover
Copy speeds	15 levels	3 levels (same as AMS TrueCopy)

Chapter 2 Overview of TrueCopy Extended Distance

This chapter presents an overview of TrueCopy Extended Distance software and discusses the functional and operational details of the product. This chapter includes the following sections:

- Typical TCE System Environment (section 2.1)
- TCE Functional Components (section 2.2)
- TCE Features (section 2.3)
- TCE Functional Overview (section 2.4)

2.1 Typical TCE System Environment

A typical TCE hardware configuration includes two TagmaStore AMS subsystems (one local subsystem and one remote subsystem), two hosts connected to each TagmaStore subsystem and two management servers. The two TagmaStore AMS subsystems are connected to each other via a WAN using two extenders. The two hosts are connected to each of the TagmaStore AMS subsystems via fibre channel connections. The two hosts communicate with each other on a LAN or WAN using the TCP/IP protocol. Each of the two management servers are connected to the TagmaStore AMS subsystems via a management LAN.

The logical configuration of the local TagmaStore AMS subsystem includes a command device, a differential management logical unit (DM-LU), two primary data volumes (P-VOLs), belonging to the same consistency group (CTG), and a logical unit for two data pools. The remote TagmaStore AMS subsystem also has a command device, a DM-LU, two secondary data volumes (S-VOLs), belonging to the same CTG, and a logical unit for two data pools.

In addition to the above components, the TCE system architecture also includes the Storage Navigator Modular (SNM) software and the Command Control Interface (CCI) software. Navigator is installed on each management server and is used to configure the TCE system environment. Command Control Interface software is installed on each host and manages TCE volume pair operations.

Note: For TCE on NAS Modular system environment, see Appendix C.

2.2 TCE Functional Components

TCE remote replication and data restoration operations are performed using the following functional components:

- P-VOLs and S-VOLs
- Command Devices
- Differential Management Logical Units (DM-LUs)
- Data pools (POOLs)
- Consistency Groups (CTGs)
- Volume Pairs

The following sections discuss the above TCE functional components. Figure 2.1 illustrates the interactions between the above functional components.

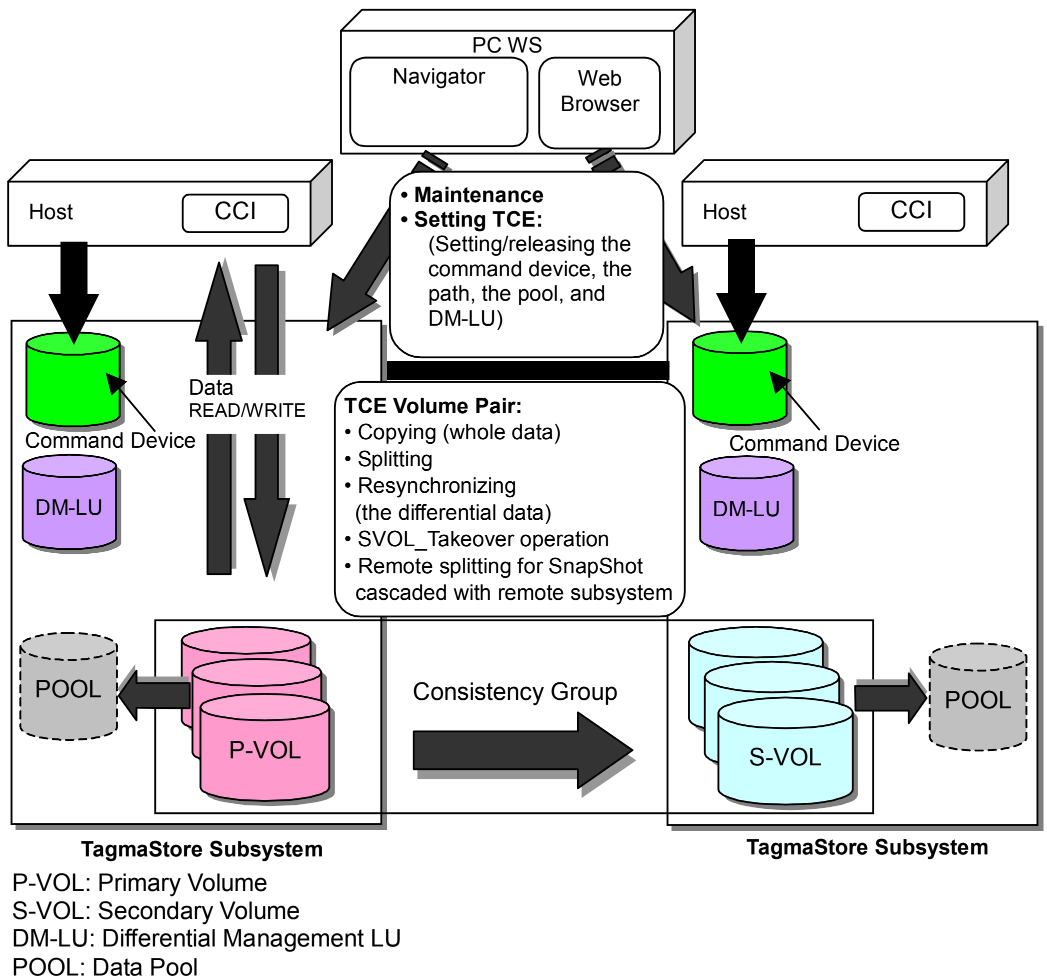


Figure 2.1 TCE Environment

2.2.1 P-VOLs and S-VOLs

Primary volumes (P-VOLs) are logical volumes in the local TagmaStore AMS subsystem that contain original application data. Secondary volumes (S-VOLs) are secondary or mirrored logical volumes on the remote TagmaStore AMS subsystem that contain backup data. In a TCE environment, every P-VOL on the local subsystem has one corresponding S-VOL on the remote subsystem. The logical unit size of an S-VOL is the same as that of the P-VOL. The capacity of P-VOL is set based on capacity required by an application and cannot exceed 2TB. The total capacity of all the P-VOLs and S-VOLs in a TCE environment must remain below 30TB.

P-VOLs and S-VOLs are paired together to execute TCE remote replication operations. TCE pairs can be formed between logical volumes with differing RAID level, number of drives, and drive type (FC or SATA). If the server is in a cluster configuration or if an application or database uses two or more logical volumes, these logical volumes are placed in the same consistency group (CTG).

2.2.2 Command Devices

Command devices are dedicated logical volumes that are used by management software such as CCI, to interface with the TagmaStore AMS subsystems. Up to two command devices can be configured per TCE subsystem. The logical unit size of a command device is set to 33 MB or more.

Command devices can be shared between several hosts. However, in such cases the data in the command devices is not shared. Command devices are not used by ordinary applications. Access restrictions are usually set to ensure that command devices can be accessed only by the host on which CCI is installed. For more information on setting command devices, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

2.2.3 Differential Management Logical Units (DM-LUs)

Differential Management Logical Units (DM-LUs) are logical units used to manage differential data in the TagmaStore AMS subsystem. The logical unit size is usually set to 5 GB or more. There may be up to two DM-LUs configured per TCE subsystem.

2.2.4 Volume Pairs

TCE remote copy operations are performed using logical volume pairs.. Each TCE pair consists of one primary volume (P-VOL) on the local TagmaStore AMS subsystem and one secondary volume (S-VOL) on the remote TagmaStore AMS subsystems. TCE supports a maximum of 1,022 pairs (AMS500) or 2,046 pairs (AMS1000).

TCE supports different pair operations for creating and managing TCE volume pairs. Pairs are created and managed using pair commands issued from the Command Control Interface (CCI) software. Table 2.1 summarizes the supported pair commands for performing TCE pair operations.

Table 2.1 TCE pair operation commands

Task	Command Name	Contents
Pair creation	paircreate	Creates a TCE pair.
Pair configuration check	pairdisplay	Checks the configuration of the created TCE pair.
PAIR status waiting	pairevtwait	Waits for transition to the PAIR status after initial copy has been completed.
Pair status monitoring	pairdisplay	Displays the pair status to confirm whether operation is normal.
Synchronization monitoring	pairsyncwait	- Checks the time difference between the P-VOL and S-VOL. - Also waits for the completion of copying of specific data written to the P-VOL to the S-VOL.
Remote snapshot	pairsplit –mscas	Creates the remote snapshot of the TCE pair P-VOL.
Pair splitting	pairsplit	Temporarily stops update copy processing for the specified TCE pair.
Pair recovery	pairresync	- Resumes the remote copy operation after the completion of maintenance. - Also resumes the remote copy operation for a TCE pair that was suspended due to a failure.
Takeover	horctakeover	Continues operation using the S-VOL.
Pair deletion (Normal)	pairsplit –S	Deletes a TCE pair if the asynchronous remote copy operation is no longer necessary.
Pair deletion (Remote Only)	pairsplit –R	Deletes the S-VOL if the S-VOL cannot be correctly deleted for a reason such as a path failure.

2.2.5 Data Pools (POOLs)

TCE remote copy operations occur as background processes and are independent of host I/O operations involving the P-VOL. As a result of the asynchronous nature of copy operations, there is always some untransferred differential data on the P-VOL, resulting from intermittent host I/O operations. Each TagmaStore AMS subsystem has disk volumes (data pools) in which this differential data is temporarily saved.

Pool capacity can be configured by the user and is based on differential data volume and update cycle considerations. The maximum pool capacity can be set to 15 TB. Up to 64 logical units can be assigned to a data pool, provided the maximum pool size is not exceeded. For details on setting pool capacities, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00*.

2.2.6 Consistency Groups (CTG)

When a file system or a logical volume which stores application data is configured from two or more logical units, TCE manages these multiple logical volumes as a consistency group (CTG). All disk volumes in a consistency group are treated as a single entity. A set of TCE volume pairs can also be managed and operated as a consistency group. A CTG can comprise of one or more volume pairs. Grouping TCE pairs allows pair operations to be concurrently performed on multiple pairs.

Consistency groups guarantee write order across the logical units during a remote copy. When a set of host I/O operations are performed on P-VOLs of a primary CTG, the write order of the host I/O operations is preserved and repeated during updates to the S-VOLs of a remote CTG. If a remote I/O operation cannot be completed on any of the S-VOLs belonging to the remote CTG, the I/O operations on all remote volumes in that group are suspended simultaneously. This ensures concurrency of data within every secondary CTG, since remote data volumes are frozen in a “time-consistent” state. The pair in which a different controller takes charge is also manageable as one CTG. Figure 2.2 illustrates write sequence guarantee within a CTG.

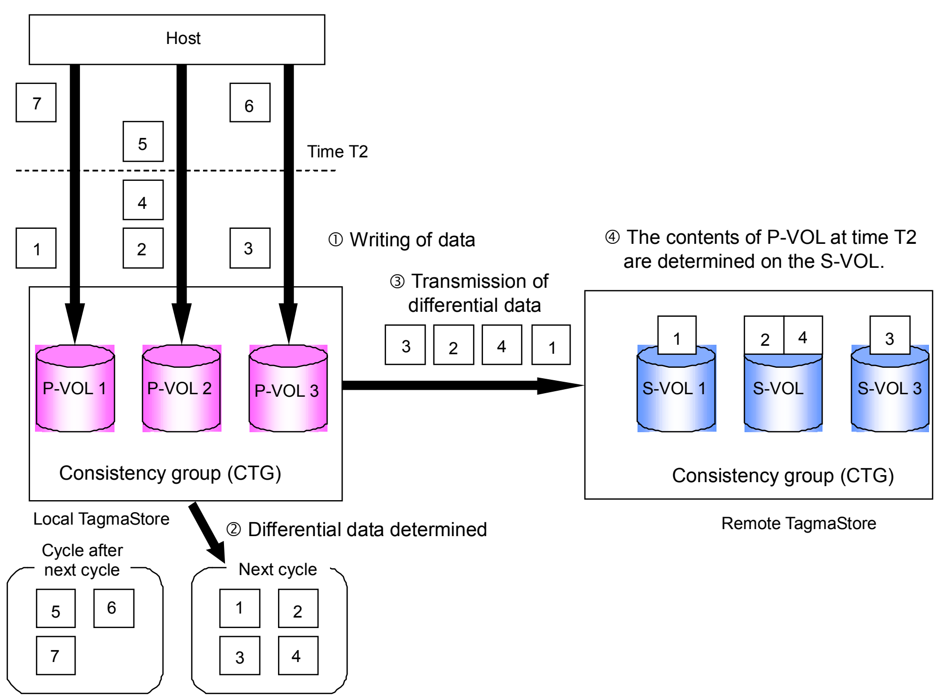


Figure 2.2 TCE Consistency Groups

When remote takeover operations are performed, internally pre-determined data from the remote subsystem data pool is restored simultaneously to the corresponding S-VOLs in the remote CTG. Since there are no data discrepancies, on the S-VOLs, host applications at the remote site will always access consistent data in the S-VOLs. Data consistency at the remote site enables quicker restart of operations upon disaster recovery.

2.3 TCE Features

TrueCopy Extended Distance offers the following differentiating capabilities:

- Write Sequence Guarantee to ensure updates to the primary volume are replicated in exactly the same sequence at the secondary site
- Consistency Group support for applications made up of several volumes, to guarantee that the replication process is consistent.

TCE functionality is based on the following feature functions.

- Local SnapShot function
- Remote SnapShot function
- Naming function
- Time Management function
- Write Consistency feature

2.3.1 Local Snapshot Creation Function

TCE uses Copy-on-Write SnapShot software, to create in-system replicas of critical data in the primary volume. Asynchronous TCE remote copy operations and SnapShot operations execute concurrently in the local TagmaStore AMS subsystem. Up to 15 (14: for micro program less than 0750/A version) snapshots of the P-VOL can be created in the local TagmaStore AMS subsystem. Snapshots images of the P-VOL are created and stored in a data pool upon receipt of a snapshot creation command from the host. These snapshots are useful for accurate data restoration of primary data volume, in the event of P-VOL destruction.

2.3.2 Remote Snapshot Creation Function

TCE also uses Copy-on-Write SnapShot software, to create snapshots of backup data in the remote S-VOL. Snapshot images of the S-VOL data are created upon receipt of a remote snapshot creation command from the local subsystem. These snapshots are saved in the remote data pool. Up to 15(14: for the micro program less than 0750/A version) snapshots can be acquired per TCE pair.

Remote snapshot creation is initiated by a command from the local host to the P-VOL on the local subsystem. The local TagmaStore AMS subsystem then sends a remote snapshot request to the remote TagmaStore AMS subsystem. Remote update processing continues undisrupted even when remote snapshots are created. Figure 2.3 illustrates the remote snapshot creation function.

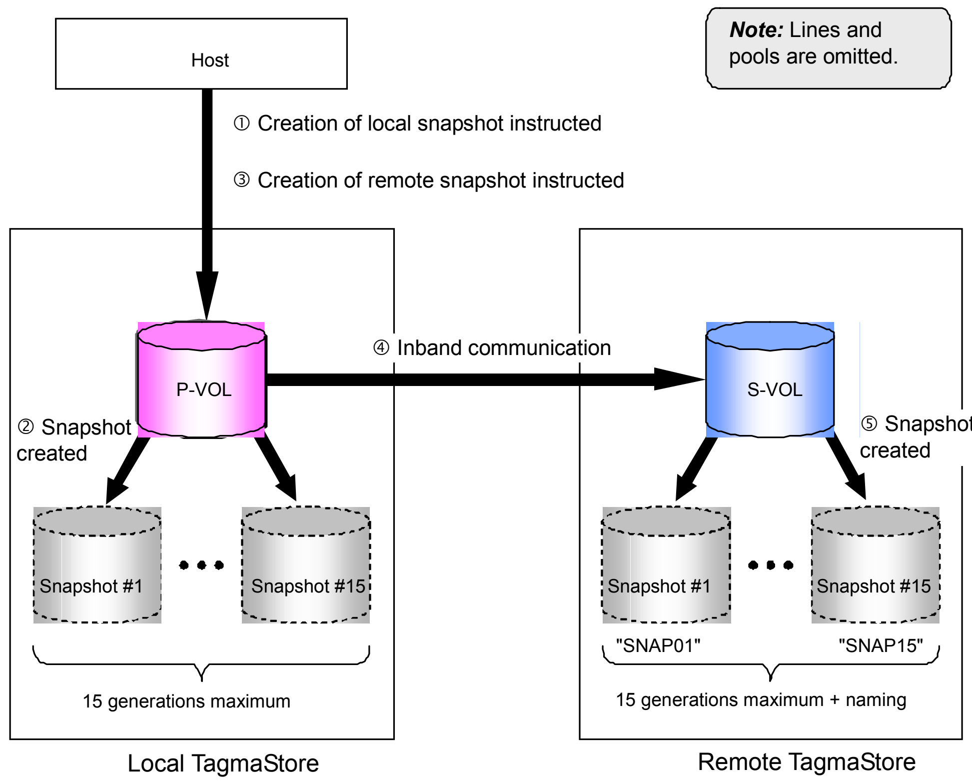


Figure 2.3 Remote SnapShot Function

2.3.3 Naming Function

This unique TCE function adds a human-readable character string of ASCII 31 characters to every remote snapshot. Using a character string rather than an LU number (logical unit number) to identify snapshots allows files to be easily searched within several generations of snapshots and reduces the risk of operator error, etc.

2.3.4 Time Management Function

This function adds a day and time string obtained from the local server, to every snapshot. Use of timestamps simplifies snapshot aging operations involving finding and deleting old snapshots. Since TagmaStore AMS subsystems have independent clocks, NTP (Network Time Protocol) is used to adjust clock times to better manage remote snapshots.

2.3.5 Write Order Guarantee Feature

In order to continue application processing at the remote site upon P-VOL failure at the local site, it is critical to maintain data consistency in the remote S-VOL. TCE ensures that data is updated in S-VOL in the same order that the host has updated data in the P-VOL. This “write sequence guarantee” feature is implemented by inserting sequence numbers in each update record. Update records are then sorted in the cache within the remote system, to assure write sequencing. When primary subsystem volumes comprise of multiple logical units, they are grouped into consistency groups. Write sequencing is guaranteed among volume pairs in each of these consistency groups.

2.4 TCE Functional Overview

TCE functions are enabled by installing TCE software on each of the local and remote subsystems. TCE functions include both remote replication and disaster recovery functions. This section explains TCE internal processes that implement remote replication and data recovery functions.

TCE remote replication involves copying of production data from the local TagmaStore AMS subsystem to the remote TagmaStore AMS subsystem, and is performed as an automated background cycle update process. The primary data volume is continuously available for host I/O operations during the replication process. The following steps outline a typical replication process in a TCE environment.

1. Copy-on-Write Snapshot in-system replication software independently creates point-in-time replicas of the P-VOL, and saves these images in the data pool on the local subsystem. Snapshot functions occur as fully automated local replication processes.
2. Data updated to the P-VOL as a result of write operations from the host (differential data) is determined on the local subsystem.
3. Differential data is transferred to the remote S-VOL, to create a recent I/O consistent recovery copy. Any further updates to the P-VOL are saved in the local data pool until the next update cycle.
4. Acknowledgement of update completion is sent from the remote subsystem to the local subsystem. Remote replication is suspended until the next update cycle.
5. Differential data in the S-VOL (resulting from a previous update) is internally pre-determined and saved in the remote data pool.
6. The remote subsystem queries the primary site to see if any disaster recovery process should be initiated. If so, the update cycle is halted and recovery processes begin at the remote site. If not, another remote update cycle is resumed.

Replication processes can be suspended during I/O-intensive batch activity such as a database reorganization, to reduce bandwidth requirements. In addition, at any point prior to refreshing the remote copy, the replication process described above can be suspended in order to make tape backups at the remote location. Figure 2.4 illustrates a typical TCE update process.

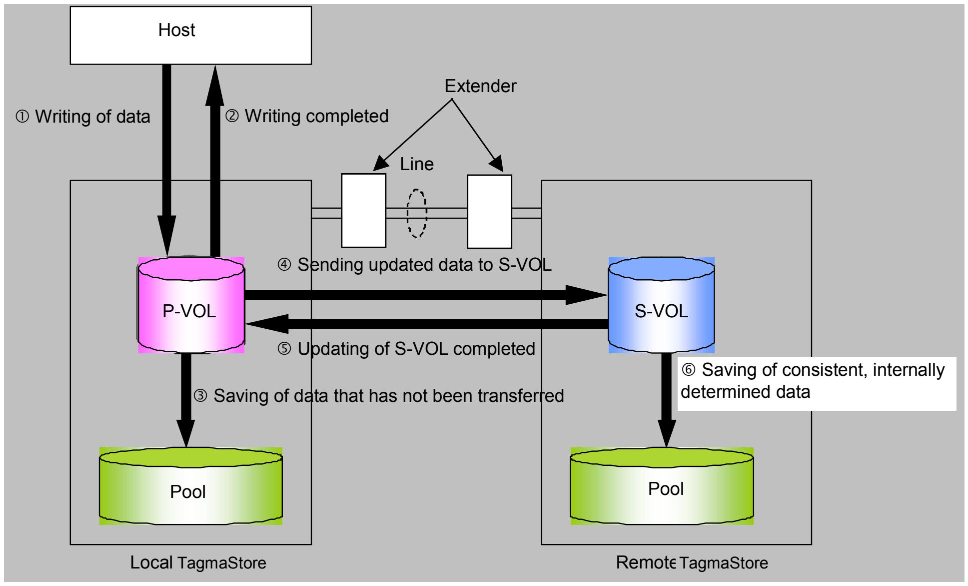


Figure 2.4 Outline of TCE Function

TCE recovery processes depend on the nature and outcome of a disaster. Disasters may result in logical data destruction (such as the erasing of a file) or physical damages to P-VOL (such as in a natural disaster). Recovery processes may involve restoring data and critical operations at the primary site using remote backup data or executing remote takeover operations at the remote recovery site. Takeover operations designate the remote S-VOL for critical application processing

When logical data destruction occurs, data is restored to the P-VOL from the local snapshot in the primary site or from the remote snapshot in the secondary site. If restoration is to be performed at a file system level, the file to be restored is determined on the remote S-VOL, and transferred to the primary site, where it is copied to a temporary disk location. Restoration is completed by copying the restored file to its correct file location and validating its contents.

In the case of physical failure of the P-VOL., the restoration procedure involves the following tasks:

1. To prepare for restoration, writing from the local server to the P-VOL is suspended, and P-VOL is un-mounted at the primary site.
2. Pre-existing P-VOL and S-VOL pairs on the local subsystem are deleted.
3. The local drive on which failure occurred, is replaced by Hitachi maintenance personnel.
4. On the remote site, if the S-VOL is mounted on the remote server, writing from the remote server to the S-VOL is suspended and the S-VOL is unmounted.
5. A configuration definition file is created on the remote server, describing the configuration for a new TCE pair to be created on the remote subsystem.

6. A new TCE pair is created using the old S-VOL for the old P-VOL, on the remote subsystem. This pair configuration enables data to be transferred from old S-VOL to P-VOL.
7. Data from the old S-VOL is restored on the old P-VOL.
8. The snapshot to be restored is determined from the snapshots available in the remote data pool. Data to be restored is found by referring to the split-marker character string and the backup time obtained from CCI or Navigator software.
9. A new TCE pair is created on the remote subsystem, using the old S-VOL and snapshot volume on the remote side.
10. Data is restored from the snapshot to the S-VOL.
11. After the completion of restoration, the old-S-VOL and old P-VOL pair are resynchronized, by recovering data from the S-VOL to the old P-VOL at the restore destination (primary site).
12. The P-VOL S-VOL pair created at the remote site is deleted. At this time, it is also established that no data has been written to the P-VOL and the S-VOL.
13. On the local side, a configuration definition file is created. It is used to describe the configuration for a new TCE pair to be created on the primary subsystem. If the logical unit number of the P-VOL on the primary side has not changed, the configuration definition file created for backup operation can be reused.
14. A new TCE pair is created on the local subsystem, involving the P-VOL on local side and the S-VOL on the remote side.
15. The P-VOL is then mounted onto the local server and the recovery process is completed.

The recovery process involves multiple pair operations and status transitions at various stages. For more information on disaster recovery processes, see Chapter 6 Using TCE for Disaster Recovery.

Chapter 3 TCE Requirements

This chapter describes TCE operational requirements and provides an overview of TCE management software. This chapter includes the following sections:

- TCE Specifications (section 3.1)
- TCE System Requirements (section 3.2)
- TCE Management Software (section 3.3)

3.1 TCE Specifications

Table 3.1 summarizes TCE specifications.

Table 3.1 TCE Specifications

Supported Arrays	AMS500 and AMS1000	
System Configuration	Bi-directional replication between two arrays	
Path Configuration	2 paths required (1 path/CTL)	
Port Modes	Fibre only, Initiator and target intermix mode One port is usable for the host I/O and the copy of TCE at the same time.	
Pair Configuration	1 PVOL to 1 SVOL	
RAID Levels	RAID 1 (1D+1P) RAID 5 (2D+1P to 15D+1P) RAID 1+0 (2D+2D to 8D+8P) RAID 6 (2D+2P to 28D+2P) # PVOL and SVOL can be different RAID levels	
Supported Drive Types	FC drives and SATA drives FC drives are recommended	
LU Size	PVOL = SVOL	
Path Failover	Yes (path or controller failures)	
Host Access	PVOL: Read & Write SVOL: Unable to read and write when PAIR or COPY	
Max Pairs /CTL	AMS500: 1,022 pairs, AMS1000: 2,046 pairs	
# of CTGs	16 / System	
# of Pairs /CTG	AMS500: 1,022 pairs, AMS1000: 2,046 pairs	
# of Pools /CTL	1 (Shared with Snapshot)	
Pool Size	15 TB/Pool, 64 LUs/Pool	
# of Command Devices	2 / System (Shared with Snapshot, ShadowImage)	
Differential Size	512Bytes (PAIR), 64KBytes (Others)	
# of Differential LUs	2 / System (Require min 5GB of LU size)	
License Model	TC and TCE can’t use same PP license key and TC and TCE can't be used together	
Co-Existence Within Same Array	Snapshot	Yes
	ShadowImage	Yes
	TrueCopy Basic	No

Supported Arrays	AMS500 and AMS1000	
	Others	Yes: LUN expansion, cache residency manager, SNMP agent support manager, password protection, LUN manager, data retention utility, cache partitioning manager
Cascading/ Combination	Snapshot	Yes: TCE PVOL can be CoW PVOL Yes: TCE SVOL can be CoW PVOL No : CoW VVOL can be TCE PVOL No : CoW VVOL can be TCE SVOL
	ShadowImage	No
	TrueCopy Sync	No
	Others	Unable to format LU of a pair Unable to change the size of LU of a pair Unable to delete RAID group or LU of a pair
Management Software	Storage Navigator Modular	- Remote Path Management - Pool Management - Differential LU and Command Device - TCE Pair Information
	Performance Monitor By SNM	- Initiator Related Performance Information Monitoring and Display - Amount of transferred data, network latency, # of errors, etc.
	RAID Manager	- Pair operations - Create/delete (3 copy speeds) - Display - Split/resync - SVOL-Takeover - Create snapshots of PVOLs at remote - Confirm synchronization (pairsyncwait) * Swap-Takeover is not supported
	Web GUI	- Information Display (paths and pairs) - Path Activation

3.2 TCE System Requirements

Table 3.2 describes the operational requirements for TCE.

Table 3.2 TCE Requirements

Parameter	TCE Requirement
User interface	CCI: Version 01-18-03/06 or later. Navigator: version 3.50 or later
Controller configuration	Dual controller configuration is required.
Supported subsystem type	AMS500 and AMS1000 subsystems are supported. Combinations of AMS500 and AMS1000 are also valid. For details of supported system configurations, see section 3.1.1
Supported Platforms	Table 3.3 shows supported subsystem combinations.
Number of subsystem	Local subsystem and remote subsystem
Max number of Command devices	It is recommended that two command devices be set for each TagmaStore AMS subsystem. The command devices must be set on both the local and remote subsystems.
Max number of Differential management LU	It is recommended that two Differential Management LUs be set according to the following conditions. ▪ To be created in different RAID group ▪ To be allocated in different controllers The Differential Management LU size must be greater than or equal to 5 GB. The Differential Management LUs must be set on both the local and remote subsystems.
Supported line bandwidth	20 Mbps or more (100 Mbps or more is recommended). However, when a low line band is adopted, it is required to make the time for time-out of the path setting, CCI command, and host I/O longer. Besides, when the line band is low, the response to the CCI command may take several seconds per pair.
Size of LU	The LU size for P-VOL & S-VOL must be the same and must be larger than or equal to 65,538 blocks (1 block = 512 bytes). The LU size of the command device must be 33MB.
Unit of pair management	TCE pairs are formed between LUs, and are managed per logical unit.
Pair structure	One copy (S-VOL) per P-VOL.
Number of paths per TCE pair	Two paths per pair must be set. One path is specified per controller. Since it requires a dual controller configuration, total of two paths are specified.
Size of TCE pairs	Size of all P-VOL and S-VOL (per subsystem) must be less than 30 TB
Max number of LUs that can be paired	AMS1000: 2,046/subsystem AMS500: 1,022/subsystem
Max number of Consistency Groups (CTGs)	Max 16 CTGs /subsystem

Parameter	TCE Requirement
	AMS1000: 2,046 pairs/CTG. AMS500: 1,022 pairs/CTG It is necessary to create a configuration in which the number of the CTGs specified correspond to the groups specified in the configuration definition file.
Supported RAID level	RAID 1 (1D+1D), RAID 5 (2D+1P to 15D+1P), RAID 1+0 (2D+2D to 8D+8D), RAID 6 (2D+2P to 28D+2P)
Combination of RAID levels	All combinations are supported. The number of data disks does not have to be the same.
Data Pool	TCE and Copy-on-Write SnapShot share a common data pool on each subsystem. The subsystem must be restarted to acquire pool resources. If the pool is already in use by Snapshot, a restart is not required.
Size of Data Pool	Max 15 TB/POOL, Max 64 LUs/POOL (Size of all P-VOL and S-VOL) ÷ 4+ (size of all POOL) must be less than 15 TB
Number of Data pools	One data pool can be set per controller (Pool number "0" is for the Controller 0 and Pool number "1", for the Controller 1.)
Types of the drive for a P-VOL/S-VOL and Data pool	LUs consisting of FC drives and SATA drives can be assigned to the P-VOL, S-VOL, and data pool. Assignment of an LU consisting of FC drives to the P-VOL, S-VOL, and data pool, is recommended. When creating a pair with the LUs configured in the SATA drive, the use conditions of the SATA drive may be different. (See Table 3.3)
Owner controller of Data pool	A data pool on the owner controller side of each LU of the P-VOL and the S-VOL is used.

***Note*: For TCE requirements on NAS Modular, see Appendix C.**

Table 3.3 Characteristics of FC Drive and SATA Drive

Items	SATA Drive		FC Drive
	ATE250R/ATE400R	Others	
Drives presumed operation environment	The Read/Write access takes 330 hours/month or less on the average.	The Read/Write access takes 720 hours/month.	The Read/Write access takes 720 hours/month.

3.2.1 Supported Platforms

Table 3.4 lists the platforms supported by TCE.

Table 3.4 List of Supported Platforms

No.	Platform	OS REV	Disaster Recovery		Backup Use	Data Migration Use
			High Availability Configuration (Automatically Switching)	Manual Switching		
1	SUN	Solaris 8	Not supported	Supported	Supported	Supported
		Solaris 9				
		Solaris 10				
2	PC Server (Microsoft)	Windows 2000	Not supported	Supported	Supported	Supported
		Windows Server 2003 (IA32)				
		Windows Server 2003 (IA64)				
		Windows Server 2003 (x64)				
3	HP	HP-UX 11.0	Not supported	Supported	Supported	Supported
		HP-UX 11i				
		Tru64 UNIX 5.1	Not supported	Supported	Supported	Supported
4	IBM	AIX 5.1	Not supported	Supported	Supported	Supported
		AIX 5.2				
		AIX 5.3				
5	Red Hat	Red Hat Linux AS2.1, AS3.0, AS4.0	Not supported	Supported	Supported	Supported
		Red Hat Linux AS3.0 EM64T				
		Red Hat Linux AS3.0 IA64				
6	SGI	IRIX 6.5.x	Not supported	Supported	Supported	Supported

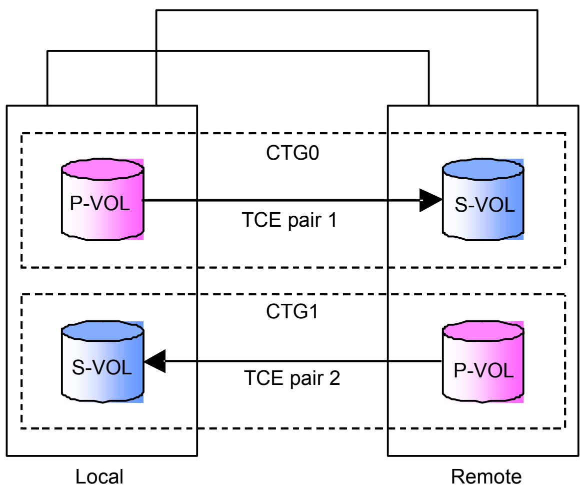
Table 3.5 shows the combinations of subsystems supported by TCE.

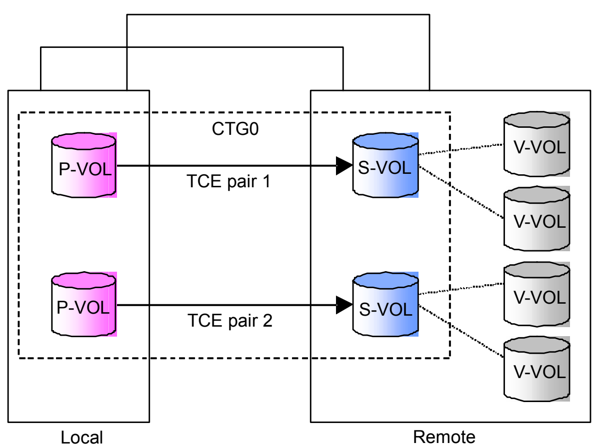
Table 3.5 Subsystem combinations supported by TCE

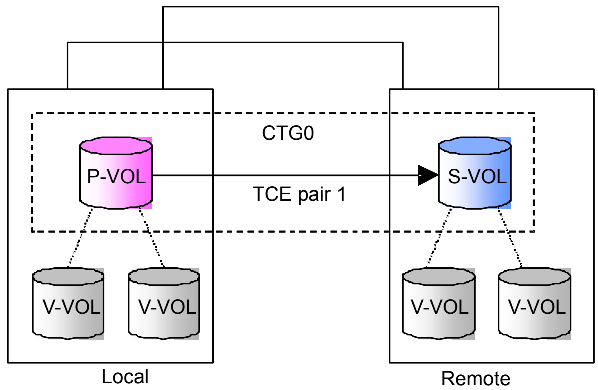
Local Side Subsystem Type	Remote Side Subsystem Type			
	WMS100	AMS200	AMS500	AMS1000
WMS100				
AMS200				
AMS500			Supported	Supported
AMS1000			Supported	Supported

Note: The combination of the 9500V series and AMS/WMS is not supported.

3.2.2 Supported Configurations

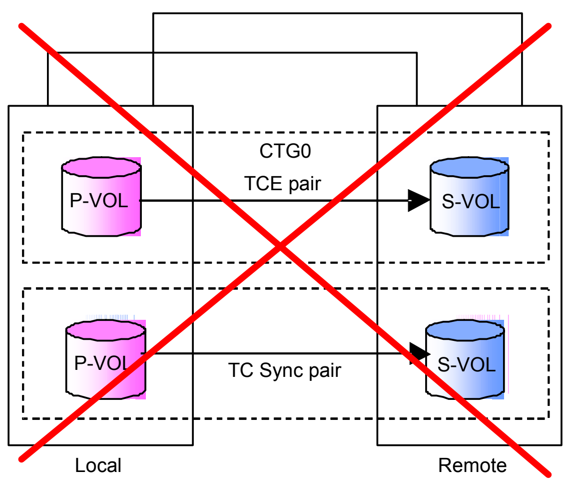




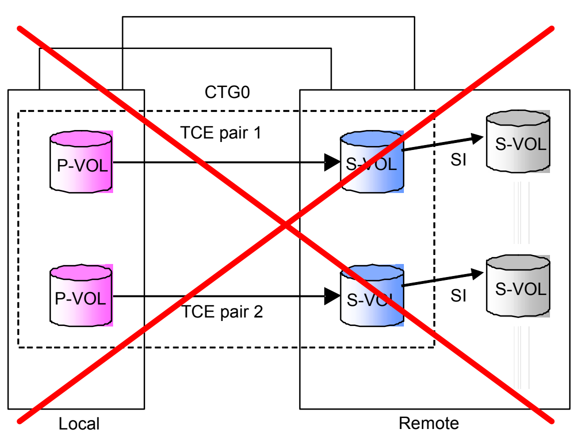


3.2.3 Unsupported Configurations

- TrueCopy and TCE are mutually exclusive



- ShadowImage cascade from TCE S-VOL is not supported



3.2.4 Licensing

License keys for TCE are issued on a capacity free basis, depending on the array on which TCE software is to be installed. TCE functionality is activated using a serialized license key, which can be used to enable TCE on a specific array determined by its serial number. License keys are distributed via a license key file on the product CD. This license key file can be executed from Storage Navigator Modular software. TCE and TrueCopy have separate license keys.

3.3 TCE Management Software

This section gives details of management software used to deploy, operate, and manage a system using TCE. The following TCE management software products are discussed:

- Command Control Interface software (CCI)
- Storage Navigator Modular (SNM)
- Web GUI
- Performance Monitor
- HiCommand Device Manager
- HiCommand Tuning Manager

Figure 3.1 illustrates Hitachi management software in a TCE system environment.

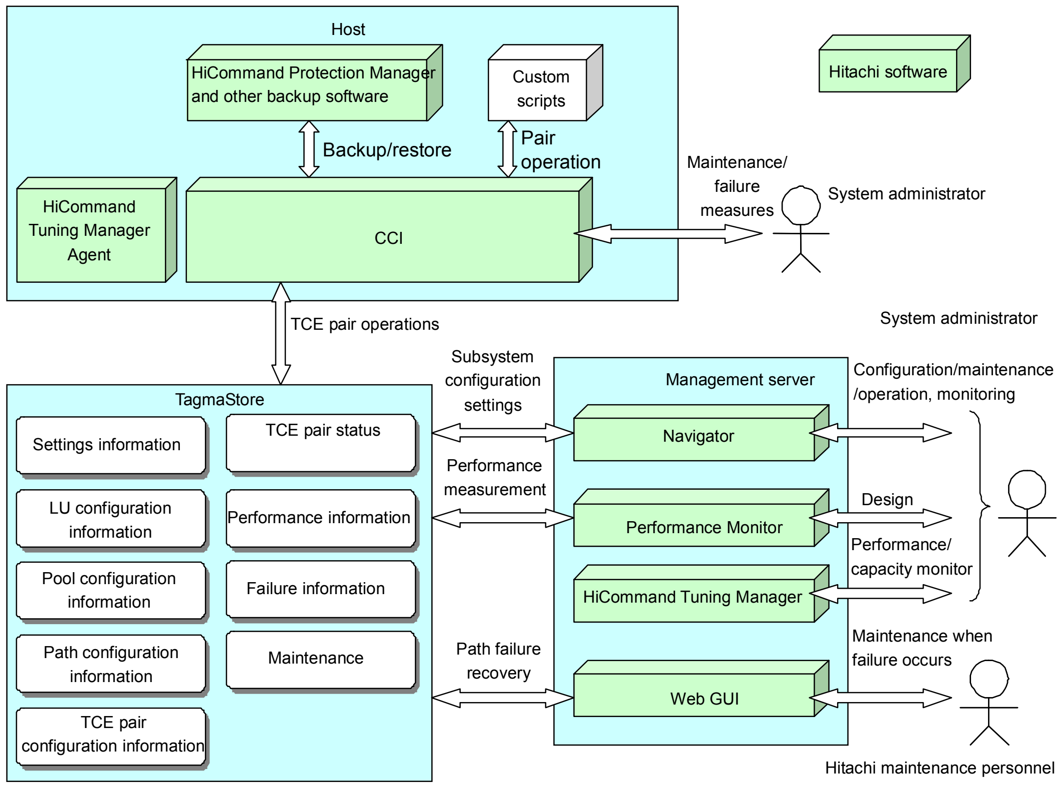


Figure 3.1 TCE Management Software

3.3.1 Command Control Interface software (CCI)

CCI software is used to display TCE volume information, create and manage TCE pairs, and issue commands for disaster recovery operations. CCI software resides on the UNIX®/ MS Windows® management server and interfaces with the TagmaStore AMS subsystems through dedicated logical volumes. CCI commands can be issued from the UNIX command line or using a script file. For more information on CCI, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

3.3.2 Storage Navigator Modular (SNM)

Storage Navigator Modular software GUI displays detailed TCE information and is used for configuring the TCE environment. SNM software communicates directly with the TagmaStore AMS subsystems via a local area network (LAN). SNM software supports both a GUI and CLI user interface to perform configuration, monitoring, maintenance and system reduction.

TCE configuration tasks include setting up data pool, cycle time, logical units and path for remote TCE operations, enabling or releasing a command device etc. SNM interfaces display important failure information for subsystems and maintenance information including the amount of differential data, the time difference between the P-VOL and the S-VOL, and the used capacity of a data pool. Navigator can be used to increase the pool capacity by adding more LUs to the pool.

In the event of a system failure or disaster at the primary site, the SNM software simplifies and expedites disaster recovery procedures. System Reduction tasks such as deleting command devices and data pools can also be performed using SNM software. Figure 3.2 illustrates a Storage Navigator Modular GUI window.

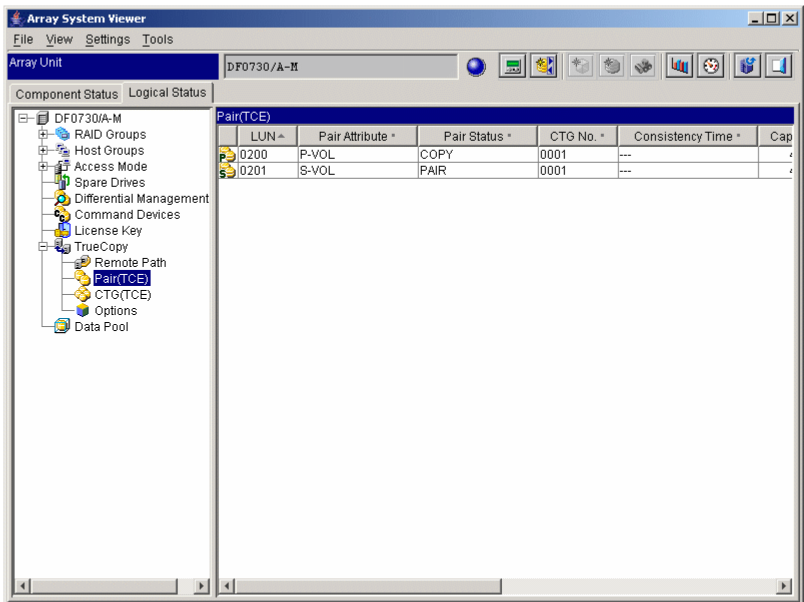


Figure 3.2 Sample Storage Navigator GUI

For more information on SNM software, see *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Graphical User Interface (GUI) User's Guide (MK-95DF711)* or *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Command Line Interface (CLI) User's Guide (MK-95DF712)*.

3.3.3 Web GUI

The Web GUI is a management tool used primarily by Hitachi maintenance personnel to monitor failure information (messages) generated during system operation. Messages may include pool over, link failures, suspending of pairs by errors, etc. Maintenance operations such as recovery from TCE remote path blockage) are also performed using the Web GUI. Figure 3.3 illustrates a typical Web GUI window.

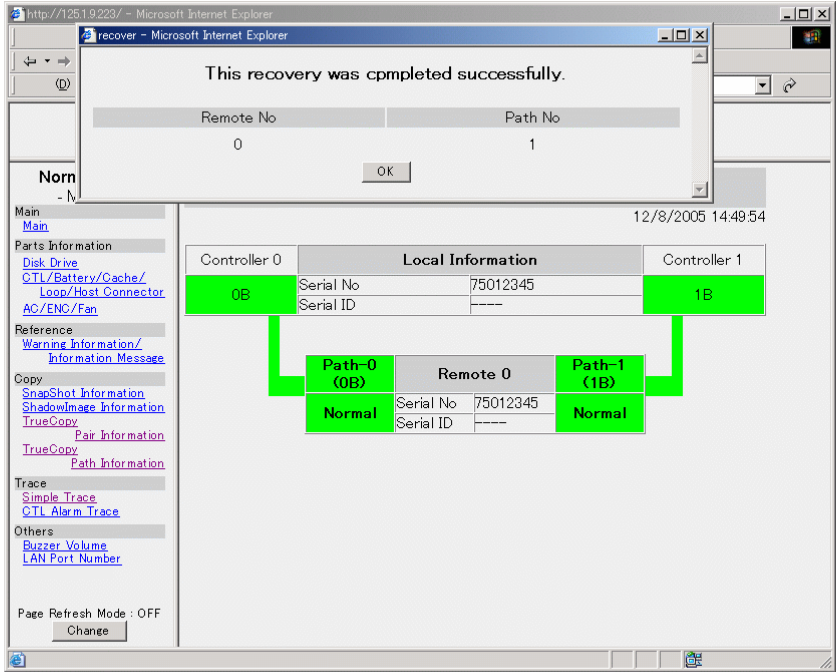


Figure 3.3 Sample Web GUI

3.3.4 Performance Monitor

Performance Monitor is a controller-based software application that acquires information on the performance of RAID groups, logical units, and other elements of the disk subsystem. Performance information may include amount of data inflow from the host, amount of data outflow to the remote subsystem, data transfer speed, and utilization rates of resources such as hard disk drives, cache and processors.

Performance information collected is used to design systems with optimal performance, and determine peak inflow, and peak inflow period. This information is also used to maintain RPO performance of the system by balancing data inflow and outflow. Performance information acquired is displayed with line graphs in the Performance Monitor UI. Figure 3.4 illustrates a typical Performance Monitor window.

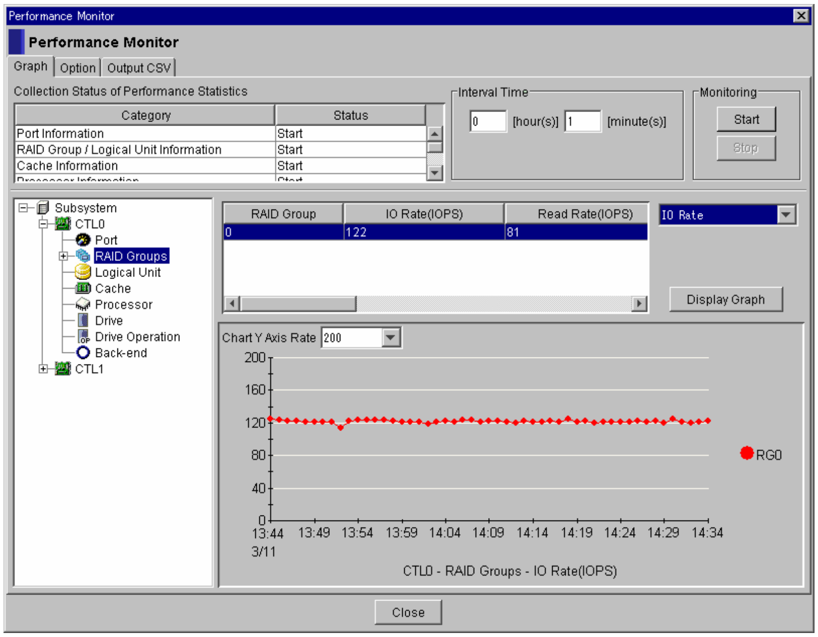


Figure 3.4 Sample Performance Monitor Window

Table 3.6 shows the types of data collected and displayed by Performance Monitor.

Table 3.6 Data Acquired by Performance Monitor

Type of information	Contents	Per Acquired			
		Per Port		Per Logical Unit	
		Control Command	Transfer Command	Control Command	Transfer Command
Command number (IOPS)	The number of commands issued during the observation period	O	O	x	O
Amount of data transfer (MB/s)	The amount of data transferred during the observation period	O	O	x	O
Response times (ms)	The average value and the maximum value of the response times during the observation period	O	O	x	△
Error count	Time out number	△		x	x

O: Graphic display and data collection are possible.
△: Only data collection is possible.
x: Neither graphic display nor data collection is possible.

For more information on Performance Monitor software, see *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Performance Monitor Software User's Guide (MK-95DF706)*.

The role of each management software product explained above is summarized in Table 3.7 below.

Table 3.7 Role of Management Software in TCE environment

Management Software Name		Navigator	CCI	Performance Monitor	Web GUI
Role overview		Sets the subsystem information	Operates on TCE pairs.	Measures performance	Tool for maintenance personal
User Interface Format		GUI, CLI	CLI, API	GUI	GUI
Design	Inflow estimation			√	
Configuration	Command device creation	√			
	DM LU creation	√			
	Logical unit creation	√			
	Pool creation	√			
	Remote path setting	√			
	Cycle time setting	√			
	TCE pair creation		√		
	Pair configuration check		√		
	PAIR status wait		√		
	Forecasting initial copy completion time	√	√		
Operation, monitoring	Monitoring pair status		√		
	Monitoring synchronization	√	√		
	Monitoring used pool capacity used	√			
	Collecting performance information			√	
	Collecting failure information	√			√
	Remote snapshot		√		
Maintenance	Pair splitting		√		
	Pair recovery		√		
	Changing of pool size	√			
	Recovery after path blocked				√
Disaster recovery	Volume restore		Δ		
	Takeover		Δ		
	Fail back		Δ		
System	Pair deletion		√		

Management Software Name		Navigator	CCI	Performance Monitor	Web GUI
reduction	Logical unit and pool deletion, etc.	√			

Δ: Pair swapping and pair re-synchronization after takeover are not supported.

3.3.5 Optional Software

The following optional management software is also used in a TCE system environment:

- HiCommand Device Manager
- HiCommand Tuning Manager

3.3.5.1 HiCommand Device Manager

HiCommand Device Manager is a storage management product which enables users to consolidate storage operations and manage capacity in systems containing multiple storage subsystems. Targeted for users managing multiple storage subsystems in open or shared environments, Device Manager quickly discovers the key configuration attributes of storage subsystems and allows storage administrators to begin proactively managing complex and heterogeneous storage environments using its easy-to-use, browser-based GUI. Device Manager's built-in report facility compiles and presents key information in preformatted reports (HTML) and as comma-separated values for export.

Device Manager provides link-and-launch integration with other software such as HiCommand Tuning Manager, HiCommand Protection Manager and Storage Navigator Modular (for Web) software. For more information on HiCommand Device Manager, see *Hitachi HiCommand Device Manager Web Client User's Guide (MK-91HC001) or Hitachi HiCommand Device Manager Command Line Interface (CLI) User's Guide* (*MK-91HC007).*

3.3.5.2 HiCommand Tuning Manager

HiCommand Tuning Manager is a storage resource management product which enables centralized management of a storage area network environment. This software monitors and analyzes the performance information and capacity information of the host, fibre channel switches, and storage devices. Changes in host I/O performance and capacity over time are displayed graphically and useful for planning future system expansion and reduction. For more information on HiCommand Tuning Manager, see *Hitachi HiCommand Tuning Manager User's Guide (MK-92HC022).*

Chapter 4 Setting up a TCE System

This chapter describes TCE software installation and provides an overview of tasks involved in configuring a TCE system environment. The following sections are included in this chapter:

- TCE System Design Considerations (section 4.1)
- Configuring a TCE System (section 4.2)

4.1 TCE System Design Considerations

A TCE system environment must be configured after taking into account target performance objectives established during the performance design process. Some of the performance objectives include: disaster scale, RPO and RTO performance and IOPS performance. Design considerations should be based on information such as average IOPS, average I/O size limit, I/O patterns, peak IOS, and peak period.

An optimal system configuration should then be designed after determining design values for copy method, cycle time, line distance and bandwidth, P-VOL and S-VOL sizes, and RAID configurations for P-VOL, P-VOL pool and S-VOL. The primary objective of a performance design process should be to build a TCE system configuration that achieves an optimal balance between the average inflow to the P-VOL and the average outflow to the S-VOL. For a detailed discussion on performance design, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00.*

TCE is an extra-cost option which must be installed and enabled with a license key. TCE functionality is not installed and activated with the initial configuration of TagmaStore AMS subsystems. You must install TCE on each subsystem, using the Storage Navigator Modular GUI or CLI. This section describes TCE installation and un-installation procedures using Storage Navigator Modular GUI. For procedures performed by using the Storage Navigator Modular Command Line Interface (CLI), see Appendix A.

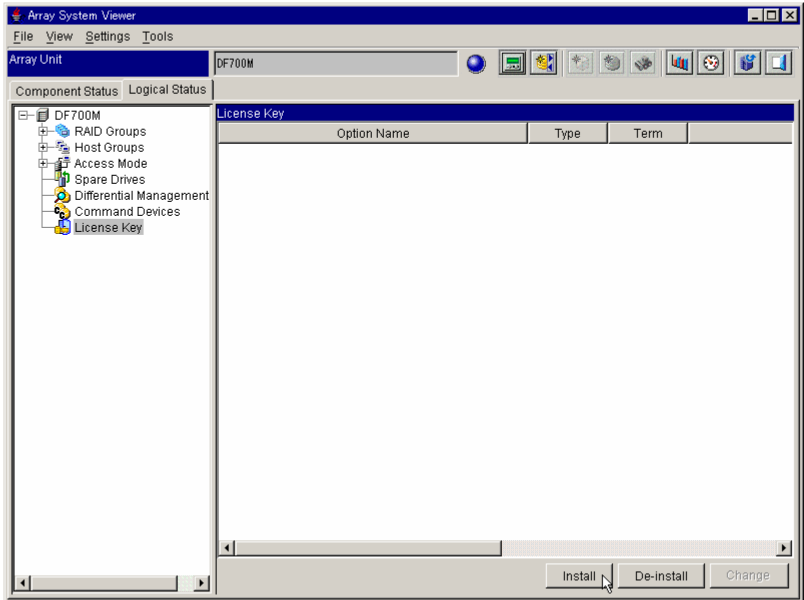
4.1.1 Installing TCE (using Navigator GUI)

TCE installation is performed using the key code or key file provided with the product. To install TCE:

1. Start the Navigator, and change the operation mode to **Management Mode** (administrator mode).
2. Register the subsystem in which TCE is to be installed, then connect to the subsystem.

   The Array System Viewer panel is displayed, and the connected subsystem is displayed.
3. Select the **Logical Status** tab.
4. Select the **License Key** icon and select the **Open** button.

Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

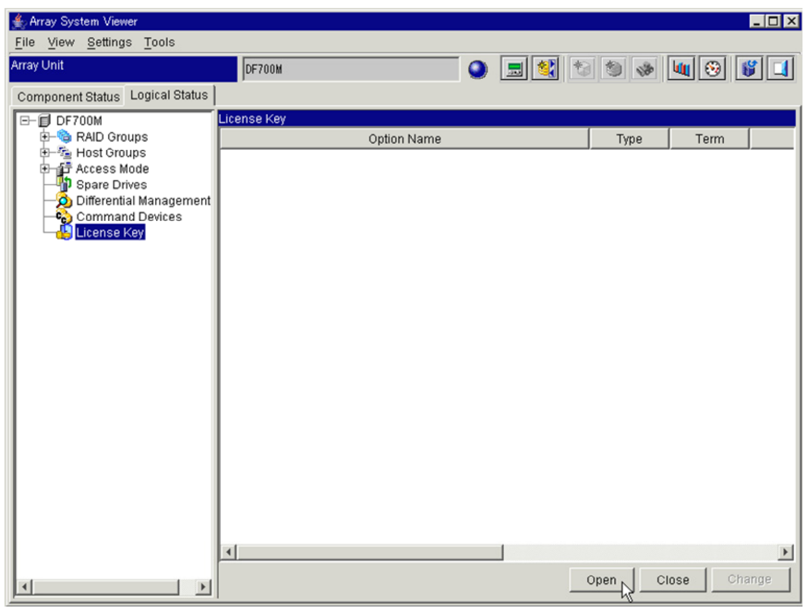


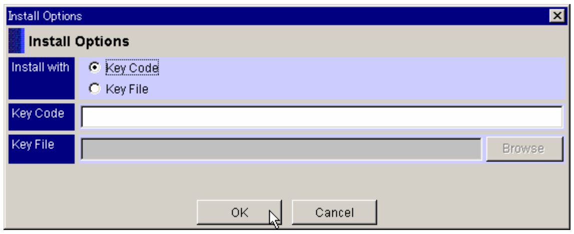
Figure 4.1 Array System Viewer Panel (Logical Status Page)

5. Select the **Install** button. (Navigator: Version 5.00 or later)

   Select the **Open** button. (Navigator: Less than 5.00 version)

   The **Install/Unlock Options** dialog is displayed. To install the option using a key code, click the **Key Code** radio button, and enter the key code in the **Key Code** text box. To install the option using a key file, click the **Key File** radio button. Insert the path to the key file in the **Key File** text box using the **Browse** button. Click the **OK** button.

Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

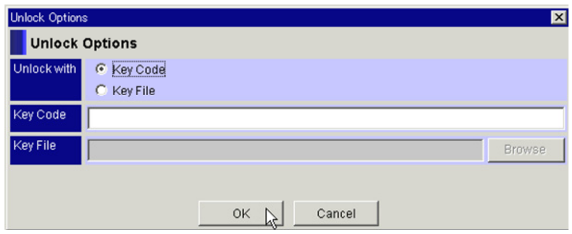
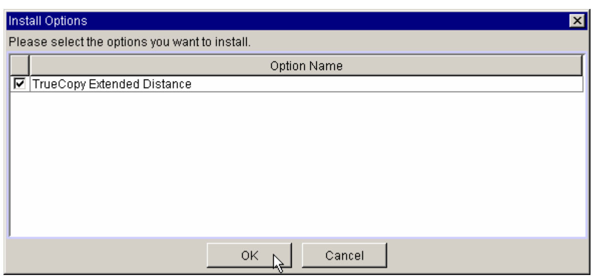


Figure 4.2 Install/Unlock Options Dialog

6. If you install the option using the key code, the **Unlock Options** dialog displays a list of available options in the **Option Name** field. Check the TrueCopy Extended Distance option and click the **OK** button.

Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

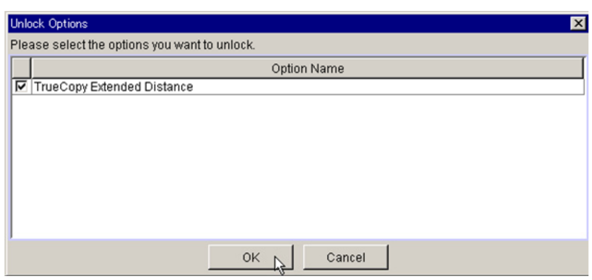
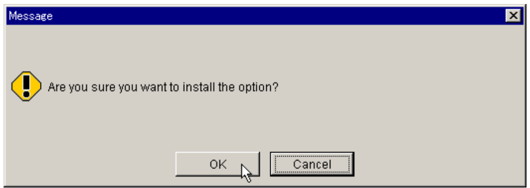


Figure 4.3 Options Selection Dialog

7. An install confirmation dialog is displayed. Select the **OK** button.

   Navigator: Version 5.00 or later



   Navigator: Less than 5.00 version

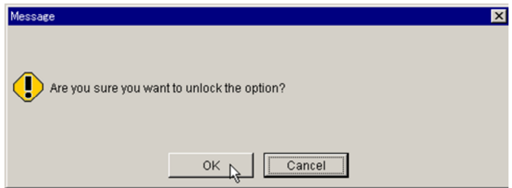
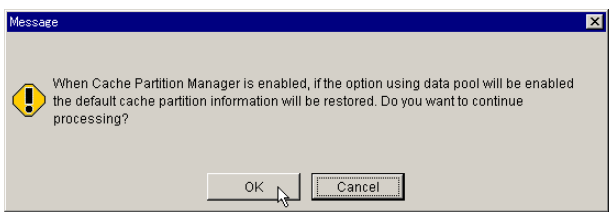


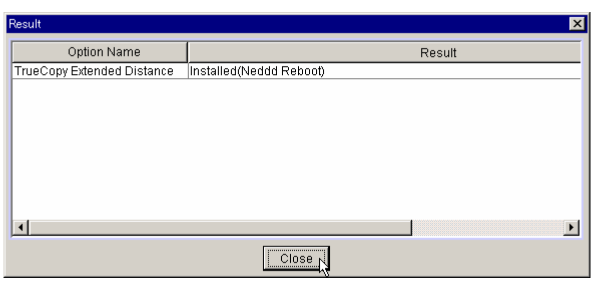
Figure 4.4 Install/Unlock Confirmation Request Message

8. When Navigator version is 3.00 or later and Cache Partition Manager is enabled, the following message is displayed. TCE uses the data pool. When the data pool is used, the information on the cache partition is returned to the default once because a part of the cache memory is used for the internal resource management. When there is no problem, click the **OK** button. When you do not want to lose the configuration information on the partition immediately, click the **Cancel** button and install the TCE function again after checking the information regarding the partition.



9. If you unlock the option using the key file, the **Result** dialog is displayed the TCE feature status as unlocked. Click the **Close** button.

   Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

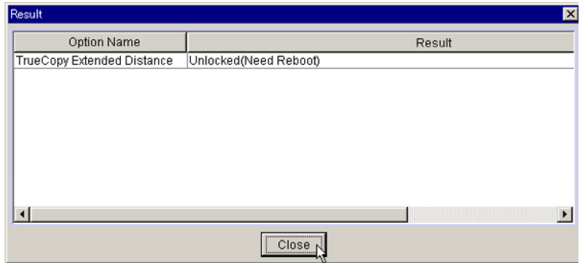
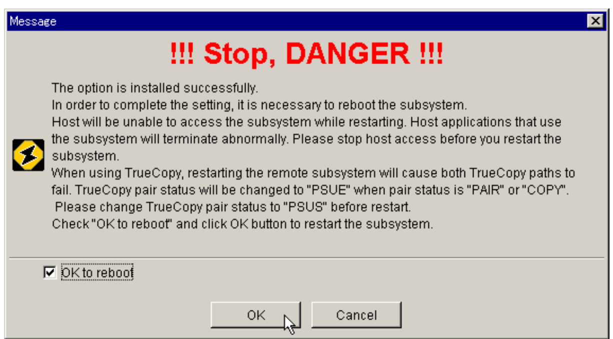


Figure 4.5 Result Dialog

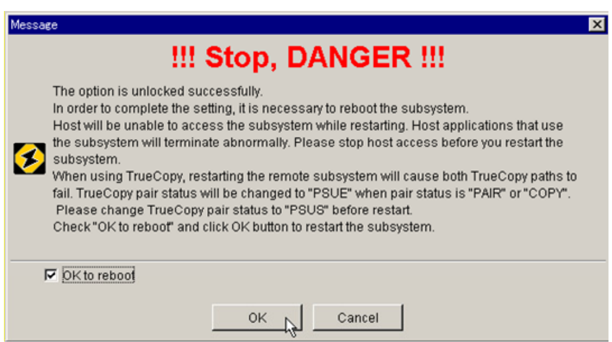
10. An install confirmation message is displayed along with an option to restart the subsystem. Select the **OK** button to restart the subsystem.

    It is necessary to ensure that the host has stopped accessing data on the subsystem before executing a restart, since the subsystem cannot access the host until the restart is completed. It is also necessary to ensure the subsystem is in a normal state before installing or un-installing TCE. Installation or un-installation cannot be performed if a failure such as a controller blockage has occurred.

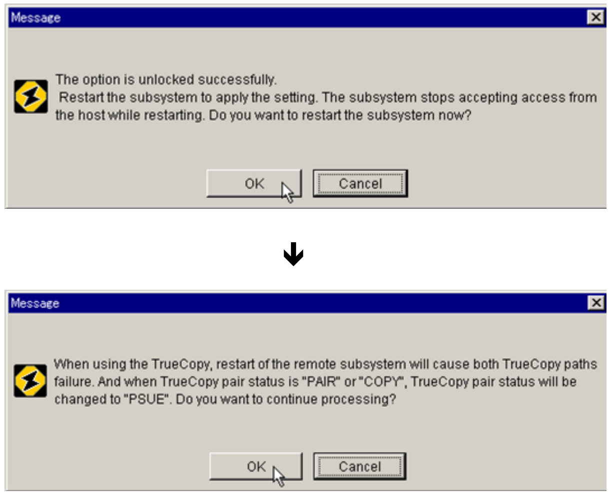
    For Navigator: Version 5.00 or later



For Navigator version 4.03 or later



For Navigator version earlier than 4.03 (unlocked using key code)



For Navigator version lower than 4.03 (unlocked with key file)

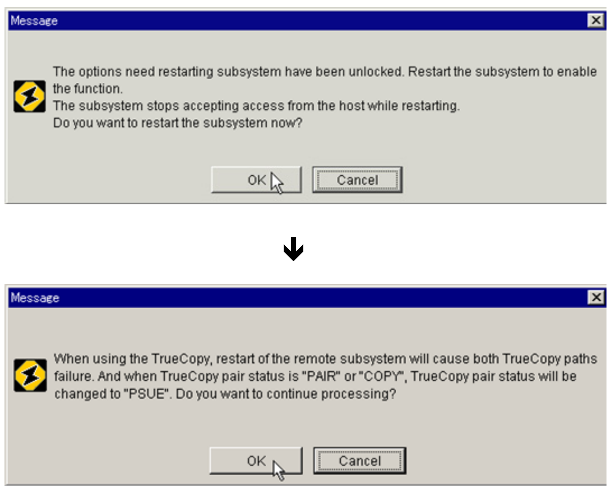
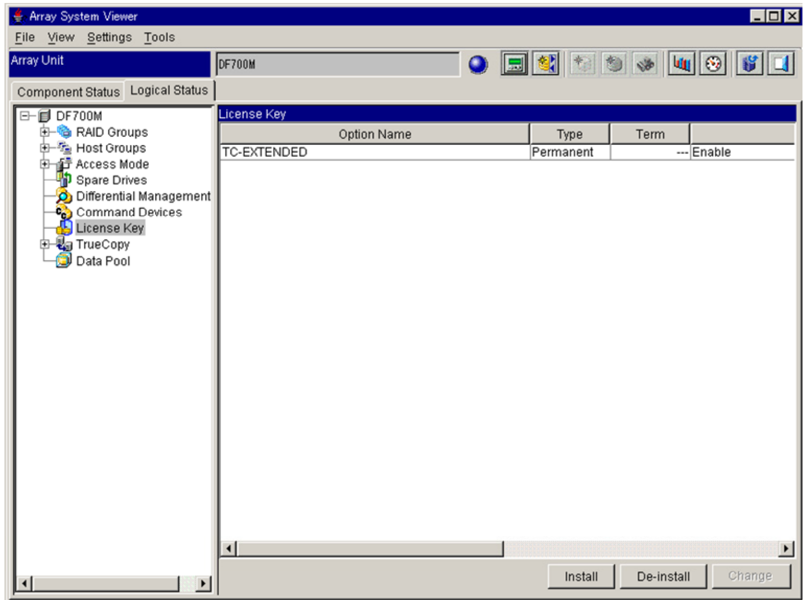


Figure 4.6 Unlock Confirmation Message

Note: If Copy-on-Write SnapShot software was installed prior to TCE installation, the restart request will not be displayed. Copy-on-Write SnapShot installation involves a subsystem reboot to allocate resources for the data pool, eliminating the need for a restart during TCE installation. In such a case subsystem restart (and validation of the TCE feature) may be deferred to a later time. Previous subsystem settings remain valid until a subsequent restart is executed.

11. When a subsystem restart is deferred, the Array System Viewer window is displayed instead of the unlock confirmation message. It displays the enabled TCE option in the License key panel.

Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

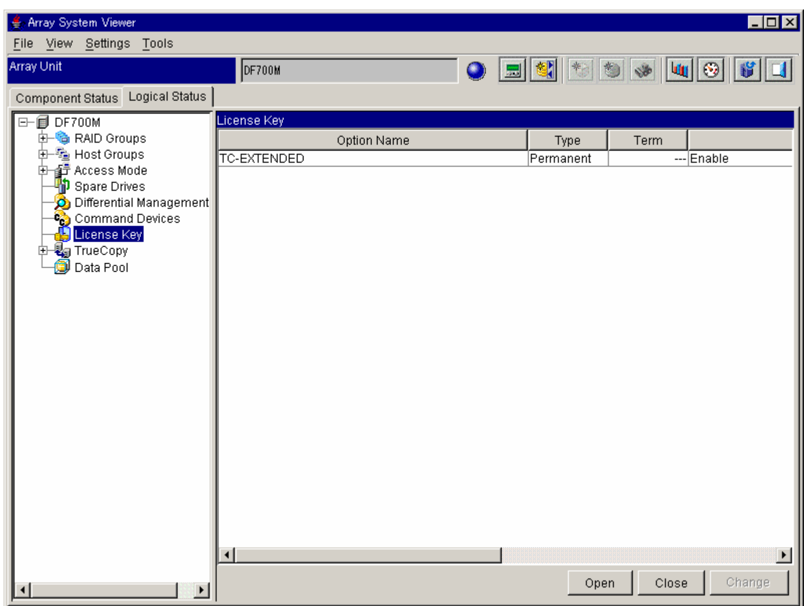


Figure 4.7 Array System Viewer Panel (Logical Status Page Option: Enable)

12. If you choose to restart the subsystem, the time the restart began is displayed in a **Reboot** dialog. Restarting takes approximately four to fifteen minutes.

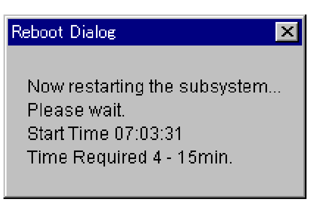


Figure 4.8 Reboot Message

Note: It may take up to 15 minutes for the subsystem to respond, depending on the configuration of the subsystem.

13. A message stating that the restart was successful is displayed. Click the **OK** button to complete TCE installation.

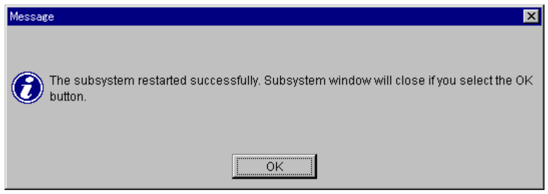


Figure 4.9 Restart Successful Message

- ***Note:*** For details on TCE installation when Cache Partition Manager is used in the system environment, see Appendix C. For detailed information of Remote Copy Connection Hardware installation, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00.*

4.1.2 Uninstalling TCE

Uninstallation of the TCE option requires the license key used for installation. The following conditions must be ensured before uninstallation:

– All TCE pairs must be released (the status of all LUs must be SMPL).

– All the Data Pools must be unallocated, except when SnapShot is still in use.

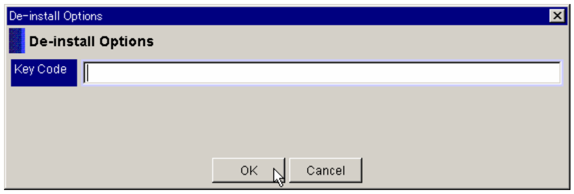
– Path settings must be released.

To uninstall TCE:

1. Start the Navigator, and change the operation mode to **Management Mode** (administrator mode).
2. Register the subsystem in which TCE is to be uninstalled, and then connect to the subsystem. The Array System Viewer panel opens displaying the connected subsystem.
3. Select the **Logical Status** tab and select the **License Key** icon (see Figure 4.1).
4. Select the **De-install** button. (Navigator: Version 5.00 or later)

   Select the **Close** button. (Navigator: Less than 5.00 version)
5. Select the **Close** button. The Lock Options dialog is displayed. Enter the license key code in the text box and click the OK button.

   Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

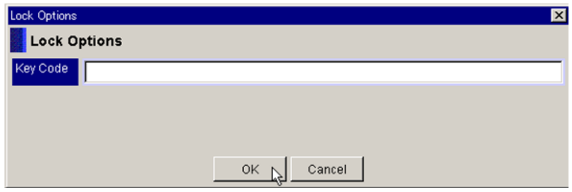
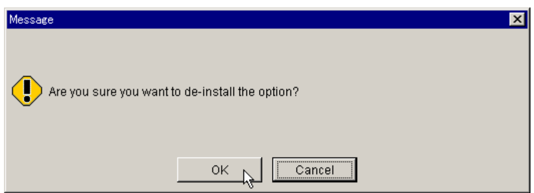


Figure 4.10 De-install/Lock Options Dialog

6. A message appears, requesting confirmation to lock the TCE option. Select the OK button.

   Navigator: Version 5.00 or later



Navigator: Less than 5.00 version

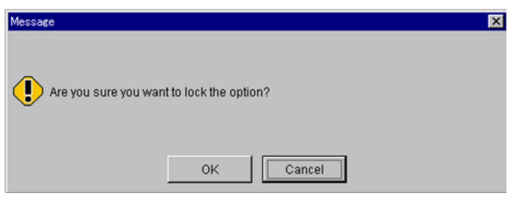


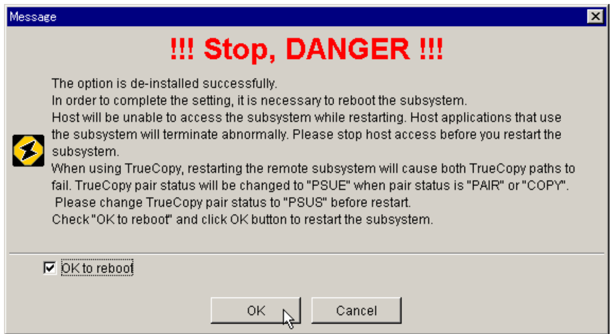
Figure 4.11 De-install/Lock Confirmation Request Message

7. A message appears, confirming that the TCE feature is locked. It also requests confirmation to restart the subsystem. A subsystem restart should be executed in order to validate the locking of the TCE feature. Click the **OK** button to restart.

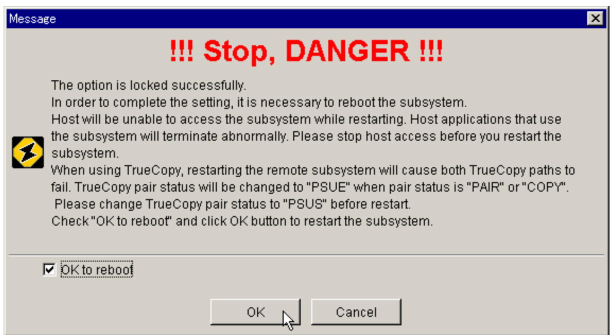
   Note: The restart confirmation dialog is not displayed under the following circumstances:

   - If a restart was performed at the time of invalidating the function, the restart confirmation dialog is not displayed.
   - If Copy-on-Write SnapShot is still in use at the time of TCE un-installation.

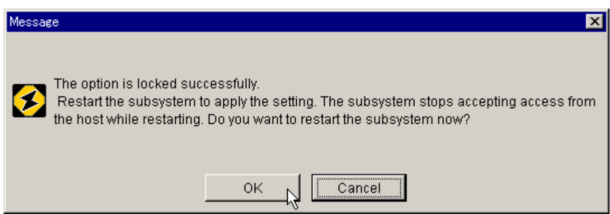
For Navigator: Version 5.00 or later



For Navigator version 4.03 or later



For Navigator version earlier than 4.03 (unlocked using key code)



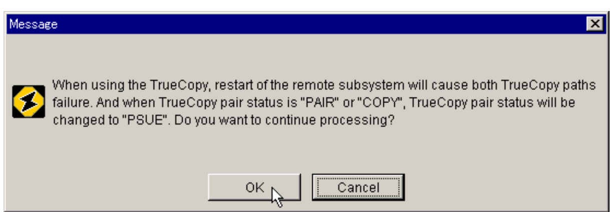


Figure 4.12 Lock Confirmation Message

When you choose to restart the subsystem, a **Reboot message** showing the time restart began is displayed. Restarting takes approximately four to fifteen minutes. It may take up to 15 minutes for the subsystem to respond, depending on the configuration of the subsystem.

When restarting is completed, another message appears stating that the restart was successful. Click the **OK** button to complete TCE uninstallation.

8. When a subsystem restart is deferred, the Array System Viewer window is displayed instead of the lock confirmation message. It displays the disabled TCE option in the License key panel. Click the **Close** button to close the window and complete TCE is uninstallation.

4.1.3 Enabling or Disabling TCE

The default status of the TCE option upon installation is set as “enable”. The status of TCE option can be explicitly set to “enable” or “disable” after installation, by selecting the TCE option and clicking the **Change** button on the Array System Viewer window. For more details on the enabling and disabling procedure, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00*.

The following conditions must be satisfied in order to disable TCE.

– All TCE pairs must be released (status of all LUs should be SMPL).

– All the Data Pools must be deleted, except when Copy-on-Write SnapShot is still in use.

– Path settings must be released.

Note: A subsystem restart is necessary to effect the changes made to TCE option status.

4.2 Configuring a TCE System

TCE system configuration is performed using Storage Navigator Modular software (SNM) and Command Control Interface software (CCI). The following TCE setup tasks should be performed using Storage Navigator Modular software:

- Unlocking the TCE option using the license key and restarting the subsystem (as discussed in Section 4.1.1)
- Setting the Host Group Option
- Setting the Command Device
- Setting the Differential Management LU (DM-LU)
- Confirming Remote Side Logical Unit
- Setting the Data Pool (POOL)
- Setting TCE Remote Path
- Setting the Cycle Time

In addition to the above tasks, the following TCE setup tasks must be performed on the host using CCI:

- Defining the Configuration Definition File
- Setting the Environment Variable

Completion of the above configuration tasks is necessary for TCE functions to become operational. The following sections describe each of the configuration tasks listed above.

4.2.1 Creating Logical Units

Creation of logical units involves specifying a LU mapping mode and setting the Target ID (LU mapping). The following sections describe procedures for the above two tasks.

4.2.1.1 Specifying Mapping Mode

To specify a mapping mode,

1. On the Array System Viewer panel, select the **Logical Status** tab.
2. Select the **Access Mode** plus signs next to the **Mapping Mode**.

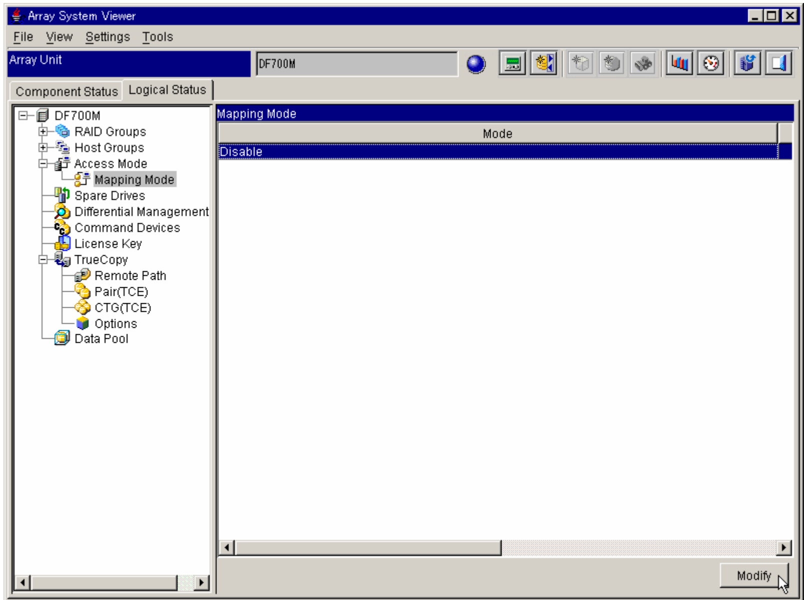


Figure 4.13 Array System Viewer Panel (Specifying Mapping Mode)

3. Select the **Disable** mode in the **Mapping Mode** list, and click the **Modify** button. The **Mapping Mode** dialog is displayed.

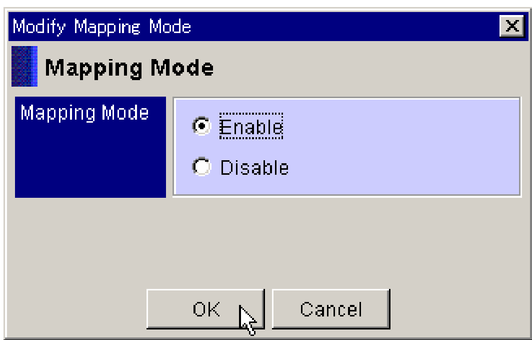


Figure 4.14 Mapping Mode Dialog

4. Select the **Enable** radio button, and click the **OK** button.
5. A confirmation message appears. Click the **OK** button thrice.

4.2.1.2 Setting Mapping Information

To set the mapping information,

1. On the Array System Viewer panel, select the **Logical Status** tab.
2. Expand the **Port 0A** folder and select **000:G000**.
3. Select the **Logical Unit** icon, and click the **Modify Mapping** button.

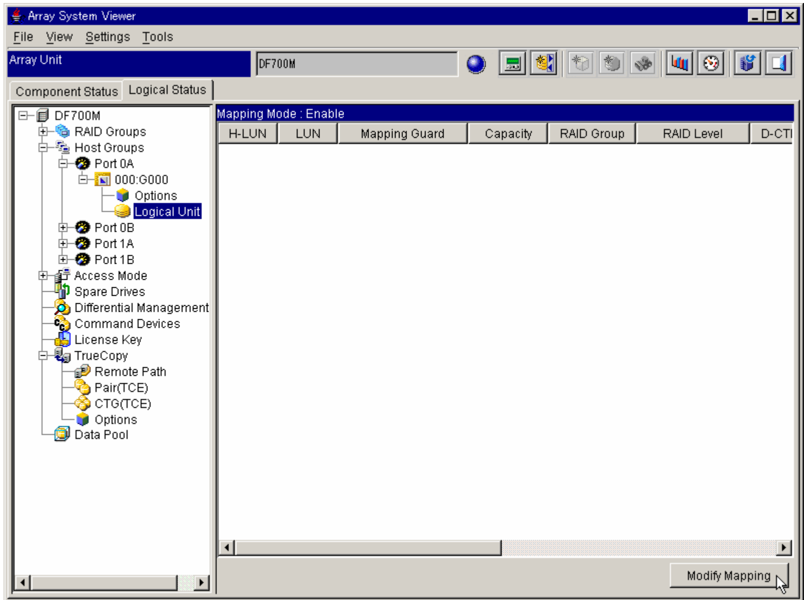


Figure 4.15 Array System Viewer Panel (Setting Mapping Information)

The Mapping Property panel is displayed.

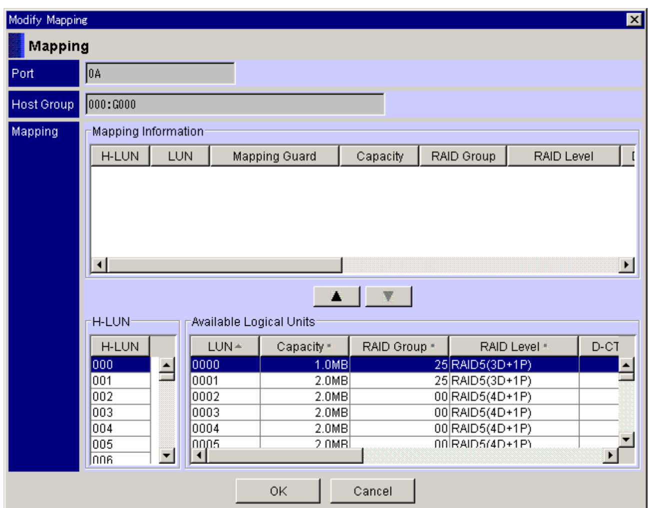


Figure 4.16 Mapping Property Panel (before mapping)

4. Select one H-LUN to be added. Select LUN, and click the ▲ button. The added contents are displayed in Mapping Information list.

   Select the following items:

   –For H-LUN, select the LU number that the host can recognize.

   –For LUN, select the LU number of the subsystem.

   Click the **OK** button.

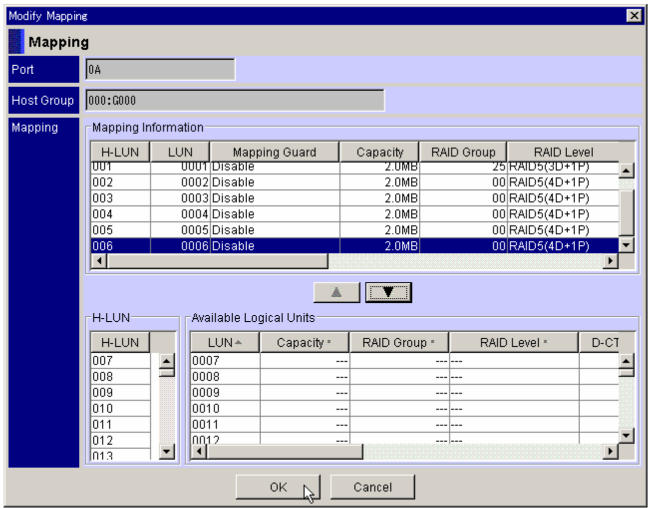


Figure 4.17 Mapping Property Panel (After Setting)

5. A confirmation message is displayed. Click the **OK** button twice.
6. The set mapping information is updated and displayed in the Array System Viewer window.

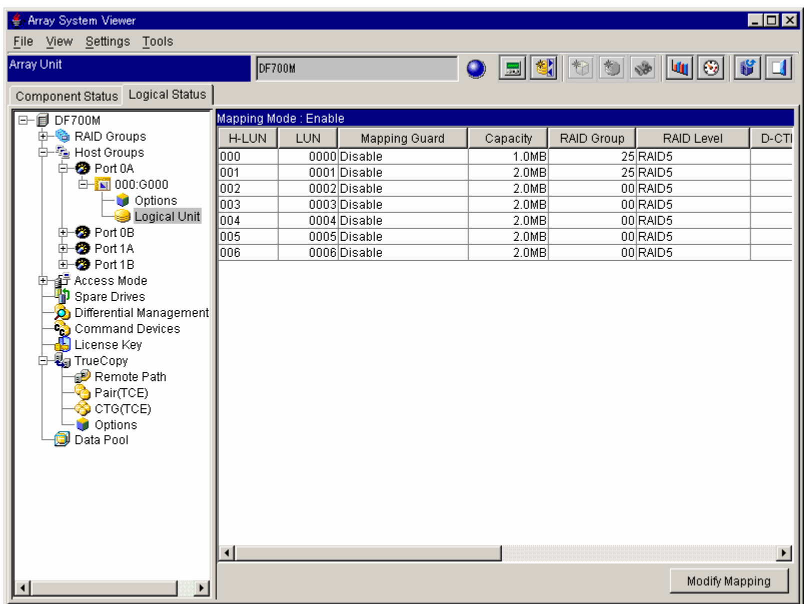


Figure 4.18 Array System Viewer Showing Mappings

4.2.2 Setting the Host Group Option

The host group option must be set when MC/Service Guard is used on a HP server. To set the host group option,

1. On the Array System Viewer panel, select the **Logical Status** tab.
2. Expand the **Port 0A** folder and select **000:G000**.

3. Select **Options** folder under **000:G000**.
4. Click the **Detail Setting** button.

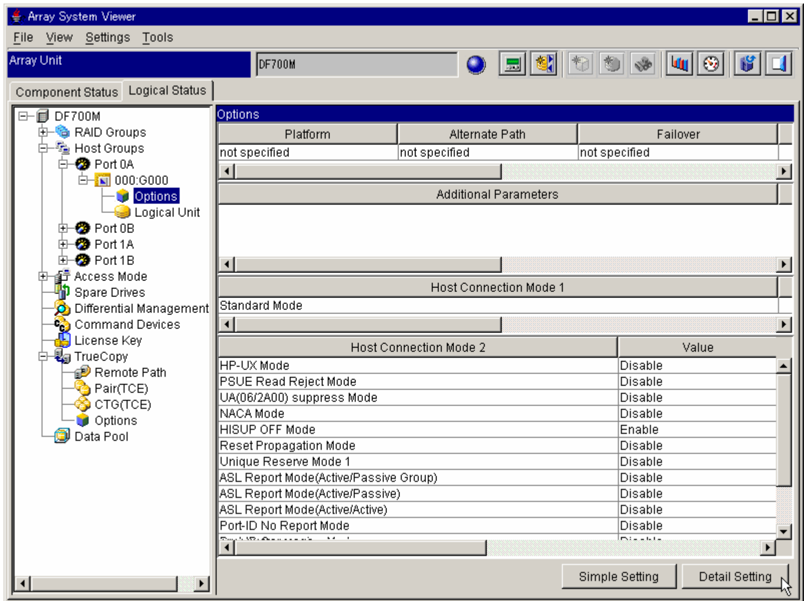


Figure 4.19 Array System Viewer Panel (Setting the Host Group Option)

5. Click the check box of the **PSUE Read Reject Mode** to enable option under the **Host Connection Mode 2** setting item. Then click the **OK** button.

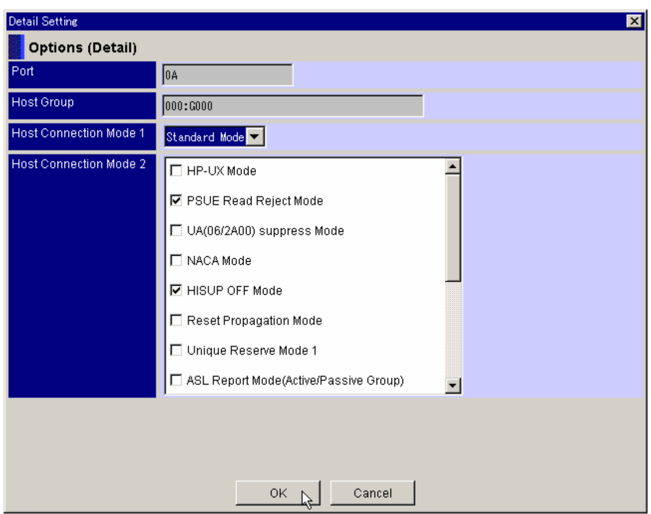


Figure 4.20 Selecting the Host Connection Mode

6. Click the **OK** button thrice in the message displayed, to complete setting the host group option.

4.2.3 Setting the Command Device

Command devices are user-specified, dedicated logical volumes on the TagmaStore AMS subsystems that function as interfaces to the CCI software. TCE commands are issued by the CCI to the TagmaStore AMS subsystem through the command devices. Command devices must be designated on both the local and remote subsystems, in order to accept read and write commands that are to be executed by the subsystem and return read requests to the UNIX®/PC host. The LUs designated as command devices must be recognized by the host. Command devices must be defined in the HORCM_CMD section of the configuration definition file for the CCI instance on the attached host. Two command devices may be set for each TagmaStore AMS subsystem.

To designate command device(s):

1. Start the Navigator, and change the operation mode to **Management Mode** (administrator mode).
2. Connect to the subsystem. The Array System Viewer panel opens and displays the connected subsystem.
3. Select the **Logical Status** tab and click on the **Command Devices** icon.

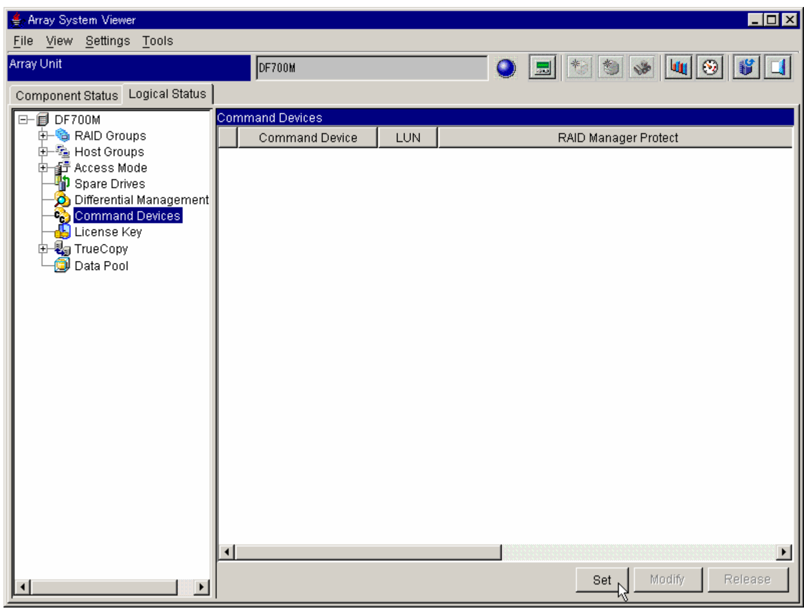


Figure 4.21 Array System Viewer Panel (Command Device Page: Before Setting)

4. Select the **Set** button. The **Command Devices Settings** dialog is displayed.

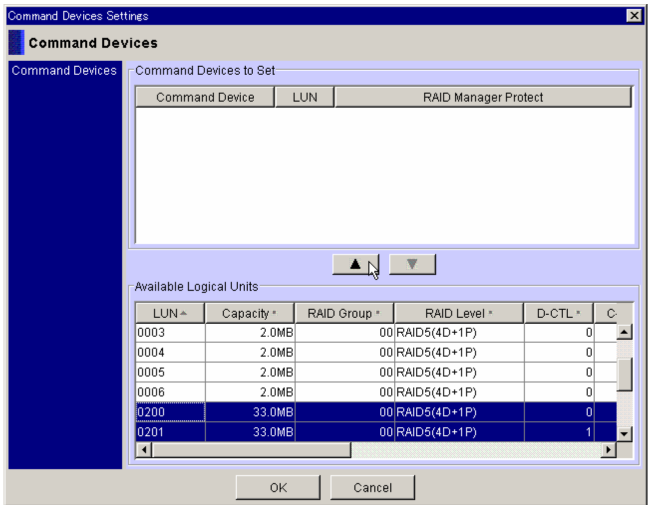


Figure 4.22 Command Devices Settings Dialog (Before Setting)

5. In **Available Logical Units** list, select the **LUN** you want to set as the command device, and click the ▲ button. The selected **LUN** moved to the **Command Devices to Set** list.

   To change logical volumes designated as command devices, select the **LUN** on the **Command Devices to Set** list, and click the ▼ button. The selected **LUN** is moved to the **Available Logical Units** list.

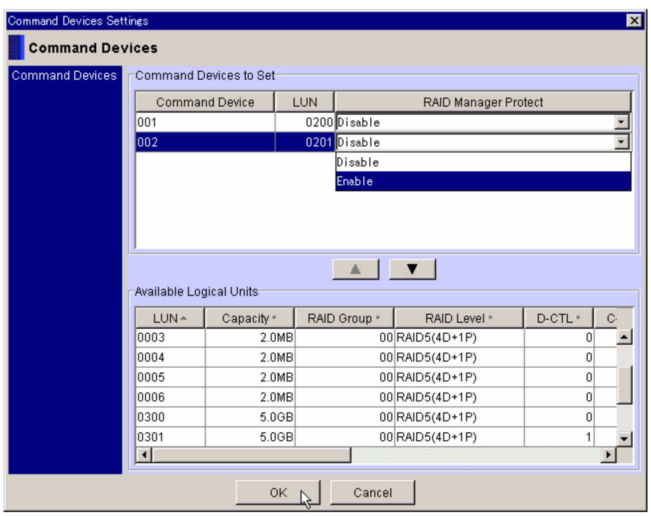


Figure 4.23 Command Devices Settings Dialog (After Setting)

6. If you want to use the protection function of CCI, in the **RAID Manager Protect** drop-down list, select **Disable** or **Enable**.

7. Select the **OK** button. The updated command device settings are displayed in the Array System Viewer panel.

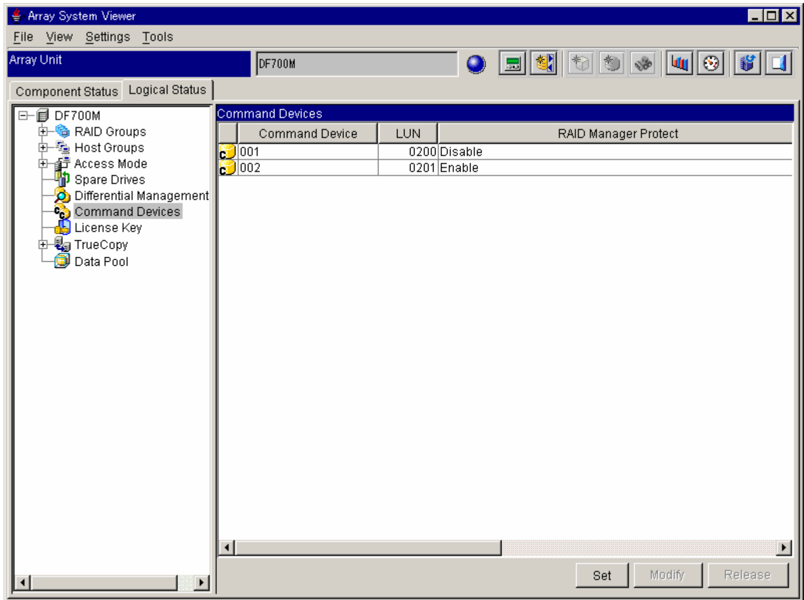


Figure 4.24 Array System Viewer Panel (Command Device Page: After Setting)

Note: To use the alternate command device function, or to avoid data loss and subsystem downtime, it is recommended that two command devices be designated for each TagmaStore AMS subsystem. If the two command devices are set within one disk subsystem, it is necessary to assign them to the different RAID groups to avoid both command devices becoming unavailable at the same time owing to drive failures. For details on the alternate command device function, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

8. A message confirming the setting of the command device is displayed. Click the **OK** button to complete designating the command device.
9. To release an already set command device, select the **LUN** of the command device you want to release, and select the **Release** button.

   Note: The following command device release restrictions apply when either ShadowImage or SnapShot coexists with TCE or when a TCE path is defined:

   - If two command devices are set, only one command device can be released.

   - If only one command device is set, that command device cannot be released.

4.2.4 Setting the Differential Management LU (DM-LU)

The Differential Management LU is an exclusive logical unit for storing the differential data at the time when the volume is copied. The Differential Management LU in the disk subsystem is treated in the same as other logical units. However, a logical unit that is set as the Differential Management LU is hidden and not recognized by a host.

Differential Management LUs must be set on both the local and remote subsystems and must be at least 5GB in size. It is recommended that two Differential Management LUs be set for each subsystem. The second Differential Management LU is used for mirroring.

To designate Differential Management LUs:

1. On the Array System Viewer panel, select the **Logical Status** tab.
2. Select the **Differential Management LU** icon.

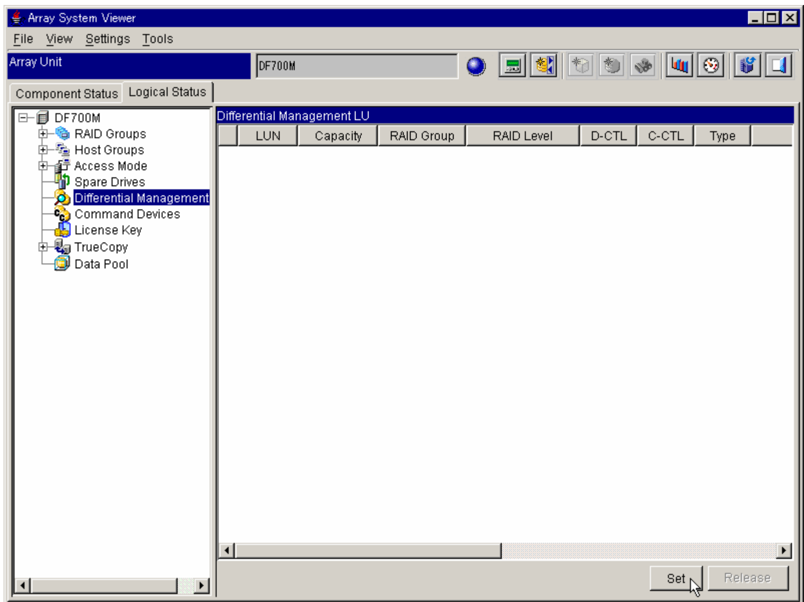


Figure 4.25 Array System Viewer Panel (prior to setting DM-LU)

3. Click the **Set** button. The **Select Logical Unit** dialog is displayed.

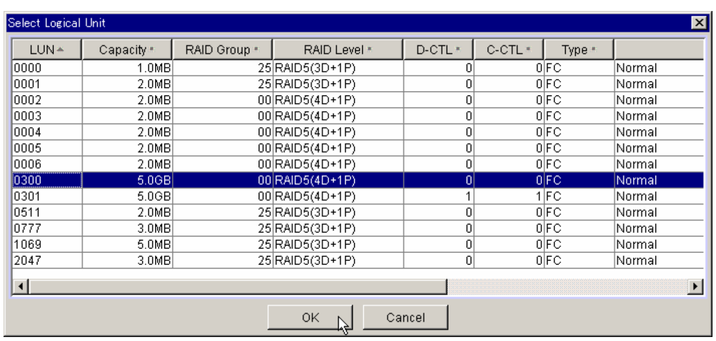


Figure 4.26 Select Logical Unit Dialog

4. Select the **LUN** you want to set as the Differential Management LU, and click the **OK** button.
5. A message confirming the designated DM-LU is displayed. Click the **OK** button. The DM-LU setting information is displayed in the Array System Viewer panel.

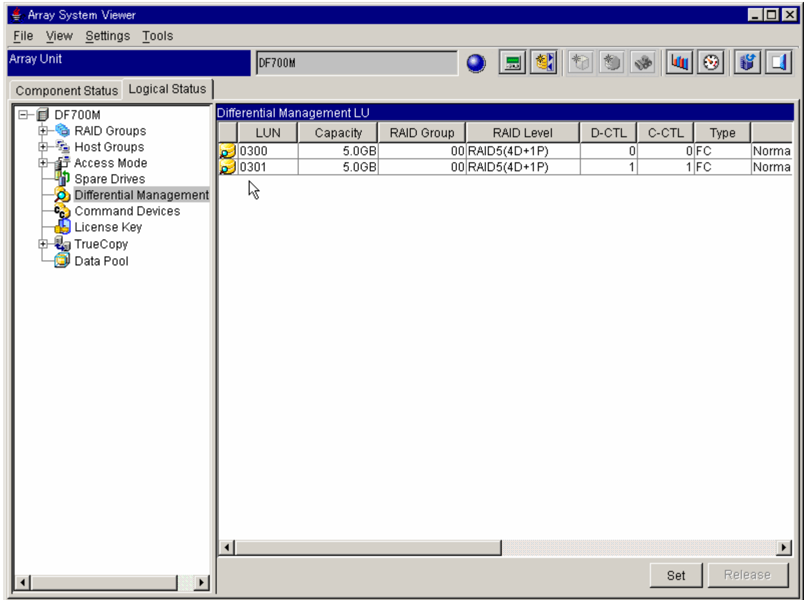


Figure 4.27 Array System Viewer Panel (after setting DM-LU)

Note: The following DM-LU release restrictions apply when ShadowImage or SnapShot coexists with TCE, or a TCE path is defined or when a SnapShot data pool is allocated:

- If two DM-LUs are set, only one DM-LU can be released.

- If only one DM-LU is set, that DM-LU cannot be released.

4.2.5 Confirming Remote Side Logical Unit

It is necessary to ascertain that the numbers of the default controller (D-CTL) and the current controller (C-CTL) controlling the logical unit to be assigned to the S-VOL are the same. A TCE pair cannot be created amongst logical units unless the D-CTL and C-CTL numbers match.

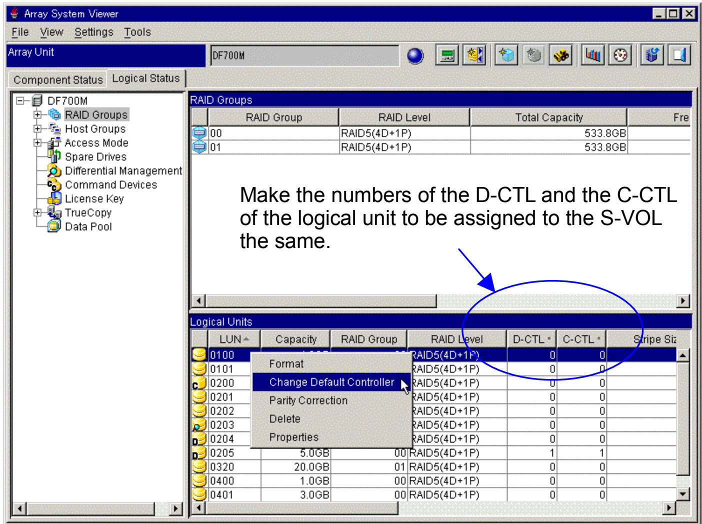


Figure 4.28   Confirming Remote Side Logical Unit

4.2.6 Setting the Data Pool (POOL)

A data pool is created on each subsystem by allocating formatted logical units of size 20GB or more. Up to 64 logical units can be assigned to each data pool. Those logical units that are assigned to a data pool become unrecognized by a host.

Accurate capacity of a newly created data pool cannot be determined immediately after allocating logical units. Pool capacities are reflected every 3 minutes per 100GB. The following restrictions apply when allocating logical units for the data pool:

- Logical units consisting of FC drives cannot coexist in a data pool with those consisting of SATA drives.
- When using TCE with Cache Partition Manager, the segment size of the LU belonging to POOL must be of the default size (16KB) or less.

To designate logical units for a Data Pool (s):

1. On the Array System Viewer panel, select the **Logical Status** tab. The **Logical Status** page will be displayed.
2. Select the **Data Pool** icon in the left pane. The **Data Pool Information** will be displayed in the right pane.
3. Select the **Data Pool 0** or **Data Pool 1**. Select **Add** button. (**Data Pool 0** is for Controller 0 and **Data Pool 1** is for Controller 1.)

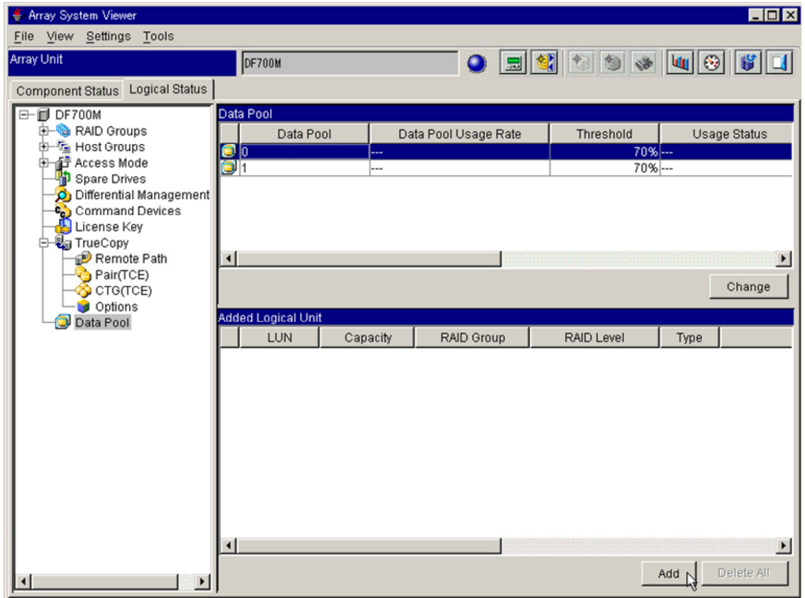


Figure 4.29 Setting the Data Pool

4. The **Add Logical Unit** dialog is displayed. Select a logical unit for the Data Pool (POOL) and click the **OK** button.

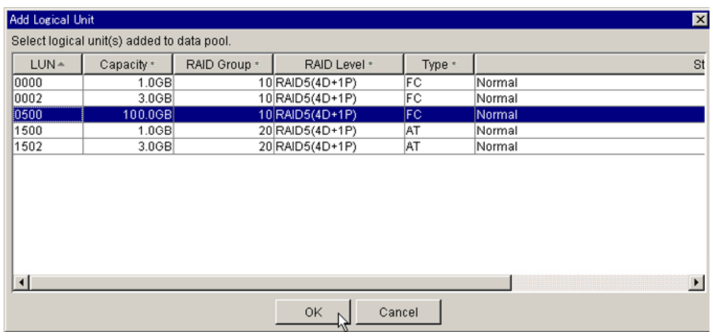
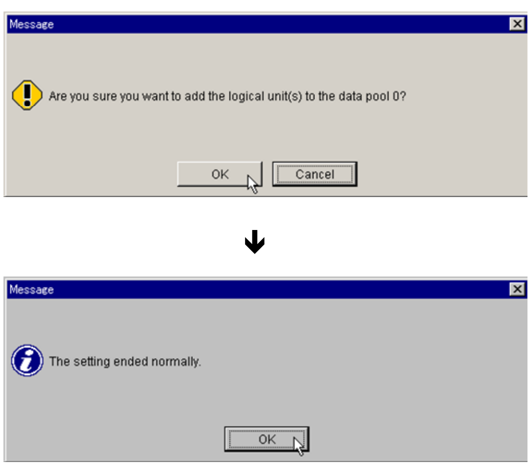


Figure 4.30 Add Logical Unit Dialog

5. A message appears, asking for confirmation to add the Data Pool (POOL). Select the **OK** button.



The added Data Pool (POOL) is updated and displayed in the Data pool panel.

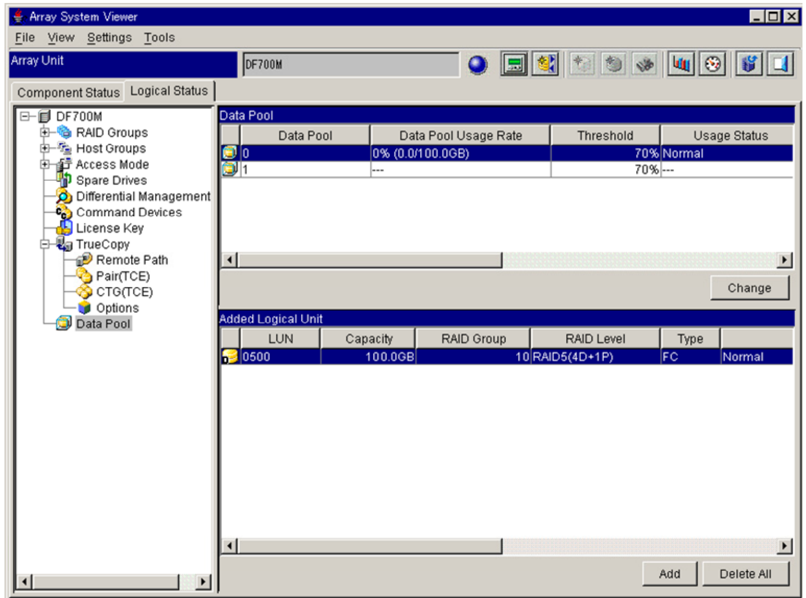
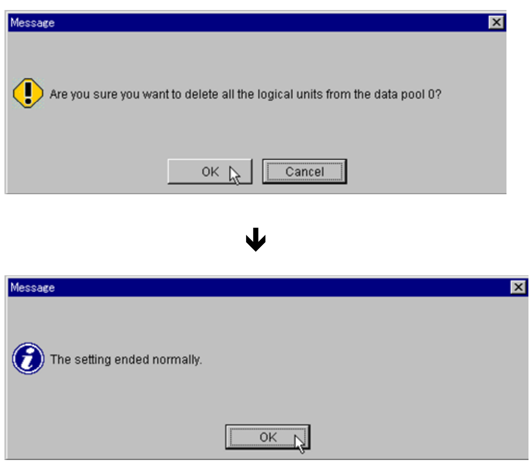


Figure 4.31 Array System Viewer Data Pool Panel

6. To allocate another logical unit to the Data Pool, repeat steps 3 to 5.
7. To delete logical units assigned to the Data Pool, select the **LUN** on the **Added Logical Unit** panel and click the **Delete All** button.
8. A message appears, asking you to confirm that you want to delete the Data Pool (POOL). Select the **OK** button.



9. To change the threshold value of the Data Pool, click the **Change** button (see Figure 4.31. The **Property** dialog is displayed. Specify a threshold value between 50 and 80 (default value is 70). Click the **OK** button.

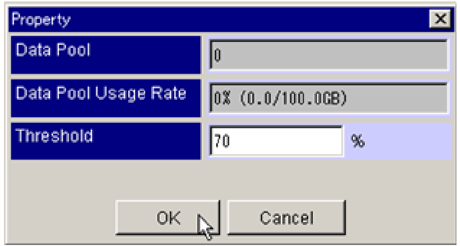


Figure 4.32 Property Dialog

4.2.6.1 Determining Pool Capacity

P-VOL pool size:

The P-VOL pool size is computed using the following formula:

> [P-VOL pool size] = ([Peak inflow speed] – [Peak outflow speed]) × [Peak period] ÷ [Average I/O size] × 64 × [Margin rate]

The peak outflow speed is the data transfer speed at which the local TagmaStore AMS subsystem can perform processing at peak IOPS. If I/O is excessive, I/O processing overtakes the data transfer. As a result, the outflow speed becomes slower than the average outflow speed.

At the peak IOPS, differential data stays in the local TagmaStore AMS subsystem at a speed of (Peak inflow speed – Peak outflow speed). The differential data increases during the peak period. Consequently, differential data stays at the local TagmaStore AMS subsystem for the duration of ([Peak inflow speed] – [Peak outflow speed]) × [Peak time].

Data that has not been transferred to a pool is saved in 64-KB units, regardless of the I/O size. As a result, more data is saved to the pool than the inflow amount. For this reason, the differential data quantity is calculated as 64 KB. A margin is provided for changes in the peak inflow speed, peak period, or outflow speed.

S-VOL pool size

The differential data generated on the P-VOL is copied to the S-VOL pool. Therefore the S-VOL pool size is determined in the same way as the P-VOL pool size.

> [S-VOL pool size] = ([Peak inflow speed] – [Peak outflow speed]) × [Peak period] ÷ [Average I/O size] × 64 × [Margin rate]

Note: Since SnapShot is used with TCE, a much larger pool capacity will be required since used capacity of a pool does not decrease for each TCE update cycle.

4.2.7 Setting TCE Remote Path

TCE operations are performed between the subsystems that are connected by a fibre channel interface. A route called a "path" connects the port from the local subsystem that executes the volume replication to the port on the remote subsystem.

Two paths must be set for each disk array subsystem (one path for each of the two controllers built in the subsystem). The use of two controllers and two paths ensure that copy operations are uninterrupted in the event of a path failure. The ports of a controller in a subsystem, may be connected only to the identical ports of the corresponding controller on the remote subsystem. Figure 4.33 shows examples of a permitted configuration and a disallowed configuration.

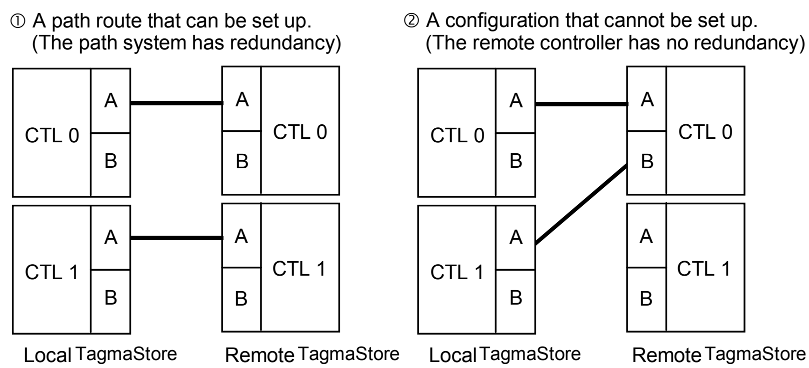


Figure 4.33 Path Configurations

To specify TCE paths:

1. Start the Navigator and change the operation mode to **Management Mode** (administrator mode).
2. Connect to the subsystem. The Array System Viewer panel opens displaying the connected subsystem.
3. Select the **Logical Status** tab.
4. Select the **Remote Path** icon under the **TrueCopy** icon. The **TrueCopy** icon will be displayed only when the TCE option is enabled.

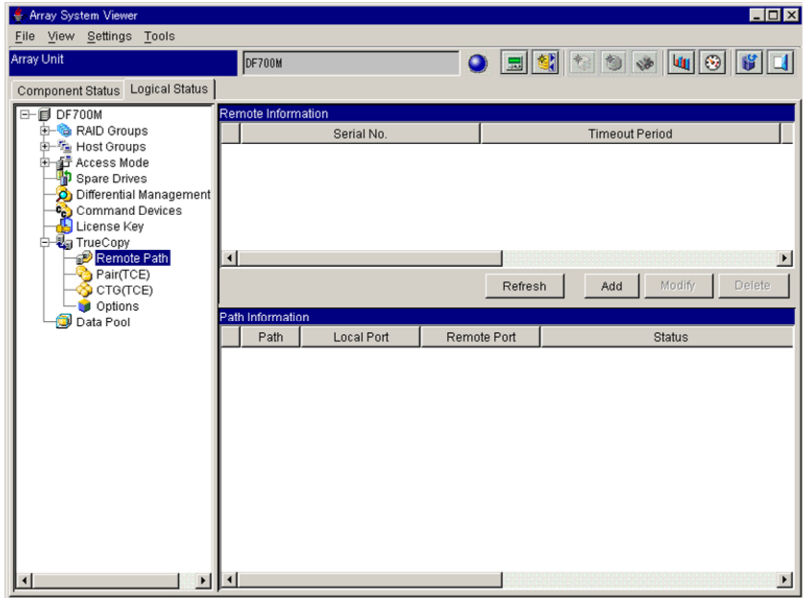


Figure 4.34 Array System Viewer Panel (Remote Path Page: Before Setting)

5. Select the **Add** button. The **Remote Path** setting panel is displayed.

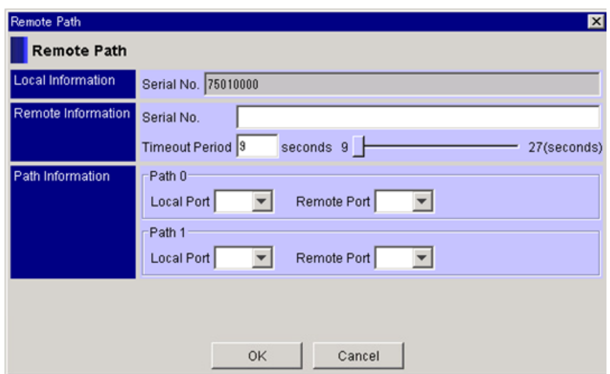


Figure 4.35 Remote Path Setting Panel

6. On the Remote Path setting panel, enter the following **Local information** and **Remote information** data:

- Enter an eight digits number in decimal notation for **Serial No.**
- Enter a value for **Timeout Period**. The default value is 9 seconds. The timeout value may be changed when the remote subsystem is at a long distance (using Extender). The following standard time out values may be set based on the frequency band of the line.

Table 4.1 Standard Timeout Periods

Frequency band of a line	Time limit (seconds)
200 Mbps or more	9
40 to 200 Mbps	18
20 to 40 Mbps	27

For recommendations on setting timeout periods, see Chapter 8

Select the port from the pull-down menus for setting **Path 0 information** and **Path 1 information.** Table 4.2 shows ports to be selected based on type of subsystems.

Table 4.2 Port settings by Subsystem Type

	Local Subsystem Type		Remote Subsystem Type	
	AMS500	**AMS1000**	**AMS500**	**AMS1000**
Path 0 Port	0A or 0B	0A, 0B, 0C, 0D	0A, 0B, 1A, 1B	0A, 0B, 0C, 0D, 1A, 1B, 1C, 1D
Path 1 Port	1A or 1B	1A, 1B, 1C, 1D	0A, 0B, 1A, 1B	0A, 0B, 0C, 0D, 1A, 1B, 1C, 1D

7. Click the **OK** button.
8. If remote path was set successfully, a panel is displayed. Click the **OK** button. The Remote Path panel displays the path that was set.

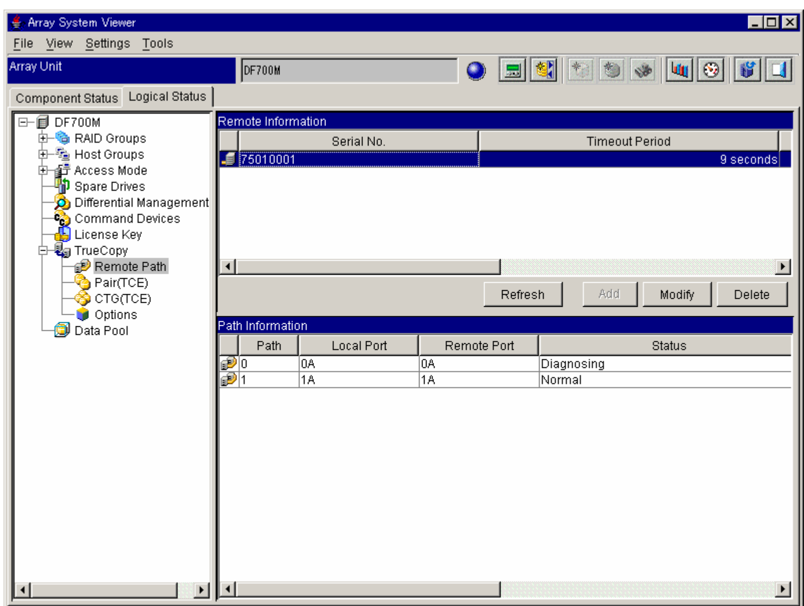


Figure 4.36 Array System Viewer Panel (Remote Path Page: After Setting)

Click the **Refresh** button to see latest path settings. Click the Delete button to delete specific path information.

The following restrictions apply when setting TCE paths:

- Command Device(s) should be designated before any TCE paths are set.
- TCE paths cannot be set when failures such as a controller blockage has occurred.

The following restrictions apply when releasing TCE paths:

- Previously set paths must be released before specifying any changes to the path information.
- All TCE pairs must be released before TCE paths can be released.

4.2.7.1 Confirming Status of Paths

TCE supports a function that periodically issues commands between subsystems and monitors the path status. When a path is blocked due to path failure, its status is reported by an LED or on the Web interface (no status is reported for temporary command error).

You can also verify the path status at any time using Navigator. Path status is displayed in the **Component Status** page of the Array System Viewer Panel. Remote path may have status set as Undefined, Normal, Detached or Diagnosing

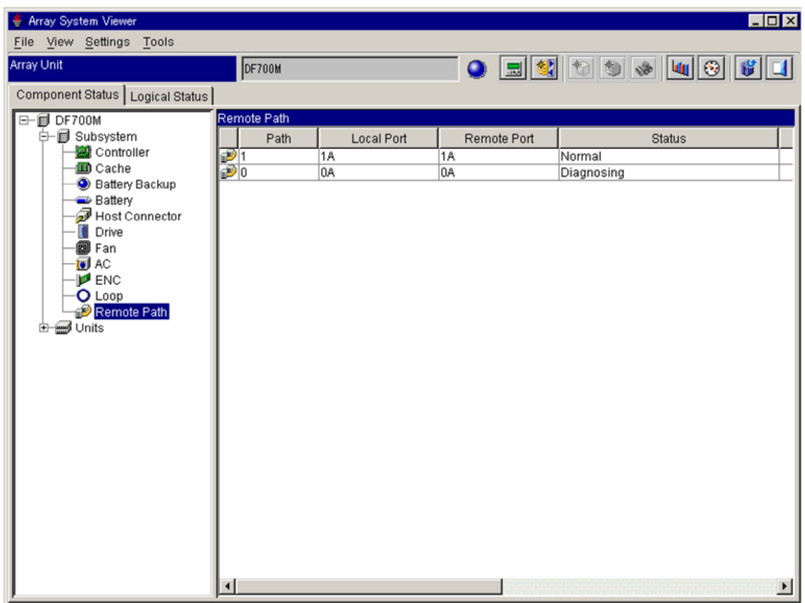


Figure 4.37 Array System Viewer Panel (Component Status Page)

4.2.8 Setting the Cycle Time

Cycle time must be set for each subsystem and determines intervals at which remote copy of differential data is performed. The shortest cycle time that can be set for a subsystem is calculated as number of CTGs × 30 seconds. Remote copy operations may take longer than the cycle time specified by a user, depending on the amount of the differential data or low line speed.

To designate the cycle time:

1. On the Array System Viewer panel, select the **Logical Status** tab.
2. Click on the **TrueCopy** icon and select the **Options** icon.

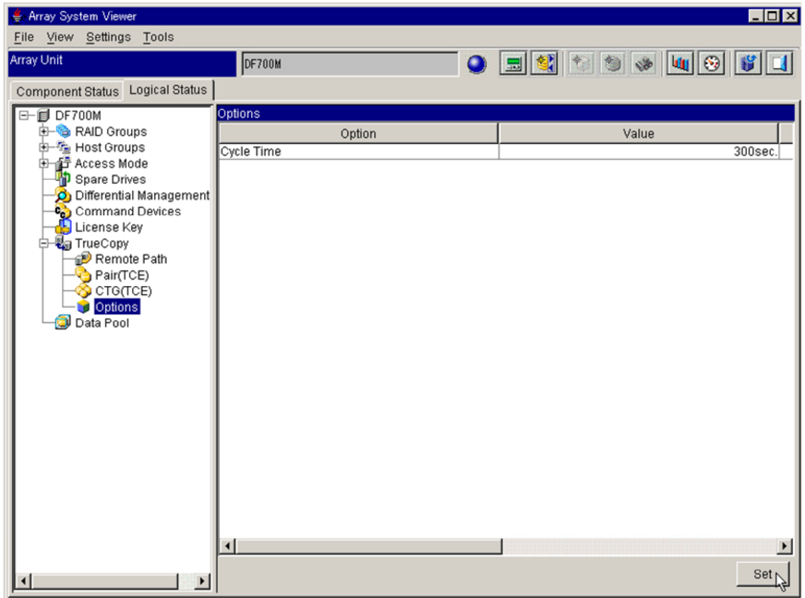


Figure 4.38 Setting the Cycle Time

3. Click the **Set** button. The **Options** dialog is displayed.

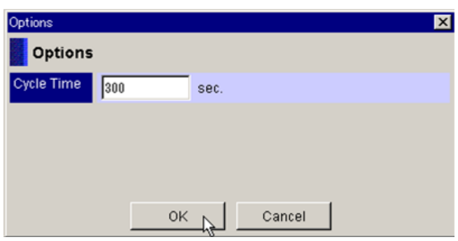
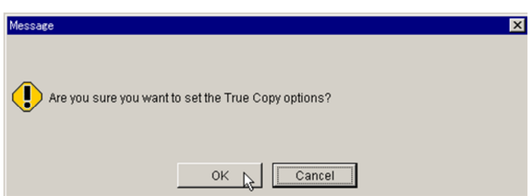


Figure 4.39 Options Dialog

4. Specify the cycle time and click the **OK** button. The cycle time is 300 seconds by default and can be specified in a range from 30 to 3600 seconds.
5. Click the **OK** button in the next message.



4.2.9 Defining the Configuration Definition File

The configuration definition file describes the system configuration for making CCI operational in a TCE environment. The configuration definition file is a text file created and/or edited using any standard text editor, and can be defined from the PC where the CCI software is installed. The sample configuration definition file (HORCM_CONF) included with the CCI software should be used as the basis for creating your custom configuration definition file(s). The sample file should be copied, customized with necessary parameters, and placed in the proper directory by a system administrator. For more details on configuration definition file, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

The configuration definition file can be automatically created using the mkconf command tool. However, the parameters such as poll (10ms) must be set manually (see step 4 below). For details on the mkconf command tool, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

The following describes an example for manually defining the configuration definition file, for a system configuration with two instances within the same server (Windows® NT/2000).

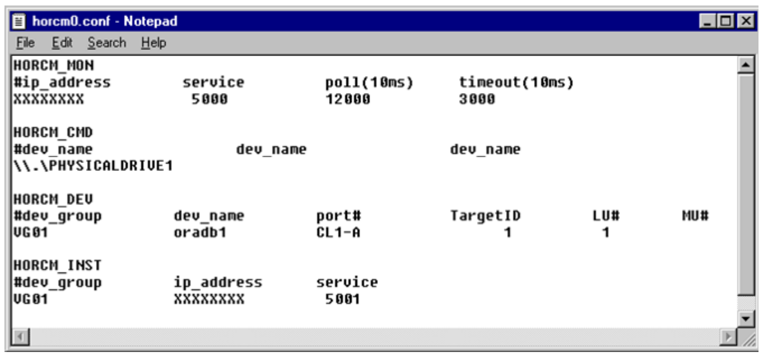
1. On the host where CCI is installed, verify that the CCI is not running. If the CCI software is still running, shut down the CCI software using the horcmshutdown command. For details on horcmshutdown, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).
2. At the command prompt, make two copies of the sample file (horcm.conf).

```
c:\HORCM\etc> copy \HORCM\etc\horcm.conf \WINNT\horcm0.conf
c:\HORCM\etc> copy \HORCM\etc\horcm.conf \WINNT\horcm1.conf
```

3. Open horcm0.conf file using the text editor.
4. In the **HORCM_MON** section, set the necessary parameters.

Note: A value more than or equal to 6000 must be set for poll (10ms). Please refer to the manual *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701) for details on calculating the poll (10ms) value. Specifying the value incorrectly may cause a conflict in the internal process, by suspending the process temporarily and stopping the internal process of the subsystem. For more details on configuration parameters, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

5. In the **HORCM_CMD** section, specify the physical drive (command device) on the subsystem.




```
HORCM_MON
#ip_address        service         poll(10ms)      timeout(10ms)
XXXXXXXX            5000           12000           3000

HORCM_CMD
#dev_name                dev_name                  dev_name
\\.\PHYSICALDRIVE1

HORCM_DEV
#dev_group          dev_name        port#           TargetID        LU#         MU#
VG01                oradb1          CL1-A                 1           1

HORCM_INST
#dev_group          ip_address      service
VG01                XXXXXXXX         5001
```

Figure 4.40 Horcm0.conf Example

6. Save the configuration definition file and use the **horcmstart** command to start the CCI software. For details on **horcmstart**, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

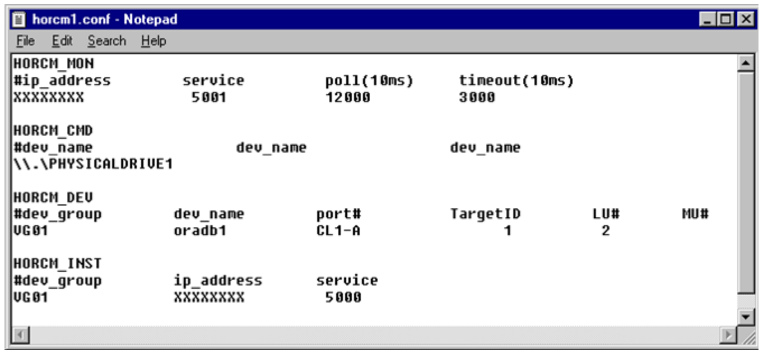
7. Execute the **raidscan** command and make note of the target ID displayed in the execution result. For details on the **raidscan** command, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

8. Shut down the CCI software and then open the configuration definition file again.

9. Set the necessary parameters in the **HORCM_DEV** section. For the target ID, set the ID noted from the **raidscan** result. Do not set the MU# parameter.

10. Set the necessary parameters in the **HORCM_INST** section, and then save (overwrite) the file.

11. Repeat step 3 to 10 for the **horcm1.conf** file (Figure 4.41)



```
horcm1.conf - Notepad
File Edit Search Help
HORCM_MON
#ip_address         service        poll(10ms)     timeout(10ms)
XXXXXXXX             5001          12000          3000

HORCM_CMD
#dev_name                 dev_name                  dev_name
\\.\PHYSICALDRIVE1

HORCM_DEV
#dev_group          dev_name       port#          TargetID       LU#       MU#
VG01                oradb1         CL1-A             1           2

HORCM_INST
#dev_group          ip_address     service
VG01                XXXXXXXX        5000
```

Figure 4.41 Horcm1.conf Example

12. Enter the following in the command prompt to verify the connection between CCI and the subsystem.

```
C:\>cd horcm\etc

C:\horcm\etc>echo hd1-3 | .\inqraid
Harddisk 1 -> [ST] CL1-A Ser =75000174 LDEV =   0 [HITACHI  ]  [DF600F-CM
]
Harddisk 2 -> [ST] CL1-A Ser =75000174 LDEV =   1 [HITACHI  ]  [DF600F
]
              HORC = SMPL  HOMRCF[MU#0 = NONE  MU#1 = NONE  MU#2 = NONE]
              RAID5[Group 1-0] SSID = 0x0000
Harddisk 3 -> [ST] CL1-A Ser =75000174 LDEV =   2 [HITACHI  ]  [DF600F
]
              HORC = SMPL  HOMRCF[MU#0 = NONE  MU#1 = NONE  MU#2 = NONE]
              RAID5[Group 2-0] SSID = 0x0000

C:\horcm\etc>
```

4.2.10 Setting the Environment Variable

To perform TCE operations, you must set the environment variable for the execution environment. The following describes an example assuming a system configuration with two instances within the same server (Windows® 2000).

1. Set the environment variable for each instance. Enter the following at the command prompt.

```
C:\HORCM\etc>set HORCMINST=0
```

2. Execute the **horcmstart** script, and then execute the **pairdisplay** command to verify the configuration.

```
C:\HORCM\etc>horcmstart 0 1
starting HORCM inst 0
HORCM inst 0 starts successfully.
starting HORCM inst 1
HORCM inst 1 starts successfully.

C:\HORCM\etc>pairdisplay -g VG01
group   PairVOL(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status, Fence, Seq#,P-
LDEV# M
VG01    oradb1(L)    (CL1-A , 1,  1 )75000174     1.SMPL ----- ------,-----   -
---  -
VG01    oradb1(R)    (CL1-B , 1,  2 )75000175     2.SMPL ----- ------,-----   -
---  -
```

Chapter 5 Remote Data Replication with TCE

This chapter describes remote data replication workflow and processes in a TCE environment. It provides details of TCE remote replication pair operations and includes sample backup and data transfer scenarios to demonstrate use of TCE functionality. The following sections are included in this chapter:

- TCE Remote Replication Workflow (section 5.1)
- TCE Remote Replication Operations (section 5.2)
- Copy-on-Write SnapShot Operations (section 5.3)
- Sample Remote Replication Scenarios (section 5.4)

5.1 TCE Remote Replication Workflow

A typical TCE remote replication workflow includes the following processes:

- Initial Copy Operation
- Cycle Update Process
- Resynchronization Process

This section discusses the above remote replication processes.

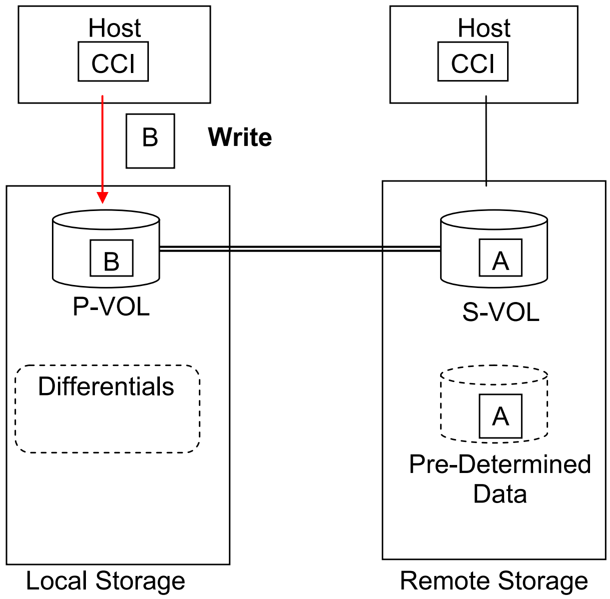
5.1.1 Initial Copy operation

An initial copy operation is performed when data volume pairs are created. The initial copy operation replicates the entire P-VOL data to the S-VOL prior to any update process. The P-VOL remains available to all hosts for read and write operations throughout the initial copy operation. Any updates received by the P-VOL from the host during the initial copy become new differential data; the new differential data is also copied after all existing data is copied. When the difference between the P-VOL and the S-VOL data reaches a pre-defined threshold, the initial copy is completed. For more information on the initial copy operation, see Section 5.2.2 Creating TCE pairs.

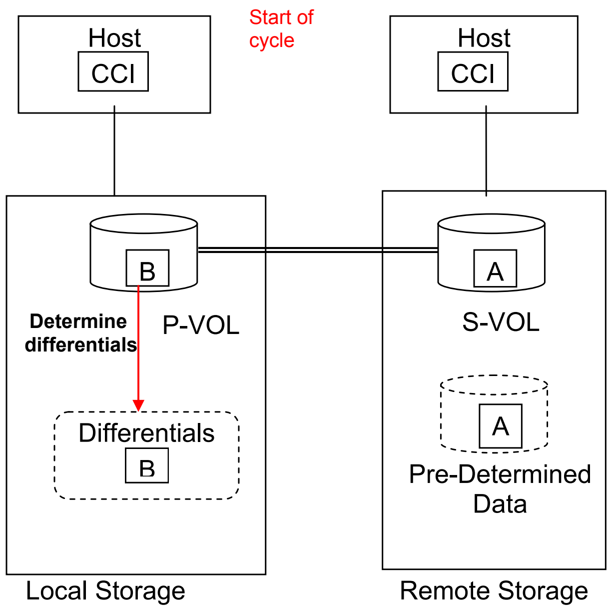
5.1.2 Cycle Update Process

Cycle update processing involves periodically transferring data updates from the P-VOL to the S-VOL. TCE remote replication processes are implemented as recurring operations executed in specific time periods (cycles). The following illustrations show the sequence of internal operations during one TCE update cycle.

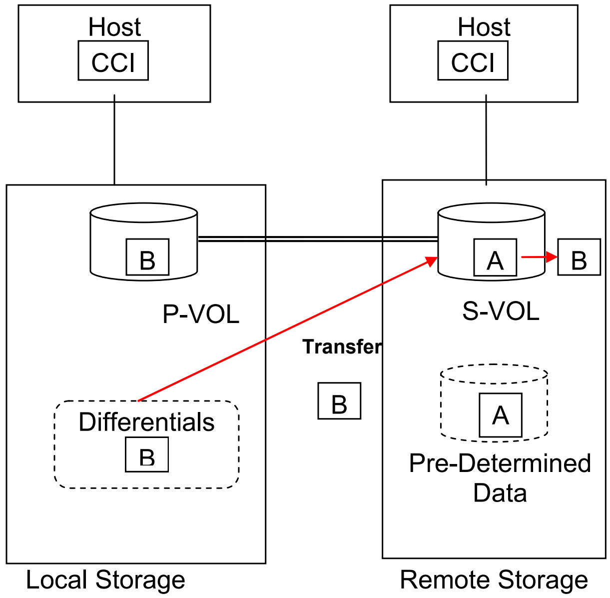
1. After initial copy or resynchronization, the P-VOL and S-VOL data are identical. A host write operation transfers new data updates (data B) to the P-VOL.



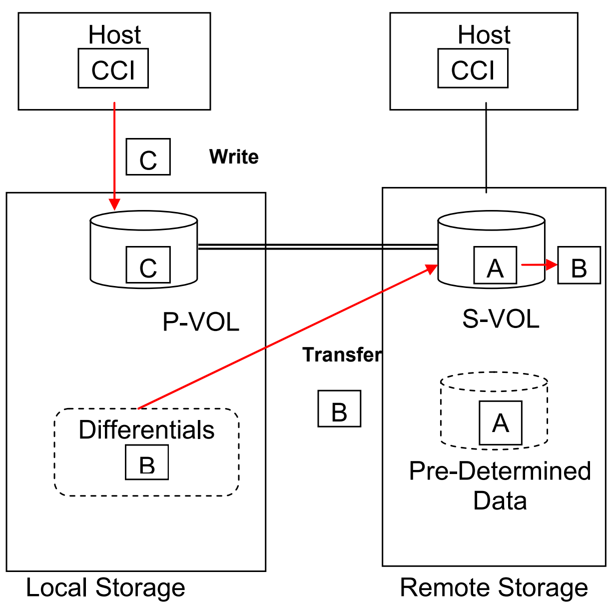
2. The local subsystem initiates a remote copy cycle and determines differential data between P-VOL and S-VOL. In this case, differential data is the data resulting from the host write operation (data B).



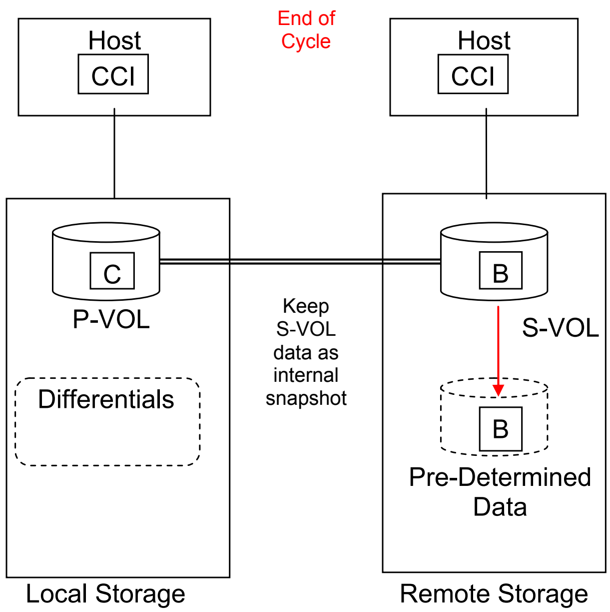
3. The local subsystem transfers differential data (data B) to the remote storage subsystem and S-VOL is updated. Before differential data is transferred to the remote subsystem, data in the S-VOL (data A) is not concurrent with data in the P-VOL (data B). After the data update is completed, a snapshot of the S-VOL is saved in an internal V-VOL.



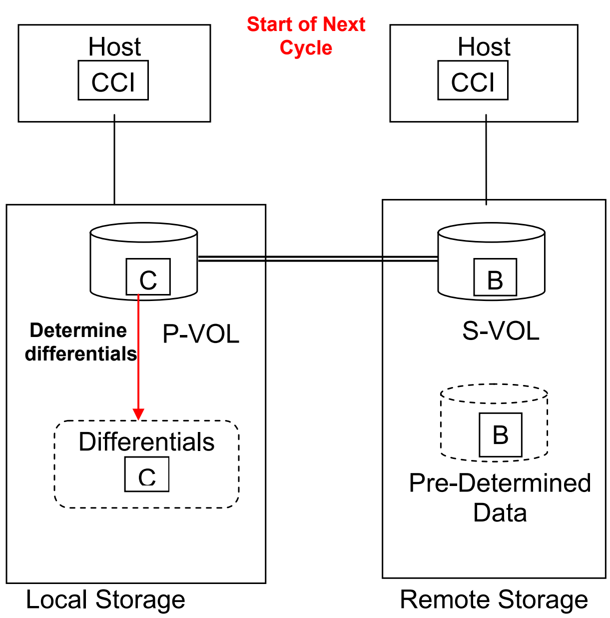
4. During the update cycle, the host may write another update (data C) to the P-VOL. The new data is not transferred to the remote site along with the differential data (data B) determined in the current update cycle. New updates are accumulated in the local data pool until the next update cycle. Thus, write order is guaranteed between each update cycle.



5. As a result of data transfer, the S-VOL is updated with differential data (data B). An S-VOL determination process creates an internal snapshot of differential data in the S-VOL. The internally pre-determined data is saved to the remote data pool and used to restore TCE subsystems in the case of a failover. For more information on data pools, see Section 5.1.2.1 Processing using Data Pools.



6. The next update cycle starts and new differential data (data C) is determined again. The update cycle repeats the same processes discussed in steps 2 to 5.



The completion of update processes in an update cycle depends on inflow of updates to the P-VOL. Cycle update times may increase if the inflow to the P-VOL (update data quantity) increases. As a result of the increase in the cycle time, the time difference between the P-VOL data and S-VOL data increases and degrading RPO value occurs. If the update inflow decreases, the cycle time reverts to the set time. Figure 5.1 shows a typical cycle update timing chart illustrating the above scenarios.

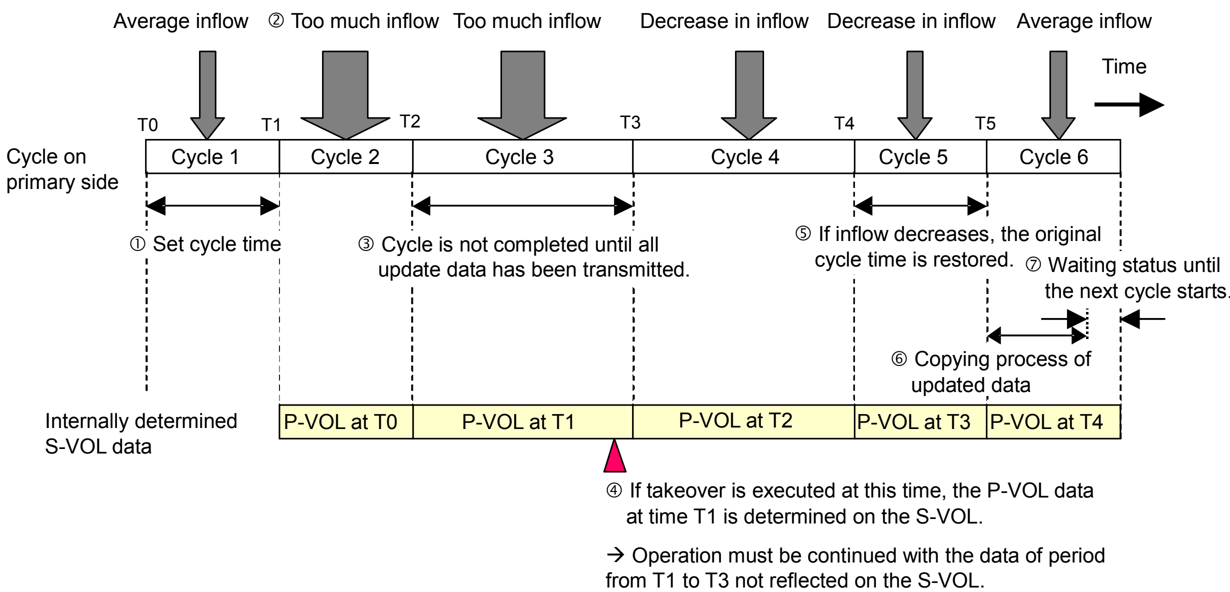


Figure 5.1 Cycle Update Timing Chart

Note: The length of an update cycle (cycle time) can be configured using Storage Navigator Modular. The default value of an update cycle is 300 seconds. The minimum value of a cycle is 30 seconds. Cycle times must always be set larger than (number of CTGs x 30) seconds.

5.1.2.1 Processing using Data Pools

The data pool of a local TagmaStore AMS subsystem has a different role from the pool of the remote TagmaStore AMS subsystem. Data pools on the local TagmaStore AMS subsystem are used to save intermittent updates to the P-VOL, when there are multiple host write operations during an update cycle. Differential data resulting from intermittent write operations to the P-VOL, are accumulated in the data pool until a subsequent update cycle (shown in Figure 2-6).

Data pools in the remote TagmaStore AMS subsystem are used to store internally pre-determined S-VOL differential data from the preceding update cycle. This differential data is used for data restoration during recovery operations. Remote snapshots created upon receipt of explicit remote snapshot creation commands from the primary host, are also saved to the remote data pool.

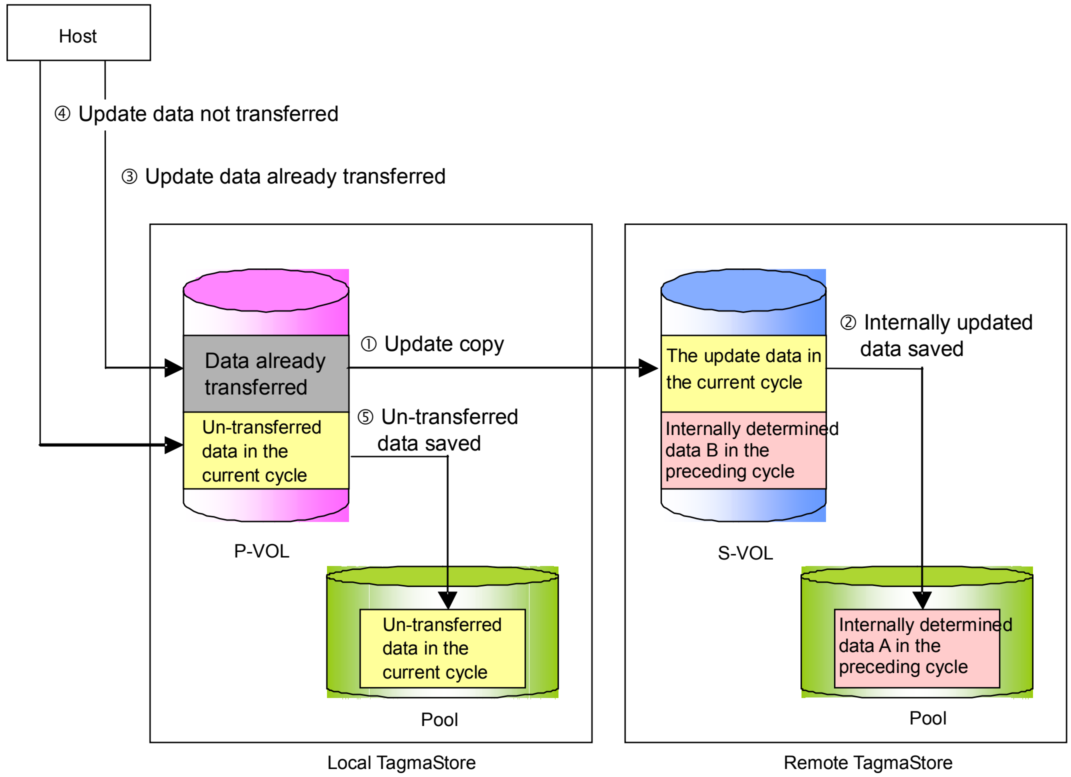


Figure 5.2 Update Processing Using Data Pools

The data pools on both subsystems are shared by TrueCopy and SnapShot functions. Data consistency policies determine how each function handles pool overflow. Remote subsystems delete aged snapshots on the remote data pool, when pool capacity is exceeded. Pre-determined S-VOL data is always kept current and used during remote takeover operations. For more information on determining pool capacity and overflow handling, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00.*

5.1.2.2 Write-Sequence Fidelity

Reconstruction of consistent file systems and databases requires that both data manager and replication technology have the critical property of write ordering. Database and file managers maintain some very complex internal data structures, including indexes, structured data tables, directories, logs, and so forth. They preserve the integrity of these internal structures by carefully sequencing every write that affects them at every stage. Careful write sequencing and strict adherence to dependencies allow file systems and databases to recover from failures, no matter what I/O activity is in progress when failure occurs.

TCE replication processes have this write sequence guarantee feature. Data is written to replicas in the same order and with the same dependencies in which it was written to the P-VOL. The remote subsystem always has consistent data from the previous update cycle, which is available for data restoration in the event of a takeover by the remote site. The following figure illustrates how writing sequencing is ensured in a TCE replication environment.

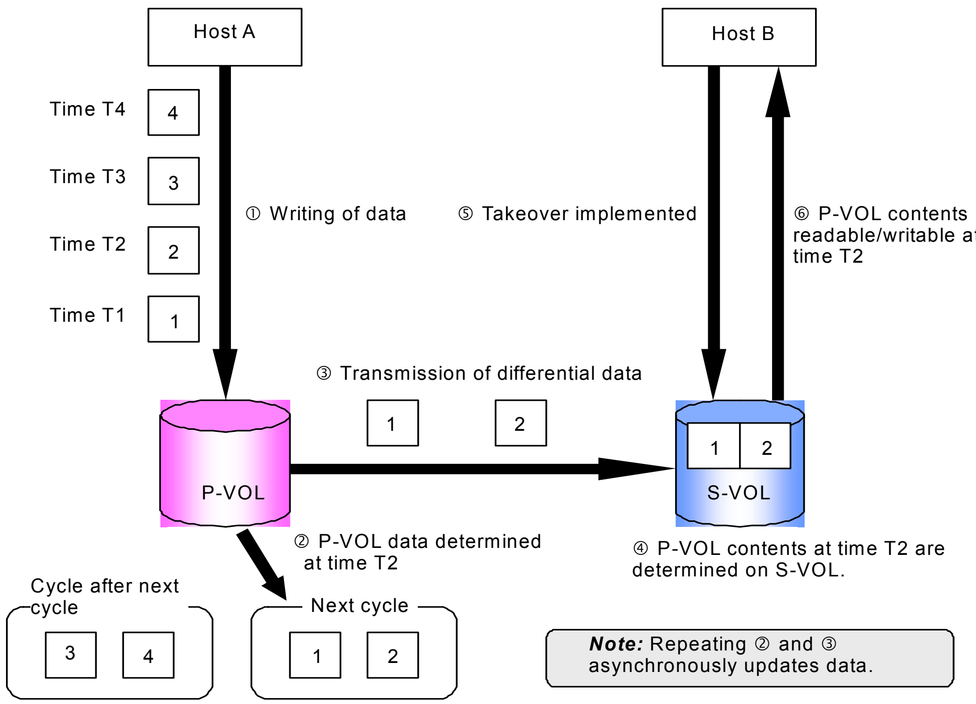


Figure 5.3 Write Sequencing

Processing such as cycle update processing, pair splitting, pair deleting and creating remote snapshot involves data transfer. Time to complete such processing depends on the order of data transfers among CTGs and the amount of data to transfer. Using TCE, data transfers can occur concurrently for up to 4 pairs per controller (8 pairs per subsystem). If there are 5 or more pairs, the additional pairs for completion of data transfer between four pairs. Initial copy and resynchronization can also be performed as concurrent data transfers. Up to 2 initial copy and resynchronization copy processes can be performed concurrently, along with the cycle update processes.

Note: Since initial copy and resynchronization copy operations involve transfer of larger volumes of data compared to cycle update processes, the completion time of cycle update processes may be delayed when they occur concurrently with the former processes.

Figure 5.4 shows concurrent processing in two CTGs with 4 pairs each. The first CTG pairs are involved in a cycle update process and the second CTG pairs are involved in a pair deletion process.

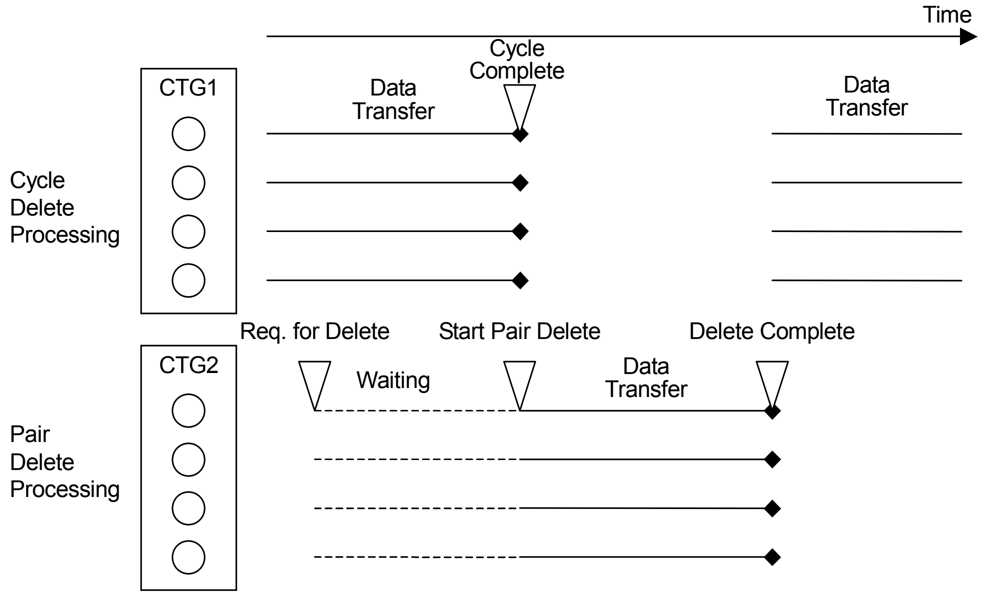


Figure 5.4 Concurrent Cycle Update and Pair Deletion Processes

In this example, there are 4 data transfers being processed concurrently for CTG1. During the processing, pair deleting of CTG2 is requested. Because 4 data transfers are already being processing, data transfers which are required for deleting CTG2 are not initiated. There is a wait time for the CTG2 pair deletion processes to begin. This wait period depends on the amount of data transferred between CTG1 pairs and the data transfer rate. Pair operations such as deletion, splitting and processes such as remote snapshot creation are not initiated as soon as they are requested, due to the time delays mentioned above.

5.1.3 Resynchronization

Resynchronization of data volume pairs is performed to make data in the S-VOL consistent with data in the P-VOL. Resynchronization copy operations are essential for restoring data consistency in previously split or suspended volume pairs, before update operations can be resumed on them. For more information on resynchronizing volume pairs, see Chapter 7 Section 7.3.2 Pair Recovery (pairresync) Operation.

5.2 TCE Remote Replication Operations

Remote replication command operations are performed from the UNIX®/PC server host using CCI software. The following pair operations are used to initiate remote replication operations:

- Pair configuration check (**pairdisplay**) operation
- Pair creation (**paircreate**) operation
- Pair status waiting (**pairevtwait**) operation
- Pair synchronization (**pairsyncwait**) operation

In addition to the above operations, the following pair monitoring, maintenance and system reduction operations are also performed, during the replication process:

- Pair status monitoring (**pairdisplay**) operation
- Pair splitting (**pairsplit**) operation
- Remote Snapshot Pair splitting (**pairsplit -mscas**) operation
- Pair recovery (**pairresync**) operation
- Pair release (**pairsplit -S**) operation
- Pair suspension (**pairsplit -R**) operation

The following sections discuss the pair operations used to initiate remote replication. For more information on pair monitoring, maintenance and system reduction operations, see Chapter 7 Section 7.2 System Monitoring and Section 7.4 System Reduction.

5.2.1 Pair Status Check (pairdisplay) operation

All pair operations require volume pairs to be in required pair statuses, before the operations can commence. For example, the **paircreate** operation requires both P-VOL and S-VOL to be in SMPL (simplex) pair status, before it can be initiated. The **pairdisplay** command is used for pair status monitoring before or after other pair operations and displays the status of volume pairs. The **-fc** option shows detailed pair status information. The following example shows the result of a **pairdisplay** command for a group named vg01.

```
c:\horcm\etc>pairdisplay -g vg01
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV#
M
vg01    oradb1(L)   (CL1-A , 1,  1 )75000174     1.SMPL ----- ------,-----  ---
-  -
vg01    oradb1(R)   (CL1-B , 1,  2 )75000175     2.SMPL ----- ------,-----  ---
-  -
```

For more information on pair statuses, see Chapter 7 Section 7.1.1 Pair Status Transitions. For more details on the **pairdisplay** command and its options, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

5.2.2 Creating TCE pairs

The **paircreate** operation establishes new volume pair(s). The P-VOL and S-VOL to be paired must be in the SMPL (simplex) state before pair creation. After the paircreate operation, the status of the volume pair is set to COPY. An initial copy operation occurs after a new volume pair is created. When initial copy is complete, the pair status changes to PAIR.

The remote update cycle uses the volume pairs to perform regular differential data updates in the cycle time specified by a user. Figure 5.5 illustrates pair statuses during the pair creation process.

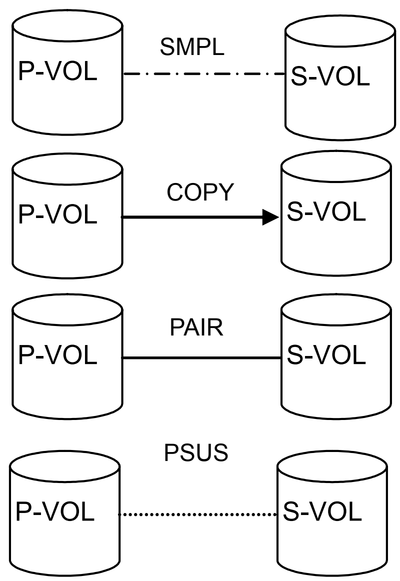


Figure 5.5 Creating a TCE Pair

Note: An attempt to create a pair in cycle intervals shorter than (number of CTGs × 30 seconds) results in an error.

5.2.2.1 Pair Creation (paircreate) Operation

To create TCE pairs,

1. Execute the **pairdisplay** command to verify that the status of the TCE volume is SMPL. In this example, the group name in the configuration definition file is vg01.

```
c:\horcm\etc>pairdisplay -g vg01
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV#
M
vg01    oradb1(L)    (CL1-A , 1,   1 )75000174     1.SMPL ----- ------,-----  ---
-  -
vg01    oradb1(R)    (CL1-B , 1,   2 )75000175     2.SMPL ----- ------,-----  ---
-  -
```

2. Execute the **paircreate** command. It is recommended to enable a fast copy by specifying a value from 11 to 15 for **-c** option which sets the copy pace. Then, execute the **pairevtwait** command to verify that the status of each volume is COPY.

```
c:\horcm\etc>paircreate -g vg01 -f async -vl -c 15
c:\horcm\etc>pairevtwait -g vg01 -s copy -t 300 10
pairevtwait : Wait status done.
```

   Note: Host I/O performance may be lowered due to a fast copy pace. If host I/O performance is to be optimized, copy pace should be set to medium or slow (1 to 10).

3. Execute the **pairdisplay** command to verify the pair status and the configuration.

```
c:\horcm\etc>pairdisplay -g vg01
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV#
M
vg01    oradb1(L)    (CL1-A , 1,  1 )75000174     1.P-VOL COPY ASYNC ,75000175
2  -
vg01    oradb1(R)    (CL1-B , 1,  2 )75000175     2.S-VOL COPY ASYNC ,-----
1  -
```

For more details on the **paircreate** command and its options, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

Upon creation, the TCE pair status is set to COPY. An initial copy operation occurs after pair creation. The pace for the initial copy operation can be set as slow, medium or fast. When the number of write operations from the host is high, the initial copy operation continues till the difference between the P-VOL and S-VOL data reaches a particular threshold. The fluctuations in the amount of differential data can be monitored using Storage Navigator Modular GUI as the synchronization rate of the pair. TCE pair status is set to PAIR after the initial copy operation.

5.2.2.2 Pair status waiting (pairevtwait) operation

The **pairevtwait** operation is used to wait for pair volumes to transition to the PAIR status, before other pair operations can be continued. Transitions to PAIR status are confirmed after completion of processes such as initial copy or pair splitting operations. The following example shows the result of a **pairevtwait** command for a group named vg01.

```
c:\horcm\etc>paircreate -g vg01 -f async -vl -c 15
c:\horcm\etc>pairevtwait -g vg01 -s copy -t 300 10
pairevtwait : Wait status done.
```

5.2.3 Synchronizing Data Volumes

Synchronization of data in the S-VOL with data in the P-VOL is essential for seamless takeover operations at the remote site in the event of a disaster. Data synchronization ensures data in the S-VOL is identical to the data in the P-VOL at the beginning of each update cycle and is implemented by internal S-VOL determination processes. Pair synchronization also involves the **pairsyncwait** operation, which confirms that S-VOL is synchronized prior to the next update cycle.

5.2.3.1 S-VOL Determination

Internal processes in the remote subsystem replicate the S-VOL independent of the update operations. This process known as **S-VOL determination** occurs at the end of each update cycle. At all times, a pre-determined copy of S-VOL data consistent with P-VOL data (as of the previous cycle) is maintained on the remote site. Figure 5.6 illustrates S-VOL determination during the cycle update process.

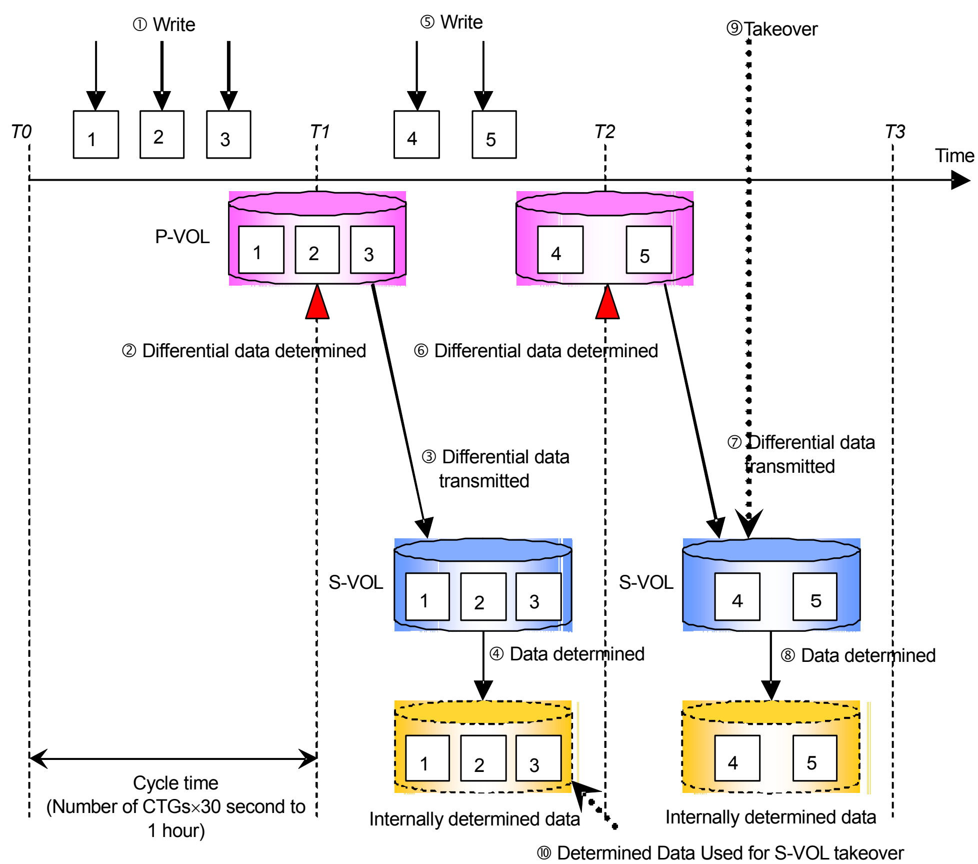


Figure 5.6 S-VOL Determination

5.2.3.2 Data Transfer Confirmation (pairsyncwait) operation

The **pairsyncwait** operation is used for synchronization of volume pairs and for monitoring occurrences of determination processes. The time difference between data determination processes is monitored and completion of copy operations is confirmed using the **pairsyncwait** operation. Succeeding operations are delayed by a wait period, until the P-VOL and S-VOL data are synchronized.

The local TagmaStore AMS subsystem manages two different Q-Markers for P-VOLs and its associated S-VOLs in a CTG. When a P-VOL in the CTG is updated by a host, the Q-Marker is incremented by one. When a cycle completes, the Q-Marker of S-VOLs in a CTG is updated to the Q-Marker of P-VOLs recorded at the time the cycle started.

```
# date       /* Obtain current time
Tue Nov 22 11:18:58 2005/11/22

# pairsyncwait –g vg01 –nowait      /* Obtain current sequence number
UnitID   CTGID   Q–Maker   Status     Q–Num
   0       3   01003408ef  NOWAIT   2

# pairsyncwait –g vg01 -t 100 –m 01003408ef  /* Wait with obtained sequence number
UnitID   CTGID   Q–Maker   Status     Q–Num
   0       3   01003408ef  DONE   0

# date
Tue Nov 22 11:21:10 2005/11/22
```

By measuring the time until that Q-Marker is reflected on the S-VOL, the time difference between the P-VOL data and S-VOL data is estimated. In this example, it takes about two minutes for the data written to the P-VOL to be reflected on the S-VOL. If S-VOLs' Q-Marker is larger than or equal to Q-Marker got at Time_0, it means all differential data updated before Time_0 has been copied to the S-VOLs and P-VOL data newer than Time_0 can be read from S-VOL by pairsplit or horctakeover.

Any abnormal synchronization can be detected by periodically executing an automated script to monitor the time difference between the P-VOL and the S-VOL. Notifications can be sent to system administrators upon detecting any abnormality. For more information on the **pairsyncwait** command, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

5.3 Copy-on-Write SnapShot Operations

Hitachi's Copy-on-Write Snapshot is a local replication solution designed to reduce backup time and provide point-in-time backup copies. Copy-on-Write Snapshot accelerates backup and recovery of production data, reducing critical application downtime from hours to seconds. This section describes TCE backup operations with Copy-on-Write Snapshot.

5.3.1 Point-in-time Snapshots

Snapshot backups involve creating point-in-time frozen replicas (snapshots) of data in both primary and secondary storage subsystems. The static images of the primary volume are used to determine data to be updated to the secondary storage subsystem during the remote backup operation. The multiple snapshots created prior to the onset of a disaster, are useful for data recovery, particularly from rolling disasters. For more information on snapshots, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Copy-on-Write SnapShot Software User's Guide (MK-95DF708).*

5.3.2 Differential Data Management

Differential data is data resulting from updates to the P-VOL or S-VOL during each update cycle. During each cycle a snapshot image of P-VOL differential data is captured and saved in an internal V-VOL (virtual volume). Each time differential data is updated to the S-VOL, a snapshot image of the differential data in the S-VOL is also determined and copied to an internal V-VOL (virtual volume) on the remote subsystem.

5.3.2.1 Granularity of Differential Data

Since only the differential data in the P-VOL is transferred to the S-VOL during an update cycle, the size of data sent to S-VOL is the same as that of data written to P-VOL. The amount of differential data that can be managed per write command is limited by the difference between the number of incoming host write operations (inflow) and outgoing data transfers (outflow). The number of I/Os managed per write command depends on models and the amount of cache memory. If the difference between the number of incoming host writes and outgoing data transfers is 1000, the differential data managed per write command is retained for 21 minutes (in a AMS500 subsystem). If the upper limit is exceeded, any remaining write data is managed as 64-KB chunks of differential data.

Table 5.1 shows differential data management information for TagmaStore AMS models.

Table 5.1 Maximum Amount of Manageable Differential Data and Time

#	Model Name	Shortest Time Upper Limit Is Reached in the Worst Case	
		1,000 IOPS	2,000 IOPS
1	AMS500	About 21 minutes	About 10 minutes
2	AMS1000	About 43 minutes	About 21 minutes

Differential data is also transmitted in 64 KB units under the following circumstances:

- Pair split
- Pair delete
- Remote path(s) failover
- Both remote paths are failed
- Controller(s) failover
- Planned shutdown
- System reset

Transmission of 64-KB differential data may extend update cycle times resulting in decreased RPO. Optimum network bandwidth selections must be made to balance inflow and outflow to optimize RPO. Size of differential data is also determined by the number of data drives in a RAID group. For more information on differential data management, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00.*

During initial copy processing, data is transferred one stripe size at a time. The size of one stripe string is determined by the number of data drives in a RAID group. When a volume pair is resynchronized data is transferred in 64 KB units since differential data is not sequentially located on the P-VOL.

5.3.3 TCE and Snapshot Product Differentiation

Table 5.2 summarizes differences between TCE and Copy-on-Write SnapShot.

Table 5.2 Comparison of TCE and Copy-on-Write SnapShot

Areas of Comparison	TCE	SnapShot
Threshold Over	Threshold for a pool usage can be defined. If threshold is over, pair statuses do not change. Pool status also changes to PFUS.	Threshold for a pool usage can be defined. If threshold is over, pair statuses change to PFUS. Pool status also changes to PFUS.
Pool Over at Local	Pair status of P-VOL with PAIR status changes to PFUS. Pair status of P-VOL with COPY status changes to PSUE.	Pair status changes to PSUE.
Pool Over at Remote	Pair status of P-VOL changes to PSUE. Pair status of S-VOL with PAIR status changes to PFUS. Pair status of S-VOL with COPY status changes to PSUS(N). Note: Meaning of PFUS in TCE is different from that in SnapShot.	Pair status changes to PSUE.
Data Consistency	S-VOL data is consistent at CTG level even if a pool is over.	V-VOL data is invalid in a case of a pool over.
Recover from PSUE	Resynchronize a pair.	Delete a pair and recreate a pair again.
Split-Maker	Split-Maker can be specified in pairsplit -mscas command issued from a local CCI. The character string can be reference when the -v smk option is specified for the pairdisplay operation.	No split-maker support.
Failures	Failures at Local: P-VOL changes to PSUE. S-VOL does not change. Data consistency is ensured if a pair status of S-VOL is PAIR. Failures at Remote: P-VOL change to PSUE. S-VOL changes to PFUS. No data consistency for S-VOL.	Pair status changes to PSUE and V-VOL data is invalid.
Specifying CTG	CTG# is specified by paircreate -f async [CTG#].	CTG # is specified by paircreate -m grp [CTG#].
# of CTG	16 * # of CTG for SnapShot and # of CTG for TCE are independent. SnapShot supports 128 and TCE supports 16.	128

5.3.4 TCE and Snapshot Cascade Configuration

A cascade configuration is a connection configuration of volume pairs in which a P-VOL or an S-VOL from a pair belonging to one copy function is used as a P-VOL or S-VOL of the other copy function. TCE supports cascading with ShadowImage and SnapShot. A TCE P-VOL or S-VOL can be used only as the P-VOL of a Snapshot volume pair. Table 5.3 shows permitted TCE cascade configurations.

Table 5.3 Cascade of Other Copy Functions

TCE	TCE		TrueCopy		ShadowImage		SnapShot	
	P-VOL	S-VOL	P-VOL	S-VOL	P-VOL	S-VOL	P-VOL	V-VOL
P-VOL							√	
S-VOL							√	

There are restrictions for performing pair operations and volume restoration, in a TCE- SnapShot cascade configuration. For detailed information on TCE- SnapShot cascade configurations, see Chapter 1.

5.3.5 Remote SnapShot Pair Splitting (pairsplit –mscas) Operation

The **pairsplit -mscas** operation is used for splitting a remote SnapShot pair which is in a cascade configuration with a TCE pair. This command to split the TCE-SnapShot pair (TCE volume & SnapShot P-VOL pair), is issued by the primary host to the CTG containing the TCE volume pair (TCE P-VOL & S-VOL pair). It cannot be directly issued to the TCE-SnapShot pair by the host on the secondary side. When this command is executed, SnapShot V-VOL data is synchronized with the P-VOL data of the TCE pair in PAIR status.

This operation requires cascaded TCE pair to be in PAIR status. The SnapShot pair to be split must be in PAIR or PSUS status and is identified by a specifying an MU number. The MU numbers of all the SnapShot pairs concerned must be the same as the MU number specified when the **pairsplit -mscas** command is executed.

All other TCE pairs in the same CTG, should be in COPY or PAIR status. Other TCE pairs in the same CTG cannot be split when the **pairsplit -mscas** operation is being executed on a cascaded TCE-Snapshot volume pair. For more details on the **pairsplit -mscas** operation, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

Figure 5.7 illustrates a **pairsplit -mscas** operation in a TCE-SnapShot cascade configuration.

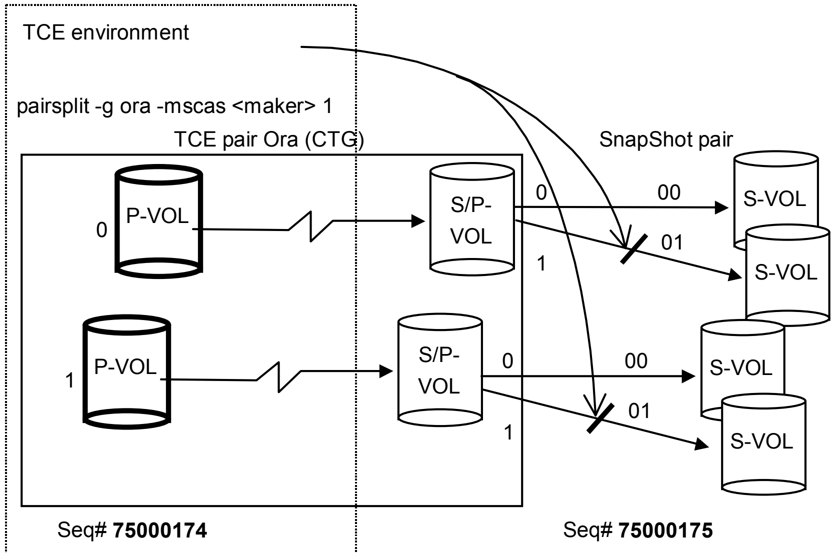


Figure 5.7 Pairsplit Operation in a Cascade Configuration

In the above example, the TCE group name in the configuration definition file is specified as **ora** and the name of the groups with the cascaded SnapShot pairs are specified as **00** and **01**. The following command with the **-mscas** option executes the pairsplit operation.

```
c:\horcm\etc>pairsplit -g ora –macas Split-Maker 1
```

Executing the **pairdisplay** command for the group **ora**, verifies that the pair is still in the required PAIR status. Pair status of the TCE pair on the primary subsystem does not change as a result of this pairsplit operation.

```
c:\horcm\etc>pairdisplay –g ora
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
vg01    oradb1(L)   (CL1-A , 1,  1 )75000174     1.PAIR ----- ------,-----  ---- -
vg01    oradb1(R)   (CL1-B , 1,  2 )75000175     2.PAIR ----- ------,-----  ---- -
```

Note: CCI adds a human-readable character string of ASCII 31 characters to each remote snapshot. The character string can be referenced using the **-v smk** option of the **pairdisplay** operation and helps to easily identify a snapshot amongst many generations of SnapShot replicas.

5.3.5.1 Confirmation of SnapShot Pair Split (Indirect)

The **pairsyncwait** command can be executed on the primary subsystem to confirm that pair splitting has occurred and P-VOL data has been transferred to the S-VOL. Each execution of the **pairsplit -mscas** command, increments Q-Marker count by 1.

```
c:\horcm\etc>pairsyncwait -g ora -t 10000
UnitID  CTGID   Q-Marker   Status  Q-Num
     0      3  00101231ef Done        2
```

Note: Completion of the pairsplit operation may be paused for one cycle in response to the **pairsyncwait** operation. Splitting of the SnapShot pair can also be confirmed directly on the remote subsystem where the SnapShot pair is stored.

5.3.5.2 Confirmation of SnapShot Pair Split (Direct)

The **pairevtwait** command can be executed on the remote subsystem to verify if pair status has changed as a result of the pair splitting. Pair status of the SnapShot pair, changes from PAIR to PSUS upon completion of the pair split operation.

```
c:\horcm\etc>pairevtwait -g o1 -s psus -t 300 10
pairevtwait : Wait status done.
```

The **pairdisplay** command (with the **-v smk** option) can also be executed to verify that the remote SnapShot pair has been split. In the following example, the **pairdisplay** command is executed for the reference group **o1**.

```
c:\horcm\etc>pairdisplay –g o1 -v smk
Group   PairVol(L/R) Serial# LDEV# P/S Status UTC-TIME -----Split-Maker-----
o1    URA_000(L)   75000175     2 P-VOL PSUS  -     -
o1    URA_000(R)   75000175     3 S-VOL SSUS  123456ef Split-Maker
```

5.4 Sample Remote Replication Scenarios

The following sections provide sample backup and data transfer scenarios using the TCE functionality.

5.4.1 Backup Scenario

This section presents a scenario in which the TCE function is applied to a backup system. Figure 5.8 shows a system configuration in which SnapShot operations at the local site and remote site are combined. This system configuration allows creation of remote snapshots using a script run on the local host.

In this configuration, the local TagmaStore subsystem and remote TagmaStore subsystem are directly coupled using a fibre channel. For long-distance connections, both the TagmaStore subsystems are connected with an extender via a WAN.

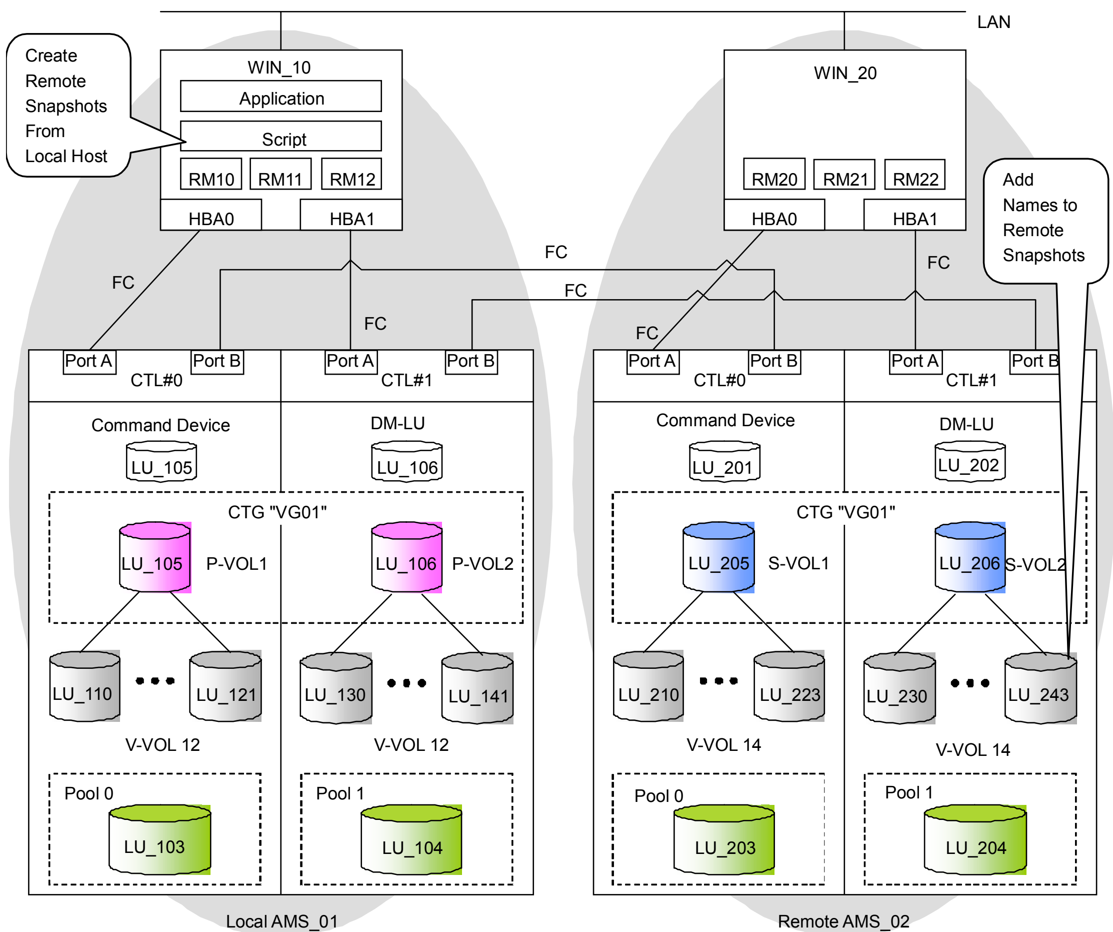


Figure 5.8 Sample Backup System Configuration

5.4.1.1 Backup System Configuration

This section explains connections setup in the sample configuration between the server, storage subsystem and logical units created in the subsystem. For details about connecting units and creating a logical unit, see *Hitachi TagmaStore™ Adaptable Modular Storage 500™ User and Reference Guide* and/or *Hitachi TagmaStore™ Adaptable Modular Storage 1000™ User and Reference Guide*.

- **WIN_10 server** runs applications and has two HBAs. One of the HBAs is connected to CTL#0 Port_A and the other is connected to CTL#1 Port_A of AMS_01. In addition, WIN_10 and WIN_20 are connected via a LAN so that TCP/IP communication can be established.

  In this example, the application data is stored on the G: drive. The LDM volume \Device\HarddiskVolumeX is mounted on the G: drive. This LDM volume consists of LU_105 and LU_106 of AMS_01.

- **Storage TAGMASTORE_01** is the local unit of this TCE configuration that stores application data. Two remote paths connect AMS_01 and AMS_02 subsystems. CTL#0 Port_A and CTL#1 Port_B of AMS_01 are connected to CTL#0 Port_A and CTL#1 Port_B of AMS_02. A path is set from the P-VOL and the control device to WIN_10. However, a drive number is not assigned to the command device.

  The following logical units are set on AMS_01:

    - LU_101 for the command device
    - LU_102 for DM_LU
    - LU_105 and LU_106 on which application data is to be stored
    - V-VOL LU_110 to LU_121 (12 LUs) to create snapshot of LU_105
    - V-VOL LU_130 to LU_141 (12 LUs) to create snapshot of LU_106
    - LU_103 assigned as pool 0 of CTL#0
    - LU_104 assigned as pool 1 of CTL#1

- **Server WIN_20** is the server on which the remote CCI operates. WIN_10 has two HBAs, one connected to CTL#0 Port_A and the other connected to CTL#1 Port_A of AMS_02.

- **Storage AMS_02** is the remote subsystem of this TCE configuration that stores a remote copy of application data and the remote snapshots of the application. A path is set from the S-VOL and the control device to WIN_20. However, a drive number is not assigned to the command device.

  The following logical units are set for AMS_02:

    - LU_201 for the command device
    - LU_202 for DM_LU
    - LU_205 and LU_206, remote copies of the application data
    - V-VOL LU_210 to LU_221 (12 LUs), cascaded from S-VOL LU_205
    - V-VOL LU_230 to LU_241 (12 LUs), cascaded from S-VOL LU_206

- LU_203, assigned as pool 0 of CTL#0
- LU_204, assigned as pool 1 of CTL#1

5.4.1.2 Backup Policy

A backup policy is the set of rules that determines which data is to be backed up and when. Table 5.4 shows an example of backup policy. According to the backup policy in Table 5.4, a snapshot is created on the volume \Device\HarddiskVolumeX at the local site once every two hours. Volume \Device\HarddiskVolumeX contains the application data of the WIN_10 server that operates at the local site. A snapshot at the remote site is created once a day. The local snapshot is retained for one day, and the remote snapshot is retained for two weeks. The snapshot is stored on an FC drive at both the local and remote sites.

Table 5.4 Sample Backup Policy

#	Item		Description	Remarks
1	Target server		WIN_10	Name of server where application runs. Windows® is used.
2	Target volume		\Device\HarddiskVolume X (G: drive)	- Name of volume managed by Windows® and storing application data - Consists of LU_105 and LU_106 of AMS_01
3	Schedule	Local snapshot	Every 2 hours	
		Remote snapshot	Every other day	
4	Aging	Local snapshot	1 day	12 V-VOLs are used.
		Remote snapshot	2 weeks	14 V-VOLs are used.
5	Backup destination	Local snapshot	AMS_01 (FC drive)	AMS_01 is the name of the storage storing application data.
		Remote snapshot	AMS_02 (FC drive)	AMS_02 is the name of the storage at the remote site.

5.4.1.3 Backup System Settings

Table 5.5and Table 5.6 show system settings used for the server and storage subsystems in this sample TCE configuration. The CCI instance number and the CCI communication port number on the server side, are required for starting CCI. These settings are necessary for both TCE and SnapShot.

Table 5.5 Server Related System Settings

#	Target	Setting Item	Setting Value	Remarks
1	WIN_10	DNS name	WIN_10	
2		IP address	192.168.1.10	
3		CCI instance number (TCE P-VOL)	10	
4		CCI instance number (SnapShot P-VOL)	11	
5		CCI instance number (SnapShot S-VOL)	12	
6		CCI communication port number (TCE P-VOL)	50001	
7		CCI communication port number (SnapShot P-VOL)	50002	
8		CCI communication port number (SnapShot S-VOL)	50003	
9	WIN_20	DNS name	WIN_20	
10		IP address	192.168.1.20	
11		CCI instance number (TCE S-VOL)	20	
12		CCI instance number (SnapShot P-VOL)	21	
13		CCI instance number (SnapShot S-VOL)	22	
14		CCI communication port number (TCE S-VOL)	50004	
15		CCI communication port number (SnapShot P-VOL)	50005	
16		CCI communication port number (SnapShot S-VOL)	50006	

Table 5.6 Storage Related System Settings

#	Target	Setting Item	Setting Value	Remarks
1	AMS_01	Serial number	75003001	
2		IP address	192.168.1.1	
3		Command device	LU_101	Path setting to WIN_10
4		DM-LU	LU_102	
5		Pool 0	LU_103	
6		Pool 1	LU_104	
7		P-VOL 0	LU_105	Path setting to WIN_10
8		P-VOL 1	LU_106	Path setting to WIN_10
9		V-VOL for P-VOL 0	LU_110 to LU_121	12 times/day
10		V-VOL for P-VOL 1	LU_130 to LU_141	12 times/day
11	AMS_02	Serial number	75003002	
12		IP address	192.168.1.2	
13		Command device	LU_201	Path setting to WIN_20
14		DM-LU	LU_202	
15		Pool 0	LU_203	
16		Pool 1	LU_204	
17		S-VOL 0	LU_205	Path setting to WIN_20
18		S-VOL 1	LU_206	Path setting to WIN_20
19		V-VOL for S-VOL 0	LU_210 to LU_223	For 2 weeks
20		V-VOL for S-VOL 1	LU_230 to LU_243	For 2 weeks

5.4.1.4 Pair Creation

Assuming the above backup system configuration has been implemented, the following section provides an example of TCE pair creation and TCE-SnapShot pair cascading.

Creating the TCE pair

A configuration definition file must be created before creating TCE pairs. In this example two files are created, one each for the WIN_10 and WIN_20 servers. The environment variable **HORCM_CONF** specifies the folder in which the files are to be created. The contents of the configuration definition files for WIN_10 (local) and WIN_20 (remote) are shown below. The **HORCM_MON** and **HORCM_INST** variables are also set as indicated in Figure 5.9.

```
# horcm10.conf

HORCM_MON
#ip_address    service       poll(10ms)  timeout(10ms)
WIN_10         50001         1000        3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1e}

HORCM_DEV
#dev_group     dev_name      port#       TargetID   LU#     MU#
vg01           dev01         CL1-A          1       105     0
vg01           dev02         CL2-A          1       106     0

HORCM_INST
#dev_group     ip_address    service
vg01           WIN_20        50004
```

```
# horcm20.conf

HORCM_MON
#ip_address    service       poll(10ms)  timeout(10ms)
WIN_20         50004         1000        3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1f}

HORCM_DEV
#dev_group     dev_name      port#       TargetID   LU#     MU#
vg01           dev01         CL1-A          1       205     0
vg01           dev02         CL2-A          1       206     0

HORCM_INST
#dev_group     ip_address    service
vg01           WIN_10        50001
```

Figure 5.9 Sample Configuration Definition Files

The command device of the **HORCM_CMD** entry can be searched using the following command.

```
# raidscan -x findcmddev hdisk0,20
cmddev of set# 75003001 = \\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1e}
```

Note: The WIN_10 server can manage a maximum of 20 drives. The number of drives managed depends on the system.

The **HORCM_DEV** entry specifies a P-VOL. LU_105 and LU_106 are managed as group **vg01**. The device names of the respective logical units are **dev01** and dev02. The output of the following commands displays the port number (Port#), Target ID (TID#), and logical unit number (LU#). CL1-A indicates a path to LU_105 extends to port A of controller 0 of AMS_01.

```
C:\HORCM\etc>raidscan -p cl1-a -CLI
Port# /ALPA/C TID#  LU# Seq#   Num   LDEV#   P/S Status Fence P-Seq# P-LDEV#
CL1-A   ef  5   1   105 75003001     1      105 SMPL     -      -      -     -
CL1-A   ef  5   1   101 75003001     1      101 SMPL     -      -      -     -

C:\HORCM\etc>raidscan -p cl2-a -CLI
Port# /ALPA/C TID#  LU# Seq#   Num   LDEV#   P/S Status Fence P-Seq# P-LDEV#
CL2-A   ef  5   1   106 75003001     1      106 SMPL     -      -      -     -
```

The environment variables **HORCMINST**, **HORCC_MRCF**, and **HORCM_CONF** must be set before starting CCI.

- **HORCMINST** is the instance number of CCI and is determined from the file name of a configuration definition file. The instance number of CCI for the TCE pair of WIN_10 is 10. Therefore, the file name of the configuration definition file is "**horcm10.conf**". Accordingly, the value of HORCMINST is set to 10.
- **HORCM_CONF** is set to the name of the folder where the configuration definition file has been created.

After the environment variables are set, CCI is started on both WIN_10 and WIN_20 servers, using the **horcmstart** command. When executing the **horcmstart** command, instance numbers 10 and 20 must be specified as options. The value 20 is the instance number of CCI on the WIN_20 side. If the **HORCC_MRCF** environment variable is defined, you must undefine it so the TCE pair can be manipulated.

Note: CCI at the local site and CCI at the remote site communicate with each other using TCP/IP. A communication port must be specified in the configuration definition file and security settings must be changed on the OS side to enable use of a specified port number. For information about changing the security setting, see your operating system documentation.

When CCI has been started, you can create the TCE pair by using the **paircreate** command. If a CTG ID is not specified when **paircreate** is executed, the local TagmaStore subsystem assigns an available CTG ID automatically. Pairs belonging to a group defined in a configuration file are added to the CTG. If CTG ID is specified, an optional CTG ID can be assigned to a group in the configuration file. The **pairdisplay** command is used to check if pairs are created correctly.

```
# paircreate -g vg01 -c 15 -f async -vl
# pairdisplay -g vg01
Group  PairVol(L/R) (Port#, TID, LU-M), Seq#, LDEV#, P/S, Status, Fence,
Seq#, P-LDEV#, M
vg01   dev01(L)   (CL1-A, 1, 0-0)    75003001,  105, P-VOL, COPY, ASYNC,
75003002,  205,    -
vg01   dev01(R)   (CL1-A, 1, 0-0)    75003002,  205, S-VOL, COPY, ASYNC,  -
---    ,  105,    -
vg01   dev02(L)   (CL2-A, 1, 0-0)    75003001,  106, P-VOL, COPY, ASYNC,
75003002,  206,    -
vg01   dev02(R)   (CL2-A, 1, 0-0)    75003002,  206, S-VOL, COPY, ASYNC,  -
---    ,  106,    -
```

A SnapShot pair can be created when the TCE pair transitions to PAIR status. Before creating a cascaded SnapShot pair from the S-VOL of the TCE pair, the status of the TCE pair must be changed to PSUS, using the following **pairsplit** command.

```
# pairsplit -g vg01
# pairdisplay -g vg01
Group  PairVol(L/R) (Port#, TID, LU-M), Seq#, LDEV#, P/S, Status, Fence,
Seq#, P-LDEV#, M
vg01   dev01(L)   (CL1-A, 1, 0-0)    75003001,  105, P-VOL, PSUS, ASYNC,
75003002,  205,    -
vg01   dev01(R)   (CL1-A, 1, 0-0)    75003002,  205, S-VOL, SSUS, ASYNC,  -
---    ,  105,    -
vg01   dev02(L)   (CL2-A, 1, 0-0)    75003001,  106, P-VOL, PSUS, ASYNC,
75003002,  206,    -
vg01   dev02(R)   (CL2-A, 1, 0-0)    75003002,  206, S-VOL, SSUS, ASYNC,  -
---    ,  106,    -
```

Creating the SnapShot pair

Two configuration definition files for the SnapShot pair must be created on the local WIN_10 server. Figure 5.10 shows the configuration definition files of a SnapShot pair that is cascaded to the P-VOL of a TCE pair.

The P-VOLs of the SnapShot pair are LU_105 and LU_106. The configuration definition file of the P-VOL of the SnapShot pair (shown at the top of Figure 5.10 creates 12 devices for one P-VOL. Each device belongs to a different device group. The configuration definition file of the S-VOL of the SnapShot pair (shown at the bottom of Figure 5.10) associates one V-VOL with each device. Figure 5.14 shows the pair configuration.

The **HORCM_MON** and **HORCM_INST** are set according to Table 5.5.

The instance of CCI is different for the TCE pair and the SnapShot pair.

```
#horcm11.conf
```

```
HORCM_MON
#ipaddress   service   poll(10ms)   timeout(10ms)
WIN_10       50002     1000         3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1e}

HORCM_DEV
#dev_group   dev_name   port#   TargetID   LU#  MU#
vg10         dev00      CL1-A   1          105  0
vg11         dev00      CL1-A   1          105  1
vg12         dev00      CL1-A   1          105  2
   :
   :
vg21         dev00      CL1-A   1          105  13

vg10         dev01      CL2-A   1          106  0
vg11         dev01      CL2-A   1          106  1
vg12         dev01      CL2-A   1          106  2
   :
   :
vg21         dev01      CL2-A   1          106  13

HORCM_INST
#dev_group   ipaddress   service
vg10         WIN_10      50003
vg11         WIN_10      50003
vg12         WIN_10      50003
   :
   :
vg21         WIN_10      50003
```

```
#horcm12.conf


HORCM_MON
#ipaddress   service   poll(10ms)   timeout(10ms)
WIN_10       50003     1000         3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1e}

HORCM_DEV
#dev_group   dev_name   port#   TargetID   LU#  MU#
vg10         dev00      CL1-A   1          110  0
vg11         dev00      CL1-A   1          111  0
vg12         dev00      CL1-A   1          112  0
   :
   :
vg21         dev00      CL1-A   1          121  0

vg10         dev01      CL2-A   1          130  0
vg11         dev01      CL2-A   1          131  0
vg12         dev01      CL2-A   1          132  0
   :
   :
vg21         dev01      CL2-A   1          141  0

HORCM_INST
#dev_group   ipaddress   service
vg10         WIN_10      50002
vg11         WIN_10      50002
vg12         WIN_10      50002
   :
   :
vg21         WIN_10      50002
```

Figure 5.10 Configuration Definition Files for SnapShot Pair (local WIN_10)

Figure 5.11 shows the configuration definition files of a SnapShot pair that is cascaded to the S-VOL of the TCE pair. These two files are created on WIN_20 server. Figure 5.16 shows the pair configuration of the SnapShot pair to be cascaded to the S-VOL.

```
#horcm21.conf


HORCM_MON
#ipaddress   service   poll(10ms)   timeout(10ms)
WIN_20       50005     1000         3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1f}

HORCM_DEV
#dev_group   dev_name   port#   TargetID   LU#  MU#
vg30         dev00      CL1-A   1          205  0
vg31         dev00      CL1-A   1          205  1
vg32         dev00      CL1-A   1          205  2
   :
   :
vg43         dev00      CL1-A   1          205  13

vg30         dev01      CL2-A   1          206  0
vg31         dev01      CL2-A   1          206  1
vg32         dev01      CL2-A   1          206  2
   :
   :
vg43         dev01      CL2-A   1          206  13

HORCM_INST
#dev_group   ipaddress   service
vg30         WIN_20      50006
vg31         WIN_20      50006
vg32         WIN_20      50006
   :
   :
vg43         WIN_20      50006
```

```
# horcm22.conf
```

```
HORCM_MON
#ipaddress   service   poll(10ms)   timeout(10ms)
WIN_20       50006     1000         3000

HORCM_CMD
#dev_name
\\.\Volume{b9b31c79-240a-11d5-a37f-00c00d003b1f}

HORCM_DEV
#dev_group   dev_name   port#   TargetID   LU#  MU#
vg30         dev00      CL1-A   1          210  0
vg31         dev00      CL1-A   1          211  0
vg32         dev00      CL1-A   1          212  0
   :
   :
vg43         dev00      CL1-A   1          223  0

vg30         dev01      CL2-A   1          230  0
vg31         dev01      CL2-A   1          231  0
vg32         dev01      CL2-A   1          232  0
   :
   :
vg43         dev01      CL2-A   1          243  0

HORCM_INST
#dev_group   ipaddress   service
vg30         WIN_20      50005
vg31         WIN_20      50005
vg32         WIN_20      50005
   :
   :
vg43         WIN_20      50005
```

Figure 5.11 Configuration Definition Files of SnapShot Pair (for remote WIN_20)

The environment variables **HORCMINST**, **HORCC_MRCF**, and **HORCM_CONF** must be set before starting CCI. **HORCMINST** is the instance number of CCI for SnapShot and is set as 11 and 12 for WIN_10, and 21 and 22 for WIN_20. **HORCM_CONF** is set to the name of the folder where the configuration definition file has been created. The **HORCC_MRCF** environment variable is set as 1, implying that the SnapShot pair is to be manipulated.

After the environment variables are set, CCI is started on both WIN_10 and WIN_20 servers, using the **horcmstart** command. The SnapShot pair cannot be split when the TCE pair is in PAIR or COPY status. So it is necessary to verify if the TCE pair is in PSUS status by using the **pairdisplay** command.

If the pair splitting operation is delayed, you can use the **pairevtwait** command to wait for the pair status transition. The amount of P-VOL difference can be checked from the TCE pair information displayed by Navigator. The time required for completion of the status transition to PSUS can be estimated from the amount of the difference and the line speed.

After confirming that the TCE pair is in PSUS, 12 SnapShot pairs for the P-VOL of the TCE pair are created on the AMS_01 side. Another 14 SnapShot pairs for the S-VOL of the TCE pair are created on the AMS_02 side. A consistency group for the SnapShot pairs can be created by specifying the **-m grp** option at the time of the SnapShot pair creation. This allows SnapShot pairs to be split simultaneously as a CTG.

```
# paircreate –g vg30 –vl
# pairdisplay -g vg30
Group  PairVol(L/R) (Port#, TID, LU-M), Seq#, LDEV#, P/S, Status, Seq#, P-
LDEV#, M
vg30   dev00(L)   (CL1-A, 1, 205-0)    75003002,  205, P-VOL, PAIR, ----
, ----,    -
vg30   dev00(R)   (CL1-A, 1, 210-0)    75003002,  210, S-VOL, PAIR, ----
, ----,    -
vg30   dev01(L)   (CL1-A, 1, 206-0)    75003002,  206, P-VOL, PAIR, ----
, ----,    -
vg30   dev01(R)   (CL1-A, 1, 230-0)    75003002,  230, S-VOL, PAIR, ----
, ----,    -
```

Figure 5.12 Creating SnapShot Pair (WIN_20)

The SnapShot pair is split using a **pairsplit** command.

```
# pairsplit –g vg30
# pairdisplay -g vg30
Group  PairVol(L/R) (Port#, TID, LU-M), Seq#, LDEV#, P/S, Status, Seq#, P-
LDEV#, M
vg30   dev00(L)   (CL1-A, 1, 205-0)    75003002,  205, P-VOL, PSUS, ----
, ----,    -
vg30   dev00(R)   (CL1-A, 1, 210-0)    75003002,  210, S-VOL, SSUS, ----
, ----,    -
vg30   dev01(L)   (CL1-A, 1, 206-0)    75003002,  206, P-VOL, PSUS, ----
, ----,    -
vg30   dev01(R)   (CL1-A, 1, 230-0)    75003002,  230, S-VOL, SSUS, ----
, ----,    -
```

Figure 5.13 Splitting SnapShot Pair (WIN_20)

5.4.1.5 Backup Operations

This section describes a script-based automated backup implementation using TCE. According to the backup policy in Table 5.4 the local snapshot of application data is obtained once every two hours and the remote snapshot is obtained once a day. The local snapshot is saved for one day, and the remote snapshot is saved for two weeks. Figure 5.15 shows an example of an automation script.

The TCE function for TagmaStore can implement local snapshot and remote snapshot on the WIN_10 server, where the application runs. The WIN_10 server executes the two types of scripts; one for creating local snapshots and the other for creating remote snapshots. The WIN_20 server is not involved when a backup operation is performed.

Creating the Local Snapshot

When the local snapshot is created, the snapshot to be obtained is determined by the device group. The configuration definition file in Figure 5.10 is set to use device groups **vg10** to **vg21**. For example, to create a snapshot on LU_110 and LU_130 of V-VOL, you should specify vg10 as the device group.

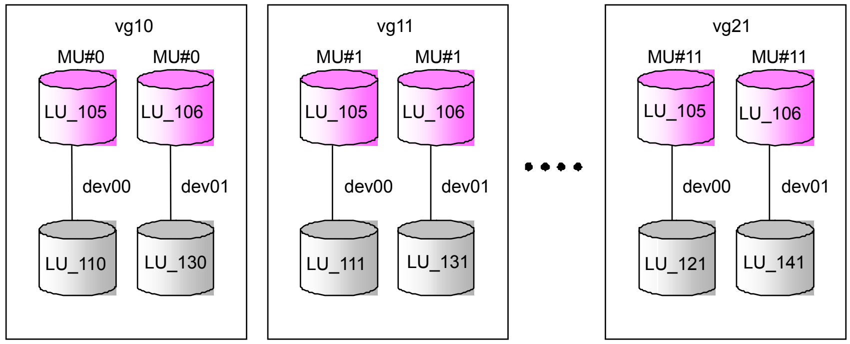


Figure 5.14 Sample Pair Configuration (Snapshot Cascaded to P-VOL)

The script for obtaining the local snapshot should be registered in the scheduler provided by the OS, so that the script is executed once every two hours. The **local_vgn_file** file is used by the script to record the last two digits of the device group to be used for the following iteration.

Before creating the snapshot, it is necessary to resynchronize the pairs and confirm that the pairs are in PAIR status. After transition to the PAIR status, the pairs are split by group. To obtain a consistent backup group, I/O to the P-VOL must be stopped during snapshot pair creation. The script should be executed after the application is stopped and the file system unmounted.

```
#!/usr/bin/perl

$mfile = "local_vgn_file";

# Set the environment variables

$ENV{'HORCC_MRCF'} = 1;
$ENV{'HORCMINST'} = 10;

# Obtain the MU number

open (IN, $mfile) or die "$!";
$local_vgn = <IN>;
close (IN);

# Obtain the local snapshot

system ("\HORCM\pairresync -g vg$local_vgn");
system ("\HORCM\pairevtwait -g vg$local_vgn -s pair");

# It is assumed that the application is stopped and the file system is
unmounted before executing the following command.

system ("\HORCM\pairsplit -g vg$local_vgn");

# The file system can be mounted and the application can restart from this
point.

# Set the next MU#.  The loop starts when 12 MUs are specified.

$local_vgn = ($local_vgn + 1) % 12;
open (OUT, $mfile) or die "$!";
print OUT local_vgn;
close (OUT);
```

Figure 5.15 Sample Script for Creating Local Snapshot

Creating the Remote Snapshot

The timing of remote snapshot creation is independent of update cycle time. At the time of remote snapshot creation, remote snapshot data is synchronized with data in the local P-VOL. Time to complete remote snapshot creation depends on the amount of the differential data in the P-VOL and data transfer rate. Completion time can be estimated using CCI or Navigator to view the amount of differential data.

When a remote snapshot is created, the snapshot is distinguished by the MU number of the S-VOL of the TCE pair associated with the snapshot. A command for remote snapshot creation is issued to the AMS_01 subsystem by the WIN_10 server on the local side. For example, MU# 0 must be specified to split V-VOL LU_210 and LU_230 in vg30 (Figure 5.16).

Note: Local snapshot creation uses a device group to determine snapshot to be generated, while remote snapshot creation uses the MU#.

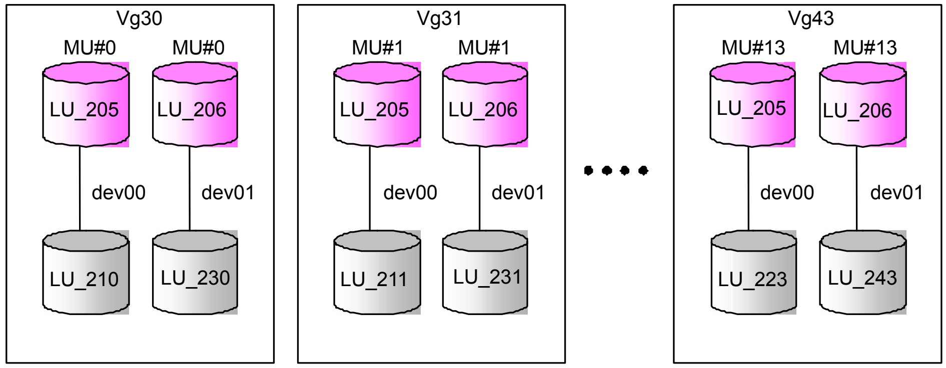


Figure 5.16 Example of Pair Configuration of Snapshot to be Cascaded with S-VOL

Figure 5.17 shows a sample remote snapshot creation script. This script should be registered in the scheduler provided by the OS, so the script is executed once a day. The local snapshot creation script and the remote snapshot creation script are both executed by the local WIN_10 server.

The remote snapshot creation command is issued to the P-VOL of a TCE pair. If a remote snapshot creation command is issued from the local server, the remote snapshot data is synchronized with P-VOL data at the time of issuing that command. The remote snapshot data is not determined from the S-VOL data at the time of issuing the command.

In this example, a remote snapshot creation command is issued to AMS_01, using the **-mscas** option. The character string and the MU# should be specified with the **-mscas** option. The pair status of the snapshot to be cascaded to the S-VOL remains PSUS.

```
#!/usr/bin/perl

$mfile = "remote_vgn_file";

# Set the environment variables

$ENV{'HORCMINST'} = 10;

# Obtain the MU number

open (IN, $mfile) or die "$!";
$remote_vgn = <IN>;
close (IN);

# Set the character string to be appended to the remote snapshot

($sec, $min, $hour, $day, $month, $year) = localtime;
$month = $month + 1;
$string = $year+'/'+$month+'/'+$day+'/'+$hour+'/'+$min+'/'+$sec;

# Obtain the remote snapshot

# Stop the application and un-mount the file system here.

system ("\HORCM\pairsplit -g vg01 -l –mscas $string $remote_mum");

# Mount the file system and start the application here.

# Set the next MU#.  The loop starts when 14 MUs are specified.

$remote_mun = ($remote_mun + 1) % 14;
open (OUT, $mfile) or die "$!";
print OUT remote_mun;
close (OUT);
```

Figure 5.17 Sample Remote Snapshot Creation Script

The remote snapshot creation script uses the **remote_mu_file** file to manage the MU# for determining subsequent remote snapshots. A character string is appended to each remote snapshot for easy identification. In the example in Figure 5.17, a character string indicating the time the snapshot was created, is appended to each remote snapshot.

To maintain a consistent backup group while the application is updating the P-VOL data, I/O to the P-VOL must be stopped while the snapshot is being created. In this scenario, the application is stopped and the file system is unmounted, before the script is executed.

5.4.2 Data Transfer Scenario

Moving data involves copying of application data to another unit at a remote site and then analyzing the copied data. Figure 5.18 shows a sample system configuration for moving data.

In this scenario, P-VOLs LU_105 and LU_106 of the local TagmaStore subsystem are grouped into the same CTG and then copied to S-VOLs LU_205 and LU_206 of the remote TagmaStore subsystem. An application running on the WIN_10 server reads and writes the application data stored in the P-VOLs. On the WIN_20 server, a data analysis application analyzes a copy of the P-VOL data created on the S-VOLs. In this example, data analysis is scheduled to occur once every hour.

The configuration of logical units on the P-VOL and S-VOL is the same as for backup usage (as explained in section 5.4.1.1 Backup System Configuration. The only exception is that the snapshot is not cascaded with the P-VOL and S-VOL. The configuration definition file of CCI used to create a TCE pair is the same as that shown in Figure 5.10.

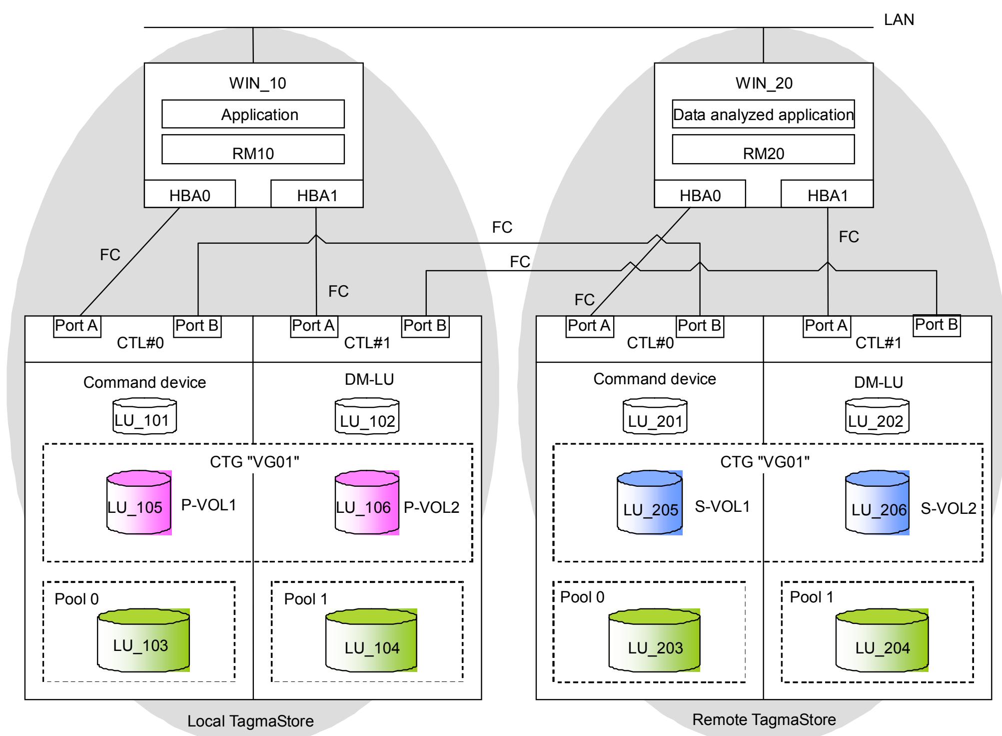


Figure 5.18 System Configuration for Moving Data

To implement data transfer, the following tasks must be performed on the local and remote subsystems.

Tasks on the local side:

1. Stop writing to the P-VOL by WIN_10 and un-mount the P-VOL.
2. Split the TCE pair to determine the S-VOL data.
3. Wait for the completion of splitting the TCE pair. Before the TCE pair can be split, updated differential data on the P-VOL must be transferred to the S-VOL. Wait for completion of the transfer processing.
4. Mount the P-VOL by WIN_10 and resume writing to the P-VOL.

Tasks on the remote side:

5. Mount the S-VOL. The contents of the S-VOL data are the contents of the P-VOL when the TCE was split.
6. Read and analyze the S-VOL data. The S-VOL data can be updated, but updated data is lost when the TCE pair is resynchronized. If updated data is necessary, be sure to save the data to a logical unit other than the S-VOL.
7. Un-mount the S-VOL.

Additional tasks on the local side:

8. Re-synchronize the TCE pair to resume re-synchronized copy.
9. Wait for the completion of re-synchronizing the TCE pair.

Table 5.7 summarizes the data transfer procedure using the above system.

Table 5.7 Data Transfer Procedure

#	Local Side	Remote Side
1	Stop writing to the P-VOL and un-mount it.	
2	Split the TCE pair. # pairsplit –g vg01 -rw	
3	Wait for the completion of splitting TCE. # pairevtwait –g vg01 –s psus	
4	Mount and resume writing to the P-VOL.	
5		Mount the S-VOL.
6		Perform data analysis.
7		Un-mount the S-VOL.
8	Recover the TCE pair. # pairresync –g vg01	
9	Waiting for the completion of recovering the TCE pair. # pairevtwait –g vg01 –s pair	

Chapter 6 Using TCE for Disaster Recovery

This chapter discusses disaster recovery procedures using TCE and provides sample scenarios to demonstrate data restoration. The following sections are included in this chapter:

- Disaster Recovery Options (section 6.1)
- Data Recovery Considerations (section 6.2)
- Disaster Recovery Processes (section 6.3)
- Sample Disaster Recovery Scenarios (section 6.4)

6.1 Disaster Recovery Options

Disaster recovery procedures involve failover operations to continue critical application processing at the remote site and fallback operations to restore data and application processing at the local site after failures have been rectified. Failover to a remote site is also required for online maintenance of a local storage subsystem in order to continue critical application processing without interruption.

If both the host and the storage subsystems on the local side are destroyed by means of a disaster, e.g. earthquake, the data recovery and failover operations are executed using the stand-by host and the storage subsystems on the remote side.

TCE systems can be classified into the following categories based on their disaster recovery and failover implementations:

- **Automated Failover:** This scenario is used for higher RTO requirements and requires server clustering or High Availability (HA) software (such as Veritas VCS). TCE systems configured for automated failover, use clustering software to monitor the occurrence of failures. By installing the High Availability software on both the local side and the remote side, critical business operations can continue by automatically switching over to the stand-by host on the remote side (failover), when a failure occurs in the host or in the subsystem. Although the input and output operations from the host on the local side cannot be executed immediately, such systems ensure reduced recovery time from failures.
- **Manual Failover:** This scenario is used for lower RTO performance. In this system implementation, failures are detected by users and recovery operations are initiated by manually issuing takeover commands. Such systems have longer recovery times from failures.

Note: TCE Version 3.x cannot be used for configuring automated failover since it does not support TCE pair swapping capability required for setting up a clustering system, DR testing and maintenance.

6.1.1 Types of System Failures

Table 6.1 shows the impact of different types of system failures and associated recovery procedures.

Table 6.1 Type of Failure and Recovery

Type of Failure	Location	Situation	Recovery procedure	Corrected by
Drive dual failure	P-VOL	Data not reflected on the S-VOL may have been lost.	Recover the pair after the drive is replaced.	Hitachi personnel
	P-VOL	Remote copy cannot be continued because S-VOL cannot be updated.		
	P-VOL pool	Remote copy cannot be continued because update copy cannot be executed.		
	S-VOL pool	Takeover to the S-VOL cannot be done because internally pre-determined data of the S-VOL is lost.		
Path blockage		Remote copy cannot be continued because communication with the remote AMS/WMS storage subsystem is not possible.	- Replace the parts. - Recover the remote line. - Recover the remote subsystem.	Hitachi personnel
Exceeding pool capacity	P-VOL pool	Remote copy cannot be continued because update copy cannot be executed. All V-VOLs of SnapShot have been lost.	- Recover the pair. - Recover the SnapShot	Customer
	S-VOL pool	All V-VOLs of SnapShot have been lost.		

Hitachi maintenance personnel correct drive failures and path blockage by replacing the parts causing the failure. Thereafter, the customer is responsible for recovering the TCE pair.

6.1.2 Restoration of Logical Data Files

Application data in the primary data volumes may be destroyed by two types of disasters:

- Logical data destruction such as the erasing of a file.
- Drive failure or physical damages to the P-VOL resulting from natural disasters.

In the event of a disaster, application data can be restored from local backup data or from remote snapshots. Data restoration from remote snapshots is discussed below. For information on restoring data from local snapshots, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Copy-on-Write SnapShot Software User's Guide (MK-95DF708)*.

Backup data can be easily restored at the file level from a remote snapshot since the P-VOL is operational after logical data destruction. First, the remote snapshot that holds the files to be restored is found from the snapshots obtained from the remote storage subsystem. Data to be restored is determined by referring to the character string associated with the snapshot and the backup time obtained from CCI or Navigator software.

The snapshot to be used for restoration is mounted on the remote server and the files to be restored are determined by accessing the data volume. The restored file is first copied to a temporary disk area on the primary server and then copied to its appropriate location. Restoration task is completed after verifying contents of the restored file and normal application processing at the primary site. For more information on logical data restoration, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy™ Extended Distance Configuration Guide, MK-95DF779-00.*

6.1.3 Recovery from P-VOL Failure

If the P-VOL is physically destroyed by a disaster, restoration procedures and backup data location depend on the scope of the disaster. Table 6.2 lists types of disasters and corresponding scope and recovery locations.

Table 6.2 Disaster Scope and Restore Destination

Disaster Type	Scope	Restore Destination
Drive failure	Logical unit	The P-VOL or logical unit of a normal RAID Group after the drive is replaced
Subsystem failure	Subsystem	Separate subsystem
Site failure	Site	Subsystem at a separate site (e.g., in the same city)
Natural disaster	Extensive	Subsystem at a geographically remote site (outside the city or state)

If a drive failure occurs P-VOL data can be restored at the volume level after the drive has been replaced. If a subsystem failure occurs, data is restored to an alternate subsystem at the same primary site. In the event of a site failure or natural disaster, data is restored to a subsystem at a separate site. The distance of the alternate subsystem from the original site, depends on the geographic impact of the disaster.

6.1.4 Recovery from Path Failure

When a remote path failure occurs it is necessary to perform maintenance tasks such as part replacement. If the path does not recover automatically after maintenance, path recovery operations are executed using Web GUI. After the path is recovered, TCE pairs with PSUE status must be recovered using CCI. For more information on recovery from path failure, see *Hitachi TagmaStore™ Adaptable Modular Storage TrueCopy Extended Distance Configuration Guide, MK-95DF779-00.*

6.1.5 Recovery from Pool Over

When a pool overflow occurs for a data pool shared by TCE and SnapShot, TCE pairs transition to PFUS status and SnapShot pairs transition to PSUE status. Recovery from pool overflow is performed by allocating additional logical units to increase unused pool capacity. TCE pairs should then be recovered by performing pair resynchronization (**pairresync** operation). SnapShot pairs should also be recovered, by first deleting the pairs (**pairsplit -S** operation) and then recreating them (**paircreate** operation).

6.2 Data Recovery Considerations

6.2.1 Concurrency of Secondary Storage Volumes

There is a possibility for the concurrency of S-VOL data and P-VOL data to be compromised since remote copy functions work independent of the caching mechanism of a host operating system. Data updated by an application or a file system is cached in the memory of a host and updated on the P-VOL using write operations. The latest data cached on the host between any two write operations, will therefore be unavailable on the P-VOL when remote copy operations are performed during that time interval. As a result, the data in the S-VOL may not be identical to the primary data in the P-VOL. For example in Figure 6.1 data "4" is cached in a host memory and not written to the P-VOL. Therefore data "4" cannot be read from the P-VOL during a remote copy to the S_VOL.

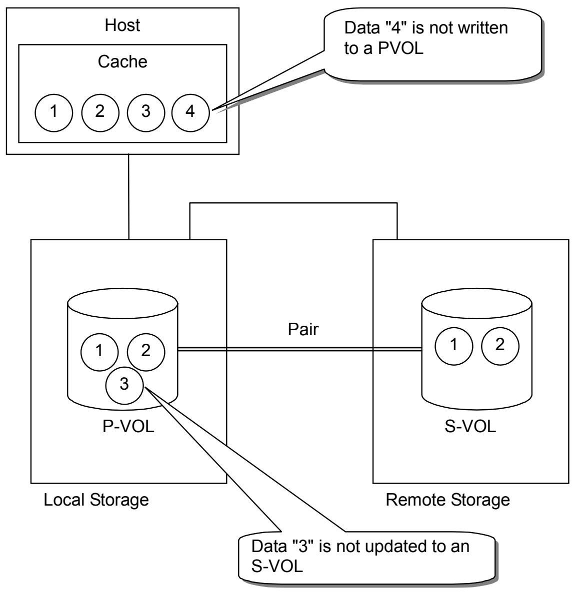


Figure 6.1 Concurrency of S-VOL

To synchronize the P-VOL and S-VOL, the remote copy function must be updated with the cached data from the host memory. To enable this, cached data in a host memory needs to be written to a P-VOL by unmounting the file system on the host. With this operation, the P-VOL is updated with latest data and subsequent remote copy operations will ensure that S-VOL is concurrent.

6.2.2 Consistency of Secondary Storage Volumes

Data consistency at the remote site enables quicker restart of operations upon disaster recovery. When remote takeover operations are performed, internally pre-determined data from the remote subsystem data pool is restored simultaneously to the corresponding S-VOLs in a remote CTG. Since there are no data discrepancies on the S-VOLs, host applications at the remote site will always access consistent data.

When pairs belonging to a CTG are split and deleted, differential data on the P-VOLs is determined and transferred to the S-VOLs, before the pairs are deleted. Pair split operations are completed only after the differential data is transferred. Data consistency across pairs in a CTG is therefore ensured during pair split operations.

Note: Data consistency of remote S-VOLs is not guaranteed when only the pair relationship is deleted at the remote TagmaStore AMS subsystem (using -R option of pairsplit command). Data discrepancies may result since S-VOL update operations are interrupted before a cycle update completes.

6.2.3 Pair Creation/Resynchronization for each CTG

In the pair creation/resynchronization performed with a specification of a certain group, renewal of the cycle is started from a pair for which the initial copy is completed first and the status of the pair above is changed to PAIR. A pair, initial copy for which is not completed before the first cycle renewal, renews the cycle taking the next occasion for the renewal. Therefore, the time when the status of the each pair is changed to PAIR may differ from the other one by the cycle time length.

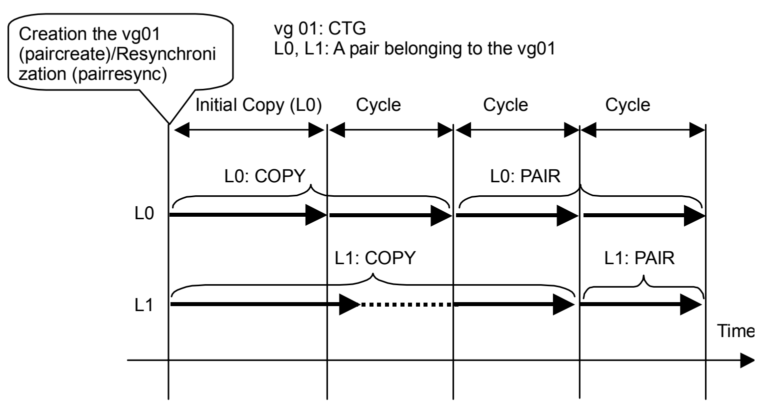


Figure 6.2 Pair Creation/Resynchronization for each CTG-1

In the pair creation/resynchronization newly performed for each CTG, the time for renewing the cycle is decided by the pair for which the initial copy is completed first. The pair, the initial copy for which is completed later than the first completion of the initial copy, employs the renewed cycle starting from the cycle after next at the earliest.

When a pair is newly formed or resynchronized for a CTG for which the cycle has already been renewed, the initial copy is made not being interlocked with the cycle but the cycle is started after the initial copy is completed in a manner that the pair employs the existing cycle.

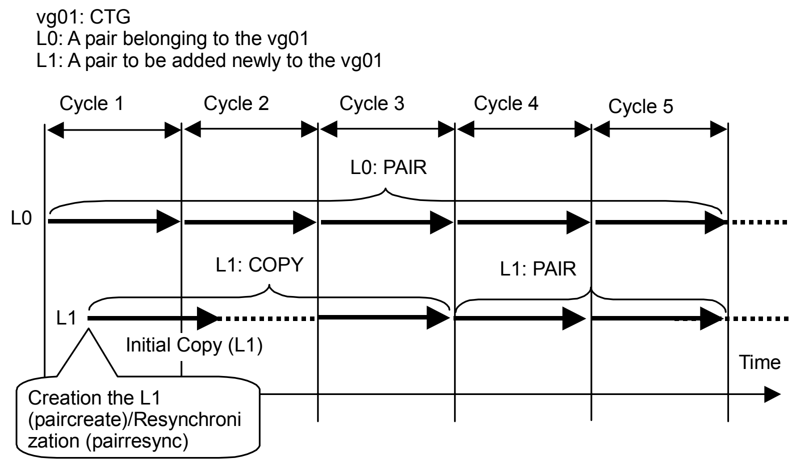


Figure 6.3 Pair Creation/Resynchronization for each CTG-2

When a pair is newly added to the CTG, the pair is synchronized with the existing cycle timing. In the example, the pair is synchronized with the existing cycle from the cycle 3 and its status is changed to PAIR from the cycle 4.

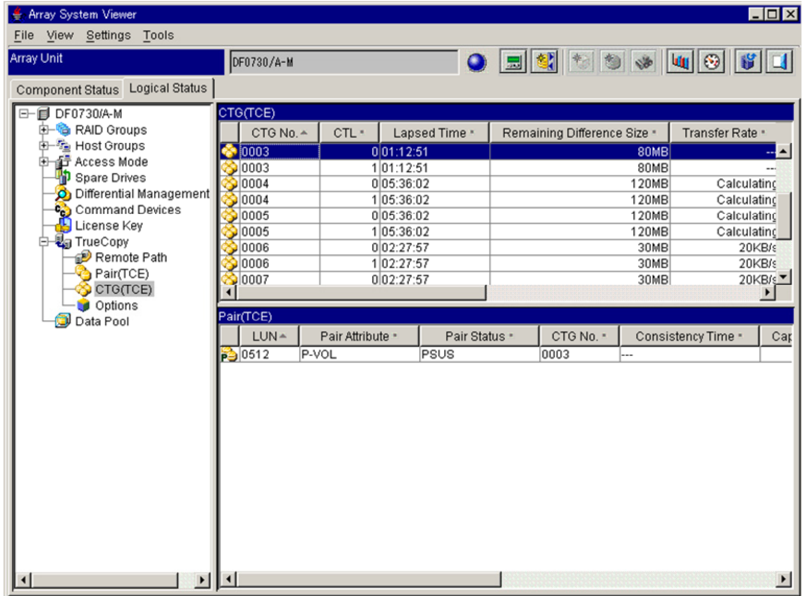
When the paircreate or pairresync command is executed, the pair undergoes the differential copy in the COPY status, undergoes the cyclic copy once, and then placed in the PAIR status. When a new pair is added to a CTG, which is already placed in the PAIR status, by the paircreate or pairresync command, the copy operation halts until the time of the existing cyclic copy after the differential copy is completed. Further, it is not placed in the PAIR status until the first cyclic copy is completed after it begins to act in time to the cycle. Therefore, the pair synchronization rate displayed by Navigator or CCI may be 100% or not changed when the pair status is COPY.

When you want to confirm the time from the stop of the copy operation to the start of the cyclic copy, check the start of the next cycle by displaying the predicted time of completing the copy using Navigator. The procedure for displaying the predicted copy completion time is discussed in section 6.2.4.

6.2.4 Confirming the Status of CTG

The CTG status can be verified in Navigator. CTG status is displayed in the **Logical Status** page of the Array System Viewer Panel.

When one of the CTG numbers displayed on the **CTG(TCE)** pane is selected, the pairs included in the selected CTG are displayed on the **Pair(TCE)** pane. The items of contents displayed on the **Pair(TCE)** pane are the same as those of the pair information shown in section 7.1.1.11.



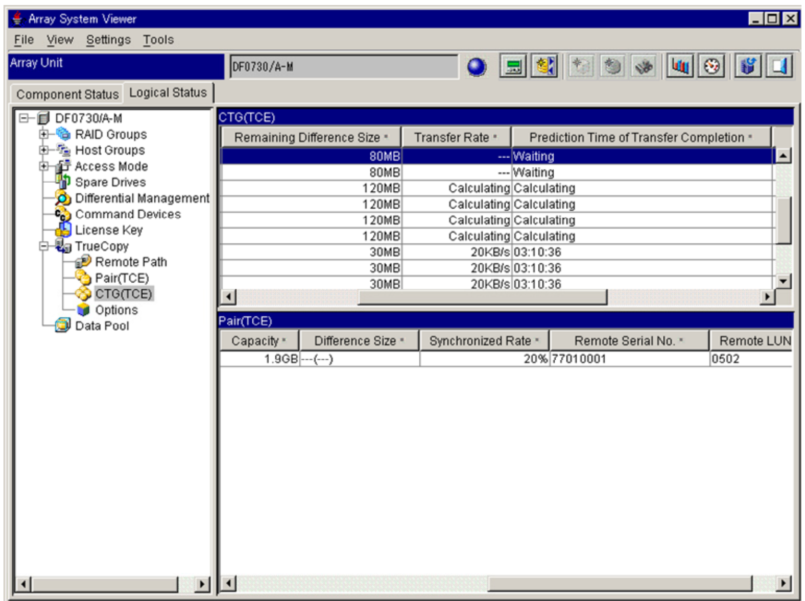


Figure 6.4 Array System Viewer Panel (Logical Status Page: TCE CTGs)

Table 6.3 CTG Information

Displayed Item	Contents
CTG No.	CTG number
CTL	Controller number
Lapsed Time	The Lapsed Time after the current cycle is started is displayed (in hours, minutes, and seconds).
Remaining Difference Size	A size of the residual differential data to be transferred in the current cycle is displayed. The size of the data, which has not been transferred to the remote subsystem and thus remains in the local subsystem among the differential data generated through the host I/O accepted by the P-VOL, is displayed for each P-VOL. When the pair status is PAIR, the size of the differential data tends to repeat increase and decrease because the differential data generated by that the P-VOL accepts the host I/O instruction at the time of PAIR is eliminated regularly by the cyclic copy. In the current cyclic operation, the size of the differential data may not be reduced to zero even when the current cycle ends because the differential data at the time of the cycle start is transferred to the remote subsystem and the differential data generated after the cycle start is transferred to the remote subsystem in the following cycle.
Transfer Rate	The Transfer Rate of the current cycle is displayed (kB/s). During a period from the start of the cycle to the execution of the copy operation or a waiting period from completion of the copy operation to the start of the next cycle, "---" is displayed. While the Transfer Rate to be output is being calculated, "Calculating" is displayed.
Prediction Time of Transfer Completion	The predicted time when the data transfer is completed (Prediction Time of Transfer Completion) for each cycle of the CTG is displayed (in hours, minutes, and seconds). If the predicted time when the data transfer is completed (Prediction Time of Transfer Completion) cannot be calculated because it is maximized temporarily, "99:59:59" is displayed. During a waiting period from completion of the cyclic operation till the start of the next cycle, "Waiting" is displayed. While the predicted time to be output is being calculated, "Calculating" is displayed.

6.3 Disaster Recovery Processes

In the event of a disaster at the primary site, the cycle update process is suspended and updating of the S-VOL stops. If the host requests an S-VOL takeover (horctakeover), the remote storage subsystem restores the S-VOL using internally determined data from the previous cycle.

6.3.1 Salvaging P-VOL Updates

In an ideal disaster recovery scenario, the remote subsystem also waits for the local subsystem to send un-transferred data updates, determined on the P-VOL at the time the takeover command was received. Provided the local subsystem and remote path are functional, latest data on the P-VOL is made available on the S-VOL (mirroring consistency) before takeover is executed.

This version of TCE does not support mirroring consistency of S-VOL even though the local storage subsystem and the remote path are functional. Therefore when a horctakeover command is issued, data available on the S-VOL is the internally pre-determined data created during the previous update cycle. P-VOL and S-VOL data are therefore not identical when takeover is executed and P-VOL data updates made during the time the takeover command was issued cannot be salvaged.

6.3.2 Takeover Processing

S-VOL takeover is performed when the horctakeover operation is issued by the secondary subsystem. The TCE pair is split and system operation can be continued with the S-VOL only. In order to settle the S-VOL data being copied cyclically, it is restored using the data that was pre-determined in the preceding cycle and saved to the data pool. The S-VOL is immediately enabled to receive the I/O instruction (placed in the SSWS status) although the performance of the host I/O execution is generally lowered because the restoration of the settled data is done internally. The P-VOL and S-VOL data are not the same after this operation is performed.

6.3.3 Swapping P-VOL and S-VOL Pair

The SWAP Takeover is an operation to continue the system operation by reversing the characteristics of the P-VOL and the S-VOL. This function swaps a relationship between P-VOL and S-VOL. After S-VOL takeover, host operations continue on the S-VOL and S-VOL data becomes updated as a result of I/O operations. When continuing application processing using the S-VOL or when restoring application processing to the P-VOL, the swap function makes P-VOL up-to-date, by reflecting updated data on the S-VOL to the P-VOL.

Note: The swap function, which sends the updated data of the S-VOL to the P-VOL by reversing the P-VOL and the S-VOL relationship, is not supported in this version of TCE. A TCE pair must be deleted and then recreated in reverse order to copy updated data from S-VOL to P-VOL.

6.4 Sample Disaster Recovery Scenarios

This section discusses sample scenarios for the following disaster recovery procedures:

- Restoration after Logical Data Destruction
- Restoration after P_VOL Failure
- Executing S-VOL Takeover (horctakeover)
- Fallback to a Local P-VOL

6.4.1 Restoration after Logical Data Destruction

When data is logically destroyed, the P-VOL itself is normal. Therefore, the backup data is restored to the P-VOL. If only a few files have been destroyed, the data is restored at the file level. Figure 6.5 outlines restoration at the file level.

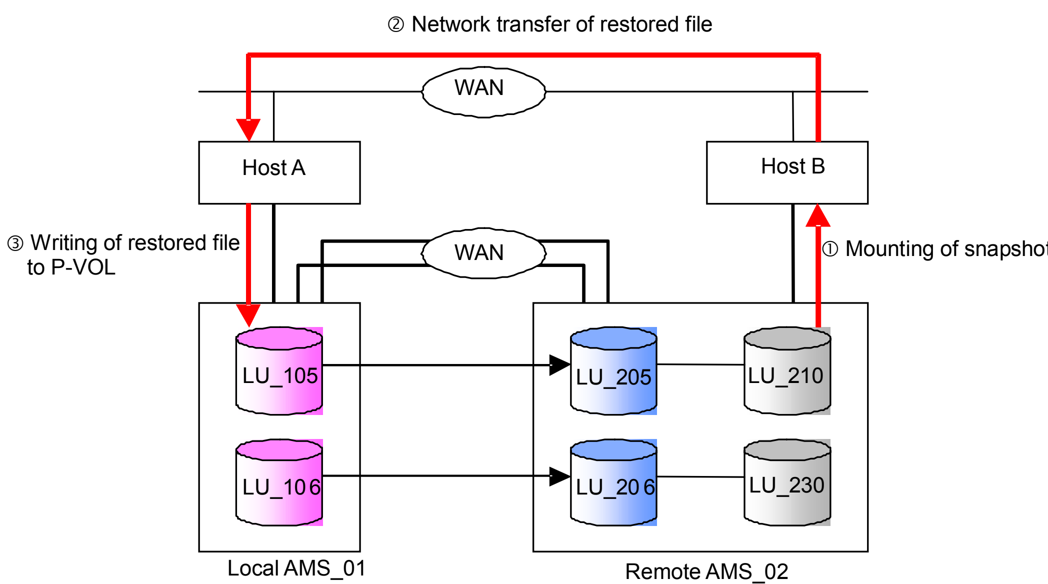


Figure 6.5 Restoring Destroyed Data at the File Level

The snapshot that holds the files to be restored is first found from the snapshots obtained from the remote AMS_02. You can find the data to be restored by referring to the character string (Split-Marker in Figure 6.6) and backup time obtained from CCI or Navigator.

```
C:\HORCM\etc>pairdisplay -g vg30 -v smk
Group   PairVol L/R   Serial# LDEV# P/S Status UTC-TIME -----Split-Marker------
vg30    dev00 L    75003002   205  P-VOL PSUS        - -
vg30    dev00 R    75003002   210  S-VOL SSUS 43d3b076             KM19781028
vg30    dev01 L    75003002   206  P-VOL PSUS        - -
vg30    dev01 R    75003002   230  S-VOL SSUS 43d3b076             KM19781028
```

```
C:\HORCM\etc>
```

Figure 6.6 Sample Remote Snapshot Information using CCI

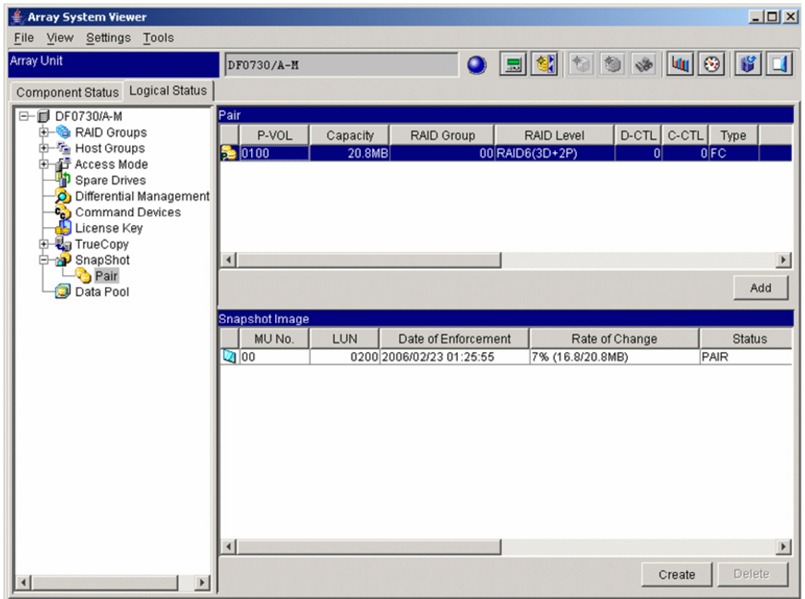


Figure 6.7 Sample Remote Snapshot Information in Navigator

After determining the snapshot that holds the data to be restored, mount that snapshot on the server WIN_20 at the remote side (① in Figure 6.5). Next, determine the file to be restored by accessing the volume mounted on the WIN_20 server. Copy the restored file to the temporary disk area of the WIN_10 server by using a file transfer tool such as FTP (②).

The WIN_10 server stops and copies the restored file into the temporary disk area to an appropriate location if the application that uses the restored file is operating (③). Check the contents of the copied file. If the application operates correctly, the restoration task has been completed.

6.4.2 Restoration after P_VOL Failure

The procedure of restoring P-VOL data if the P-VOL is hit by a disaster is explained below. To simplify the explanation, the restoration of data to the logical unit of normal RAID Group of #1 is discussed in the following example.

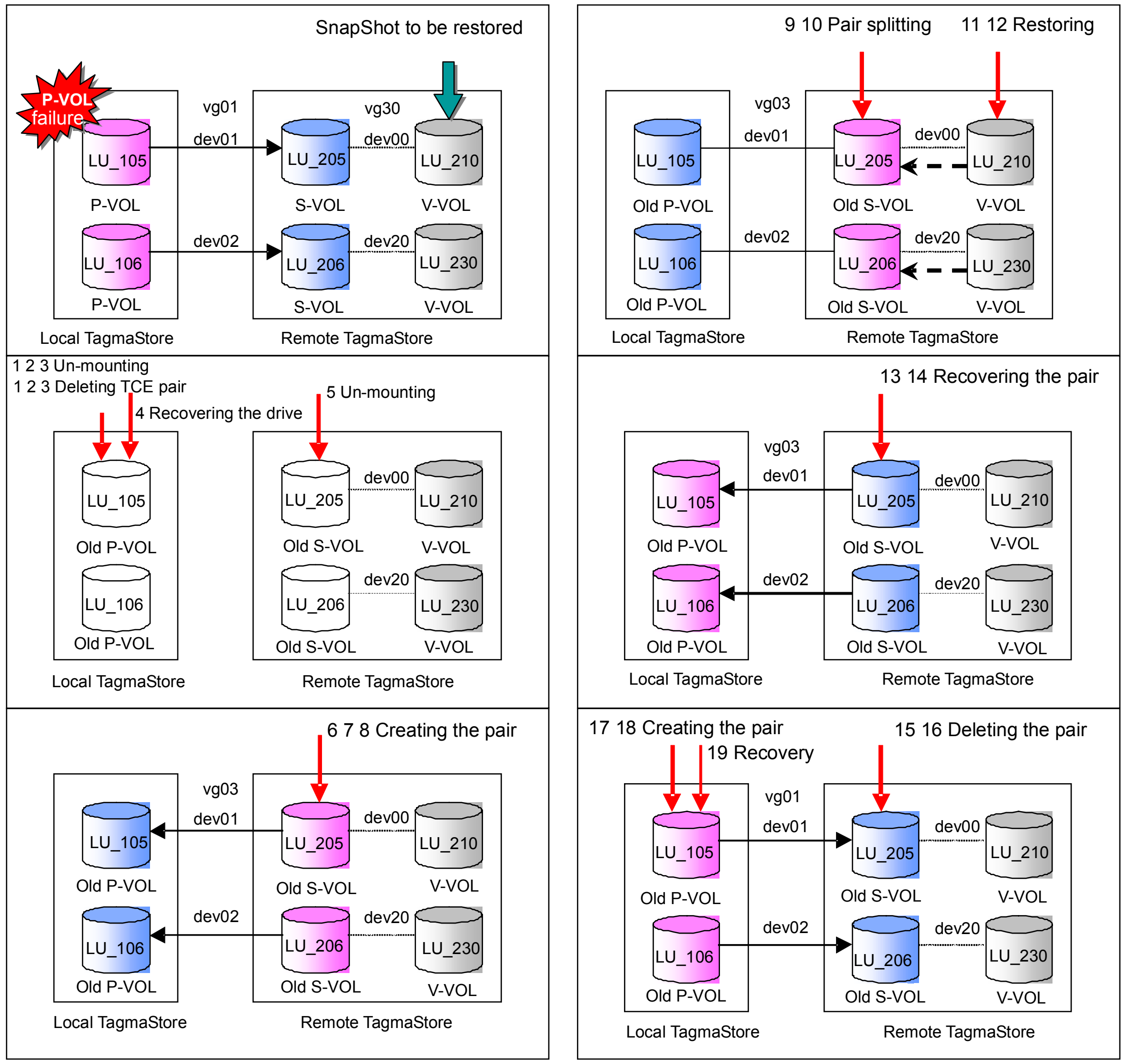


Figure 6.8 Restoration Procedure (outline)

Table 6.4 Restoration Procedure (detailed steps)

#	Local Side	Remote Side
1	Stop writing to and un-mount the old P-VOL.	Stop writing to the old S-VOL.
2	Delete the TCE pair. # pairsplit –S –g vg01	
3	Wait for completion of deleting the TCE pair. # pairevtwait –g vg01 –s smpl	
4	Recover the drive that has failed (done by Hitachi personnel).	
5		Un-mount the S-VOL (only when mounted).
6		Create a configuration definition file. The TCE pair is created from the old S-VOL for the old P-VOL.
7		Create the TCE pair. # paircreate –g vg03 –vl Note: If the contents of the old P-VOL and the old S-VOL are the same, initial copy processing can be omitted by using -nocopy.
8		Wait for the PAIR status of the TCE pair. # pairevtwait –g vg03 –s pair
9		Split the TCE pair. # pairsplit –g vg03
10		Wait for the PSUS status of the TCE pair. # pairevtwait –g vg03 –s psus
11		Restore the SnapShot. # pairresync –restore –g vg30
12		Wait for the completion of restoring. # pairevtwait –g vg30 –s pair
13		Recover TCE pair. # pairresync –g vg03
14		Wait for the PAIR status of the TCE pair. # pairevtwait –g vg03 –s pair
15		Delete the pair. # pairsplit –S –g vg03
16		Wait for the completion of deleting the TCE pair. # pairevtwait –g vg03 –s smpl
17	Create a configuration definition file. The file existing before the occurrence of failure is used.	
18	Create the pair (without copy). # paircreate –nocopy –g vg01 –vl	
19	Mount the P-VOL and perform recovery processing.	

The following explanation uses the information in Figure 6.8 and Table 6.4.

Tasks at the local site:

1. To prepare for restoration, stop writing from the WIN_10 server to the P-VOL, and un-mount the P-VOL.
2. Delete the TCE pair from WIN_10 (vg01).
3. Wait for the completion of deleting the TCE pair by WIN_10.
4. Replace the drive on which the failure has occurred. Hitachi maintenance personnel replace the drive.

Tasks at the remote site:

5. If the S-VOL (LU_205 and LU_206) is mounted, stop writing from the WIN_20 server to the S-VOL, and un-mount the S-VOL.
6. Create a configuration definition file for a new TCE pair (vg03) on WIN_20. In this definition file, describe the configuration of TCE for LU_205 and LU_206 of the remote AMS_02 for LU_105 and LU_106 of the local AMS_01.
7. Create a new TCE pair from WIN_20. If it is ensured that P-VOL data and S-VOL data are identical, no initial copy is required.
8. Wait until the new TCE pair transits to the PAIR status.
9. After the transition to the PAIR status, split the new TCE pair to restore data from the snapshot (LU_210 and LU_230) to the P-VOL (LU_205 and LU_206) of the new TCE pair.
10. Wait until the TCE pair enters the PSUS status.
11. After the transition to the PSUS status, restore the snapshot to the P-VOL of the new TCE pair.
12. Wait for the completion of restoration.
13. After the completion of restoration, re-synchronize the TCE pair, and copy the restored data to the S-VOL of the new TCE pair (volume at the restore destination, LU_105 and LU_106).
14. Wait until the new TCE pair enters the PAIR status.
15. Delete the new TCE pair (vg03). At this time, confirm that no data has been written to the P-VOL and the S-VOL.
16. Wait for the completion of deleting the new TCE pair.

Tasks at the local site:

17. To create a TCE pair for the volume of the remote AMS_02 from the restored volume, create a configuration definition file. If the logical unit number of the P-VOL of the local AMS_01 has not changed, the configuration definition file for backup operation can be used as is.

18. Create a TCE pair (vg01). At this time, initial copy processing can be omitted by using the -nocopy option if no data has been written to the P-VOL and the S-VOL from step 15 to step 18.
19. Finally, mount the P-VOL (LU_105 and LU_106) onto WIN_10, and recover the data. This completes the processing.

Note: If a large number of files have been corrupted because of a virus infection or if most of the files on the drive have been erased, the P-VOL is still normal but the entire file system should be restored. In this case, restore the file system at the volume level according to the procedure for restoration when the P-VOL has failed. At this time, initial copy in step 7 above is unnecessary, since the contents of the P-VOL and the S-VOL match as a result of deleting the pair in step 2.

6.4.3 Executing S-VOL Takeover (horctakeover)

The following method explains the takeover process for continuing the operation with only the S-VOL. The example shows CCI commands issued by a remote host for takeover operations.

1. In this example, the group name in the configuration definition file is vg01. Check that there is a TCE pair you want to change the operation to that uses the S-VOL only.

```
c:\horcm\etc>pairdisplay –g vg01
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV#
M
vg01    oradb1(L)   (CL1-A , 1,  1 )75000174     1.P-VOL PSUE ASYNC ,75000175
2  -
vg01    oradb1(R)   (CL1-B , 1,  2 )75000175     2.S-VOL PAIR ASYNC ,-----
1  -
```

2. Change the pair operation to use only the S-VOL by executing the horctakeover command.

```
c:\horcm\etc>horctakeover –g vg01
```

3. Execute the pairdisplay command to verify that the pair status changed to SSWS.

```
c:\horcm\etc>pairdisplay –g vg01
Group   PairVol(L/R) (Port#,TID, LU) ,Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV#
M
vg01    oradb1(L)   (CL1-A , 1,  1 )75000174     1.P-VOL PSUE ASYNC ,75000175
2  -
vg01    oradb1(R)   (CL1-B , 1,  2 )75000175     2.S-VOL SSWS ASYNC ,-----
1  -
```

The operation was changed to use the S-VOL only. For details on the **horctakeover** command, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

When the SVOL_Takeover is executed, the operation can be started at the S-VOL immediately, but data restoration processing from a POOL of the secondary site to the S-VOL is performed in the background. During the period from the execution of the SVOL_Takeover until the completion of the data restoration processing, performance of the host I/O for the S-VOL is deteriorated by approximately 60%.

The completion of the data restoration processing is known through a change of the pair status, which is displayed by Navigator, from SSWS(R) to SSWS. Besides, a percentage of the size of the data, which is not yet copied from the POOL to the S-VOL, is displayed as the “differential data size” in the pair information, and it is used as a standard value to estimate the time until the data restoration processing is completed.

6.4.4 Fallback to a Local P-VOL

Fallback process involves restarting business operations at local site after the storage subsystems have been recovered. A typical fallback process is shown in Figure 6.9

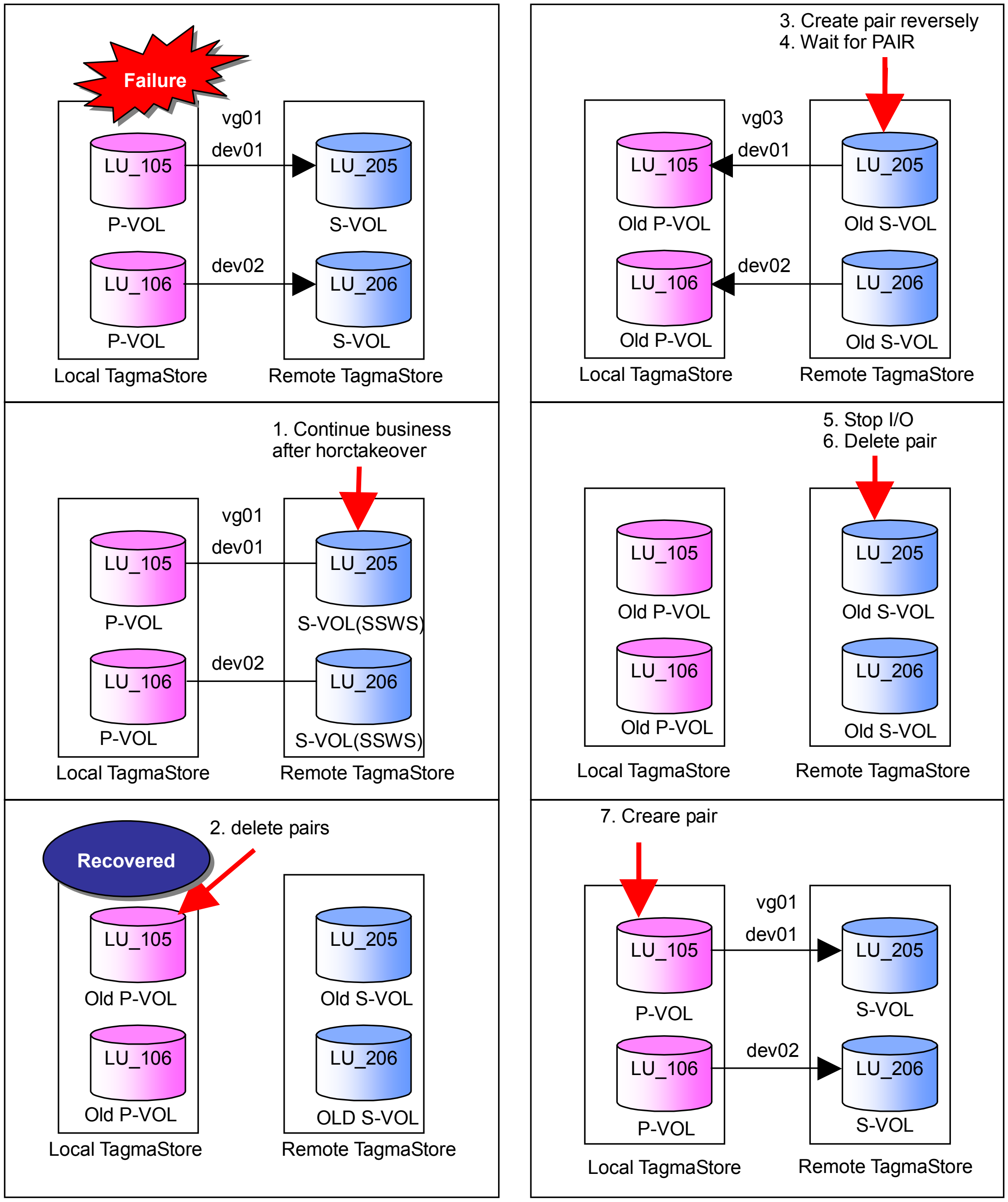


Figure 6.9 Fallback Procedure

Table 6.5 Fallback Procedure (detailed steps)

#	Local Side	Remote Side
1		Execute horctakeover and continue businesses at remote
2	Delete TCE pair # pairsplit -S -g vg01	
3		Create TCE pair reversely # paircreate -f async -g vg03 -vl
4		Wait for PAIR status transition # pairevtwait -g vg03 -s pair
5		Stop I/O and un-mount the S-VOL
6		Delete TCE pair # pairsplit -R -g vg03 Wait for SMPL status transition # pairevtwait -g vg03 -s smpl
7	Create a new pair # paircreate -f async -nocopy -g vg01 -vl	

Tasks at the remote site:

1. Issue horctakeover command from a remote host and make S-VOL available for the remote host. If succeeded, businesses can continue on the S-VOL.

Tasks at the local site:

2. When a local site is recovered after the disaster and continuing businesses at local is required, perform fallback process. First delete the TCE pair from the local host.

Tasks at the remote site:

3. Create a new TCE pair reversely. So old S-VOL and old P-VOL are paired.
4. Wait for a completion of an initial copy.
5. After transiting to PAIR status, stop I/O to S-VOL and un-mount it. At this point, all data updated on old S-VOL after Takeover is reflected to old P-VOL.
6. Delete the pair and wait for a completion.

Tasks at the local site:

7. Create a TCE pair from old P-VOL to old S-VOL again. In this case, no initial copy is required.

Chapter 7 Managing a TCE System

This chapter provides an overview of TCE pair operations and discusses TCE system monitoring, maintenance and system reduction tasks and associated TCE pair operations. This chapter includes the following sections:

- TCE Pair Operations (section 7.1)
- System Monitoring (section 7.2)
- System Maintenance (section 7.3)
- System Reduction (section 7.4)

7.1 TCE Pair Operations

TCE pair operations are used to build and operate a TCE system. Pair operations are performed by commands issued from CCI. Table 7.1 lists TCE pair operations and their corresponding CCI commands. The command issued depends on the tasks to performed, such as system configuration, operation, monitoring, maintenance, or recovery.

Table 7.1 System Management Tasks and CCI Commands

Tasks	TCE Operations	Command Name	Contents	Remarks
Configuration and expansion	Pair creation	paircreate	Creates a TCE pair.	
	Pair configuration check	pairdisplay	Checks the configuration of the created TCE pair.	
	PAIR status waiting	pairevtwait	Waits for transition to the PAIR status after initial copy has been completed.	Waits for completion of processing, such as initial copy and pair splitting, takes time.
Operation and management	Pair status monitoring	pairdisplay	Displays the pair status to confirm whether operation is normal.	The -fc option shows the detailed information of the pair status.
	Synchronization monitoring	pairsyncwait	- Checks the time difference between the P-VOL and S-VOL. - Also waits for the completion of copying of specific data written to the P-VOL to the S-VOL.	The larger the time difference between the P-VOL and S-VOL, the poorer the RPO performance.
	Remote snapshot	pairsplit -mscas	Creates the remote snapshot of the TCE pair P-VOL.	
Maintenance	Pair splitting	pairsplit	Temporarily stops update copy processing for the specified TCE pair.	A pair is split after the P-VOL and S-VOL become identical.
	Pair recovery	pairresync	- Resumes the remote copy operation after the completion of maintenance. - Also resumes the remote copy operation for a TCE pair that was suspended due to a failure.	
Disaster recovery	Takeover	horctakeover	Continues operation using the S-VOL.	- Untransmitted data that is on the P-VOL when the command is issued is not reflected on the S-VOL. - Takeover can be executed only when

Tasks	TCE Operations	Command Name	Contents	Remarks
				the pair status of the S-VOL is PAIR, PSUE, PFUS, or PSUS.
System reduction	Pair deletion (Normal)	pairsplit -S	Deletes a TCE pair if the asynchronous remote copy operation is no longer necessary.	
	Pair deletion (Remote Only)	pairsplit -R	Deletes the S-VOL if the S-VOL cannot be correctly deleted for a reason such as a path failure.	

Pair operations for configuring and initializing TCE functions are discussed in Chapter 5 Remote Data Replication with TCE.

7.1.1 Pair Status Transitions

Pair status transitions occur as a result of TCE pair operations and can be used to monitor TCE operations and detect system failures. Figure 7.1 shows pair status transitions during normal functioning of a TCE system.

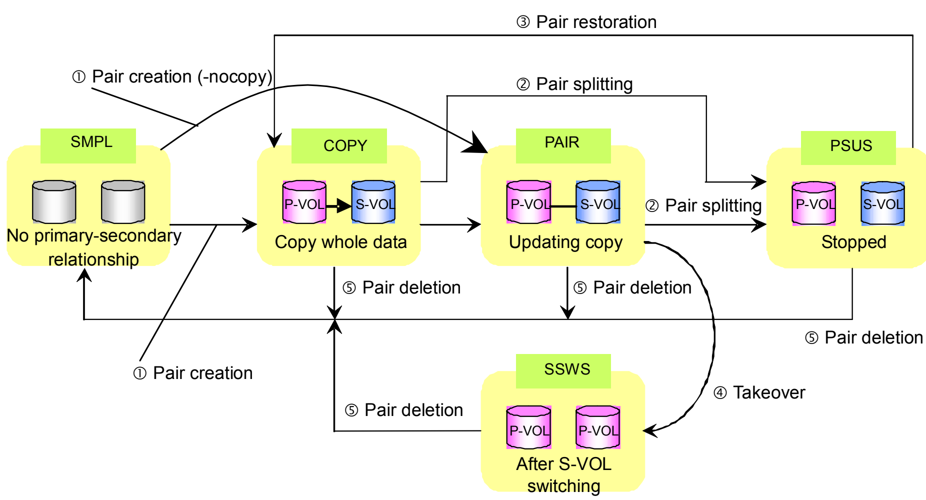


Figure 7.1 Pair Status Transitions

7.1.1.1 SMPL → COPY Transition

A data volume is initially in the SMPL status, before it is paired with another data volume. When a TCE pair is created, the pair status changes to COPY (① in the above figure).

The COPY status indicates that an initial copy (copying of all data already in the P-VOL to the S-VOL) is in progress. If there is no update to a P-VOL during an initial copy, P-VOL data and S-VOL data are identical after a completion of the initial copy. If there are any updates to the P-VOL, there is untransferred differential data in the P-VOL and S-VOL data will not be identical to the data in the P-VOL.

When a **paircreate** command is executed with a **-nocopy** option, the new TCE pair created transitions directly from SMPL status to PAIR status since no initial copy is performed. This option can be used if P-VOL data and S-VOL data are identical.

7.1.1.2 COPY → PAIR Transition

When initial copy has been completed, the pair status changes to PAIR. When a TCE pair is in PAIR status, there is data consistency in the CTG and takeover of the S-VOL is possible. Cycle update processing can be performed only when a TCE pair is in PAIR status.

Note: When multiple TCE pairs are created simultaneously in a CTG, each initial copy is performed in separate cycles. TCE pairs in the CTG transition to PAIR status a minimum of one cycle after the first TCE pair transitions to PAIR status.

7.1.1.3 PSUS Transition

TCE pairs transition to the PSUS status when an initial copy or cycle update copy operation is stopped as a result of pair splitting. The **pairsplit** command changes pair status from COPY or PAIR to PSUS (②). If a pair is split before completion of an initial copy, S-VOL data is inconsistent with P-VOL data. If a pair is split during a cycle update copy, data consistency is guaranteed across pairs in a CTG. S-VOL data becomes identical to P-VOL data at the time the split command was issued to the P-VOL.

The local subsystem transfers all P-VOL data to the S-VOL, before the **pairsplit** command is processed. The completion of a pair split operation may therefore be delayed depending on the volume of differential data in the P-VOL. In the PSUS status, the host can read and write the S-VOL. The local subsystem and the remote subsystem use snapshots (bitmap tables) to record data updates to the P-VOL and S-VOL, when the pair is in PSUS status. These snapshots are used to resynchronize the TCE pair when they transition to the COPY status during pair resynchronization.

7.1.1.4 PSUS → COPY Transition

TCE pairs transition from the PSUS to COPY status, when a pair resynchronization command is executed (Figure 7.1 ③). The remote subsystem tracks updates made to the S-VOL when it is in PSUS status. The local subsystem gets these updates from the remote subsystem, and executes a resynchronization copy operation to restore the P-VOL data to the S-VOL. After resynchronization, data updates that were written to the S-VOL during the PSUS period are lost.

7.1.1.5 SSWS Transition

TCE pairs transition to SSWS (Suspend for Swapping with S-VOL side only) status when an S-VOL takeover is executed. Update copy processing is stopped on transition to SSWS status. The S-VOL takeover **(horctakeover)** command can be executed only when the pair status of the S-VOL is PAIR, PSUE, PSUS or PFUS.

When the **horctakeover** command is executed at the remote subsystem, internally pre-determined data is restored to the S-VOL. An application can read and write consistent data from the S-VOL when it is in SSWS status (④). SSWS pair status can be confirmed by using the **-fc** option of the **pairdisplay** command.

Even if pairs within the same CTG have different pair statuses, consistency of each S-VOL is guaranteed. However since S-VOLs are determined at different points of time, write order is not guaranteed across the CTG.

7.1.1.6 SMPL Transition

A TCE pair transitions to SMPL status when a pair deletion command (**pairsplit -S**) is performed to delete the pair relationship. When a pair transitions from PAIR to SMPL status, the local subsystem transfers all data updated before the **pairsplit -S** command was received, to the S-VOL. The completion time of a **pairsplit -S** command therefore depends on the amount of differential data at P-VOL.

A TCE pair will not be deleted by a **pairsplit -S** command, if there is a remote path failure during execution or if the S-VOL is in SSWS(R) pair status. In such a case a **pairsplit -R** command is necessary to delete the pair. If a pair in COPY status is deleted, S-VOL data is inconsistent because the initial copy operation is interrupted. If a pair in PAIR status is deleted, data consistency across pairs in a CTG is guaranteed and S-VOL data becomes identical to P-VOL data at the time the split command was issued to the P-VOL.

7.1.1.7 PSUE Transition

A TCE pair transitions to PSUE status when the subsystem has detected a failure and can no longer continue copy processing with the TCE pair. For example, when the line between the local and remote subsystem is disconnected or if a dual drive failure has occurred on the P-VOL or S-VOL. All pairs in a CTG transition to PSUE status simultaneously, even if one pair in the CTG has failed.

If the TCE pair transitioned from PAIR to PSUE status, data consistency across pairs in a CTG is guaranteed. If the TCE pair transitioned from COPY to PSUE status, data consistency is not guaranteed. To recover from the PSUE status and restore the PAIR status, the failure needs to be rectified and the TCE pair must be resynchronized.

7.1.1.8 PFUS Transition

A TCE pair transitions to PFUS status when the used capacity of the data pool on the local or remote subsystem has reached to its upper limit. When pool capacity is exceeded, any un-transferred data in the local subsystem or internally pre-determined S-VOL data in the remote subsystem cannot be copied to the pool. Therefore update copy processing cannot be continued. For a pair to recover from the PFUS status to the PAIR status, it is necessary to increase pool capacity and resynchronize the pair.

If a pair transitions from PAIR status to PFUS status, data consistency across pairs in a CTG is guaranteed. If a pair transitions from COPY status to PFUS status, no data consistency is guaranteed. All pairs using a common data pool transition to PFUS status simultaneously, when pool capacity is exceeded. When a pair (using the exceeded data pool) belongs to more than one CTG, all CTGs associated with it transition to PFUS status, even if they use other exceeded data pools.

The PFUS status can be checked by using the **-fc** option of the **pairdisplay** command. Unless the **-fc** option is specified, CCI displays the PFUS status as the PSUS status.

7.1.1.9 PSUS(N) Transition

A TCE pair transitions to PSUS(N) status as a result of some system failure such as when a dual drive failure occurs on a pool when the pair is in SSWS(R) status. The PSUS(N) status differs from the PSUE status which also results from system failure. Data consistency in the S-VOL is guaranteed in the PSUE but not in the PSUS(N) status.

For a pair to recover from the PSUS(N) status and be restored to the PAIR status, it is required to eliminate the cause of the system failure and resynchronize the volume pair. In this case, the resynchronization copy needs to be a full volume copy since the S-VOL data is inconsistent.

The PSUS(N) status can be checked by using the **pairdisplay** command. PSUS(N) pair status is confirmed if the result shows PSUS status and the item “N”.

7.1.1.10 SSWS(R) Transition

A TCE pair transitions to SSWS(R) status when internally pre-determined data is being restored to the S-VOL, during an S-VOL takeover. When restoration is complete the pair changes to SSWS status. Restoration time can be estimated based on the amount of differential data between data volumes in a pair.

The SSWS(R) pair status disables deletion of S-VOL when a pair deletion operation is executed during S-VOL restoration. In such a case, the P-VOL transitions to SMPL status while the S-VOL remains in SSWS(R) status. A **pairsplit -R** command is necessary to delete the S-VOL in the pair, after the S-VOL transitions to SSWS status. Read or write operations can be performed on the S-VOL only when it is in SSWS status.

7.1.1.11 Confirming Status of TCE Pairs

TCE pair statuses can be verified using Navigator or CCI. In CCI, pair status can be obtained by executing the **pairdisplay** command. For more information on **pairdisplay** command, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

In the Navigator GUI, pair status information can be displayed by clicking the **Pair(TCE)** icon under the **TrueCopy** folder on the **Logical Status** tab. Figure 7.2 and Figure 7.3 illustrate TCE pair information displayed on the **Array System Viewer** panel.

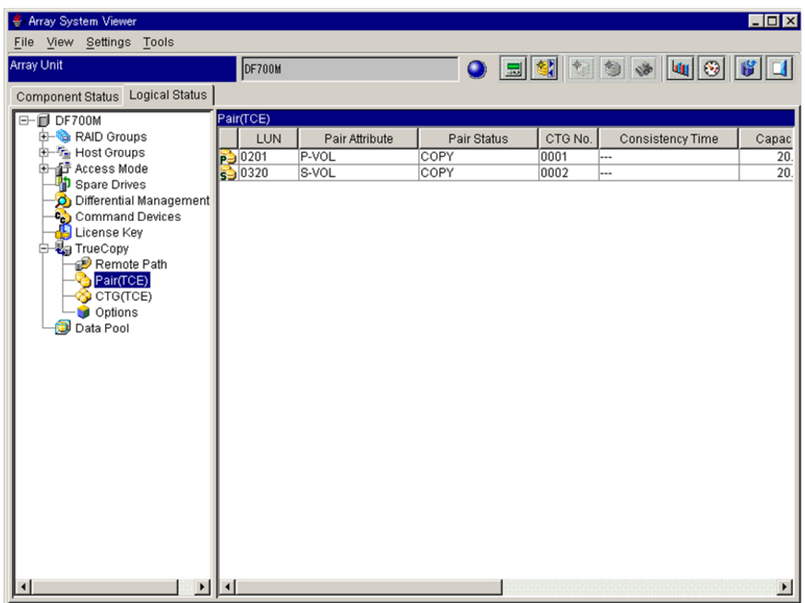


Figure 7.2 Array System Viewer Panel (TCE Pairs)

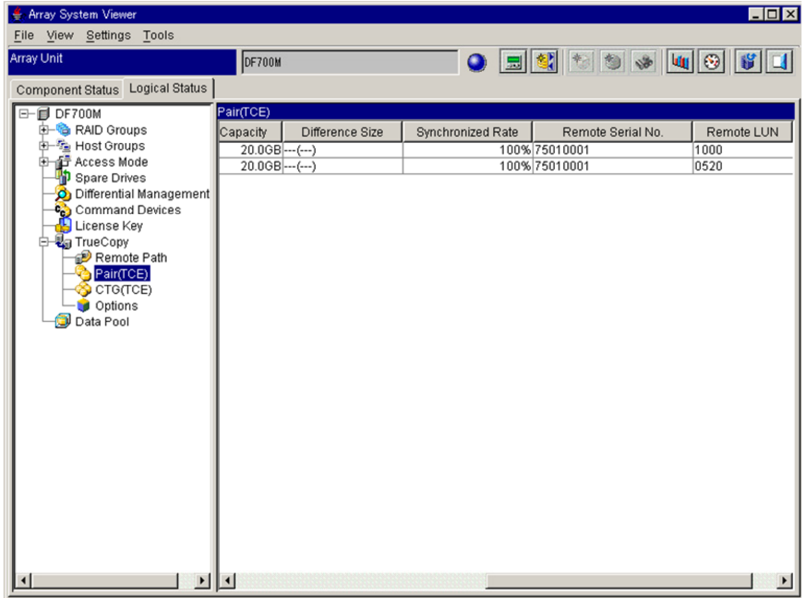


Figure 7.3 Array System Viewer Panel (TCE Pairs contd.)

Table 7.2 lists the TCE pair information displayed on the Array System Viewer panel.

Table 7.2 TCE Pair Information

Displayed Item	Contents
TCE pair icon	Icons denoting pair status.
LUN	Local subsystem logical unit number
Pair Attribute	Type of volume, P-VOL or S-VOL
Pair Status	Internal status of individual P-VOL or S-VOL of the TCE pair. The internal status is different from the TCE pair status controlled by CCI. Internal pair status is the status of each data volume. Internal pair status is essential to check for conditions such as whether the S-VOL is being restored during an SVOL takeover operation. This status information is used to determine which pair operations (such as pairsplit) can be performed on the data volume.
CTG No.	A number identifying the CTG to which the P-VOL or the S-VOL belongs.
Consistency Time	The time when the most-recently transferred differential data was created on the P-VOL. This time is used to determine the interval between P-VOL and S-VOL data updates and estimate the degree of coincidence between the P-VOL and the S-VOL data. The time at which the S-VOL was updated is valid only when the S-VOL status is PAIR, PFUS, or PSUE.
Capacity	The P-VOL or S-VOL capacity
Difference Size	For a P-VOL, difference size is the size of differential data resulting from host I/O operation. For an S-VOL, difference size is amount of residual data to be restored from the remote data pool. Difference size is displayed only when the P-VOL status is PAIR and the S-VOL status is SSWS(R). Difference size increases and decreases alternately since differential data is eliminated after each update cycle. The size of the data, which has not been transferred to the remote subsystem and thus remains in the local subsystem among the differential data generated through the host I/O accepted by the P-VOL, is displayed for each P-VOL. When the pair status is PAIR, the size of the differential data tends to repeat increase and decrease because the differential data generated by that the P-VOL accepts the host I/O instruction at the time of PAIR is eliminated regularly by the cyclic copy. In the current cyclic operation, the size of the differential data may not be reduced to zero even when the current cycle ends because the differential data at the time of the cycle start is transferred to the remote subsystem and the differential data generated after the cycle start is transferred to the remote subsystem in the following cycle.
Synchronized Rate	Rate at which the data volumes in a TCE pair are synchronized. Synchronization rate is used to estimate the degree of coincidence between the P-VOL and the S-VOL data. This information is displayed when the data volumes are in the following statuses: P-VOL: Valid when status is COPY, PSUS, PFUS, or PSUE. S-VOL: Valid when status is PSUS, PSUS(N), SSWS, or SSWS(R).
Remote Serial No.	Remote subsystem serial number
Remote LUN	Remote subsystem logical unit number

Pair status icons are displayed with the pair information in the Array System Viewer panel and denote status of paired data volumes for different pair statuses.

Table 7.3 shows the available volume status icons.

Table 7.3 Volume Status Icons

Volume Status				**Paired Status**	**Description**
Formatted	Unformatted	Regressed	Blockaded		
Yellow	Gray	Pink	Red		
				PAIR	Indicates the paired volume.
				COPY	Indicates a paired status, but initial copy or resynchronization operation is not complete.
				PSUS	In paired status, but updates to the S-VOL data are suspended due to a user-requested pairsplit operation.
				PSUE	In paired status, but updates to the S-VOL data are suspended forcibly or due to an error condition.

Availability of paired data volumes for read or write operations and data consistency in paired data volumes differs for each TCE pair status.

Table 7.4 summarizes the above information for different TCE pair statuses and indicates which management software tool can be used to confirm each pair status.

Table 7.4 Pair Status Information

Pair Status	Description	Consistency		Read/Write		Conditions	Checking Tool	
		Pair	CTG	P-VOL	S-VOL		CCI	SNM
SMPL	A TCE pair has not been created.	—	—	R/W	R/W		√	√
COPY	Initial copy or recovery copy processing is in progress.	x	x	R/W	R/W		√	√
PAIR	Update copy processing is in progress.	OK	OK	R/W	R/W		√	√
PSUS	The pair is suspended, and update copy processing is in progress.	OK	x	R/W	R/W or RO	Internal status is not PSUS(N)	√	√
SSWS	Takeover of the S-VOL has been executed.	OK	OK	R/W	R/W		√	√
PSUE	Update copy processing has been stopped because a failure occurred.	OK	OK	R/W	R/W		√	√
PFUS	The pool capacity of the local subsystem or remote subsystem has been exceeded, and update copy processing has stopped.	OK	OK	R/W	R/W		√	√
PSUS(N)	Indicates whether the S-VOL consistency is guaranteed while the pair status is PSUS.	x	x	R/W	R/W		√	√
SSWS(R)	The internally determined volume is being restored to the S-VOL after takeover of the S-VOL.	OK	OK	R/W	R/W		—	√

OK: Consistency R/W: Read/Write is possible
x: Inconsistent ~~R/W~~: Read/Write is impossible
RO: Read Only
SNM: Storage Navigator Modular

7.2 System Monitoring

A TCE system environment must be monitored periodically to confirm that all replication operations are functioning normally and the system operates as designed. If any problems relating to performance and/or disk capacity are detected, the system may have to be re-designed or necessary actions such as system upgrade and load balancing must be performed.

TCE System monitoring tasks include monitoring pair statuses, pair synchronization, and used pool capacity as well as collecting performance and failure information and creating remote snapshots. Monitoring operations can be performed using CCI, Performance Monitor or the Web GUI.

Table 7.5 summarizes the various monitoring tasks performed using different TCE management software:

Table 7.5 Monitoring with TCE Management Software

Monitoring Task	Navigator	CCI	Performance Monitor	Web GUI
Monitoring function	Sets the subsystem information	Operates on TCE pairs.	Measures performance	Tool for maintenance personal
	GUI, CLI	CLI, API	GUI	GUI
Monitoring pair status	√	√		
Monitoring synchronization	√	√		
Monitoring used pool capacity used	√			
Collecting performance information			√	
Collecting failure information	√			√
Remote snapshot		√		

Figure 7.4 shows the workflow for system operation and monitoring.

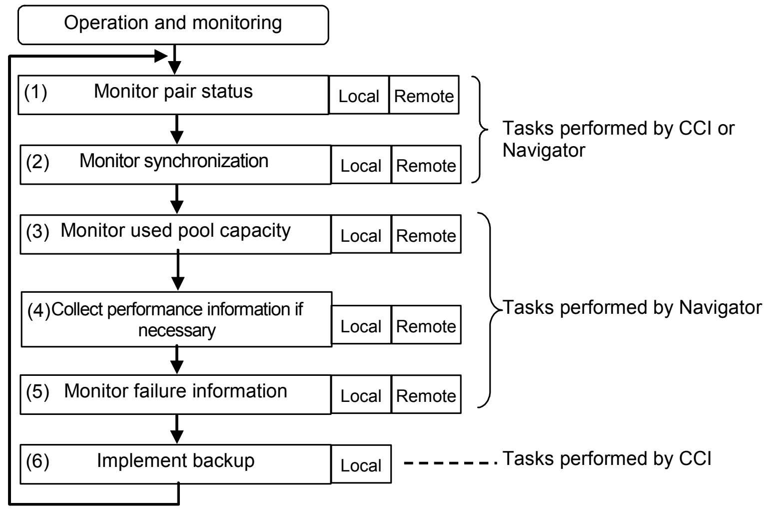


Figure 7.4 Typical Operation and Monitoring Cycle

Using relatively short monitoring cycles, allows failures and problems to be discovered as soon as they occur, resulting in a quicker recovery time and improved RTO performance.

7.2.1 Monitoring Pair Status

The CCI **pairdisplay** command can be used to check TCE pair statuses. The PAIR status denotes volume pairs are operating normally. When a pair is suspended due to any system error, its pair status changes to PSUE and the failure occurrence is recorded in the system log of the host. When the pair is in PSUE status, update processing must be stopped to analyze the cause of failure. After the failure is rectified the pairs can be resynchronized using the CCI **pairresync** command. For more information on **pairdisplay** command, see Chapter 5 Section 5.2.1 Pair Status Check (pairdisplay) operation. For more information on **pairresync** command, see Section 7.3.2 Pair Recovery (pairresync) Operation.

7.2.2 Monitoring Synchronization

Monitoring synchronization involves finding the time difference between data update processes in the P-VOL and S-VOL and confirming completion of copy operation. A larger time difference implies a decreased RPO performance. This may indicate a system failure or a performance bottleneck in the system environment. By detecting the abnormality immediately and by taking appropriate corrective measures, helps reduce risks associated with failures such as critical data losses.

The CCI **pairsyncwait** command can be used to check whether the primary and secondary time differences are within the target values. For more information on the **pairsyncwait** command, see Chapter 5 Section 5.2.3.2 Data Transfer Confirmation (pairsyncwait) operation.

Synchronization monitoring can also be performed using Storage Navigator Modular software. The time difference between the P-VOL and S-VOL updates can be determined by subtracting the S-VOL data determination time from the current time of the local AMS/WMS subsystem. Figure 7.5 illustrates the current time of a local TagmaStore AMS subsystem displayed in Navigator's **Configuration settings** window, and Figure 7.6 illustrates the determination time of the S-VOL displayed in Navigator's **Array System Viewer** panel. The time difference between the P-VOL and S-VOL can be calculated by subtracting the determination time from the current time.

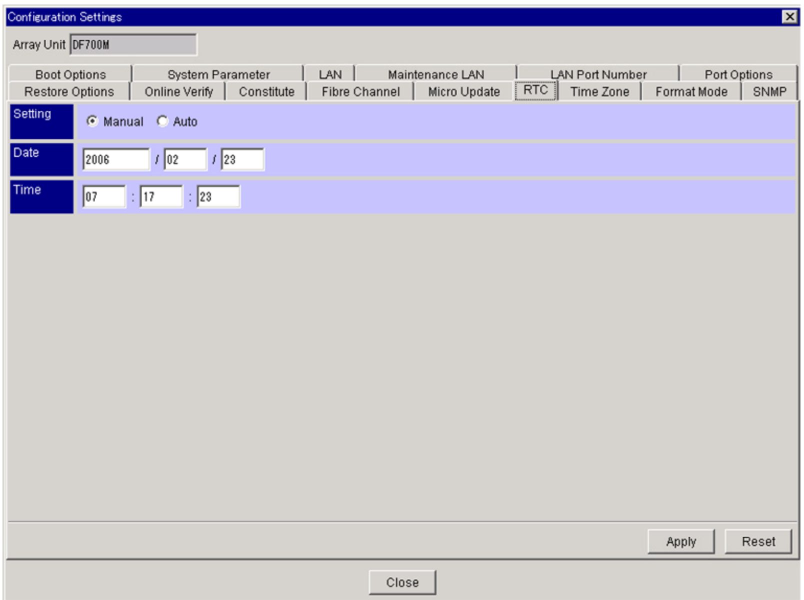


Figure 7.5 Checking Current Time of Local TagmaStore using Navigator

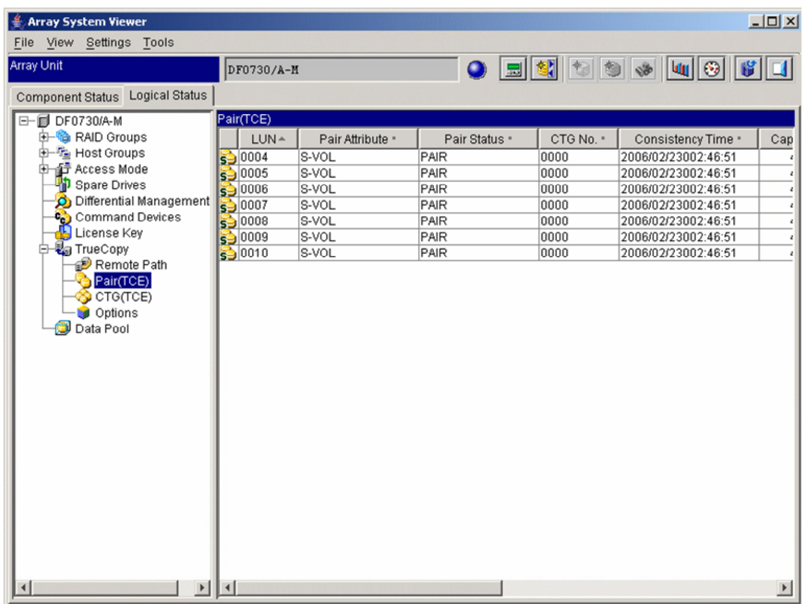


Figure 7.6 Checking Determination Time of S-VOL using Navigator

Navigator can also be used to monitor update cycle times and volume of data copied from the P-VOL to the S-VOL. This information is useful to predict cycle completion times.

Figure 7.7 illustrates cycle time and update data volume information displayed in Navigator **Array System Viewer** panel.

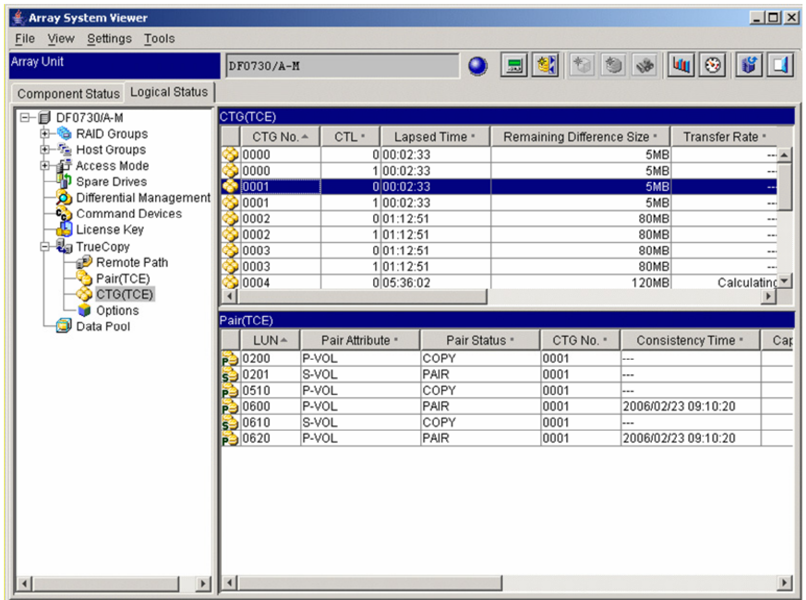


Figure 7.7 Prediction of Cycle Completion Time from Navigator

7.2.3 Monitoring Pool Capacity

Periodic monitoring of data pool capacity reduces the risk of disrupting update cycles and subsequent loss of data in the event of a disaster when update copy is suspended. A TCE pair is suspended and update copy processing is stopped, when unused pool capacity decreases below a specific threshold. By monitoring data pool usage, logical units can be added to the pool to prevent data pool overflow conditions.

Figure 7.8 shows how data pool capacity can be monitored using pool information displayed in the Navigator Array System Viewer panel.

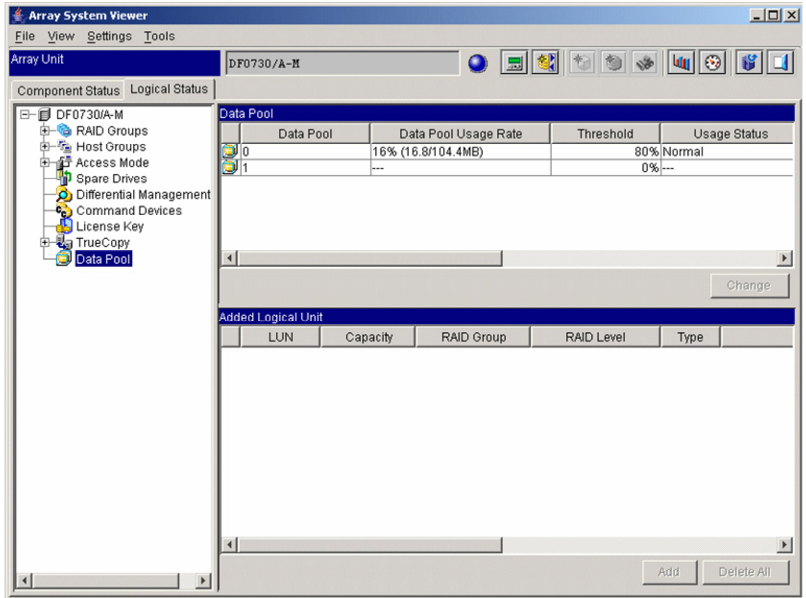


Figure 7.8 Checking Free Pool Space with Navigator

Data pool monitoring can be automated using Navigator CLI, by periodically executing a script to monitor unused pool capacity. Figure 7.9 shows a sample pool capacity monitoring script.

```
#!/usr/bin/perl

open(IN, " aupool -unit [unit name of AMS] -refer -poolno 0 |");
   /* Obtain pool usage by Navigator (CLI) aupool command */
   /* In this example, check local DF700 pool 0 */

$Threshold = 80;     /* Threshold of pool */

while(<STDIN>){
   if (/.+: (.+)%.+/) {
      if ($1 > $Threshold) {
         print "Pool is close to over!\n";
         /* You can write code here to take action for pool closing to over. */
      }
   }
}
```

Figure 7.9 Sample Pool Capacity Monitoring Script

7.2.4 Monitoring Performance

Performance Monitor is used to periodically check if the characteristics of the inflow from the host that was estimated when the system was designed match the actual operating environment. For stable operation of a TCE system, the amount of incoming data to the P-VOL of the TCE pair and the amount of outgoing data to the S-VOL must be balanced. Imbalances in data flow can result in TCE pair suspension and data pool overflow. Performance monitoring allows these problems to be detected so that drop in RPO values can be prevented.

Table 7.6 shows information that can be monitored using Performance Monitor.

Table 7.6 Performance Parameters monitored using Performance Monitor

Performance parameters	Description	Units	Level at Which Acquired
Inflow from host	The number of I/Os issued from the host to a local subsystem and the amount of data transferred from the host to a local subsystem can be measured. Read operations can be distinguished from Write operations.	IOPS, MB/s	LU, port
Outflow to remote	The number of I/Os issued to the remote subsystem and the amount of data transferred from the local subsystem to the remote subsystem can be measured.	IOPS, MB/s	LU, port
Processor utilization rate	The processor utilization rate of a TagmaStore subsystem can be measured.	%	Controller
Cache utilization rate	The amount of pending data to be written to the cache can be measured.	%	Controller
Drive utilization rate	The drive utilization rate of each drive can be measured.	%	
Response time	Response time of a remote subsystem to communication from a local subsystem.	ms	LU, port Maximum, average

7.2.5 Monitoring Failures

Navigator can be used to monitor a TCE system to detect failures and abnormal operations. Messages and alerts indicating failures are displayed and automatically communicated to Hitachi service centers using SNMP Agent, HITRACK, and Web message notifications. For more information about disaster recovery, see Chapter 1.

7.2.6 Backup with Remote Snapshots

Remote backup involves creating point-in-time frozen replicas (snapshots) of data in both primary and secondary storage subsystems. The static images of the primary volume are used to determine data to be updated to the secondary storage subsystem during the remote backup operation. The multiple snapshots created prior to the onset of a disaster, are useful for data recovery, particularly from rolling disasters.

The CCI **pairsplit** command supports new options to create a remote snapshot for backup. For more information on pairsplit operations, see Section 7.3.1 Pair Splitting (pairsplit) Operation.

7.3 System Maintenance

TCE System maintenance tasks involve pair splitting, recovery of pairs, and modifying pool size. Maintenance operations can be performed using CCI, Navigator or the Web GUI.

Table 7.7 summarizes the various maintenance tasks performed using different TCE management software:

Table 7.7 Maintenance using TCE Management Software

Maintenance Task	Navigator	CCI	Performance Monitor	Web GUI
Management software function	Sets the subsystem information	Operates on TCE pairs.	Measures performance	Tool for maintenance personal
Pair splitting		CLI /GUI		
Pair recovery		√		
Changing of pool size	CLI/GUI			
Recovery after path blockage				√

TCE system maintenance involves the following tasks at the local and remote sites. At the local site,

1. Host I/O operations affecting the P-VOL at the local site are first stopped.
2. The volume pair is then split and the P-VOL and S-VOL data is synchronized.
3. After synchronization the TCE pair is deleted.

The following tasks are then done at the remote site:

4. The remote CCI instance is used to create a pair in the reverse direction. An initial copy is not required at this time since the data in the P-VOL and S-VOL have already been synchronized.
5. Processing is paused until the TCE pair status transitions to PAIR. Since no initial copy is required, the status transitions to PAIR within one cycle.
6. The TCE pair is then split. While the pair is being split, the remote and local storage subsystems keep track of data updates to the P-VOL and S-VOL. This allows data updated in the remote subsystem during maintenance to be reflected on the local storage when pairs are resynchronized later.
7. S-VOL operation is resumed and writing to the old S-VOL is possible.

Maintenance tasks are now executed at the local site on the primary storage subsystem. After maintenance tasks at the local site are completed, the following tasks are executed at the remote site:

8. I/O operations to the remote storage subsystem are stopped at the remote site.

9. Data updated on the remote subsystem during the maintenance is reflected on the old P-VOL by re-synchronizing the pair.
10. Processing is paused until the pair status changes to PAIR.
11. The pair is split and the P-VOL and S-VOL data are synchronized.
12. The reverse pair is deleted.

The following tasks are then performed at the local site to complete system maintenance:

13. The TCE pair is created again directed from the local storage subsystem to the remote storage subsystem. Again initial copy is not required since the data of the P-VOL and S-VOL have been previously synchronized.
14. Remote replication operations are resumed at the local site.

The CCI **pairsplit** and **pairesync** commands are used to perform pair splitting and pair resynchronization maintenance operations. Some maintenance tasks, such as installing the micro-program of the remote TagmaStore AMS subsystem, are performed after update copy processing of TCE pairs has been stopped by pair splitting. Table 7.8 lists TCE system maintenance pair operations. The following sections discuss these operations in further detail.

Table 7.8 CCI Pair Commands for Maintenance

Maintenance Operations	CCI Pair Command	Description	Comments
Pair splitting	pairsplit	Temporarily stops update copy processing for the specified TCE pair.	A pair is split after the P-VOL and S-VOL are synchronized.
Pair recovery	pairresync	Resumes remote copy operation after the completion of maintenance. Also resumes the remote copy operation for a TCE pair that was suspended due to a failure.	

7.3.1 Pair Splitting (pairsplit) Operation

The pair splitting operation suspends all pending S-VOL updates (issued prior to the split command and recorded in the P-VOL track map) and releases the paired data volumes. The CCI **pairsplit** command can be used to split TCE pairs. It can be used together with the **paircreate** operation to split newly created pairs.

The **pairsplit** command can be executed using the following options:

- The **pairsplit -S** option stops the copy operations to the S-VOL of the pair and releases the volumes in paired status. The S-VOL is made identical to the state of the P-VOL if the TCE pair was previously in PAIR status. Both volumes transition to the SMPL status after this operation. For more information on **pairsplit -S** command, see Section 7.4.1 Pair Release (pairsplit –S) Operation.
- The **pairsplit -R** option is issued to the S-VOL of the TCE pair from host connected to the remote subsystem. This option forcibly interrupts copy operations irrespective of current status of the TCE pair. After this operation, only the S-VOL transitions to SMPL status. For more information on **pairsplit -R** command, see Section 7.4.2 Pair Suspend (pairsplit –R) Operation.
- The **pairsplit -mscas** command is used for splitting a remote SnapShot pair which is in a cascade configuration with a TCE pair. This command to split the TCE-SnapShot pair (TCE volume & SnapShot P-VOL pair), is issued by the primary host to the CTG containing the TCE volume pair (TCE P-VOL & S-VOL pair). It cannot be directly issued to the TCE-SnapShot pair by the host on the secondary side. When this command is executed, SnapShot V-VOL data is synchronized with the P-VOL data of the TCE pair in PAIR status. For more information on **pairsplit -R** command, see Chapter 5 Section 5.3.5 Remote SnapShot Pair Splitting (pairsplit –mscas) Operation.
- When differential data is transferred to the S-VOL to make it consistent with the P-VOL, time taken to complete the splitting operation will vary depending on the amount of the differential data. Table 7.9 summarizes response times to various options of the pairsplit command, for different TCE pair statuses.

Table 7.9 Response Time of CCI Commands

Command	Options	Status	Response	Next Status	Remarks
pairsplit	-S Delete pair	PAIR	Depend on Differential Data	SMPL	S-VOL data consistency ensured
		COPY	Immediate	SMPL	No S-VOL data consistency
		Others	Immediate	SMPL	No S-VOL data consistency
	-R Delete pair	PAIR	Immediate	SMPL (S-VOL only)	No S-VOL data consistency
		COPY	Immediate	SMPL (S-VOL only)	No S-VOL data consistency Can not be executed for SSWS(R) status
		Others	Immediate	SMPL (S-VOL only)	No S-VOL data consistency Can not be executed for SSWS(R) status
	-mscas Create remote snapshot (See note)	PAIR	Immediate	No Change	A completion time depends on the amount of differential data. A completion can be check by Split-Marker and a creation time. Cycle updating process stops during creating a remote snapshot.
		Others	—	—	—
	Others Split pair	PAIR	Depend on Differential Data	PSUS	S-VOL data consistency ensured
		COPY	Immediate	PSUS	S-VOL data consistency ensured
		Others	Immediate	No Change	S-VOL data consistency ensured

Note: Only -g option is valid. The -d option is not accepted. All S-VOLs in PAIR status must have corresponding cascading V-VOLs and the MU# of these snapshot pairs must match the MU# specified in a **pairsplit -mscas** command option.

When the split operation is complete, the pair transitions to PSUS status, and full read/write access to the split S-VOL is available. The P-VOL also remains fully accessible during the pairsplit operation. While the pair is split, the storage subsystems establish a track map for the split P-VOL and S-VOL and record all updates to both volumes. For more details on the **pairsplit** operation, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

7.3.2 Pair Recovery (pairresync) Operation

The CCI **pairresync** command is used to resynchronize split pairs (in PSUS or PFUS status) or suspended pairs (in PSUE status). During resynchronization, P-VOL data is copied to the S-VOL and the pair transitions to COPY status. When the pair is in COPY status, host I/O operations are disabled for the S-VOL but can be continued on the P-VOL. The pair transitions to PAIR status when the **pairresync** operation is completed.

When the **pairresync** operation is performed on a split pair (in PSUS status), the pair status changes to COPY and the storage subsystem merges the S-VOL track map into the P-VOL track map. Differential data from the P-VOL is then copied to the S-VOL. Updates received from the host when the pair is in the COPY status become new differential data. The pair status transitions from COPY to PAIR when the difference between the P-VOL and S-VOL data reaches a certain threshold. When the TCE pair is in PAIR status, differential data is updated using cyclic copy operations. For more information on the **pairresync** operation, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

Note: The swap function to send the differential data from the S-VOL to the P-VOL by reversing the P-VOL and the S-VOL pair is not currently supported by the resynchronization operation.

7.4 System Reduction

Changes in performance and capacity over time may warrant TCE system expansion or reduction. This section summarizes reduction tasks performed using Navigator and associated CCI pair operations. The following steps outline the workflow for system reduction:

1. TCE pair deletion is performed by issuing CCI pair deletion commands from the host at the local site. Volume pairs can be released or suspended.

2. If the remote path is blocked or the S-VOL is in SSWS(R) pair status, the S-VOL cannot be deleted by deleting the TCE pair. The pair status of the S-VOL of the TCE pair can be checked by issuing a pair status display command of CCI from the host at the remote site. If an S-VOL is in SSWS(R), it can be deleted by issuing the **pairsplit -R** command from the host at the remote site.

3. The remote path is deleted on the local site and the remote site using Navigator.

4. Next, if SnapShot is not used, the data pools are deleted at the local and remote site using Navigator. If SnapShot is still in use the data pools cannot be deleted and system reduction ends after TCE pairs and remote path are deleted.

5. Next, if ShadowImage is not used, the Command device and DM-LU are deleted at the local and remote site using Navigator. If ShadowImage is still in use the command device and DM-LU cannot be deleted and system reduction ends after data pools are deleted.

6. Finally, logical units allocated for TCE data volumes is deleted at the local and remote site using Navigator. If the logical units are in use by SnapShot or ShadowImage, they cannot be deleted.

7.4.1 Pair Release (pairsplit –S) Operation

The **pairsplit -S** operation is the normal pair deletion operation which stops the copy operations to the S-VOL of the pair and releases the two volumes in paired status. When initial copying is started (i.e., TCE pair is in any status except SMPL), the instruction to split a pair is received and processed irrespective of the pair status. As a result of the pairsplit –S operation, the status of both TCE volumes changes to SMPL.

Before a TCE pair is split, the P-VOL and S-VOL are synchronized by copying differential data determined at the point of time when the pairsplit command was issued. Therefore the time taken to split a volume pair varies depending on the amount of the differential data. Table 7.10 lists pair status transitions resulting from the **pairsplit -S** command.

Table 7.10 Pair Status After Pairsplit -S Command

Pair status before accepting a command	Result of execution of the pairsplit -S command	
	Status as a result	Command operation
PAIR	SMPL	The status is changed after the differential data of the P-VOL is copied to the S-VOL, and a response is then returned to the host.
Other PAIR	SMPL	The P-VOL data and the S-VOL are not synchronized. The response is returned immediately since there is no differential data copying delay.

If the execution of the **pairsplit -S** command fails when the TCE pair is in PAIR status, only the P-VOL status is changed to SMPL, and the operation to transfer the differential data of the P-VOL to the S-VOL is suspended. A response is returned after the P-VOL status is changed to SMPL.

If a failure occurs when the transition to SMPL status is taking place, the data copy is suspended immediately and the P-VOL status is changed to PSUE. A **pairresync** command cannot be executed while the pair status is being changed to PSUE. The pair status transitions from PSUE to SMPL, after cyclic copy is completed by all associated CTGs. For more information on the **pairsplit -S** command, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

7.4.2 Pair Suspend (pairsplit –R) Operation

A TCE volume pair can be suspended using the **pairsplit -R** command. The pairsplit –R command is a remote side pairsplit command, issued to the S-VOL of a TCE pair from the host connected to the remote subsystem. This command interrupts the copy operation forcibly and places the S-VOL in the SMPL status. The pairsplit –R operation is performed irrespective of current TCE pair status.

S-VOL data consistency is not guaranteed when a pair is forcibly suspended. In order to restore a suspended S-VOL, it is required to first execute the **horctakeover** operation. A TCE pair is in the SSWS(R) status during the **horctakeover** operation, and transitions to SSWS pair status upon completion of takeover. The **pairsplit -R** command must be executed only after checking that the restoration has been completed and the pair status has changed to SSWS. The time for completion of restoration and change of status to SSWS can be estimated from the amount of differential data between the paired volumes.

TCE pairs are also suspended when the storage subsystem detects an error condition related to a copy operation. When a volume pair in COPY status is suspended, the storage subsystem aborts the copy operation, changes the status of the P-VOL and S-VOL to PSUE, and keeps track of all subsequent read and write I/Os to the P-VOL. The details of the error condition are output to the system log or event log file. For more information on the **pairsplit -R** operation, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

Chapter 8 Best Practices and Recommendations

This chapter outlines best practices and recommendations for construct a TCE system. The following sections are included in this chapter:

- Pair Configuration (section 8.1)
- Remote Processing (section 8.2)
- Command Devices and Differential Management LU (DM-LU) (section 8.3)
- Concurrent Use of LUN Expansion (section 8.4)
- Maximum Supported Capacity of TCE (section 8.5)
- Cascade Restrictions with SnapShot P-VOL (section 8.6)
- TCE Path Timeout settings (section 8.7)
- Cautions and Restrictions (section 8.8)

8.1 Pair Configuration

- Use a small number of volumes within the same RAID group.

  When volumes are assigned to the same RAID group and used as pair volumes, there may be a case where a pair creation or resynchronization for one of the volumes causes restrictions to be placed on performance of a host I/O, pair creation, resynchronization, etc. for the other volume(s) because of contention between drives. It is recommended to assign small number of (one or two) volumes to be paired to the same RAID group. When creating two or more pairs within the same RAID group, standardize the controllers that control LUs in the same RAID group and pay attention to make the pair creation or resynchronization timely.

- Use FC drives for the primary volume.

  When a P-VOL is located in a RAID group consisting of the SATA drives, performance of a host I/O, pair creation, and pair resynchronization, etc. is lowered because of the lower performance of the SATA drive. Therefore, it is recommended to assign a primary volume to a RAID group consists of the FC drives. Regular PAIR status with a configuration of the SATA drives is not recommended, in order to avoid system operations with working hours exceeding 330 hours a month.

- Use FC drives for the secondary volume.

  When the SATA drive is used for the S-VOL or the primary or secondary data pool, there is a higher possibility that a suspension failure could occur in the TCE pair. Data transfer is also slower compared to when an FC drive is used. Therefore, it is recommended to use the FC drive. Regular PAIR status with a configuration of the SATA drives is not recommended, in order to avoid system operations with working hours exceeding 330 hours a month.

- Perform pair creation and resynchronization when I/O load is minimal.

  When a pair is newly created or resynchronization of a pair is done when a TCE pair is in the PAIR status, the volume of differential data may become larger because the transfer of the differential data between the TCE pair is delayed. When the delay continues, the data in the data pool could overflow and the TCE pair is split and placed in the PFUS status. It is recommended to perform the system operation such as pair creation or pair resynchronization when I/O load is minimal (in the night or on a holiday). It is also recommended to limit the number of the pairs in the COPY status, to one or two at a time.

- Assign a data pool volume to a distinct RAID group.

  When another volume is assigned to the same RAID group to which a data pool volume belongs, the load on drives increases and their performance is reduced .Therefore, it is recommended to assign a data pool volume to an exclusive RAID group. When there are multiple data pool volumes in a subsystem, different RAID groups should be used for each data pool volume.

- Assign four or more disks to the data disks.

  When the data disks that compose a RAID group are not sufficient, it affects the host performance and/or copying performance adversely by restricting read/write operations to the drives. Therefore, when operating pairs with TCE, it is recommended to use an LU consisting of four or more data disks.

- For bi-directional TCE pairs, perform a pair creation/resynchronization in each of the directions one after the other.

  In a configuration where each site serves as both local and remote site (as shown in Figure 8.1), creation and resynchronization of a pair may be performed from the both sites at the same time. In such a case, time interval between pair creation until completion of the resynchronization becomes longer and the influence on performance of the other operation becomes greater. This is due to reading and writing of data in parallel at the both sites. Therefore, it is recommended to perform pair creation and resynchronization in each of the two directions one after the other.

  It is required to construct a configuration considering that there may be a case where the cycles in the both directions overlap each other because the differential data of the TCE pair is regularly copied in the cycle even when the pair is in the PAIR status.

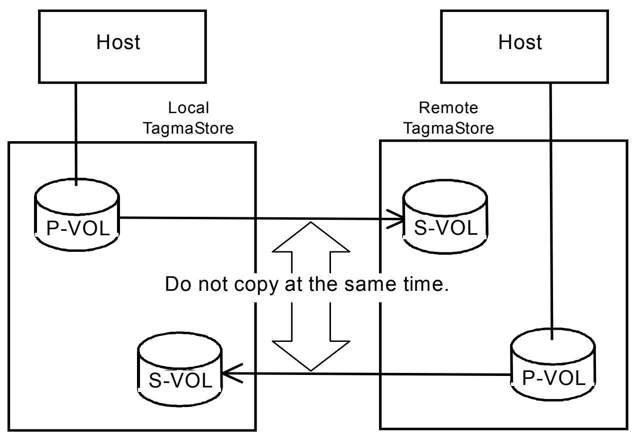


Figure 8.1 Bidirectional TCE Operation

8.2 Remote Processing

- I/O load balance

  I/O load balance may be lost when the data pool has no free capacity to hold differential data resulting from I/O operations at the local site. It is necessary to monitor the load on the remote site to avoid loss of differential data.

- Avoiding path failure

1. If a subsystem is turned on after a path has already been set, it is necessary to turn on the remote subsystem first. The local subsystem should be turned on after the remote subsystem is READY. If a subsystem is turned off after a path has already been set, it is necessary turn off the local subsystem first, followed by the remote subsystem.
2. When you restart the subsystem, check if the subsystem is on the remote side of TCE. If the subsystem on the remote side is restarted, both paths are blocked.
3. If the remote subsystem is powered off or restarted, when the TCE pair is in PAIR or COPY status, pair status is changed to PSUE. It is recommended to perform necessary power off or restart of the subsystem, after changing TCE pair status to PSUS.
4. If the remote subsystem is unavailable, you will encounter an error/blockage. When a path is blocked, a notification (TRAP) is sent to the SNMP Agent support function. It is necessary to notify the departments concerned of the error beforehand. The TrueCopy path is automatically recovered from the blockade when the subsystem is restarted. If the path is not recovered when the subsystem is READY, contact Hitachi Support personnel. The subsystem status changes to READY after power on in about five minutes or less, even when the subsystem has maximum configuration. The time required varies depending on the subsystem configuration.
5. The micro program must not be changed when the TCE pair is in COPY or PAIR status. Micro program can be changed after splitting the pair.

   If the local subsystem is directly connected with the remote subsystem the fibre transfer rate setting must not be modified while the subsystem power is on. If the setting of the fibre transfer rate is modified, it will result in a path blockage.

8.3 Command Devices and Differential Management LU (DM-LU)

- Assign multiple command devices to distinct RAID groups

  If two command devices are set within one disk subsystem, assign each of them to distinct RAID groups. If command devices are assigned to the same RAID group, both the command devices may become unavailable when disasters such as a drive failure occur.

- Assign multiple DM-LUs to distinct RAID groups

  If two DM-LUs are set within one disk subsystem, assign each of them to distinct RAID groups. If they are assigned to the same RAID group, both DM-LUs may become unavailable owing to disasters such as a drive failure.

8.4 Concurrent Use of LUN Expansion

In TCE, a unified LU can be used as a P-VOL or an S-VOL. The unified LU cannot be assigned to a data pool. The following preconditions and restrictions exist for concurrent use of TCE and LUN Expansion.

- Capacities of paired P-VOL and S-VOL must be the same
  - The P-VOL and S-VOL can be paired as long as their capacities are the same even if numbers of LUs composing them (unified LUs) are different from each other.

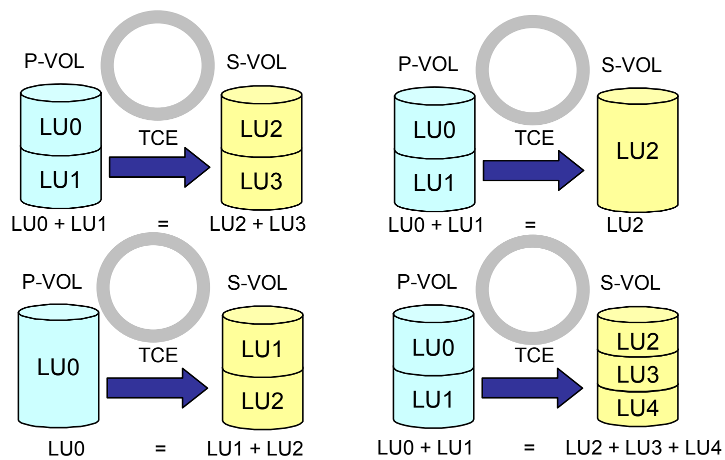


Figure 8.2 Assigning Unified LU to the P-VOL or S-VOL

- A unified LU composed of 17 or more LUs cannot be assigned to the P-VOL or S-VOL.

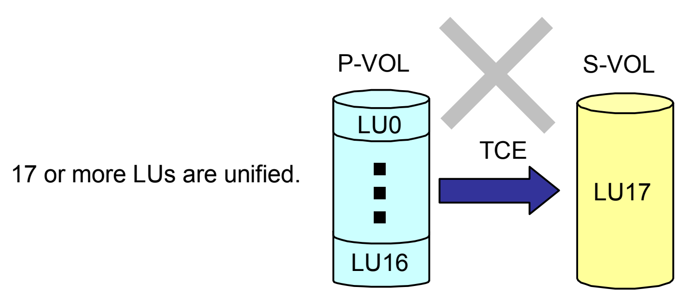


Figure 8.3 Restriction on Number of LUs in a Unified LU

- RAID levels of P-VOL and S-VOL

  All LUs, including the unified LU(s), assigned to the P-VOL and S-VOL do not have to be on the same RAID level. They also do not need to have the same number of data disks.

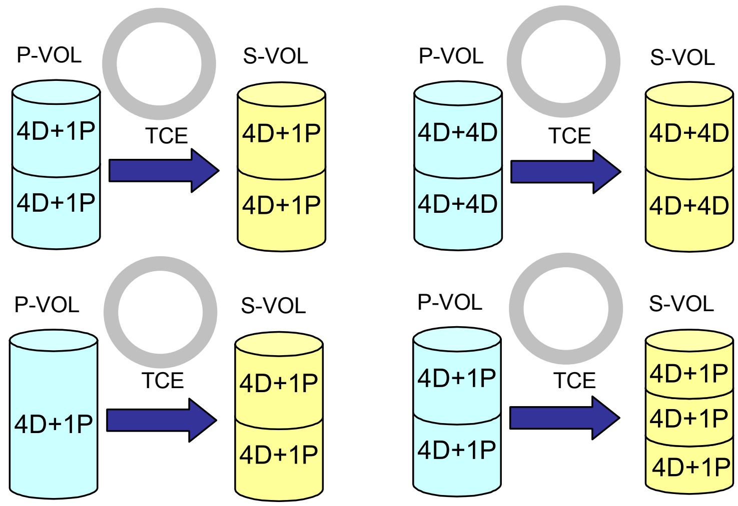


Figure 8.4 Allowed Combination of RAID Levels (TCE)

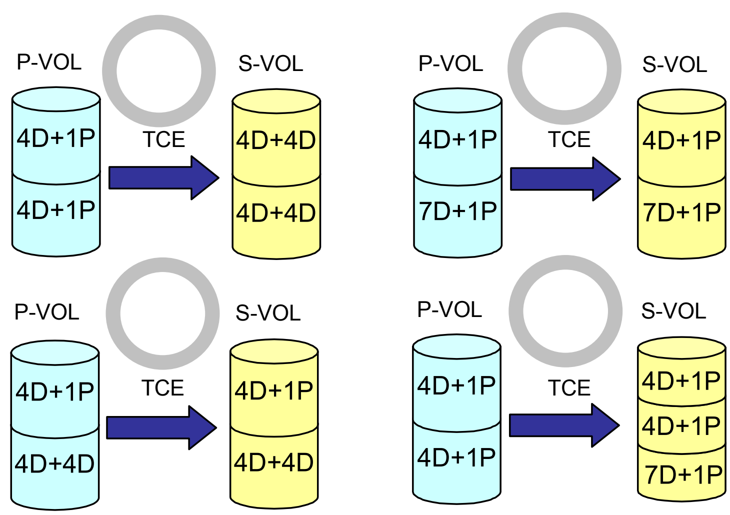


Figure 8.5 Disallowed Combination of RAID Levels (TCE)

- Other Restrictions
  - A P-VOL or an S-VOL belonging to a TCE pair cannot be used to compose a unified LU or separating unified LU. Unify or separating unified LU after splitting the pair once.
  - The total capacity of all the LUs to be unified must be 1 GB or more.
  - An LU consisting of the FC drive and that consisting of the SATA drive cannot be unified. Unify LUs consisting of drives of the same type.
  - Although an LU assigned to a data pool cannot be unified, but the data pool capacity can be increased through a setting of up to 64 usual LUs. The LU(s) can be added to the data pool online. However, an LU consisting of the FC drive and that consisting of the SATA drive cannot coexist in a data pool.

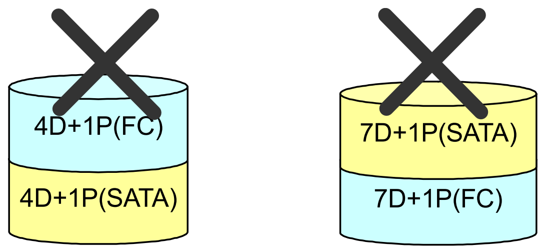


Figure 8.6 Restriction on Unification of LUs with Different Drive Types (TCE)

8.5 Maximum Supported Capacity of TCE

The P-VOL and S-VOL of the TCE function and P-VOL of the SnapShot function share common pool resources. Therefore, the pool capacity for both the functions is limited. The maximum capacity supported by a TCE pair varies depending on the P-VOL capacity of SnapShot, data pool capacity, and cache memory capacity. When using other copy system functions and TCE together, the maximum supported capacity of the P-VOL may be restricted further. Therefore, the supported capacity of TCE needs to meet the following two conditions.

- Must be less than or equal to the maximum supported capacity calculated by the capacity ratio with pool. (Refer to Table 8.2 and Table 8.3)
- Must be less than or equal to the maximum supported capacity at the time of the combined use with other copy system functions.

The maximum capacity supported by TCE is calculated using the following formula. The single maximum capacity of TCE is shown in Table 8.1.

TCE: Maximum supported capacity value of P-VOL and S-VOL (TB)

= Maximum TCE single capacity
- (2 × Total S-VOL capacity of ShadowImage) ÷ 51
- (Total SnapShot P-VOL capacity) ÷ 3

Table 8.1 Single Maximum Supported Capacity of TCE (TB)

Equipment Type	Mounted Memory Capacity	Single Maximum Supported Capacity (TB)
AMS500	1 GB/CTL	–
	2 GB/CTL	9
	3 GB/CTL	11
	4 GB/CTL	16
AMS1000	2 GB/CTL	15
	4 GB/CTL	23
	6 GB/CTL	28
	8 GB/CTL	30

Table 8.2 and Table 8.3 shows the relational expressions of the maximum supported capacities of P-VOL of SnapShot and P-VOL (S-VOL) of TCE, for specific pool capacities.

Table 8.2 Formula for Maximum P-VOL/POOL Supported Capacity (AMS500)

Capacity of Cache Memory Installed	Differential Information Used Capacity (Shared by TCE and SnapShot)
1 GB/CTL	Not supported.
2 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷4 + Total POOL capacity < 1.8 TB
3 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷ 4 + Total POOL capacity < 3.7 TB
4 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷ 4 + Total POOL capacity < 7.0 TB

Table 8.3 Formula for Maximum P-VOL/POOL Supported Capacity (AMS1000)

Capacity of Cache Memory Installed	Differential Information Used Capacity (Shared by TCE and SnapShot)
2 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷4 + Total POOL capacity < 1.8 TB
4 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷4 + Total POOL capacity < 5.5 TB
6 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷ 4 + Total POOL capacity < 11.0 TB
8 GB/CTL	P-VOL (S-VOL) of TCE and P-VOL of SnapShot capacity ÷ 4 + Total POOL capacity < 15.0 TB

Table 8.4 through Table 8.10 show the maximum supported capacity of P-VOL and S-VOL of the TCE function and P-VOL of the SnapShot function, calculated for different pool capacities using the expressions in the above table.

Table 8.4 P-VOL/POOL Supported Capacity (Cache Memory 2 GB/CTL: AMS500)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	5.1	0.51
1:0.3	3.2	0.98
1:0.5	2.4	1.2

Table 8.5 P-VOL/POOL Supported Capacity (Cache Memory 3 GB/CTL: AMS500)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	10.0	1.0
1:0.3	6.7	2.0
1:0.5	4.9	2.4

Table 8.6 P-VOL/POOL Supported Capacity (Cache Memory 4 GB/CTL: AMS500)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	20.0	2.0
1:0.3	12.7	3.8
1:0.5	9.3	4.6

Table 8.7 P-VOL/POOL Supported Capacity (Cache Memory 2 GB/CTL: AMS1000)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	5.1	0.51
1:0.3	3.2	0.98
1:0.5	2.4	1.2

Table 8.8 P-VOL/POOL Supported Capacity (Cache Memory 4 GB/CTL: AMS1000)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL(TB)
1:0.1	15.7	1.57
1:0.3	10.0	3.0
1:0.5	7.3	3.6

Table 8.9 P-VOL/POOL Supported Capacity (Cache Memory 6 GB/CTL: AMS1000)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	31.4	3.1
1:0.3	20.0	6.0
1:0.5	14.6	7.3

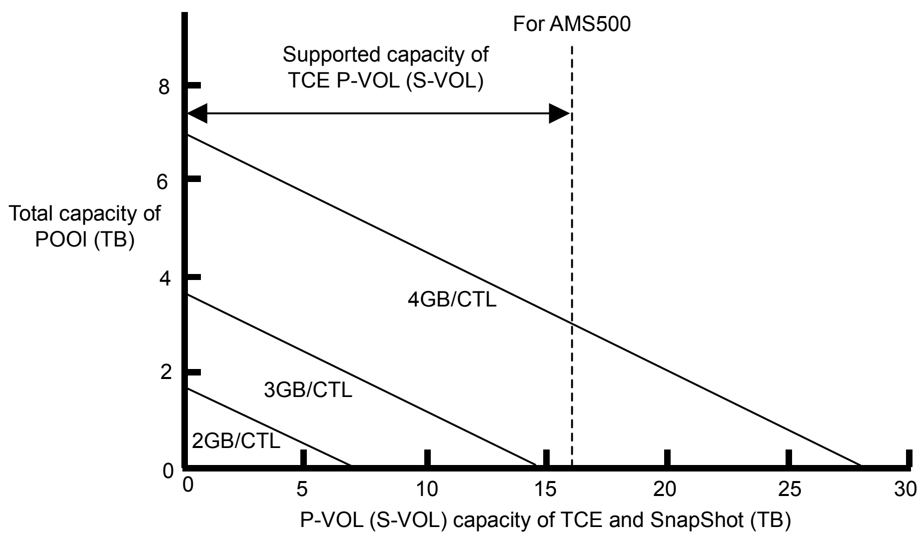
Table 8.10 P-VOL/POOL Supported Capacity (Cache Memory 8 GB/CTL: AMS1000)

Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot: Total POOL Capacity	Capacity of P-VOL (S-VOL) of TCE and P-VOL of SnapShot (TB)	Total Supported Capacity of POOL (TB)
1:0.1	42.8	4.2
1:0.3	27.2	8.1
1:0.5	20.0	10.0

The capacity of each P-VOL (S-VOL of TCE) is managed in units of 15.75 GB. When the P-VOL (S-VOL of TCE) capacity is 17 GB, the P-VOL is regarded as using a capacity of 31.5 GB. When there are two P-VOLs, each has a capacity of 17 GB. They use a total capacity of 63 GB (31.5 GB×2), though the actual capacity is 34 GB (17 GB×2).

The capacity of each LU, which has been registered with the data pool, is managed in units of 4 GB. When the LU capacity is 5 GB, the LU is regarded as using a capacity of 8 GB.
When there are two LUs registered with the data pool, each has a capacity of 5 GB. They use a total capacity of 16 GB (8 GB×2), though the actual capacity is 10 GB (5 GB×2).

The following graph shows supported volume of each cache memory, for P-VOL (S-VOL of TCE)/POOL.



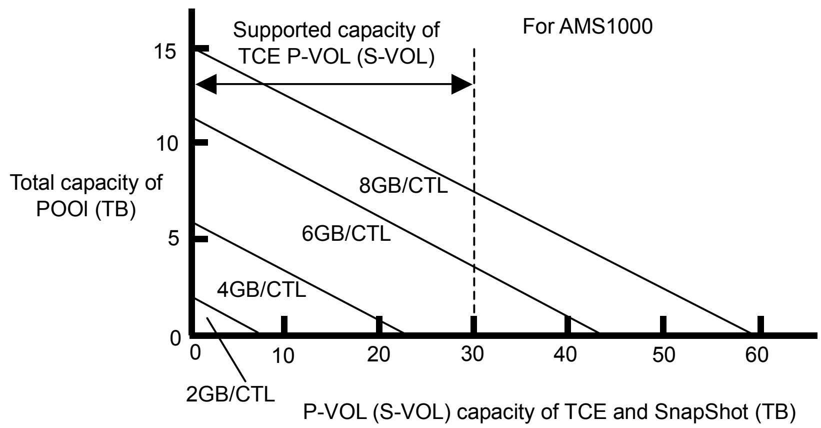


Figure 8.7 Supported Cache Memory Volume

8.6 Cascade Restrictions with SnapShot P-VOL

- LU Shared with P-VOL on SnapShot and P-VOL on TCE

  Table 8.11 shows when read/write operations are permitted from/to a SnapShot P-VOL on the local side, if a SnapShot P-VOL and a TCE P-VOL share the same LU.

Table 8.11 Read/Write Operations to a local SnapShot P-VOL

TCE P-VOL	SnapShot P-VOL				
	PAIR	RCPY	PSUS	PSUE	PSUE (Restore)
PAIR	○ R/W	x	○ R/W	○ R/W	x
COPY	○ R/W	x	○ R/W	○ R/W	x
PSUS	○ R/W	○ R/W	○ R/W	○ R/W	Δ ~~R/W~~
PFUS	○ R/W	○ R/W	○ R/W	○ R/W	Δ ~~R/W~~
PSUE	○ R/W	Δ R/W	○ R/W	Δ R/W	Δ ~~R/W~~

○: Possible, ×: Impossible
Δ: Pair operation causes an error
R/W: Read/Write by a host is possible.
R: Read by a host is possible but write is impossible.
W: Write by a host is possible but read is impossible.
~~R/W~~: Read/Write by a host is impossible.

- Single LU used for P-VOL on SnapShot and S-VOL on TCE

Table 8.12 shows whether read/write operations are permitted from/to a SnapShot P-VOL on the remote side, if a SnapShot P-VOL and a TCE S-VOL share the same LU.

Table 8.12 Read/Write Operations to a remote SnapShot P-VOL

TCE S-VOL		SnapShot P-VOL				
		PAIR	RCPY	PSUS	PSUE	PSUE (Restore)
PAIR		○ R	x	○ R	○ R	x
COPY		○ R	x	○ R	○ R	x
PSUS	R/W	○ R/W	○ R/W	○ R/W	○ R/W	Δ ~~R/W~~
	R	○ R	x	○ R	○ R	x
PSUS(N)	~~R/W~~	Δ ~~R/W~~	x	Δ ~~R/W~~	Δ ~~R/W~~	x
SSWS		○ R/W	○ R/W	○ R/W	○ R/W	Δ ~~R/W~~
SSWS(R)		○ R/W	x	Δ R/W	○ R/W	x
PFUS		○ R	x	○ R	Δ R	x
PSUE		○ R	x	○ R	Δ R	x

○: Possible, ×: Impossible
Δ: Pair operation causes an error
R/W: Read/Write by a host is possible.
R: Read by a host is possible but write is impossible.
W: Write by a host is possible but read is impossible.
~~R/W~~: Read/Write by a host is impossible.

- Number of Snapshot V-VOL generations

  Up to 15 (14: the micro program less than 0750/A version) generations of V-VOLs can be made, irrespective of whether the SnapShot P-VOL is cascaded with the TCE P-VOL and TCE S-VOL or not. Table 8.13 through Table 8.15 show TCE command and SnapShot command acceptance conditions for different TCE-SnapShot configurations.

Table 8.13 TCE Command Acceptance (TCE S-VOL-SnapShot P-VOL Cascade)

TCE S-VOL Status		Command Acceptability issued to a TCE pair						
		paircreate	pairresync	pairsplit	pairsplit -mscas *1	pairsplit -S	pairsplit -R	horctakeover
SMPL		○	x	x	x	○	○	x
COPY		○	x	○	x	○	○	x
PAIR		○	x	○	○ *2	○	○	○
PSUS	R/W	x	○	x	x	○	○	○
	R	x	○	x	x	○	○	○
PSUE		x	○	x	x	○	○	○
SSWS		x	x	x	x	○	○	○
SSWS(R)		x	x	x	x	○	x	○
PFUS		x	○	x	x	○	○	○
PSUS(N)	~~R/W~~	x	○	x	x	○	x	x

○: Command can be accepted, x: Incapable of accepting the command
*1: Can be specified for each group only
*2:Due to the cascade configuration, command can be accepted if the SnapShot pair is in the PSUS or PAIR status.

Table 8.14 SnapShot Command Acceptance (Same TCE P-VOL and SnapShot P-VOL)

TCE P-VOL Status	Command Acceptability issued to a SnapShot pair					
	paircreate	pairresync	pairresync -restore	pairsplit	paircreate -split	pairsplit -S
SMPL	○	○	○	○	○	○
COPY	○	○	x	○	○	○
PAIR	○	○	x	○	○	○
PSUS	○	○	○	○	○	○
PSUE	○	○	x	○	○	○
PFUS	○	○	○	○	○	○

○: Capable of accepting the command, x: Incapable of accepting the command

Table 8.15 SnapShot Command Acceptance (Same TCE S-VOL and SnapShot P-VOL)

TCE S-VOL Status		Command Acceptability issued to a SnapShot pair					
		paircreate	pairresync	pairresync -restore	pairsplit	paircreate -split	pairsplit -S
SMPL		○	○	○	○	○	○
COPY		○	○	x	x	x	○
PAIR		○	○	x	x	x	○
PSUS	R/W	○	○	○	○	○	○
	R	○	○	x	○	○	○
PSUE		○	○	x	○	○	○
SSWS		○	○	○	○	○	○
SSWS(R)		○	○	x	○	○	○
PFUS		○	○	○	○	○	○
PSUS(N)	~~R/W~~	○	○	x	x	x	○

○: Capable of accepting the command, x: Incapable of accepting the command

When the target LU of a TCE pair is a SnapShot P-VOL and the SnapShot pair is being restored or has been placed in the PSUE status during restoration, the TCE pair cannot be created (paircreate) or resynchronized (pairresync). Therefore, it is required to restore the SnapShot pair.

When the TCE pair is in PAIR status and the SnapShot pair cascaded with the TCE S-VOL is in PAIR or PSUS status, the pair can be split only using the pairsplit -mscas command.

When a failure occurs during restoration of the S-VOL of the TCE pair, using the settled data saved in the data pool, the SnapShot pair used in the same subsystem is placed in the PSUE status. The TCE pair and the SnapShot pair cannot be recovered unless they are split (pairsplit -S) and recreated (paircreate).

8.7 TCE Path Timeout settings

- Setting a larger time limit based on rate of errors

  It is recommended to set a larger time limit based on the rate of errors (such as the loss of data packets). The time limit must be specified shorter than the CCI monitoring time. When you specify the time limit as 10 seconds or longer, specify the time-out threshold value of the HORCM_MON to be six times as large as the time limit value. The HORCM_MON variable is set in the CCI configuration definition file for the TCE pair. For example, when you specify the time limit as 27 seconds, which is the longest allowable value, the time-out threshold value must be 27×6 = 162 seconds or longer.

8.8 Cautions and Restrictions

- SYNCHRONIZE CACHE command
  - When a SYNCHRONIZE CACHE command is received from a host, it usually writes the entire write pending data stored in the cache memory to respective drives. However if TCE is installed, the write pending data is not written to drives when the SYNCHRONIZE CACHE command is received.
  - To write the entire write pending data stored in the local cache memory to the drives when TCE is installed, it is required to turn on the Synchronize Cache Execution Mode through Navigator (in the local site). It is necessary to turn off the Synchronize Cache Execution Mode in remote site, to avoid TCE path failure.

Table 8.16and Table 8.17 list other TCE restrictions and cautions.

Table 8.16 TCE Restrictions

Parameter	TCE Restrictions
Coexistence with TrueCopy	TCE pairs and TrueCopy pairs cannot coexist in the same subsystem.
Cooperation with iSCSI	A configuration in which the connection between the host and the subsystem is made via iSCSI is allowed, but iSCSI cannot be used for the connection between the subsystems.
Restrictions on host access during a copy operation	When the pair status is PAIR or COPY, reading from/to an S-VOL is allowed. However writing from/to an S-VOL and mounting is not allowed. Reading/writing from/to a P-VOL is allowed always.
Restrictions for concurrent use LUN Expansion	It is not allowed to create a TCE pair that unifies LUs with 1 GB or less capacity.
Restrictions for formatting	The pair operation instructed by the CCI command cannot be performed for the LU under the quick formatting.
Restrictions on the cascade connection between TCE volume and SnapShot P-VOL	▪ When the target LU for the TCE pair is a SnapShot P-VOL and it is in the PSUE status during restoration, creation or resynchronization of the pair cannot be done. ▪ When the S-VOL of the TCE pair is a SnapShot P-VOL and is being restored, the horctakeover does not work. ▪ When the S-VOL of the TCE pair is a SnapShot P-VOL and in the SSWS(R) status, the SnapShot pair cannot be restored.
Restrictions for pair splitting	▪ When the instruction to split a pair is accepted, an operation to reflect the P-VOL data on the S-VOL data (splitting operation) is performed. A change status instruction cannot be accepted until the splitting operation is completed. The split processing is continued unless status becomes PSUE or PFUS. The processing is started from the continuation at the time of the next start even if the main switch of the primary subsystem is turned off during the processing. ▪ TCE pair split operation cannot be done when the pairsplit -mscas processing is being executed for the CTG. ▪ When a command to split pairs in each CTG is issued while the pairaplit -mscas processing is being executed for the cascaded SnapShot pair, the splitting cannot be executed for all the pairs in the CTG. ▪ When a command to split each pair is issued and the target pair is under the completion processing, it cannot be accepted if the PAIR to be split is undergoing the end operation. ▪ When a command to split each pair is issued and the target pair is under the completion processing, it cannot be accepted if the PAIR to be split is undergoing the splitting operation. ▪ When a command to split pairs in each group is issued, it cannot be executed if even a single pair that is being split exists in the CTG concerned. ▪ When a command to terminate pairs in each group is issued, it cannot be executed if even a single pair that is being split exists in the CTG concerned. ▪ The pairsplit -P command is not supported.
Restriction for swap	The SWAP Takeover operation that reverses the S-VOL and the P-VOL is not supported. After the SVOL_Takeover operation is performed, delete the TCE pair, and then restore the TCE pair by creating it again. In this case, a full copy of the volume is required.
Restrictions on execution of the horctakeover (SVOL_Takeover) command	▪ When the SVOL_Takeover operation is performed for a pair by the horctakeover command, the operation to restore the S-VOL data to the data saved in the data pool. Therefore, the instruction to change the status cannot be accepted until this end operation is completed. The restoration of up to four LUs can be done in parallel for each controller. When restoration of four or more LUs is required, the first four LUs are selected according to an order given in the requirement, but the following LUs are

Parameter	TCE Restrictions
	selected in ascending order of the LU numbers. ▪ Since SVOL_Takeover operation is performed on the secondary side only, the differential data of the P-VOL that has not been transferred is not reflected on the S-VOL data even when the TCE pair is operating normally. ▪ The SVOL_Takeover operation cannot be executed when the S-VOL of the pair, to which the SVOL_Takeover operation is issued, is in the PSUS(N) status. ▪ When a command specifies the target as a group, it cannot be executed for all the pairs in the CTG, even if a single pair in the PSUS(N) status exists in the CTG. ▪ When a command specifies the target as a pair, it cannot be executed if the target pair is in the SMPL or COPY status.
Restrictions on execution of pairsplit –mscas command	▪ The pair splitting instruction cannot be issued by the host on the secondary side, to the SnapShot pair cascaded with the TCE S-VOL pair, when the pair is in the COPY or PAIR status. ▪ The command cannot be executed even if a single pair is under the end operation. ▪ When the pairsplit -mscas command is being executed for any SnapShot pair that is cascaded with a pair in the specified CTG, the command cannot be executed. The pairsplit -mscas processing is continued after pair status becomes PSUE or PFUS. The processing is started from the continuation at the time of the next start even if the main switch of the primary subsystem is turned off during the processing.
Restrictions on the performance of end operation for a pair	▪ When the request to end a pair is accepted, the operation to reflect the P-VOL data on the S-VOL data (the end operation) is performed. Therefore, the request to change the status cannot be accepted until the splitting operation is completed. The end processing is continued unless it becomes PSUE or PFUS. The processing is started from the continuation at the time of the next start even if the main switch of the primary subsystem is turned off during the processing. ▪ The command to end a pair cannot be accepted when the target pair is being split. When the command to end a pair is issued to each group, it cannot be accepted if even a single pair in the CTG is being split. ▪ The command cannot be executed when the pairsplit -mscas command is being executed for the SnapShot pair cascaded with the target pair. ▪ When the command is issued to each group, it cannot be executed if the pairsplit -mscas command is being executed for the SnapShot pair cascaded with a pair in the CTG. ▪ If a pairsplit -R command requires the secondary subsystem to delete a pair, the differential data of the P-VOL that has not been transferred is not reflected on the S-VOL data in the same way as the case of the SVOL_Takeover operation. ▪ The pairsplit -R command cannot be executed during the restoration of the S-VOL data through the SVOL_Takeover operation. ▪ The pairsplit -R command cannot be issued to each group when a pair, whose S-VOL data is being restored through the SVOL_Takeover operation, exists in the CTG.
Initial copy	When receiving the paircreate command, the entire copy is executed in the direction of P-VOL to S-VOL. The write operation received by the P-VOL from the host during the initial copy becomes new differential data to the S-VOL; and the new differential data is also copied after all copies are completed. When the difference between the P-VOL and the S-VOL data reaches a certain amount, the initial copy is completed.
Copying pace	The copy pace can be adjusted to three stages by the option –c of the paircreate command and pairresync command.
Differential management	When the P-VOL and S-VOL is in split status (PSUS), write operations received individually will be managed as differential data of P-VOL and S-VOL.
Write sequencing	The differential data is transferred from the P-VOL to the S-VOL in a cycle specified by a user. Therefore, order of the transferred data, which is to be reflected on the S-VOL, in each cycle, is assured.

Parameter	TCE Restrictions
Cycle time is variable	The cycle time of the update can be changed as needed. The default cycle time is 300 seconds; and the cycle can be specified by the second up to 3,600 seconds. The shortest limit value that can be set is calculated as number of CTGs × 30 seconds.
Resynchronization	When the pairresync command is accepted, the differential data is copied from the P-VOL to the S-VOL. The differential data that is newly produced during the copy is also copied. The copy is completed when the new differential data is copied up to the certain amount.
Management of LUs while using TCE	Formatting, extending, and deleting LUs are not available. When formatting, extending, and deleting LUs, split TCE pair(s) using the pairsplit command.
Concurrent use of LUN Expansion	Neither the expansion nor the expansion release can be done for an LU belonging to a TCE PAIR. The LU can be specified as a TCE pair after the expansion or the expansion release.
Concurrent use of Data Retention Utility	The pair operation of the S-VOL for which the S-VOL Disable is set cannot be performed. The S-VOL Disable can be set for the LU, which has already been specified as an S-VOL, only when the pair status of the LU is PSUS.
Concurrent use of Cache Residency Manager	The LU specified for Cache Residency (LU cache residence) cannot be set to S-VOL or P-VOL.
Concurrent use of TrueCopy	TCE volume cannot be cascaded or used together with TrueCopy volume.
Concurrent use of ShadowImage	TCE and ShadowImage can be used concurrently. TCE volume cannot be cascaded with ShadowImage volume.
Concurrent use of SnapShot	TCE volume can be cascaded only with SnapShot P-VOL.
Concurrent use of Volume Migration	Available. However, a P-VOL, an S-VOL, and reserved LU of Volume Migration cannot be specified as a P-VOL or an S-VOL of TCE.
Relationship with number of SnapShot pairs	The total number of Snapshot P-VOLs and TCE P-VOLs and S-VOLs is limited as follows: Number of Snapshot P-VOLs + Number of TCE P-VOLs + Number of TCE S-VOLs ≤ 2,046/subsystem (AMS1000) ≤ 1,022/subsystem (AMS500) ▪ The number of TCE P-VOLs and the number of TCE S-VOLs indicate the number of TCE P-VOLs and S-VOLs in the target disk subsystem (one P-VOL or S-VOL per pair). ▪ If a SnapShot P-VOL and a TCE P-VOL or S-VOL are common (cascade configuration), they are counted as one. Example: If SnapShot has P-VOL (LU1)-V-VOL (LU2) and TCE has P-VOL (LU1)-S-VOL (LU3 remote subsystem), only one common P-VOL (LU1) is counted.
Concurrent use of Cache Partition Manager	▪ TCE can be used together with Cache Partition Manager and they can be used being combined each other. Cache partition information is initialized when SnapShot is installed in the status where Cache Partition Manager is already in use. For details, see Appendix C. ▪ When the cache partition is used together, a disk drive failure may extend not in the RAID group units but in the partition units.
Use of Cache Partition Manager for the LU belonging to data pool	▪ The segment size can be changed by using the Cache Partition Manager license. ▪ When using TCE with Cache Partition Manager, the segment size of the LU belonging to data pool must be the default size (16 kB) or less.
License	TCE becomes usable through entry the key code. TCE cannot be activated with a TrueCopy license.

Parameter	TCE Restrictions
Failures	▪ When a copy operation from P-VOL to S-VOL fails, TCE suspends the pair (PSUE). Because TCE copies data to the remote S-VOL regularly, data of the S-VOL restored from the PSUE status is consistent with P-VOL data from immediately before the occurrence of the failure. ▪ If a drive failure occurs, the TCE pair status is not affected because of the RAID architecture.
RAID group deleting, LU deleting, and formatting for a paired P-VOL or S-VOL	RAID group deletion, LU deletion, formatting, and data pool deletion for a paired P-VOL or S-VOL cannot be performed.
Restriction for formatting the volumes	Do not execute TCE operations while formatting the volume. Formatting takes priority and the TCE operations will be suspended.
Reduction of the memory	The memory cannot be reduced when the ShadowImage, SnapShot, TrueCopy, or TCE function is validated. Make the reduction after invalidating the function.

Table 8.17 TCE Cautions

Parameter	TCE Cautions
Controller check required before creating a TCE pair	▪ The TCE pair cannot be created unless the default owner controller (D-CNT) and the current owner controller (C-CNT) of the LU to be defined as the S-VOL are made to be the same. ▪ Check that the D-CTL value and the C-CTL value are the same using Navigator. For details, see section4.2.5. ▪ To make the P-VOL and the S-VOL completely identical, necessary to perform the pairsplit operation from the primary side.
Notices on the pair deletion	After the SVOL_Takeover operation is instructed, the command to delete the pair cannot be accepted until the restoration of the S-VOL data using the data pool data is completed. The progress of the restoration can be checked using Navigator.
Suspension failure inflow > outflow	If the amount of I/Os instructed to the primary subsystem from the host (inflow) is more than that of the data transferred from the primary subsystem to the secondary subsystem (outflow), a data pool overflow can occur and the TCE pair may be split.
Suspension failure When the load on the controller is heavy	▪ When the controller of the primary or secondary subsystem is overloaded due to the host I/O, etc., the performance of the data transfer is lowered. ▪ When the overload continues, the balance between the inflow and outflow data is lost; and when the amount of data in the data pool exceeds its threshold value, the pair is suspended.
Suspension failure path or the controller is switched, suspended (PFUS), extent of the influence of the pair splitting	When the path or the controller is switched, the transfer of the differential data of the TCE pair in the PAIR status delays because of a load of transfer. If the delay continues, the possibility that the pair is made to be suspended (placed in the PFUS status) becomes higher. When the data in the data pool overflows, all the TCE pairs, whose internal states conform to the following, among those that use the data pool are suspended. When a failure occurs in the data pool in the status of PAIR, SSWS(R), Suspending, Deleting, or being restored, V-VOLs of all the pairs of SnapShot being used together with TCE are placed in the PSUE status.
Delay in the settling of the S-VOL data (1)	▪ When data inflow into the primary subsystem from the host is larger than data outflow from the primary subsystem to the secondary subsystem, it takes some time to complete settling of the S-VOL data because the amount of data to be transferred increases. ▪ When S-VOL data settlement is delayed, the potential data loss increases when a failure occurs in the primary subsystem.
Delay in the settling of the S-VOL data	The differential data in the primary subsystem that has not been transferred increases under the following circumstances: Time when the load on the controller is heavy Time when initial/resynchronization copy is done SATA drives are used Path/controller is switched When the settlement of the S-VOL data is delayed. As a result, the data loss increases when a failure occurs in the primary subsystem because the time until completing the settlement of the S-VOL data increases.
Cascade with SnapShot	To split the SnapShot pair to be cascaded with the S-VOL of TCE with an instruction from the primary subsystem, it is required that the TCE pair is in the PAIR status and the SnapShot pair is in the PAIR or PSUS status at the same time.
Host I/O	▪ When the amount of inflow data is larger than that of outflow data, restriction for reducing the amount of the inflow data cannot be reduced by delaying the response to the host I/O instruction.

Parameter	TCE Cautions
	▪ It is required to construct a system considering the possibility that the amount of data in the data pool exceeds its threshold value depending on the state of the excess of the inflow data amount.
Notices on the bidirectional communication	When CCI commands are performed in the two directions at the same time, the command execution may terminate abnormally.
Notices on execution of the horctakeover (SVOL_Takeover) command	▪ The SVOL_Takeover operation can be executed for a pair in the status of PAIR, PSUE, or PSUS other than PSUS(N). ▪ When the instruction to perform the SVOL_Takeover operation is issued to each CTG, the operation is executed on TCE pairs in the CTG in the PAIR, PSUE, PFUS, or PSUS status (status other than PSUS(N)).
Notices on the execution of the pair splitting command other than pairsplit -mscas	To issue an instruction to split the SnapShot pair from the secondary subsystem, it is required to place the TCE pair in the PSUS or SMPL status.
Notices on execution of the pairsplit -mscas command	▪ The instruction is accepted only by each group and executed for each CTG. The TCE pair to be the target must be in the PAIR status. ▪ When a pair in a CTG is in a status other than COPY or PAIR, the instruction cannot be executed. ▪ The SnapShot pair cascaded with the target pair must be in the PAIR or PSUS status. ▪ The MU numbers of all the SnapShot pairs concerned must be the same as those specified when the pairsplit -mscas command was executed.
Notice on the operation with two or more CTGs	When the operation is performed with two or more CTGs, a copy operation for a certain CTG may be performed later than that for the other CTG.
Notice on a status change to PSUE	When a failure occurs in the secondary subsystem or the P-VOL accepts a SVOL_Takeover or pairsplit -R command, requiring a status change only for the secondary subsystem in the PAIR status, the P-VOL changes its status to PSUE by performing the next communication. Therefore, the P-VOL status may seem to be normal for some time while the communication or the data transfer is not performed.

8.9 Restrictions about Connection with the Host

8.9.1 Host Timeout

It is recommended that you set more than 60 seconds for the I/O timeout from the host to the TagmaStore. When you set the time limit until the time-out of the remote path to ten seconds or more, set that of a host I/O to six times as long as that of the remote path. For example, when you set the time limit until the time-out of the remote path to the maximum value of 27 seconds, set that of the host I/O to 162 (= 27×6) seconds or more.

8.9.2 VxVM and TCE Configuration

- **Volumes to be recognized by the same host**

  If you have set the P-VOL and the S-VOL to be recognized by the same host, the VxVM will not operate properly. Set only the P-VOL to be recognized by the host and let another host recognize the S-VOL.

8.9.3 AIX® and TCE Configuration

- **Volumes to be recognized by the same host**

  If you have set both the P-VOL and the S-VOL to be recognized by the same host, AIX® will not operate properly. Set only the P-VOL to be recognized by the host and let another host recognize the S-VOL.

8.9.4 Windows Server™ 2003 and TCE Configuration

- Volume mount:

  When mounting a volume, use Volume{GUID} as an argument of the mount command of CCI. The Volume{GUID} can be used by CCI 01-13-03/00 or later. (For more detail, see the *Hitachi TagmaStore® Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

- Command devices:
  - When describing a command device in the configuration definition file, specify it as Volume{GUID}. For more detail, see the *Hitachi TagmaStore® Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).
  - When a path detachment, which is caused by a controller detachment or fibre channel failure, continues for longer than one minute, the command device may be unable to be recognized when recovery from the path detachment is made. To make the recovery, execute the "re-scanning of the disks" of Windows. When Windows cannot access the command device, although CCI becomes able to recognize the command device, restart CCI.

8.9.5 Linux® and LVM Configuration

- **Volumes to be recognized by the same host**

  If you have set the P-VOL and the S-VOL to be recognized by the same host, the LVM will not operate properly. Set only the P-VOL to be recognized by the host and let another host recognize the S-VOL.

8.9.6 Windows® 2000/Windows Server 2003 and Dynamic

8.9.6.1 Environments

In an environment of the Windows Server 2000, you cannot make a P-VOL and an S-VOL into a dynamic disk. In an environment of the Windows Server 2003, you can use a P-VOL and an S-VOL as a dynamic disk.

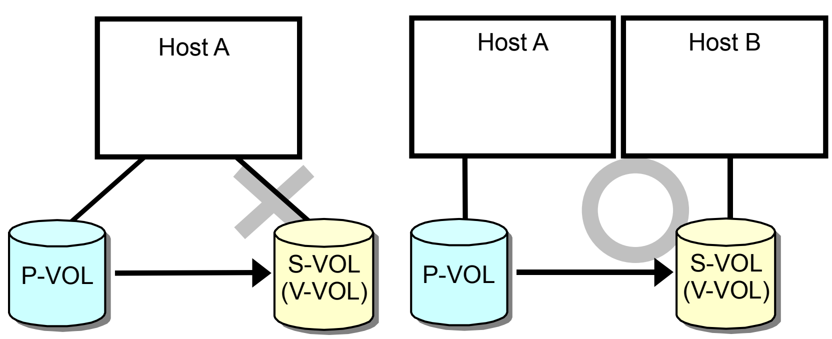
Note: A dynamic disk is a function of Windows 2000/Windows Server 2003.

8.9.6.2 Restrictions

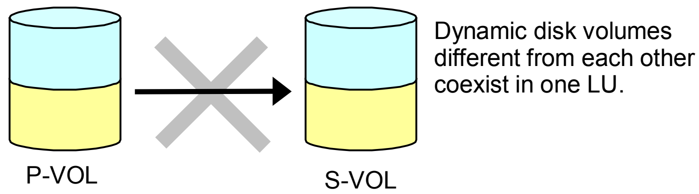
Observe the following when you use a dynamic disk with Windows Server 2003.

Follow the restrictions of the copy functions (ShadowImage, SnapShot, TrueCopy, and TCE) in addition to the following restrictions. For the restrictions of the each copying function, refer to the corresponding manual.

- When a secondary host uses an S-VOL, verify that the host recognizes it after making sure that the pair status is PSUS after the pair is created.
- One host cannot recognize both a P-VOL and an S-VOL.



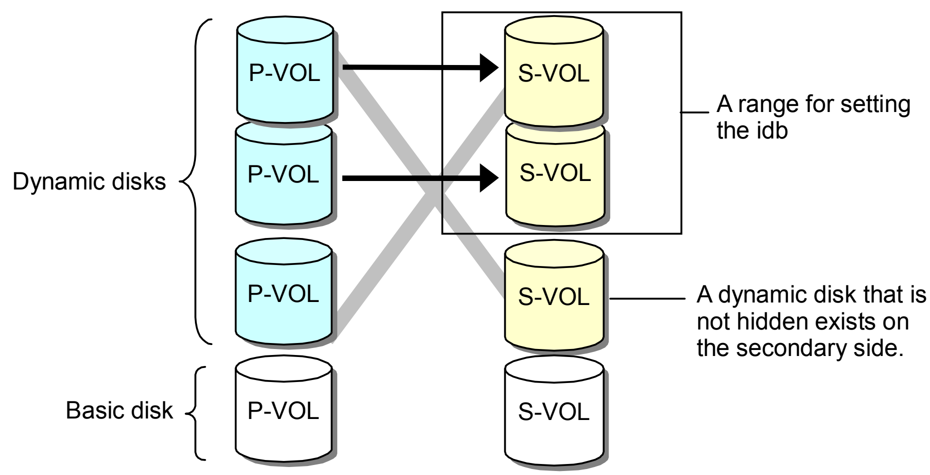
- An LU, in which two or more dynamic disk volumes coexist, cannot be copied.

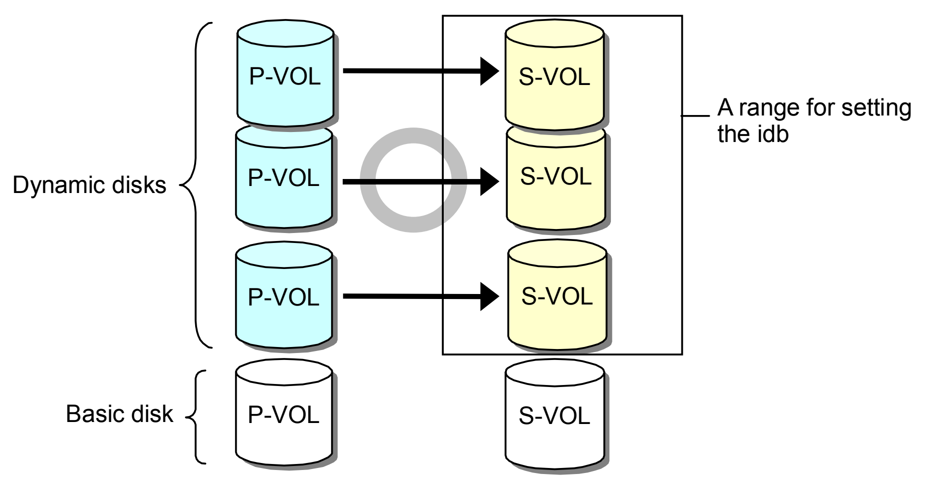

Dynamic disk volumes
different from each other
coexist in one LU.
P-VOL
S-VOL

- Do not use a dynamic disk function for volumes other than an S-VOL on the secondary host side.

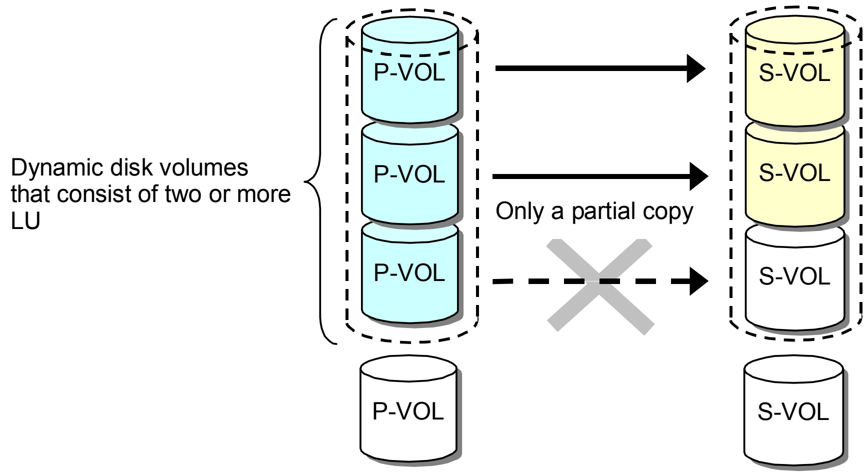
  When copying, hide all the dynamic disks exist on the primary side using `raidvchkset –vg idb`. No restriction is placed on the primary side. (Hide all the dynamic disk volumes to be restored on the primary side at the time of restoration.)

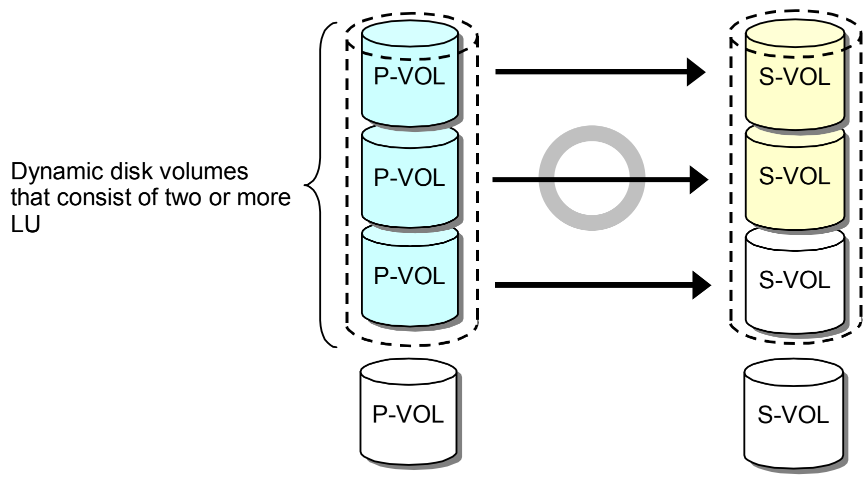
  If any one of the dynamic disks is left unhidden, a **Missing** drive occurs. When this occurs, delete it manually using the `diskpart delete` command.





- When copying dynamic disk volumes that consist of two or more LUs, do this after hiding all LUs composing the dynamic disk at the same time. After the copy is completed, release LUs from being hidden and have them recognized by a host.





- A dynamic disk cannot be used with a cluster (MSCS, VCS, and etc).
- A dynamic disk cannot be used with VxVM and HDLM.

Chapter 9 Troubleshooting

This chapter describes TCE troubleshooting procedures and details about error messages and error codes. This chapter includes the following sections:

- Troubleshooting Procedures (section 9.1)
- TCE Error Messages and Error Codes (section 9.2)
- Calling the Hitachi Data Systems Support Center (section 9.3)

9.1 Troubleshooting Procedures

If a hardware error occurs while you are operating TCE, both of the following tasks are necessary:

- A CCI user intervention
- Assistance from a Hitachi Customer Service representative

For example, when formatting is needed to resolve an LU error and that LU is used for TCE, the pair must be released by the user (CCI operation by the user) before the LU can be formatted. Therefore, contact Hitachi personnel because maintenance requires the user to issue CCI commands. Note that the Hitachi personnel can only remove errors which result from hardware. An operation such as recovering a TCE pair status (for example, re-synchronizing) must be done by the user.

If you have a problem with TCE operation, see the *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701). TCE error messages and procedures for recovering the PSUE errors are described in detail in the Troubleshooting section.

9.2 TCE Error Messages and Error Codes

This section describes error reporting processes and error message locations.

9.2.1 Error Reporting during Pair Resynchronization

The resynchronization of a TCE pair (pairresync) returns a response at the time when the status of the P-VOL is changed to COPY at the execution. Therefore, as for a failure that occurs in the process after the GOOD response is returned from the command, a cause of the error cannot be specified from the command response.

The cause of the failure can be specified by the error code for pairresync displayed in the information message in Navigator and the Web window.

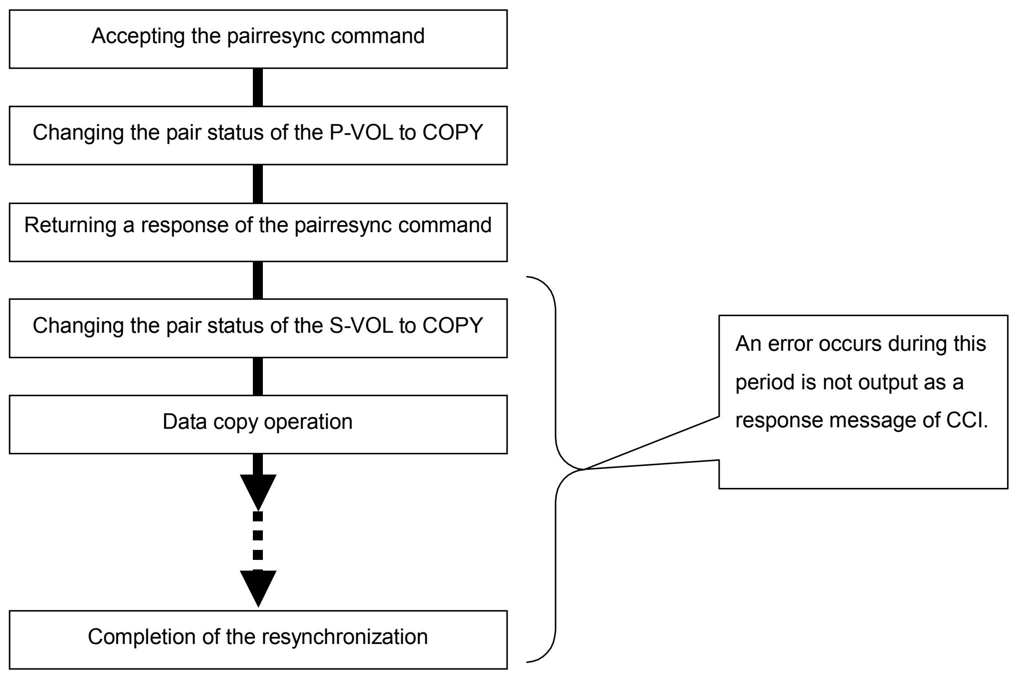


Figure 9.1 Arranging Log and Trace

9.2.1.1 Error Code Locations

The error code for pairresync can be checked by the information message in Navigator and the Web window.

The error code for pairresync is displayed in the Web window as shown in Figure 9.2. It can also be displayed when the **Information Message** is selected from the **Tools** menu in the **Array System Viewer** of Navigator (Figure 9.3). Figure 9.4 shows an example of a message display window in Navigator.

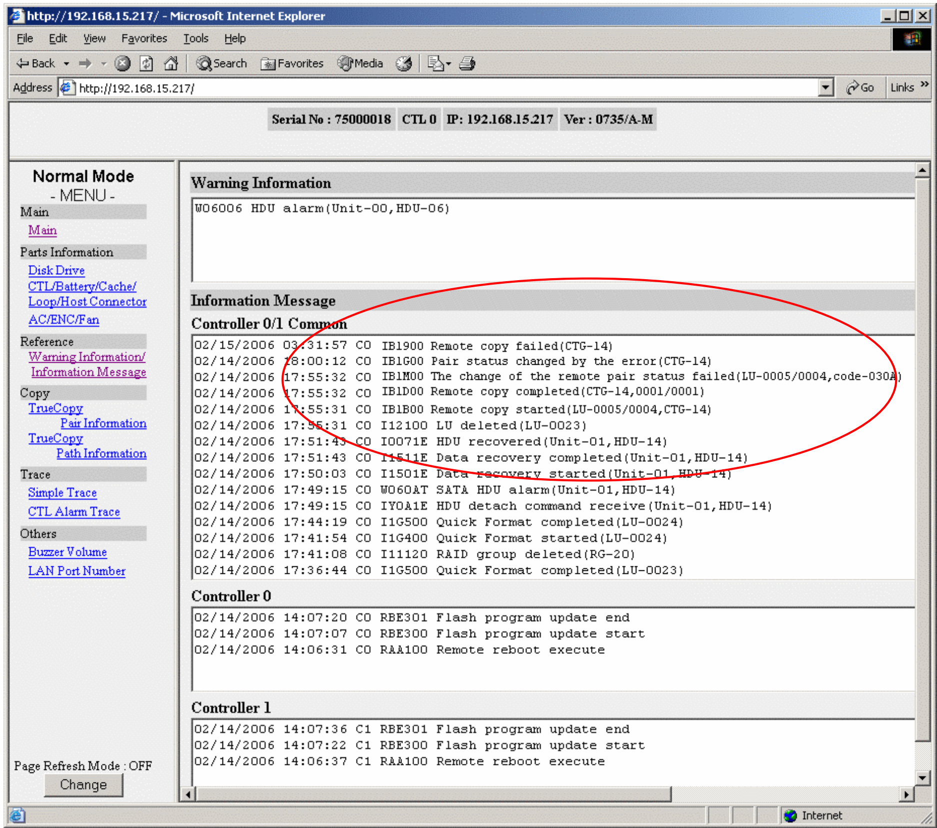


Figure 9.2 Information Message (Web)

Select the **Information Message** from the **Tools** menu in the **Array System Viewer** of Navigator.

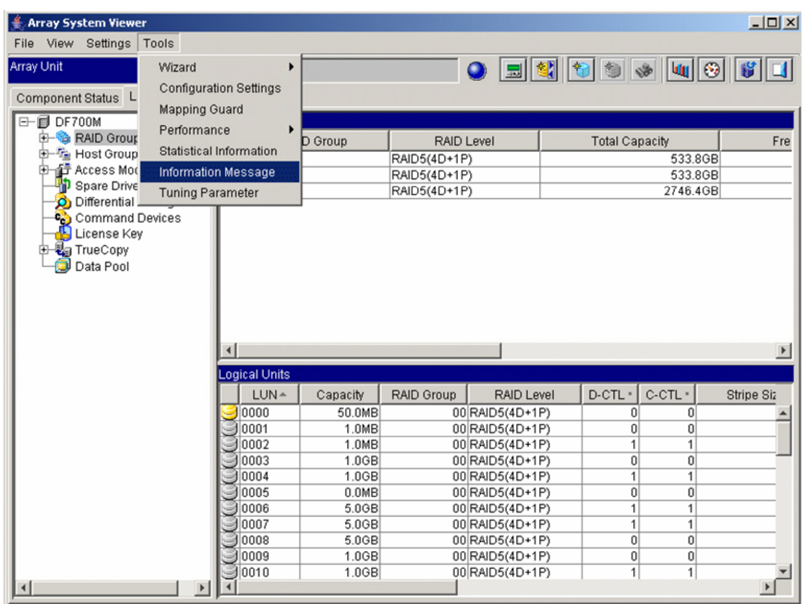


Figure 9.3 Tools Menu (Navigator)

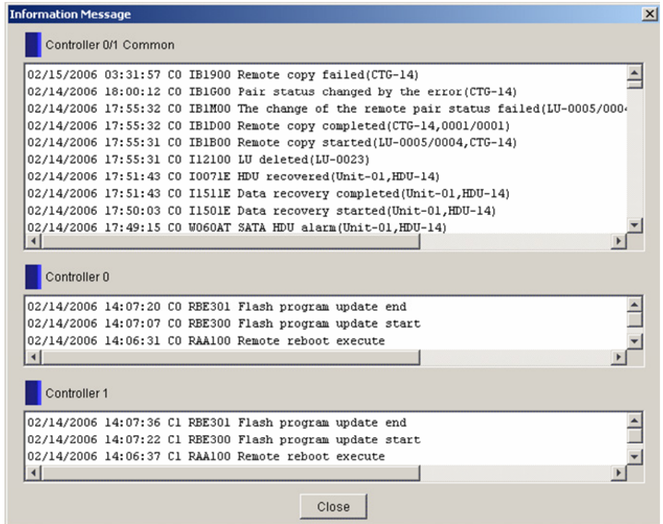


Figure 9.4 Information Message (Navigator)

9.2.1.2 Error Message Contents

Error messages contain the following information:

- Time when an error occurred
- A controller number
- A explanatory message
- An error code

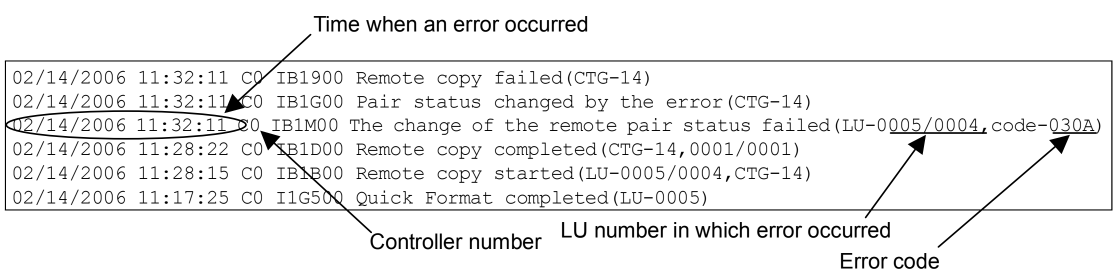


Figure 9.5 Sample Error Message

Note: The error message for pairresync is "The change of the remote pair status failed". The error code for pair resynchronization can be checked in the information message in Navigator and the Web window.

Table 9.1 Error Codes and Corrective Actions

Error Code	Error Contents	Actions to be Taken
0307	The serial number of the remote subsystem cannot be specified.	Check the serial number of the remote subsystem.
0308	The LU assigned to a TCE pair cannot be specified.	The resynchronization cannot be performed. Create a pair again after deleting the pair.
0309	Restoration from the data pool is in progress.	Retry after waiting for a while.
030A	The target S-VOL of TCE is a P-VOL of SnapShot. Besides, the SnapShot pair is being restored or reading/writing is not allowed.	When the SnapShot pair is being restored, execute it after the restoration is completed. When reading/writing is not allowed, execute it after enabling the reading/writing.
030C	The TCE pair cannot be specified in the CTG.	The resynchronization cannot be performed. Create a pair again after deleting the pair.
0310	The status of the TCE pair is SSWS.	
0311	The status of the TCE pair is SMPL.	
031F	The LU of the S-VOL of the TCE is S-VOL Disable.	Check the LU status of in the remote subsystem, release the S-VOL Disable, and execute it again.
0320	The target LU in the remote subsystem is undergoing the parity correction.	Retry after waiting for a while.
0321	The status of the target LU in the remote subsystem is other than normal or regressed.	Execute it again after restoring the target LU status.
0322	The number of unused bits is insufficient.	Retry after waiting for a while.
0323	The LU status of the data pool is other than normal or regressed.	Execute it again after restoring the LU status of the data pool.
0324	The LU of the data pool is undergoing the parity correction.	Retry after waiting for a while.
0325	The expiration date of the temporary key is expired.	The resynchronization cannot be performed because the trial time limit is expired. Purchase the permanent key.
0326	The disk drives that configure a RAID group, to which a target LU in the remote subsystem belongs have been spun down.	Perform the operation again after spinning up the disk drives that configure the RAID group.

9.3 Calling the Hitachi Data Systems Support Center

If you need to call the Hitachi Data Systems Support Center, make sure to provide as much information about the problem as possible, including:

- The configuration definition files (spreadsheets) downloaded (exported) using the Configuration File Loader function (see *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Graphical User Interface (GUI) User's Guide (MK-95DF711)* or *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Command Line Interface (CLI) User's Guide (MK-95DF712)*.
- The circumstances surrounding the error or failure
- The exact content of any error messages displayed on the host system(s)
- The service information messages (SIMs) displayed on the Storage Navigator Modular and the reference codes and severity levels of the recent SIMs

The worldwide Hitachi Data Systems Support Centers are:

- Hitachi Data Systems North America/Latin America
  San Diego, California, USA
  1-800-446-0744
- Hitachi Data Systems Europe
  Buckinghamshire, United Kingdom
  011-44-175-361-8000
- Hitachi Data Systems Asia Pacific
  North Ryde, Australia
  61-2-9325-3300

Appendix A TCE Operations Using SNM CLI

This section describes TCE operational procedures using the SNM CLI. The following sections are included:

- Installing TCE
- Uninstalling TCE
- Enabling or Disabling TCE
- Setting the Command Device
- Setting the Differential Management LU (DM-LU)
- Setting the POOL
- Setting the Cycle Time
- Setting the Target ID
- Setting the Path
- Confirming the Status of TCE Pairs
- Confirming the Status of CTG

For details on Navigator, see *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Graphical User Interface (GUI) User's Guide (MK-95DF711)* or *Hitachi TagmaStore Adaptable Modular Storage and Workgroup Modular Storage Storage Navigator Modular Command Line Interface (CLI) User's Guide (MK-95DF712)*.

A.1 Installing TCE

Since TCE is an extra-cost option, the TCE cannot usually be selected (locked) before it can be used. To make TCE available, you must install TCE and set its function selectable (unlock).

Note: Before installing/uninstalling TCE, make sure that the subsystem to be operated is in normal state. If a failure such as a controller blockage has occurred, installation/uninstallation cannot be performed.

To install TCE, the key code or key file provided with the optional feature is required. The following describes the installation procedure.

To install TCE:

1. From the command prompt, register the subsystem in which TCE is to be installed, and then connect to the subsystem.
2. Execute the **auopt** command to install TCE. The example is shown below.

   Navigator version is 5.00 or later and Cache Partition Manager is enabled

```
% auopt -unit subsystem-name -lock off -keycode manual-attached-keycode
Password: manager-password
Are you sure you want to install the option? (y/n[n]): y
When Cache Partition Manager is enabled, if the option using data pool will be
e
nabled the default cache partition information will be restored.
Do you want to continue processing? (y/n [n]): y
The option is installed successfully.
In order to complete the setting, it is necessary to reboot the subsystem.
Host will be unable to access the subsystem while restarting. Host applications
that use the subsystem will terminate abnormally. Please stop host access
before
 you restart the subsystem.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
When using TrueCopy, restarting the remote subsystem will cause both TrueCopy
pa
ths to fail.
TrueCopy pair status will be changed to "PSUE" when pair status is "PAIR" or
"CO
PY". Please change TrueCopy pair status to "PSUS" before restart.
Do you agree with restarting? (y/n [n]): y
Are you sure you want to execute? (y/n [n]): y
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Navigator version is 3.00 to 4.07 and Cache Partition Manager is enabled

```
% auopt -unit subsystem-name -lock off -keycode manual-attached-keycode
Password: manager-password
Are you sure you want to unlock the option? (y/n[n]): y
When Cache Partition Manager is enabled, if the option using data pool will be
e
nabled the default cache partition information will be restored.
Do you want to continue processing? (y/n [n]): y
The option is unlocked.
In order to complete the setting, it is necessary to reboot the subsystem.
Host will be unable to access the subsystem while restarting. Host applications
that use the subsystem will terminate abnormally. Please stop host access
before
 you restart the subsystem.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
When using TrueCopy, restarting the remote subsystem will cause both TrueCopy
pa
ths to fail.
TrueCopy pair status will be changed to "PSUE" when pair status is "PAIR" or
"CO
PY". Please change TrueCopy pair status to "PSUS" before restart.
Do you agree with restarting? (y/n [n]): y
Are you sure you want to execute? (y/n [n]): y
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Navigator version is less than 3.00 and Cache Partition Manager is enabled

```
% auopt -unit subsystem-name -lock off -keycode manual-attached-keycode
Password: manager-password
Are you sure you want to unlock the option? (y/n [n]): y
The option is unlocked.
The subsystem stops accepting access from the host while restarting.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
Do you want to restart the subsystem now? (y/n [n]): y
When using the TrueCopy, restart of the remote subsystem will cause both
TrueCop
y paths failure.
And when TrueCopy pair status is "PAIR" or "COPY", TrueCopy pair status will be
changed to "PSUE".
Now restarting the subsystem.  Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Note: It may take time for a subsystem to respond, depending on the condition of the subsystem. If there is no response after 15 minutes or more, check the condition of the subsystem.

3. Execute the **auopt** command to confirm that TCE has been installed. The example is shown below.

```
% auopt -unit subsystem-name -refer
Password: manager-password
Option Name                   Type      Term     Status
TC-EXTENDED                   Permanent ---      Enable
%
```

TCE is installed and the feature status is “Enable”. Installation of TCE is now complete.

A.2 Uninstalling TCE

To uninstall TCE, the key code provided with the optional feature is required. Once uninstalled, TCE cannot be used (locked) until it is again installed using the key code or key file.

Note: The following conditions must be satisfied in order to uninstall TCE:

– All TCE pairs must be released (the status of all LUs are SMPL).

– All the Data Pools must be deleted. However, the deletion of the Data Pools is not needed when SnapShot is to be used continuously.

– The path settings must be released.

The following describes the uninstallation procedure.

To uninstall TCE:

1. From the command prompt, register the subsystem in which the TCE is to be uninstalled, and then connect to the subsystem.
2. Execute the **auopt** command to lock TCE. The example is shown below.

   Navigator: Version 5.00 or later

```
% auopt -unit subsystem-name -lock on -keycode manual-attached-keycode
Password: manager-password
Are you sure you want to de-install the option? (y/n[n]): y
The option is de-installed successfully.
In order to complete the setting, it is necessary to reboot the subsystem.
Host will be unable to access the subsystem while restarting. Host applications
that use the subsystem will terminate abnormally. Please stop host access
before
 you restart the subsystem.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
When using TrueCopy, restarting the remote subsystem will cause both TrueCopy
pa
ths to fail.
TrueCopy pair status will be changed to "PSUE" when pair status is "PAIR" or
"CO
PY". Please change TrueCopy pair status to "PSUS" before restart.
Do you agree with restarting? (y/n [n]): y
Are you sure you want to execute? (y/n [n]): y
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Navigator: Version 4.03 or later

```
% auopt -unit subsystem-name -lock on -keycode manual-attached-keycode
Password: manager-password
Are you sure you want to lock the option? (y/n[n]): y
The option is locked.
In order to complete the setting, it is necessary to reboot the subsystem.
Host will be unable to access the subsystem while restarting. Host applications
that use the subsystem will terminate abnormally. Please stop host access
before
 you restart the subsystem.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
When using TrueCopy, restarting the remote subsystem will cause both TrueCopy
pa
ths to fail.
TrueCopy pair status will be changed to "PSUE" when pair status is "PAIR" or
"CO
PY". Please change TrueCopy pair status to "PSUS" before restart.
Do you agree with restarting? (y/n [n]): y
Are you sure you want to execute? (y/n [n]): y
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Navigator: Versions less than 4.03

```
% auopt –unit subsystem-name –lock on –keycode manual-attached-keycode
Password: manager-password
Are you sure you want to lock the option? (y/n [n]): y
The option is locked.
Restart the subsystem to apply the setting.
The subsystem stops accepting access from the host while restarting.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
Do you want to restart the subsystem now? (y/n [n]): y
When using the TrueCopy, restart of the remote subsystem will cause both
TrueCop
y paths failure.
And when TrueCopy pair status is "PAIR" or "COPY", TrueCopy pair status will be
changed to "PSUE".
Now restarting the subsystem.  Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Note: It may take time for the subsystem to respond, depending on the condition of the subsystem. If there is no response after 15 minutes or more, check the condition of the subsystem.

3. Execute the **auopt** command to confirm that TCE has been uninstalled. The example is shown below.

```
% auopt –unit subsystem-name –refer
Password: manager-password
DMEC002015: No information displayed.
%
```

TCE uninstallation is now complete.

A.3 Enabling or Disabling TCE

The TCE feature can be set to enable or disable during or after installation. By default, the feature is enabled upon installation.

Note: The following conditions must be satisfied in order to disable TCE:

– All TCE pairs must be released (the status of all LUs are SMPL).

– All the Data Pools must be deleted. However, the deletion of the Data Pools is not needed when SnapShot is to be used continuously.

– The path settings must be released.

The following describes the enabling/disabling procedure.

1. From the command prompt, register the subsystem in which the status of the feature is to be changed, and then connect to the subsystem.
2. Execute the **auopt** command to change the status (enable or disable). The following is an example of changing the status from enable to disable. If you want to change the status from disable to enable, enter **`enable`** after the **`-st`** option.

   Navigator: Version 4.03 or later

```
% auopt -unit subsystem-name -option TC-EXTENDED -st disable
Password: manager-password
Are you sure you want to disable the option? (y/n[n]): y
The option has been set successfully.
In order to complete the setting, it is necessary to reboot the subsystem.
Host will be unable to access the subsystem while restarting. Host applications
that use the subsystem will terminate abnormally. Please stop host access
before
 you restart the subsystem.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
When using TrueCopy, restarting the remote subsystem will cause both TrueCopy
pa
ths to fail.
TrueCopy pair status will be changed to "PSUE" when pair status is "PAIR" or
"CO
PY". Please change TrueCopy pair status to "PSUS" before restart.
Do you agree with restarting? (y/n [n]): y
Are you sure you want to execute? (y/n [n]): y
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Navigator: Versions less than 4.03

```
% auopt -unit subsystem-name -option TC-EXTENDED -st disable
Password: manager-password
Are you sure you want to disable the option? (y/n [n]): y
The option has been set successfully.
Restart the subsystem to apply the setting.
The subsystem stops accepting access from the host while restarting.
Also, if you are logging in, the login status will be canceled when restarting
b
egins.
Do you want to restart the subsystem now? (y/n [n]): y
When using the TrueCopy, restart of the remote subsystem will cause both
TrueCop
y paths failure.
And when TrueCopy pair status is "PAIR" or "COPY", TrueCopy pair status will be
changed to "PSUE".
Now restarting the subsystem. Start Time hh:mm:ss Time Required 4 - 15min.
The subsystem restarted successfully.
%
```

Note: It may take time for the subsystem to respond, depending on the condition of the subsystem. If there is no response after 15 minutes or more, check the condition of the subsystem.

3. Execute the **auopt** command to confirm whether the status has been changed. The example is shown below.

```
% auopt -unit subsystem-name -refer
Password: manager-password
Option Name                   Type      Term     Status
TC-EXTENDED                   Permanent ---      Disable
%
```

Enabling or disabling TCE is now complete.

A.4 Setting the Command Device

The command device is a user-selected, dedicated logical volume on the TagmaStore AMS subsystem that functions as the interface to the CCI software. TCE commands are issued by the CCI (HORCM) to the TagmaStore command device.

A command device must be designated in order for a TagmaStore AMS subsystem to accept read and write commands and return read requests to the UNIX®/PC host. The command device must be defined in the HORCM_CMD section of the configuration definition file for the CCI instance on the attached host. Two command devices can be designated for the TagmaStore AMS subsystem. You can designate command devices using the Navigator.

Note:

- LUs set as command devices must be recognized by the host.
- The Command Device LU size must be greater than or equal to 33 MB.
- For TCE, the command devices must be designated on both the remote and local subsystems.

The following restrictions apply when either pair of ShadowImage, SnapShot, or TCE exists or a TCE path is defined.

–When two command devices are set, only one command device can be released.

–When only one command device is set, the command device cannot be released.

To designate command device(s):

1. From the command prompt, register the subsystem to which you want to create the command device. Connect to the subsystem.
2. Execute the `aucmddev` command to create a command device.
   First, displays LUs to be assignable command device, and later create a command device.
   The following is an example of specifying LU 2 for command device 1. When you want to use the protection function of CCI, enter "enable" following the **-dev** option.

```
% aucmddev –unit subsystem-name –availablelist
Password: manager-password
Available Logical Units
  LUN  Capacity    RAID Group RAID Level  D-CTL C-CTL Type Status
    2   35.0 Mbyte          0  5( 4D+1P)      0     0 FC   Normal
    3   35.0 Mbyte          0  5( 4D+1P)      1     1 FC   Normal
%
% aucmddev –unit subsystem-name –set –dev 1 2
Password: manager-password
Are you sure you want to set the command devices? (y/n [n]): y
The command devices have been set successfully.
%
```

3. Execute the `aucmddev` command to verify that the command device has been created. The following shows the example.

Note: Two command devices should be designated to set the alternate command device function or to avoid data loss and subsystem downtime. When two command devices are set within the one disk subsystem, they should be assigned to separate RAID groups to avoid both command devices becoming unavailable owing to a trouble such as a drive failure. For details on alternate command device function, see *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

```
% aucmddev –unit subsystem-name –refer
Password: manager-password
Command Device    LUN  RAID Manager Protect
1                   2  Disable
%
```

4. To release an already set command device, specify as follows.
   The following is an example of releasing command device 1.

```
% aucmddev –unit subsystem-name –rm –dev 1
Password: manager-password
Are you sure you want to release the command devices? (y/n [n]): y
The command devices have been released successfully.
%
```

5. To change an already set command device it is essential to release the already set command device first, and then change the LU number. The following is an example of specifying LU 3 for command device 1.

```
% aucmddev –unit subsystem-name –set –dev 1 3
Password: manager-password
Are you sure you want to set the command devices? (y/n [n]): y
The command devices have been set successfully.
%
```

A.5 Setting the Differential Management LU (DM-LU)

The Differential Management LU is an exclusive logical unit for storing the differential data at the time when the volume is copied. The Differential Management LU in the disk subsystem is treated the same as the other logical units. However, the Differential Management LU is not recognized by a host (it is hidden).

It is required to set a logical unit with a size of at least 5 GB as the Differential Management LU. Up to the two Differential Management LUs can be set. The second one is used for the mirroring. It is recommended to set two Differential Management LUs.

The Differential Management LUs must be set on both the local and remote subsystems. To designate Differential Management LUs,

1. From the command prompt, register the subsystem on which you want to create the differential management LU and connect to that subsystem.
2. Execute the `audmlu` command to create a differential management LU.

   This command first displays LUs that can be assigned as Differential Management LUs, and later creates a Differential Management LU.

```
% audmlu  -unit subsystem-name -availablelist
Password: manager-password
Available Logical Units
  LUN Capacity     RAID Group  RAID Level  D-CTL  C-CTL Type Status
    0  5.0 GByte            0   5( 4D+1P)      0      0 FC   Normal
%
% audmlu  -unit subsystem-name -set -lu 0
Password: manager-password
Are you sure you want to set the DM-LU? (y/n [n]): y
The DM-LU has been set successfully.
%
```

3. To release an already set differential management LU, specify the –rm and –lu options in the `audmlu` command.

```
% audmlu  -unit subsystem-name -rm -lu 0
Password: manager-password
Are you sure you want to release the DM-LU? (y/n [n]): y
The DM-LU has been released successfully.
%
```

The following restrictions apply when ShadowImage, SnapShot, or TCE pairs exist, the TCE path is defined, or SnapShot data pool is defined.

– When two differential management LUs are set, only one differential management LU can be released.

– When only one differential management LU is set, the differential management LU cannot be released.

A.6 Setting the POOL

A logical unit which is formatted and larger than 20GB can be assigned to a data pool. Only a single data pool can be created for each controller. Up to 64 logical units can be assigned to each data pool. An LU consisting of an FC drive and a SATA drive cannot coexist in a data pool. When using TCE with Cache Partition Manager, the segment size of the LU belonging to data pool must be of default size (16 kB) or less. A logical unit assigned to the data pool cannot be recognized by a host.

When an LU is set to a data pool, the accurate capacity of the data pool is not displayed immediately. The data pool capacity can be confirmed after approximately 3 minutes per 100 GB.

The following is the procedure for creating a data pool for storing differential data. To designate data pool(s):

1. From the command prompt, register the subsystem on which you want to create the Data Pool, and then connect to that subsystem.
2. Execute the `aupool` command to create a Data Pool.
   First, displays LUs to be assignable Data Pool, and later create a Data Pool.
   The following is the example of specifying LU 100 for Data Pool 0.

```
% aupool –unit subsystem-name –availablelist –poolno 0
Data Pool    : 0
Available Logical Units
  LUN Capacity     RAID Group RAID Level  Type Status
  100   30.0 Gbyte          0  5( 4D+1P)  FC   Normal
  200   35.0 Gbyte          0  5( 4D+1P)  FC   Normal
%
% aupool –unit subsystem-name –add –poolno 0 -lu 100
Password: manager-password
Are you sure you want to add the logical units to the data pool 0?
 (y/n[n]): y
The logical unit has been successfully added.
%
```

3. Execute the `aupool` command to verify that the Data Pool has been created.

```
% aupool –unit subsystem-name –refer -poolno 0
Data Pool    : 0
  Data Pool Usage Rate: 6% (2.0/30.0 Gbyte)
  Threshold           : 70%
  Status              : Normal
      LUN  Capacity      RAID Group  RAID Level   Type Status
      100    30.0 Gbyte           0    5( 4D+1P)  FC   Normal

  Addable LU
      LUN  Capacity      RAID Group  RAID Level   Type Status
%
```

4. To delete an existing Data Pool, execute the `aupool` command with the -rm and -poolno options. The example shows deletion of Data Pool 0.

```
% aupool –unit subsystem–name –rm -poolno 0
Password: manager–password
Are you sure you want to delete all logical units from the data pool 0?
 (y/n[n]): y
The logical units has been successfully deleted.
%
```

5. To change an existing threshold value of the Data Pool, execute the `aupool` command with the -cng, -poolno and -thres options.

```
% aupool –unit subsystem–name –cng -poolno 0 -thres 70
Password: manager–password
Are you sure you want to change the threshold of usage rate in the data pool?
 (y/n[n]): y
The threshold of the data pool usage rate has been successfully changed.
%
```

A.7 Setting the Cycle Time

Remote copy cycle time can be set for pairs in PAIR status. The cycle time has to be set for each subsystem. The shortest time limit that can be set is calculated as (number of CTGs × 30 seconds).

To designate the cycle time:

1. From the command prompt, register the subsystem to which you want to set the cycle time, and then connect to that subsystem.
2. Execute the `autruecopyopt` command to confirm a cycle time. The cycle time is 300 seconds by default and can be specified in a range from 30 to 3600 seconds.

```
% autruecopyopt  -unit subsystem-name -refer
Cycle Time[sec.]   : 300
%
```

3. Execute the `autruecopyopt` command to set a cycle time.

```
% autruecopyopt  -unit subsystem-name -set -cycletime 300
Password: manager-password
Are you sure you want to set the TrueCopy options? (y/n [n]): y
The TrueCopy options have been set successfully.
%
```

Note: A copy operation may exceed the cycle time specified by a user. Copy times depend on the volume of differential data or line speed.

A.8 Setting the Target ID

The following is the procedure for specifying the target ID. The target ID is specified using the Navigator.

1. From the command prompt, register the subsystem to which you want to set the target ID, and then connect to that subsystem.

2. Execute the `auhgmap` command to set the target ID. In following example LU0 settings are to be recognized as 6 by the host. The port is connected via host group 0 of port 0A on controller 0.

```
% auhgmap -unit subsystem-name -MappingMode on
Password: manager-password
Are you sure you want to set the mapping mode? (y/n [n]): y
when setting starts, the subsystem stops accepting access to the controller
from the host.
Before setting, stop access to the controller from the host.
Do you want to continue processing? (y/n [n]): y
The mapping mode has been set successfully.
%

% auhgmap -unit subsystem-name -add 0 A 0 6 0
Password: manager-password
Are you sure you want to add the mapping information? (y/n [n]): y
The mapping information has been set successfully.
%
```

3. Execute the `auhgmap` command to verify that the target ID has been set.

```
% auhgmap -unit subsystem-name -refer
Mapping mode = ON
Port  Group  H-LUN    LUN
  0A      0      6      0
%
```

A.9 Setting the Path

It is necessary to designate Command Device(s) before a path setting. If a failure such as a controller blockage has occurred, a path setting cannot be performed. The SNM CLI can be used to specify, release, and reference path status. To set a path,

1. From the command prompt, register the subsystem in which you want to set the system parameter, and then connect to that subsystem.
2. The following is an example to reference the path status where path information is not yet specified.

```
% aurmtpath -unit subsystem-name -refer
Password: manager-password
Timeout Period   : 9 sec

Local Information
    Serial Number : 75010021
Remote Information
    Serial Number :

Path  Local  Remote  Status
   0     --      --  undefined
   1     --      --  undefined
%
```

3. Execute the `aurmtpath` command to set the path.

   The following example uses:

   –path 0 for 0A and path 1 for 1A port on the local subsystem

   –path 0 for 0A port

   –path 1 for 1B port on the remote subsystem with serial Number 75010027

```
% aurmtpath -unit subsystem-name -set -remote 75010027  -path0 0A 0A -path1 1A
1B
Password: manager-password
%
```

4. Execute the `aurmtpath` command to confirm whether the path has been set.

```
% aurmtpath –unit subsystem-name –refer
Password: manager-password
Timeout Period   : 9 sec

Local Information
    Serial Number : 75010021

Remote Information
    Serial Number : 75010027

Path  Local  Remote  Status
   0     0A      0A  normal
   1     1A      1B  normal
%
```

5. It is necessary to release all TCE pairs before path releasing a path. To release a previously set path, execute the `aurmtpath` command with the -rm option.

```
% aurmtpath –unit subsystem-name –rm
Password: manager-password
%
```

6. The default timeout value is 9 seconds. When remote subsystem is in a long distance using Extender, the timeout value may be changed. Standard time limit values listed in the tale below are based on frequency band of a line. However, in some cases, it is better to specify a larger value based on a rate of errors such as a packet loss of the line.

Frequency band of a line	Time limit (seconds)
200 Mbps or more	9
40 to 200 Mbps	18
20 to 40 Mbps	27

The time limit must be specified shorter than the time for monitoring by CCI. When you specify the time limit to be 10 seconds or longer, specify the time-out threshold value of the HORCM_MON variable to be six times as large as the value that has been set. The HORCM_MON variable is set in the configuration definition file for the TCE pair. For example, when you specify the time limit as 27 seconds, which is the allowable longest value, the time-out threshold value must be 27×6 = 162 seconds or longer.

```
% aurmtpath –unit subsystem-name –set –remote 75010027 -timeout 27
Password: manager-password
%
```

A.10 Confirming the Status of TCE Pairs

Paired logical units can be displayed using SNM CLI. To confirm the status of TCE pairs,

1. From the command prompt, register the subsystem to which you want to display the status of paired logical volumes and then connect to that subsystem.
2. Execute the `autruecopy` command to display the pair status.

```
% autruecopy –unit subsystem-name –refer
Password: manager-password
Pair(TCE)

                                         Remote Information
    LUN  Pair Attribute  Pair Status  CTG No.  Consistency Time     Capacity
  Difference Size     Synchronized Rate  Serial No.        LUN
    201  P-VOL           COPY              1  ---                     20.0
Gbyte
  ---       (--- )  100%              75010001        1000
    320  P-VOL           COPY              2  ---                     20.0
Gbyte
  ---       (--- )  100%              75010001        520
%
```

A.11 Confirming the Status of CTG

CTG status can be displayed using SNM CLI. To display CTG status,

1. From the command prompt, register the subsystem to which you want to display the status of CTG and connect to that subsystem.
2. Execute the `autruecopy` command to display the CTG status.

```
% autruecopy –unit subsystem-name –refer -ctg
Password: manager–password

                               Remaining
Prediction Time of
CTG No.    CTL  Lapsed Time   Difference Size[MB]  Transfer Rate[KB/s]
Completion
    1        0  05:36:02                    120           Calculating
Calculating
    1        1  05:36:02                    120           Calculating
Calculating
   :
   :
%
```

Appendix B Supported Platforms

Table B.1 shows platforms supported by TCE.

Table B.1 Supported Platforms

No.	Platform	OS REV	Disaster Recovery		Backup Use	Data Migration Use
			High Availability Configuration (Automatically Switching)	Manual Switching		
1	SUN	Solaris 8	Not supported	Supported	Supported	Supported
		Solaris 9				
		Solaris 10				
2	PC Server (Microsoft)	Windows 2000	Not supported	Supported	Supported	Supported
		Windows Server 2003 (IA32)				
		Windows Server 2003 (IA64)				
		Windows Server 2003 (x64)				
3	HP	HP-UX 11.0	Not supported	Supported	Supported	Supported
		HP-UX 11i				
		Tru64 UNIX 5.1	Not supported	Supported	Supported	Supported
4	IBM	AIX 5.1	Not supported	Supported	Supported	Supported
		AIX 5.2				
		AIX 5.3				
5	Red Hat	Red Hat Linux AS2.1, AS3.0, AS4.0	Not supported	Supported	Supported	Supported
		Red Hat Linux AS3.0 EM64T				
		Red Hat Linux AS3.0 IA64				
6	SGI	IRIX 6.5.x	Not supported	Supported	Supported	Supported

Appendix C Using TCE or SnapShot with Cache Partition Manager

TCE and SnapShot use a part of the cache area to manage the internal resources. Therefore, the cache capacity that Cache Partition Manager can use becomes smaller.

Cache partition information should be initialized as shown below when TCE or SnapShot are installed after Cache Partition Manager.

- All the logical units should be moved to the master partitions on the side of the default owner controller.
- All the sub-partitions must be deleted and the size of each master partition should be reduced to half of the user data area after installation of TCE or SnapShot.

Figure C.1 shows an example of Cache Partition Manager usage. Figure C.2 shows an example where TCE/SnapShot is installed when Cache Partition Manager is already in use.

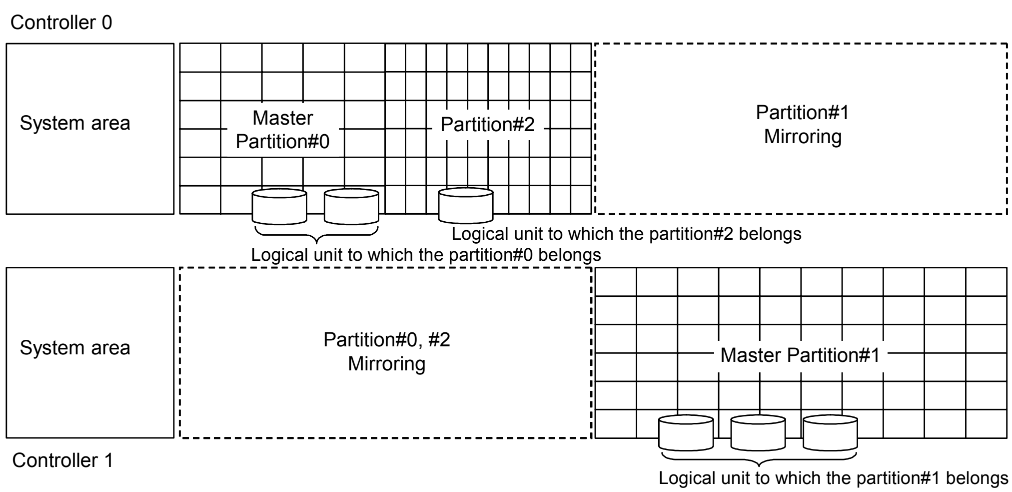


Figure C.1 Cache Partition Manager Usage

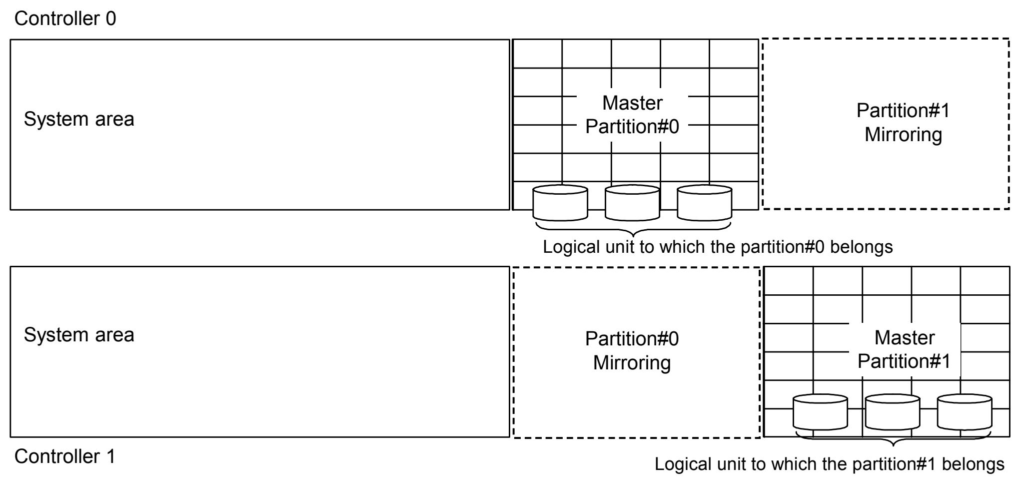


Figure C.2 TCE or SnapShot Installation with Cache Partition Manager

Appendix D TCE on NAS Modular

This appendix describes operating procedures for using TCE on NAS Modular. The following sections are included:

- Overview of TCE Operation in NAS Modular
- Scope of TCE Function
- Requirements
- Other Requirements and Restrictions
- Preparing for TCE Operations
- Overview of TCE Operations
- CCI Log Files Operations
- Commands
- Messages

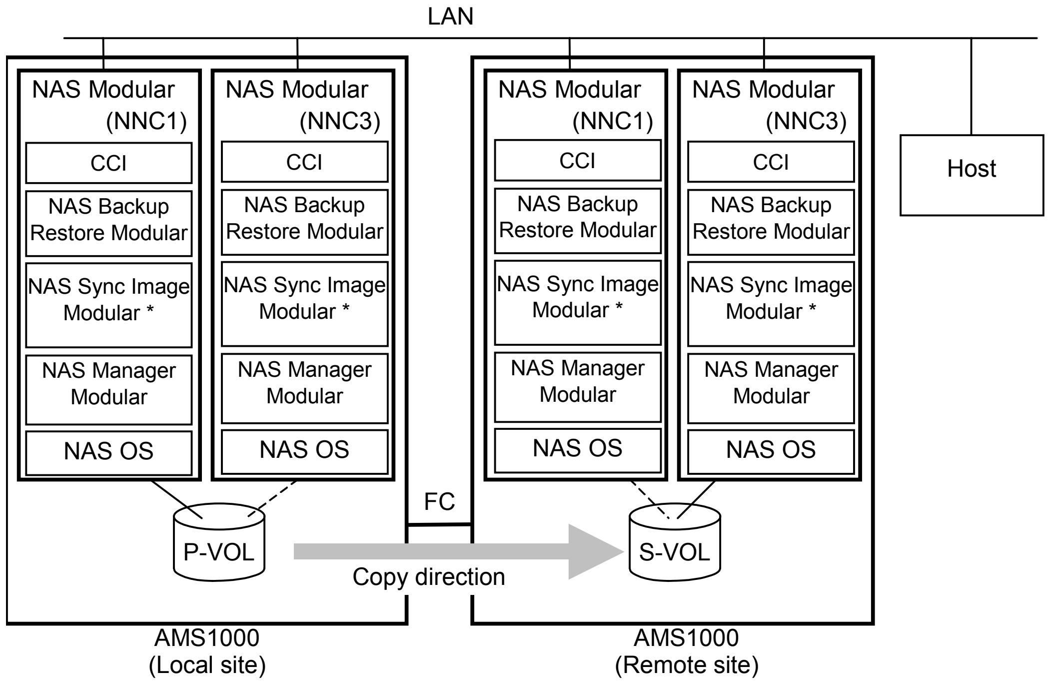
Expected readers of this appendix are:

- Persons who are familiar to use Linux computers or UNIX computers.
- Persons who understand well about basic operation, requirements and limitations of TCE, by reading chapters 1 to 4 of this document.
- Persons who understand well about operation of CCI, by reading the *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701), if they use NAS Backup Restore Modular remote copy function.
- Persons who understand well about operation of NAS Modular system, by reading the *NAS Backup Restore Modular User's Guide*.

Table D.1 lists and describes the operational requirements for TCE on NAS Modular.

Table D.1 TCE Requirements (Difference of FC and NAS)

Parameters	FC	NAS
User interface	CCI: used for issuing commands to create and split TCE pairs. Navigator: used for specifying and releasing command devices.	CCI: used for issuing commands to create and split TCE pairs. To operate the CCI, the following interfaces can be used. • Operation of the CCI commands with SSH Navigator: used for specifying and releasing command devices.
Max number of command devices	AMS1000: 2	AMS1000: 1 (NAS) AMS1000: 2* (FC) * One command device shares with NAS
Max number of pairs	AMS1000: 2,046	AMS1000: Up to 256 pairs for the NAS AMS1000: Up to 2,046 pairs for the FC (Fibre Channel) including the 256 pairs for the NAS
Target LUs	All LUs	NAS: NAS user LUs only. The NAS system LUs are out of the target. The LU, for which a file system has been created, cannot be assigned to an S-VOL
Acceptance to P-VOL during restoration	Acceptable	Not acceptable
Concurrent use of Data Retention Utility	Can be used together at the same time.	Can be used together at the same time. NAS user LUs and NAS system LUs are out of the target. When a file system is created, an attribute of S-VOL Disable, is added automatically.
Concurrent use of SnapShot	Can be used together at the same time.	Can be used together at the same time. NAS user LUs and NAS system LUs are out of the target.
CCI protection function	Supported	Not supported



* Required when copying a file system managed by NAS Sync Image Modular.

Figure D.1 TCE Operation in the NAS Modular System

D.1 Overview of TCE Operation in NAS Modular

NAS Modular system enables integrated management of the data in the disk subsystem. NAS Modular utilizes the LAN environment already in place and enables the data within a disk subsystem to be shared across heterogeneous platforms. NAS Modular also allows you to copy and maintain the data stored in the disk subsystem using the remote copy feature TCE in the AMS1000. With TCE, you can copy data utilizing the NAS server (even in the remote center), back up data to guard against a main volume failure, and perform disaster recovery at the local site.

TCE operations are performed in the NAS Modular system, using CCI for Linux in AMS1000. CCI for Linux gets installed on the NAS Modular system, if NAS Backup Restore Modular is installed. This appendix describes the applicable scope and specific procedures for the TCE function.

D.2 Scope of TCE Function

This section describes the applicable scope and specific procedures for using the TCE function.

D.2.1 Scope Related to TCE

- Volume type

  Only the user LUs may be used as TCE P-VOLs and S-VOLs in a NAS Modular system. The NAS system LUs (NAS OS LU, Cluster Management LU, and dump LU) cannot be used as TCE P-VOLs or S-VOLs.

- Platforms which can access the TCE P-VOLs or S-VOLs

  Only clients that are connected to a network via Hitachi Network Attached Storage products or a NAS Modular can access the TCE P-VOLs and S-VOLs created in a NAS Modular system. Hosts connected through a Fibre-channel port cannot access them.

- File systems which can be allocated in the TCE P-VOL

  You cannot allocate a file system that consists of 257 or more LUs and that uses Logical Volume Manager (LVM) on the NAS OS. You cannot allocate a volume group that consists of 257 or more LUs if it contains a file system managed by NAS Sync Image Modular.

- Writing data to S-VOL when a volume pair is split

  When you split a TCE volume pair in the NAS Modular system, you must change the
  S-VOL to write-enable. It is also necessary to always specify the **-rw** option instead of the **-r** option when issuing the CCI pairsplit command.

  To protect the data in the split S-VOL from being written to by the client, use the **enas_fsmount** command with **-r** option in the NAS Manager Modular, when mounting the file system in S-VOL.

D.2.2 Scope Related to CCI

- Relationships of failover and fail-back on CCI and NAS Modular system

  CCI is not subject to failover or fail-back by the NAS Modular system. Therefore on failover destination node, prepare the same configuration definition file as on failover orientation node except ip_address in the **HORCM_MON** section. If a failure occurs in the NAS unit and causes failover while CCI is operating, start CCI on failover destination node and continue operating TCE.

- Protect feature

  When you use TCE in the NAS Modular system, you create or resynchronize pairs using CCI while the NAS Modular system does not recognize the S-VOL. You therefore cannot use the protect facility in CCI since this facility prohibits the volume pair operations that the system does not recognize.

- Executing the CCI commands

  To execute the CCI commands using the **nasroot** account for logging in to NAS unit using SSH, use the **sudo** command as follows.

```
$ sudo pairdisplay -g VG01
```

Figure D.2 Sample pairdisplay Command of CCI using sudo Command

```
$ ssh –2 nasroot@123.456.78.51 sudo pairdisplay -g VG01
```

Figure D.3 Sample pairdisplay Command using SSH Client Shell Script (SSH Protocol Version 2)

Note: Create a shell script on OS that the system administrator uses. For details about creating a shell script, see the appropriate documentation for that OS.

- Special file names given from standard input device to the command in CCI

  When giving special file names from the standard input to the **raidscan**, **inqraid**, or **mkconf.sh** commands, specify **cat /home/nasroot/horc_devfile** instead of **ls/dev/sd***.

Figure D.4 Example of Executing the ‘inqraid’ Command using ‘sudo’ Command

```
$ cat /home/nasroot/horc_devfile | sudo inqraid -CLI
```

Figure D.5 Example of Executing the ‘inqraid’ Command using Shell Script created on SSH client (for SSH protocol version 2)

```
$ ssh -2 nasroot@123.456.78.51 "cat /home/nasroot/horc_devfile | sudo inqraid -CLI"
```

- Notes for scanning on NAS Modular ports using raidscan command

  When scanning on NAS Modular ports using the **raidscan** command, there may be cases that some entries such as **/dev/sda ~ sdi** which stand for NAS system LU (NAS OS LU, NAS Cluster Management LU, Backup of NAS Cluster Management LU, Dump LU, Error Information LU, or Command device) are shown. These devices must not be defined at **HORCM_DEV** section of configuration definition file since these devices cannot be defined as P-VOL or S-VOL for TCE.

D.2.3 Scope Related to NAS Backup Restore Modular

- **Limitations on the functionality of the backup management software**

  If you resynchronize a TCE pair defined by Remote Copy function after performing a backup of the S-VOL by using the backup management software, a full backup occurs next time.

D.2.4 Scope Related to Other Products of Hitachi Network Attached Storage Series

- NAS Sync Image Modular

  When you copy a file system managed by NAS Sync Image Modular using TCE, you must copy both the LUs that constitute the file system and the LUs which constitute the differential-data storage devices. If you copy only the LUs which constitute a file system, you cannot connect them to the NAS Modular system on the remote site.

  The setting of an automatic creation schedule for NAS Sync Image Modular is not copied. The differential-data snapshot that is mounted by the automatic creation schedule is not mounted at the copy destination.

D.3 Requirements

D.3.1 Hardware Requirements

You need a workstation or PC to log in to the NAS Modular system using secure shell (SSH) in addition to the prerequisite hardware for TCE, CCI, and the NAS Modular system described in the following guides:

– *TrueCopy Extended Distance User's Guide (DF700)* (this document)

– *Command Control Interface (CCI) User and Reference Guide (DF700)*

– *NAS Backup Restore Modular User's Guide*

D.3.2 Software Requirements

All of the following AMS1000 storage subsystem program products (PP) must be installed in the NAS Modular, and their licenses must be valid.

– TCE

– NAS Data Control Modular

– NAS File Sharing Modular

– NAS Manager Modular

– NAS Backup Restore Modular

– NAS Sync Image Modular (required when TCE is applied to the file system operated by NAS Sync Image Modular)

D.4 Other Requirements and Restrictions

- **UID for user registration and GID for group registration in Primary/Secondary sites**

  With TCE, the information on the file owner (UID and GID in the primary site), and the quota information and ACL information for each user (UID in the primary site) and group (GID in the primary site) included in the source file system are copied to the secondary site. When sharing the NFS/CIFS, you should note the UID for user registration and the GID for group registration in the primary and secondary sites before setting the quota or the ACL in the target file system. The following paragraphs describe some user name and group name registration scenarios:

  - Register the same user name (or group name) in both sites, and the quota settings and the ACL settings are also the same.

    For local authentication, assign the same UID (or GID) to the same user name (or group name) in both sites. For NIS, share the NIS server in both sites. For domain authentication, share the domain controller.

  - Register the same user name (or group name) in both sites. (The quota settings or the ACL settings are different for each site.)

    Set the quotas or the ACLs again in the secondary site each time the TCE volume pair is split.

  - Register the user (or group) in only one site.

    For local authentication, do not assign the target UID (GID) in another site to ensure that the UID is not overlapped with the UID (or GID) of the another user in the another site.

    For NIS, share the NIS server in both sites. For domain authentication, share the domain controller.

- **Mounting the NFS client for a file system whose data is backed up online**

  If you perform online backup for a file system accessed by the NFS client by remote copy, you must specify NFS, Version 3 before mounting a NAS Modular file system on the NFS client. If you specify the NFS Version 2, you must specify the hard option before mounting a NAS Modular file system on the NFS client.

- **Limitations on TCE operations due to the status of cluster, nodes, and resource groups**

  Connecting the device file or creating and mounting a file system is restricted, when a cluster is not configured, the cluster is stopped, the nodes are stopped, or when the resource groups are offline. Due to these restrictions, the operations performed by the target TrueCopy operation will also end in error. You should not therefore operate the cluster, nodes, and resource groups during the TCE operation. Problems arising within clusters, nodes or resource groups should be fixed immediately. The following is a list of prohibited operations when TCE operations are being performed:

  - Un-mount and mount the source file system during the splitting of a TCE pair volume.

– Connect the target file system to the NAS unit.

– Un-mount and delete the target file system before the TCE pair is resynchronized.

- **Changing system configuration during the TCE operations**

  When changing the following system configurations during TCE operations, you must change the CCI configuration definition files in NNC1 and NNC3 NAS units to which the TCE P-VOL is connected, and in NNC1 and NNC3 NAS units where the S-VOL is used:

  – Changing a fixed IP address

  – Expanding or deleting a source file system

  – Setting up, expanding, or releasing a differential-data storage device for NAS Sync Image Modular

  When the host name is specified in the CCI configuration definition file and you change the following system configurations, you must change the configuration definition file:

  – Editing the /etc/hosts file (when resolving the host name using the /etc/hosts file)

  – Changing registration information on NIS server, or changing a setting for NIS server (when resolving the host name using NIS)

  – Changing registration information on DNS server, or changing a setting for DNS server (when resolving the host name using DNS)

  – Changing the CHA (host name)

- **Splitting a TrueCopy volume pair using the online backup method**

  When it takes a time from the execution of the **horc_pvol_freeze** command until the execution of the **horc_pvol_unfreeze** command, a time-out of the access may occur depending on a client. Besides, in the case where NAS Sync Image Modular is operated in the copy source file system, the time-out tends to occur easily because the execution of the **horc_pvol_freeze** command may take a long time.

- **Restrictions on TCE at the time of the NAS OS boot**

  When TCE is used in the NAS system, you should not execute pair creation or resynchronization during the operations such as installation, boot, shutdown, node stop, and cluster stop of the NAS OS. Conversely, you should not perform operations such as installation, boot, shutdown, node stop, and cluster stop of the NAS OS during TCE pair creation or resynchronization.

  If the creation or resynchronization of a TCE pair and NAS OS operations (such as installation, boot, shutdown, node stop, and cluster stop) are performed at the same time, a failure may occur because the NAS OS is overloaded.

D.5 Preparing for TCE Operations

D.5.1 Preparing for TCE Volume Pair Operation

Preparation procedures for creating TCE volume pairs are described in Chapter 3.

D.5.2 Registration of Public Key Used for SSH

Before issuing the commands described in this document, the public key used for SSH needs to be registered in the NAS unit to which the TCE P-VOL is connected and also in the NAS unit which will use the S-VOL. For more information on registering the public key, see the *Hitachi TagmaStore® Adaptable Modular Storage and Workgroup Modular Storage NAS Manager Modular User's Guide* (MK-95DF757).

D.5.3 Configuring the CCI Environment

- Logging in to the NAS unit using SSH

  You can log into both the NAS unit in which the TCE P-VOL is connected and the NAS unit which will use the S-VOL, using the **nasroot** account via SSH. (For information about logging in to the NAS unit, see the appropriate documentation for each tool in the platform.)

- Configuring the CCI configuration definition files

  To control a TCE pair using CCI, you should first define the TCE pair using the CCI configuration definition file. The installation of NAS Backup Restore Modular creates the templates of the CCI configuration definition file. Only the **HORCM_MON** section and the **HORCM_CMD** section are described in this template.

  – The template of the CCI configuration definition file for instance number 16 is:

    `/home/nasroot/horcm16.conf`

  – The template of the CCI configuration definition file for instance number 17 is:

    `/home/nasroot/horcm17.conf`

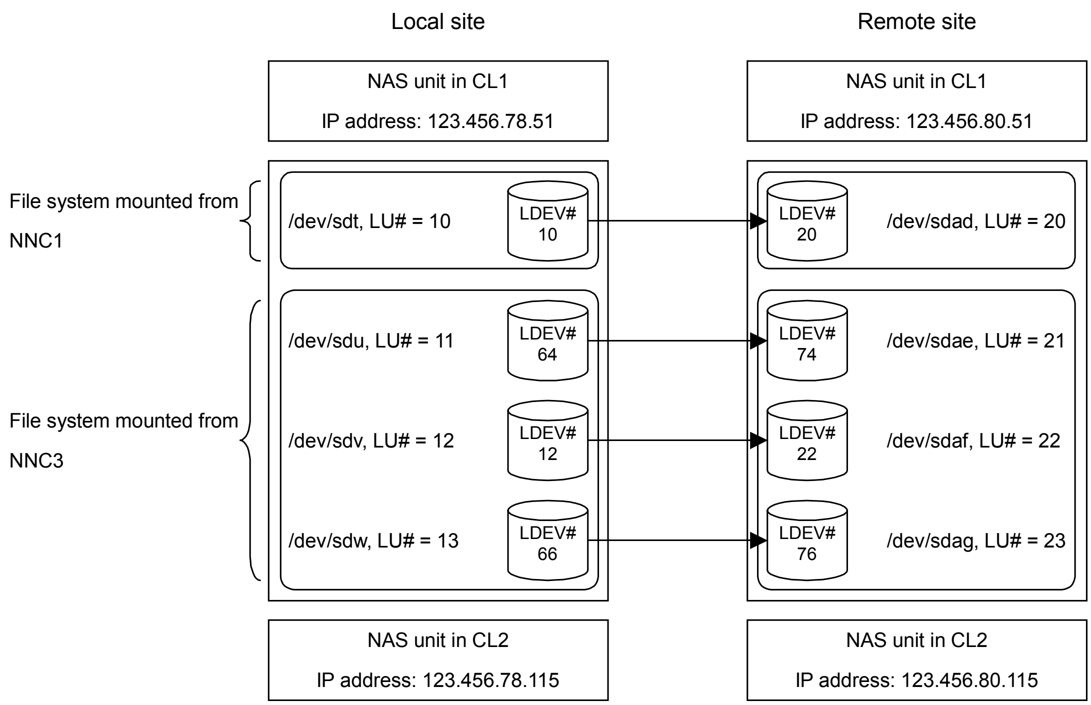
    To complete the CCI configuration definition file:

    1. Add the **HORCM_DEV** section and the **HORCM_INST** section to a template for the CCI configuration definition file using the **mkconf.sh** command in CCI.

    2. Edit the file and create the CCI configuration definition file.

These operations should be performed in each NAS unit in NNC1 and NNC3 where the TCE P-VOL is connected and in each NAS unit in NNC1 and NNC3 that will use the S-VOL. You should prepare a total of four CCI configuration files assuming one instance per NAS unit. For two instances, you need a total of eight configuration files.

In NAS Modular, you can define one or two CCI instances per NAS unit. To operate only the TCE pairs in CCI, you need one instance. The following section describes how to create a CCI configuration definition file using the NAS Modular system with LUs as an example (see Figure D.6).



Note: To enable 2 NAS units to communicate over the LAN, you must define the network configuration. Use fixed IP addresses of all eth1 of each unit. Do not mix eth1 addresses.

Figure D.6 Example of Configuration of Pair LU

- Adding **HORCM_DEV** section and the H**ORCM_INST** section using the **mkconf.sh** command of CCI.

  Add **HORCM_DEV** section and the **HORCM_INST** section to the template of the CCI configuration definition file using the **mkconf.sh** command in CCI.

```
$ cat /home/nasroot/horc_devfile | sudo mkconf.sh -gg <device group name> -i 16 -a
```

Figure D.7 Adding HORCM_DEV, HORCM_INST Sections (for instance number 16)

Note: Be sure to specify the **-gg** option for the **mkconf.sh** command. Through the specification of the **-gg** option, a pair is defined with an LU number allocated to the host group. When the **-gg** option is not specified, the copying is made to an LU other than the required one because a pair cannot be defined with the LU number allocated to the host group. When the **-gg** option is specified for the **mkconf.sh** command, the port is set as one of the CL1-C-1 and CL2-C-1. When the **-gg** option is not specified, -1 is omitted as CL1-C or CL2-C. Therefore, do not perform the pair operation and make the configuration definition file over again specifying the **-gg** option.

Note: If the **mkconf.sh** command is executed without specifying the -a option, the **HORCM_MON** section and the **HORCM_CMD** section in the CCI configuration definition file automatically created when NAS Backup Restore Modular is installed will be overwritten. In this case, delete all the lines above the **HORCM_MON** section, and change the entry service in the **HORCM_MON** section to an appropriate value based on Table D.2.

To execute the **mkconf.sh** command, the LU that constitutes the file system to be paired must be created. The LU configuration (number and size) in the file system to be paired needs to be to be exactly the same in both which the TCE P-VOL is connected and which the TCE S-VOL is connected.

To create the correct CCI configuration definition file, it is recommended that you create a file system to be temporarily paired before executing the **mkconf.sh** ***command. After creating the CCI configuration definition file using*** the **mkconf.sh** ***command,*** you can continue using the file system, which the TCE P-VOL is connected. Alternatively, you may also delete it and create a file system with the same configuration using the same LU according to the procedure described in section D.5.4. However, you must delete the temporarily created file system, which the TCE P-VOL is connected before creating a file system pair.

The sample in the following configuration explains how to create a CCI configuration definition file in the NAS unit in NNC1 to which the TCE P-VOL is connected. Create CCI configuration definition files for the other NAS units using the same procedure.

```
$ cat /home/nasroot/horc_devfile | sudo mkconf.sh -gg VG -i 16 -a
starting HORCM inst 16
HORCM inst 16 starts successfully.
HORCM Shutdown inst 16 !!!
A CONFIG file was successfully completed.
starting HORCM inst 16
HORCM inst 16 starts successfully.
DEVICE_FILE             Group    PairVol     PORT   TARG  LUN M   SERIAL  LDEV
/dev/sdj               VG       VG_000      CL1-C-1   0   17 -  77062486    70
/dev/sdk               VG       VG_001      CL1-C-1   1   18 -  77062486    18
        :
        :
/dev/sdt                VG       VG_010      CL1-C-1   0   10 -  77062486    10
/dev/sdu                VG       VG_011      CL1-C-1   0   11 -  77062486    64
/dev/sdv                VG       VG_012      CL1-C-1   0   12 -  77062486    12
/dev/sdw                VG       VG_013      CL1-C-1   0   13 -  77062486    66
/dev/sdx                VG       VG_014      CL1-C-1   0   14 -  77062486    14
/dev/sdy                VG       VG_015      CL1-C-1   0   15 -  77062486    68
/dev/sdz                VG       VG_016      CL1-C-1   0   16 -  77062486    16
HORCM Shutdown inst 16 !!!
Please check '/home/nasroot/horcm16.conf','/home/nasroot/log16/curlog/horcm_*.log',
and
modify ‘ip
_address & service'.
#
```

Figure D.8 Example of mkconf.sh Execution Result (for instance number 16)

- Editing the template for the CCI configuration definition file

  Table D.2 lists the values that are specified for the items included in the CCI configuration definition file in the NAS Modular system.

Table D.2 Configuration Definition File Settings (HORCM_MON, HORCM_CMD) and Specified Values in NAS Modular System

Section Name	Item	Specified Values in NAS Modular System
HORCM_MON	ip_address	Unique IP address for eth1 in the local NAS unit
	service	20331 (Instance number 16) or 20332 (Instance number 17)
HORCM_CMD	dev_name	/dev/sdf

Note: The host name can be specified instead of the fixed IP address if the fixed IP address and the corresponding host name are registered into **/etc/hosts**, an NIS server, or a DNS server. Refer to the *NAS Manager Modular User’s Guide* for information on how to register the fixed IP address and the corresponding host name into **/etc/hosts**, and for information on how the NAS Modular system can search the host name using NIS or DNS.

Based on Table D.2, change each entry of ip_address in **HORCM_MON** to an appropriate value. For the entry service in **HORCM_MON** and the entry in **HORCM_CMD**, appropriate values are specified when NAS Backup Restore Modular is installed.

If you have created the configuration definition file using the **mkconf** command tool, remember to change the value of the poll (10ms) manually. You must always set a value more than or equal to 6000.

```
HORCM_MON
#ip_address        service         poll(10ms)     timeout(10ms)
123.456.78.51      20331                 6000              3000

HORCM_CMD
#dev_name               dev_name                dev_name
#UnitID 0 (Serial# 77062486)
/dev/sdf

HORCM_DEV
#dev_group         dev_name         port#     TargetID     LU#     MU#
# /dev/sdj            SER =  77062486  LDEV =   70 [ FIBRE FCTBL = 3 ]
VG                 VG_000          CL1-C-1          0      0
# /dev/sdk            SER =  77062486  LDEV =   18 [ FIBRE FCTBL = 3 ]
VG                 VG_001          CL1-C-1          0      1
        :
        :
# /dev/sdw             SER =  77062486  LDEV =   66 [ FIBRE FCTBL = 3 ]
VG                 VG_013          CL1-C-1          0      13
# /dev/sdx             SER =  77062486  LDEV =   14 [ FIBRE FCTBL = 3 ]
VG                 VG_014          CL1-C-1          0      14
# /dev/sdy             SER =  77062486  LDEV =   68 [ FIBRE FCTBL = 3 ]
VG                 VG_015          CL1-C-1          0      15
# /dev/sdz             SER =  77062486  LDEV =   16 [ FIBRE FCTBL = 3 ]
VG                 VG_016          CL1-C-1          0      16

HORCM_INST
#dev_group        ip_address      service
VG                127.0.0.1       20332
```

Figure D.9 Example of CCI Configuration Definition File - 1 (for instance number 16)

Next, at **HORCM_DEV** section, remove all unnecessary LU entry (line) other than those to be managed by CCI.

LU which makes up a file system and its LDEV ID can be obtained by using the following command.

```
$ sudo horc_devlist | grep ':filesystem name'
```

For Figure D.10, LU 11, 12, 13 shown in left make up a file system **sample**, and their LDEV ID are 64, 12, 66 shown on the third column to the left.

```
$ sudo horc_devlist | grep ':sample$'
  11(  17)  64 FC   3.906GB   -- 0  Normal Own   File:sample
  12(  18)  12 FC   3.906GB   -- 0  Normal Own   File:sample
  13(  19)  66 FC   3.906GB   -- 0  Normal Own   File:sample
```

Figure D.10 Example of an Explorer Command of LU which Makes up a File System Sample

Next, modify the device file name and the device name, which are to be handled by CCI. There are some restrictions on modifying the device file name or device name.

– Volumes on both P-VOL side and S-VOL side must be assigned the same device group name and same device name.

– Volumes making up one file system must be assigned the same device group name.

– When NAS Sync Image Modular is utilized, volumes for file system and differential-data storage device must be assigned the same device group name.

```
HORCM_MON
#ip_address        service         poll(10ms)     timeout(10ms)
123.456.78.51      20331                6000              3000

HORCM_CMD
#dev_name               dev_name                dev_name
#UnitID 0 (Serial# 77062486)
/dev/sdf

HORCM_DEV
#dev_group         dev_name        port#     TargetID     LU#     MU#
# /dev/sdt             SER =  77062486  LDEV =   10 [ FIBRE FCTBL = 3 ]
VG                 VG_010          CL1-C-1         0      10
# /dev/sdu             SER =  77062486  LDEV =   64 [ FIBRE FCTBL = 3 ]
VG                 VG_011          CL1-C-1         0      11
# /dev/sdv             SER =  77062486  LDEV =   12 [ FIBRE FCTBL = 3 ]
VG                 VG_012          CL1-C-1         0      12
# /dev/sdw             SER =  77062486  LDEV =   66 [ FIBRE FCTBL = 3 ]
VG                 VG_013          CL1-C-1         0      13

HORCM_INST
#dev_group        ip_address      service
VG                127.0.0.1       20332
```

Figure D.11 Example of CCI Configuration Definition File - 2 (for instance number 16)

Next, specify the IP address of the instance to be paired in the secondary site for the **HORCM_INST** section (see Table D.3 and Figure D.12). Specify the IP addresses of the instances in both NNC1 and NNC3 for a failover between NAS unit in the secondary site.

Table D.3 Configuration Definition File Settings (HORCM_INST) and Specified Values in NAS Modular System

Section	Item	Values for NAS Modular System
HORCM_INST	ip_address	Fixed IP address of eth1 or eth2 in the NAS unit in the secondary site of TCE (But if you specify eth1 in HORCM_MON ip_address, specify eth1. If you specify eth2 in HORCM_MON ip_address, specify eth2).
	service	20331 (for instance number 16) or 20332 (for instance number 17)

Note: The host name can be specified instead of the fixed IP address if the fixed IP address and the corresponding host name are registered into **/etc/hosts**, an NIS server, or a DNS server. Refer to the *NAS Manager Modular User's Guide* for information on how to register the fixed IP address and the corresponding host name into **/etc/hosts**, and for information on how the NAS Modular system can search the host name using NIS or DNS.

```
HORCM_MON
#ip_address        service         poll(10ms)     timeout(10ms)
123.456.78.51      20331                 6000              3000

HORCM_CMD
#dev_name               dev_name                 dev_name
#UnitID 0 (Serial# 77062486)
/dev/sdf

HORCM_DEV
#dev_group         dev_name        port#     TargetID     LU#     MU#
# /dev/sdt             SER =  77062486  LDEV =   10 [ FIBRE FCTBL = 3 ]
VG                 VG_010          CL1-C-1          0      10
# /dev/sdu             SER =  77062486  LDEV =   64 [ FIBRE FCTBL = 3 ]
VG                 VG_011          CL1-C-1          0      11
# /dev/sdv             SER =  77062486  LDEV =   12 [ FIBRE FCTBL = 3 ]
VG                 VG_012          CL1-C-1          0      12
# /dev/sdw             SER =  77062486  LDEV =   66 [ FIBRE FCTBL = 3 ]
VG                 VG_013          CL1-C-1          0      13

HORCM_INST
#dev_group        ip_address      service
VG                123.456.80.51    20332
VG                123.456.80.115   20332
```

Figure D.12 Example of CCI Configuration Definition File - 3 (for instance number 16)

- Checking the contents of the CCI configuration definition file

  By combining the following commands, you can check whether an appropriate LU is specified in the **HORCM_DEV** section in the CCI configuration definition file.

  First, start CCI in both the NAS unit in which the TCE P-VOL is connected and in the NAS unit which will use the S-VOL (see Figure D.13).

```
$ sudo horc_setenv HORCMINST 16      (For instance number 16)
  or sudo horc_setenv HORCMINST 17   (For instance number 17)
$ sudo horc_unsetenv HORCC_MRCF
When you logged in to the target NAS unit using SSH and performed the above setup,
confirm the setup by once logging out of the NAS unit and relogging in.
$ sudo horcmstart.sh
```

Figure D.13 Procedure for Starting CCI

By issuing the **pairdisplay** command in the NAS unit in which TCE P-VOL is connected or in which S-VOL is used, you can see the LDEV numbers of LUs specified in the **HORCM_DEV** section (see Figure D.14).

```
$ sudo pairdisplay –g device group name
```

Figure D.14 How to Check LDEV Number of LUs Specified in the HORCM_DEV Section

You can check the device file numbers, and the LDEV numbers for the device files that constitute a file system, by issuing the **horc_devlist** command in the NAS unit in which P-VOL is connected (see Figure D.15). Compare this with Figure D.14. For information about the **horc_devlist** command, see the *NAS Backup Restore Modular User's Guide*.

```
$ sudo horc_devlist | grep ':sample$'
  0A(  10)   120 FC   20.000GB     -- 0  Normal Own   File:sample
  0B(  11)   121 FC   20.000GB     -- 0  Normal Own   File:sample
  0C(  12)   122 FC   20.000GB     -- 0  Normal Own   File:sample
```

Figure D.15 How to Check Device File Numbers and LDEV Numbers of P-VOL

You can also check the device file numbers and the LDEV numbers for the unused device files which can be S-VOL by issuing the **horc_devlist** command in the NAS unit in which S-VOL is used. Compare this with Figure D.14.

```
$ sudo horc_devlist | grep ' Free$'
  01(   1)    21 FC   20.000GB     -- 0  Normal --    Free
  02(   2)    22 FC   20.000GB     -- 0  Normal --    Free
  03(   3)    23 FC   20.000GB     -- 0  Normal --    Free
```

Figure D.16 How to Check Device File Numbers and LDEV Numbers for Unused Device Files Which Can be S-VOLs

Table D.4 lists the port names described in the **HORCM_DEV** section in the CCI configuration definition file and the locations of the NAS units.

Table D.4 Locations of NAS Unit and Port Names Specified in CCI Configuration Definition File

Location of NAS Unit	Port Name
NNC1	CL1-C-1
NNC3	CL2-C-1

After checking the CCI configuration definition file, stop CCI in both the NAS unit in which TCE P-VOL is connected and in the NAS unit in which S-VOL is used (see Figure D.17).

```
$ sudo horcmshutdown.sh
```

Figure D.17 Stopping CCI (for instance 16)

Save the NAS OS LU manually to save the configured CCI configuration definition files. For more information, see the *NAS Manager Modular User's Guide*.

- Setting the CCI user environmental variable

  In the following procedure, the environment variables **HORCMINST** and **HORCC_MRCF** are corrected corresponding to a system configuration. This setup is performed on the four nodes (CL1 and CL2 of the cluster to which TCE P-VOL is connected, and CL1 and CL2 of the cluster in which S-VOL is used).

1. At both sites, set up the environment variable of the CCI instance:

   –In case of instance number 16:

   ```
   $ sudo horc_setenv HORCMINST 16
   ```

   OR

   ```
   $ export HORCMINST=16
   ```

   –In case of instance number 17:

```
$ sudo horc_setenv HORCMINST 17
```

OR

```
$ export HORCMINST=17
```

2. At both sites, set up the HOMRCF command environment variable of the CCI as TCE:

```
$ sudo horc_unsetenv HORCC_MRCF
```

OR

```
$ export HORCC_MRCF
```

3. If procedure 1 or 2 are done through SSH with the **sudo horc_unsetenv**, log out from the NAS unit and log in to the NAS unit again to validate these settings. If there are done through SSH with the **export**, these settings validate immediately.

```
$ exit
ssh {-1 | -2} nasroot@ <fixed IP address of NAS unit>
```

Enter the command shown in the following example to check the result of setting up the environment variable:

```
$ sudo horc_printenv
```

The value of the environmental variable immediately after the installation of NAS Backup Restore Modular is shown below.

Table D.5 Environment Variable Value Immediately after Installation of NAS Backup Restore Modular

Environment Variable	Value
HORCMINST	16
HORCC_MRCF	It is not set.

D.5.4 Creating File System in TCE P-VOL

Create a file system in TCE P-VOL using the Create New File System window in File System Management in NAS Manager Modular or by using the **enas_fscreate** command. Even if you create and split the TCE volume pair without creating a file system in the TCE P-VOL, you cannot access the TCE S-VOL in the secondary site.

D.6 Overview of TCE Operations

This section describes the overview of TCE operations, CCI commands, and the commands provided by the Hitachi Network Attached Storage (NAS) series products, for using TCE for remote copy operation.

We describe only the arguments of the CCI commands, which are required for the basic TCE operations. For other arguments, please refer to *Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide* (MK-95DF701).

For the specifications of the commands provided by the Hitachi NAS series products and notes, see section D.8.

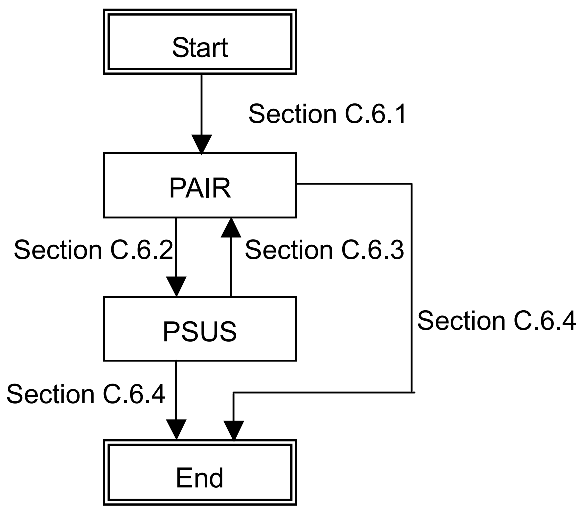


Figure D.18 Overview of Remote Copy Operations and Sections

D.6.1 Starting TCE Operations and Creating a TCE Pair

If the TCE S-VOL contains a file system, you must delete the file system using the **enas_fsdelete** command in NAS Manager Modular before starting TCE operations. If the TCE S-VOL is used for the Snapshot feature in NAS Backup Restore Modular, you must delete the Snapshot definitions using the **snapvoldel** command before starting the TCE operations. If the differential-data storage device of NAS Sync Image Modular is in the TCE S-VOL, you must release it using the **syncstop** command before starting the TCE operations.

To start TCE operations and create a TCE volume pair:

1. At the main site and the remote site, start CCI:

```
$ sudo horcmstart.sh       (1-instance configuration)
```

OR

```
$ sudo horcmstart.sh 16 17   (2-instance configuration)
```

2. At the remote site, reserve device file used as a target file system.

   –Non-LVM

```
$ sudo horc_svol_define -d  device file number
```

   –LVM

```
$ sudo horc_svol_vmdefine -d  device file number, ...
```

3. At the main site, create TCE volume pair:

```
$ sudo paircreate {-g group name | -d volume name} -f async -vl
```

4. At the main site, check the completion of TCE volume pair creation:

```
$ sudo pairvolchk {-g group-name | -d volume name}
```

   You can also use the **pairevtwait** command, which waits for the volumes to be paired.

```
pairvolchk: Volstat is P-VOL.[status = COPY] → Creating
pairvolchk: Volstat is P-VOL.[status = PAIR] → Created
```

D.6.2 Splitting a TCE Volume Pair

When splitting a TCE volume pair, you may choose either offline backup or online backup. With offline backup, you can split the pair by un-mounting the P-VOL. With online backup, you may split the pair by temporarily saving the updates to a file system without un-mounting the P-VOL.

During offline backup, a TCE volume pair is split after completely stopping access from clients by deleting CIFS/NFS shares. Since an I/O error is reported to the application if CIFS/NFS shares are deleted while the application writes data in P-VOL, or if the application tries to write data in P-VOL after deleting CIFS/NFS shares, it can be distinguished by the application data reflected in the TCE volume pair. It is therefore applicable to most applications.

During online backup, however, a TCE volume pair is split without deleting CIFS/NFS shares. Since an I/O error is not reported to the application when writing data to P-VOL, it cannot be distinguished of which time data is reflected in S-VOL. For this reason, it is applicable only to the application that can identify where data was updated with a journal file etc.

- Offline backup

  To split the TCE volume pair with offline backup:

1. At the main site, stop the program that accesses P-VOL.
2. At the main site, delete NFS/CIFS shares in the P-VOL using the **enas_nfsdelete** command and the **enas_cifsdelete** command in NAS Manager Modular. Un-mount the
   P-VOL using the **enas_fsumount** command.
3. At the main site, prevent NAS Sync Image Modular from performing operations on the
   P-VOL:

```
$ sudo horc_pvol_freeze –f source file system name
```

4. At the main site, which the P-VOL is connected, split the TCE volume pair.

```
$ sudo pairsplit {-g group-name | -d volume name} -rw
```

5. At the main site, verify that the TCE volume pair is split:

```
$ sudo pairvolchk {-g group name | -d volume name }
```

You can also use the **pairevtwait** command, which waits for the pair to be split (PSUS).

```
$ sudo pairvolchk : Volstat is P-VOL.[status = COPY] → Splitting
$ sudo pairvolchk : Volstat is P-VOL.[status = PSUS] → Split
```

6. At the main site, enable the operations from NAS Sync Image Modular on the P-VOL:

```
$ sudo horc_pvol_unfreeze -f source file system name
```

7. At the main site, mount P-VOL using the **enas_fsmount** command in NAS Manager Modular, and create CIFS/NFS shares using the **enas_nfscreate** command and the **enas_cifscreate** command.
8. At the main site, restart the program that accesses the P-VOL.
9. At the remote site, connect the target file system to the NAS unit.

–Non-LVM

```
$ sudo horc_svol_import -f target file system name -d device file number
```

–LVM

```
$ sudo horc_svol_vmimport -f target file system name -d device file number, ...
```

10. At the remote site, mount S-VOL using the **enas_fsmount** command. Create a shared-directory using the **enas_dircreate** command, or change the attributes of a shared-directory using the **enas_diredit** command. Create CIFS/NFS shares using the **enas_nfscreate** command and the **enas_cifscreate** command.

Note: When the S-VOL is the file system managed by NAS Sync Image Modular, mount the differential-data snapshots using the **syncmount** command, and create the CIFS/NFS shares using the **enas_nfscreate** command and the **enas_cifscreate** command, if necessary.

11. At the remote site, start the program that accesses the S-VOL.

- Online backup

To split the TCE volume pair with online backup:

1. At the main site, prevent NAS Sync Image Modular from performing operations on P-VOL, hold the write request to P-VOL, and write the un-reflected data to the disk:

```
$ sudo horc_pvol_freeze -f source file system name
```

2. At the main site, split the TCE volume pair:

```
$ sudo pairsplit {-g group name | -d volume name } -rw
```

3. At the main site, verify that the TCE volume pair has been split:

```
$ sudo pairvolchk {-g group name | -d volume name }
```

You can also use the **pairevtwait** command, which waits for the pair to be split (PSUS).

```
$ sudo pairvolchk : Volstat is P-VOL.[status = COPY] → Splitting
$ sudo pairvolchk : Volstat is P-VOL.[status = PSUS] → Split
```

4. At the main site, permit the write request to P-VOL and enable the operations from NAS Sync Image Modular on P-VOL:

```
$ sudo horc_pvol_unfreeze -f source file system name
```

5. At the remote site, connect the target file system to the NAS unit.

   –Non-LVM

```
$ sudo horc_svol_import -f target file system name -d device file number
```

   –LVM

```
$ sudo horc_svol_vmimport -f target file system name -d device file number, ...
```

6. At the remote site, mount the S-VOL using the **enas_fsmount** command.
7. Create a shared-directory using the **enas_dircreate** command or change the attributes of a shared-directory using the **enas_diredit** command.
8. Create CIFS/NFS shares using the **enas_nfscreate** command and the **enas_cifscreate** command.

   Note: When the S-VOL is the file system managed by NAS Sync Image Modular, mount the differential-data snapshots using the **syncmount** command, and create the CIFS/NFS shares using the **enas_nfscreate** command and the **enas_cifscreate** command, if necessary.
9. At the remote site, start the program that accesses the S-VOL.

D.6.3 Re-synchronizing a TCE Volume Pair

You must delete the target file system using the **enas_fsdelete** command before re-synchronizing the volume pair. Deleting the target file system, however, clears only the management area of the file system. Re-synchronizing the volume pair is not the same as performing initial copy.

Note: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

To resynchronize the TCE volume pair:

1. At the remote site, delete the CIFS/NFS shares using the **enas_cifsdelete** command and the **enas_nfsdelete** command, and un-mount all file systems in the pair using the **enas_fsumount** command.

   Note: In the NAS Sync Image Modular target file system, delete the CIFS/NFS shares for the differential-data snapshot using the **enas_cifsdelete** command and the **enas_nfsdelete** command, un-mount the differential-data snapshot using the **syncumount** command, and release the differential-data storage device using the **syncstop** command.

2. At the remote site, delete the file system in the S-VOL using the **enas_fsdelete** command.

3. At the remote site, reserve device file to be used as a target file system.

   –Non-LVM

   ```
   $ sudo horc_svol_define -d device file number
   ```

   –LVM

   ```
   $ sudo horc_svol_vmdefine -d device file number, ...
   ```

4. At the main site, resume the TCE volume pair:

   ```
   $ sudo pairresync {-g group name | -d volume name }
   ```

5. At the main site, verify that the TCE volume pair has been resumed:

   ```
   $ sudo pairvolchk {-g group name | -d volume name }
   ```

   You can also use the **pairevtwait** command, which waits for the volumes to be paired.

   ```
   $ pairvolchk : Volstat is P-VOL.[status = COPY] → Resuming
   $ pairvolchk : Volstat is P-VOL.[status = PAIR] → Resumed
   ```

D.6.4 Deleting a TCE Volume Pair

Completing TCE operations by deleting the TCE volume pair in the PSUS state differs depending upon whether you continue using the file system in S-VOL, or if you destroy the file system.

When deleting the TCE volume pair in a state other than PSUS, you cannot use that file system because the consistency of data in the S-VOL is not guaranteed.

To delete the TCE volume pair in PSUS state, and to continue using the S-VOL file system:

1. At the main site, delete the TCE volume pair:

   ```
   $ sudo pairsplit {-g group name | -d volume name} -S
   ```

2. At the main site, verify that the TCE volume pair has been deleted:

   ```
   $ sudo pairvolchk {-g group name | -d volume name}
   ```

   You can also use the **pairevtwait** command, which waits for the pair to be deleted.

   ```
   $ pairvolchk : Volstat is P-VOL.[status = COPY] → Deleting
   $ pairvolchk : Volstat is P-VOL.[status = SMPL] → Deleted
   ```

3. At the main site and the remote site, stop CCI:

   ```
   $ sudo horcmshutdown.sh                (1-instance configuration)
   ```

   OR

   ```
   $ sudo sudo horcmshutdown.sh  16  17     (2-instance configuration)
   ```

Note: If you start the split pair operation with offline backup, but have not yet connected the target file system to the NAS unit at the remote site, finish connecting the target file system to the NAS unit, and then start the previously described operations. If you start the split pair operation with online backup, but do not connect the target file system to the NAS unit at the remote site, finish connecting the target file system to the NAS unit, and then start the above operations (see section D.6.2).

To delete the TCE volume pair in PSUS state without using the S-VOL file system:

1. At the remote site, terminate the program that accesses the S-VOL.
2. At the remote site, delete CIFS/NFS shares using the **enas_nfsdelete** command and the **enas_cifsdelete** command, and un-mount S-VOL using the **enas_fsumount** command.

   Note: When the S-VOL is the file system managed by NAS Sync Image Modular, delete the CIFS/NFS shares under the differential-data snapshots using the **enas_nfsdelete** command and **enas_cifsdelete** command, un-mount the differential-data snapshots using the **syncumount** command, and release the differential-data storage devices using the **syncstop** command.
3. At the remote site, delete the S-VOL file system using the **enas_fsdelete** command.
4. At the main site, delete the TCE volume pair:
   ```
   $ sudo pairsplit {-g group name | -d volume name} -S
   ```
5. At the main site, verify that the TCE volume pair has been deleted:
   ```
   $ sudo pairvolchk {-g group name | -d volume name}
   ```
   You can also use the **pairevtwait** command, which waits for the pair to be deleted.
   ```
   $ pairvolchk : Volstat is P-VOL.[status = COPY] → Deleting
   $ pairvolchk : Volstat is P-VOL.[status = SMPL] → Deleted
   ```
6. At the main site and the remote site, stop CCI:
   ```
   $ sudo horcmshutdown.sh                (1-instance configuration)
   ```
   OR
   ```
   $ sudo sudo horcmshutdown.sh  16  17     (2-instance configuration)
   ```

Note: If you start the split pair operation with offline backup, but have not yet connected the target file system to the NAS unit at the remote site, finish connecting the target file system to the NAS unit, and then start the previously described operations. If you start the split pair operation with online backup, but do not connect the target file system to the NAS unit at the remote site, finish connecting the target file system to the NAS unit, and then start the above operations (see section D.6.2).

To delete the TCE volume pair in a state other than PSUS state:

1. At the main site, delete the TCE volume pair:
   ```
   $ sudo pairsplit {-g group name | -d volume name} -S
   ```
2. At the main site, verify that the TCE volume pair has been deleted:

```
$ sudo pairvolchk {-g group name | -d volume name}
```

You can also use the **pairevtwait** command, which waits for the pair to be deleted.

```
$ pairvolchk : Volstat is P-VOL.[status = COPY] → Deleting
$ pairvolchk : Volstat is P-VOL.[status = SMPL] → Deleted
```

3. At the remote site, release the device file used in the target file system.

–Non-LVM

```
$ sudo horc_svol_delete -d device file number
```

–LVM

```
$ sudo horc_svol_vmdelete -d device file number, ...
```

4. At the main site and the remote site, stop CCI:

```
$ sudo horcmshutdown.sh                  (1-instance configuration)
```

OR

```
$ sudo sudo horcmshutdown.sh  16  17     (2-instance configuration)
```

D.7 CCI Log Files Operations

This section explains the format of the CCI log files in the NAS Modular system, and the method of downloading log files to the SSH client, and provides notes on the operations of the CCI log files.

D.7.1 CCI Log Files Format

CCI log file consists of start-up logs, error log, trance and core files and is stored in the format shown in Table D.6. "*" in Table D.6 denotes the CCI instance number (16 or 17), Host denotes the name of the node name of the target NAS unit, and **PID** denotes the process ID of the CCI command, **CMD** shows the process name (horcmgr in case of CCI, or a command name in case of the CCI command), and **TIME** shows the creation time of the core file.

Table D.6 CCI Log Files Format in the NAS Modular System

Log File Classification	Log File Name	Log Files Format
CCI logs under operation	CCI startup log	/home/nasroot/log*/curlog/horcm_HOST.log
	Command log	/home/nasroot/log*/horcc_HOST.log
	CCI error log	/home/nasroot/log*/curlog/horcmlog_HOST/horcm.log
	CCI trace log	/home/nasroot/log*/curlog/horcmlog_HOST/horcm_PID.trc
	Command trace	/home/nasroot/log*/curlog/horcmlog_HOST/horccc_PID.trc
	Core file	/var/core/core-PID-CMD-TIME
CCI logs saved automatically	CCI startup log	/home/nasroot/log*/tmplog/horcm_HOST.log
	Command log	/home/nasroot/log*/horcc_HOST.log
	CCI error log	/home/nasroot/log*/tmplog/horcmlog_HOST/horcm.log
	CCI trace log	/home/nasroot/log*/tmplog/horcmlog_HOST/horcm_PID.trc
	Command trace	/home/nasroot/log*/tmplog/horcmlog_HOST/horccc_PID.trc
	Core file	/var/core/core-PID-CMD-TIME

D.7.2 Downloading Log Files to the SSH Client

You can download or delete core files output by CCI on the List of RAS Information (List of core files) window provided in NAS Manager Modular. Besides, when you examine the detailed factor of the abnormal termination of the CCI command through the log file (excluding the core file) of the CCI or want to save the log file before it is deleted, you can download the log using the List of RAS Information (Batch-download) window of NAS Manager Modular specifying the Backup log groups. For details, please refer to *NAS Manager Modular User's Guide*.

Note: To view the CCI log files using Windows®, you must ensure that the text viewer you are using can display the text containing the line feed code LF (Line Feed).

D.7.3 Checking Log File Capacity

Since the CCI log files are output to NAS OS LU, if TCE operations still show an error, log files will become large and the available space on NAS OS LU will decrease. You should periodically check the capacity of the log files (except the trace files) to which CCI restricts capacity, stop CCI, and delete log files using the **horc_logremove** command before exceeding 1 MB. Check the capacity of log files using the following command.

```
$ ls -l -R /home/nasroot/log*   (* means an instance number 16 or 17)
```

D.8 Commands

This section describes the commands that are used in TCE operations in the NAS Modular system and that are provided by Hitachi Network Attached Storage series products. For details of these commands, please refer to *NAS Backup Restore Modular User's Guide*, *NAS Manager Modular User's Guide*, and *NAS Sync Image Modular User's Guide*.

- Commands that NAS Backup Restore Modular provides:
  - horc_devlist
  - horc_logremove
  - horc_printenv, horc_setenv, horc_unsetenv
  - horc_pvol_freeze
  - horc_pvol_unfreeze
  - horc_svol_define, horc_svol_vmdefine
  - horc_svol_delete, horc_svol_vmdelete
  - horc_svol_import, horc_svol_vmimport
- Commands that NAS Manager Modular provides:
  - enas_cifscreate, enas_cifsdelete
  - enas_dircreate, enas_diredit
  - enas_fscreate, enas_fsdelete, enas_fslist, enas_fsmount, enas_fsumount
  - enas_nfscreate, enas_nfsdelete
- Commands that NAS Sync Image Modular provides:
  - syncdel, synclist, syncmount, syncstop, syncumount

D.9 Messages

For further information about the messages output by the commands described in section D.8, please refer to *NAS Backup Restore Modular User's Guide*, *NAS Manager Modular User's Guide*, and *NAS Sync Image Modular User's Guide*.

Appendix E Command Operations on NAS Modular

This appendix describes command operations on NAS Modular. The following sections describe commands and procedures for creating pairs, splitting pairs, resynchronizing pairs and deleting pairs.

During the fail-over of the NAS unit, the commands shown in Table E.1 cannot be used. When the fail-over occurs because of a failure caused by a controller blockade or an NAS unit failure, perform the pair operation from the beginning after returning all the pairs in the group to the status as before the procedure start by executing the fail-back.

Table E.1 Commands that Cannot be Used During the Fail-over

Category	Command
NAS Backup Restore Modular commands	horc_svol_import, horc_svol_vmimport
NAS Manager Modular commands	enas_cifscreate, enas_cifsdelete
	enas_fscreate, enas_fsdelete
	enas_fsmount, enas_fsumount
	enas_nfscreate, enas_nfsde

E.1 Creating Pairs

E.1.1 Commands for Creating Pairs

Table E.2 lists commands used for creating pairs.

Table E.2 Commands for Creating Pairs

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
2		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
3	CCI commands	sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
4		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair creation.
5		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.

E.1.2 Procedure for Creating Pairs

The pair is usually created in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation.

Table E.3 Procedure for Creating Pairs

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processin g	Command
1			Public	Mount	SMPL	SMPL	–	–		
2									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
3			Public	Mount	SMPL	SMPL	–	–		
4	pairdisplay	Confirm the configuration definition file before crating pair								
5			Public	Mount	SMPL	SMPL	–	–		
6	paircreate	Creates a volume pair								
7			Public	Mount	COPY	COPY	–	–		
8	pairdisplay	Confirm a volume pair status								
9			Public	Mount	COPY	COPY	–	–		
10	pairvolchk	At this point a pair status is COPY								
11	Execute pairvolchk several times		Public	Mount	COPY	COPY	–	–		
12	pairvolchk	When a pair status change to PAIR, crating pair is completed								
13			Public	Mount	PAIR	PAIR	–	–		

E.2 Splitting Pairs

E.2.1 Commands for Splitting Pairs

Table E.4 Commands for Splitting Pairs

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
2		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
3		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
4		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
5		sudo enas_fsumount file system name	Un-mount a file system.
6		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
7		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
8		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
9		sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
10		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
11		enas_dircreate -u owner -g group name -m mode of directory to be changed path of subdirectory	Create a subdirectory.
12		enas_diredit -u owner -g group name -m mode of directory to be changed path of subdirectory	Change directory attributes.
13	CCI commands	sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
14		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair splitting.

E.2.2 Procedure for Splitting Pairs with P-VOL Mounted

The following section describes the procedure for splitting a pair when P-VOl is mounted. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.5 Procedure for Splitting Pairs with P-VOL Mounted

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processin g	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	horc_pvol_freeze	Hold the access								
3			Public	Mount	PAIR	PAIR	–	–		
4	pairsplit -rw	Split a pair								
5			Public	Mount	PSUS	SSUS	–	–		
6	horc_pvol_unfreeze	Cancel suppression of operations								
7	Execute pairvolchk several times		Public	Mount	PSUS	SSUS	–	–		
8	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
9			Public	Mount	PSUS	SSUS	–	–		
10									Connect the file system	horc_svol_vmimpo rt (when LVM is not used, horc_svol_import)
11			Public	Mount	PSUS	SSUS	Un-mount	Non-public		
12									Mount	enas_fsmount
13			Public	Mount	PSUS	SSUS	Mount	Non-public		
14									Create a subdirectory / change a directory attributes	enas_dircreate/ enas_diredit
15			Public	Mount	PSUS	SSUS	Mount	Non-public		
16									Share	enas_nfscreate/ enas_cifscreate
17			Public	Mount	PSUS	SSUS	Mount	Public		

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processin g	Command
18						Start a program that accesses the file system at the remote site.				

Note: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

E.2.3 Procedure for Splitting Pairs with P-VOL Un-mounted

The following section describes the procedure for splitting a pair when P-VOl is unmounted. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.6 Procedure for Splitting Pairs with P-VOL Un-mounted

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Comma nd
1	Stop a program that accesses the file system at the main site.									
2			Public	Mount	PAIR	PAIR	–	–		
3	enas_nfsdelete/ enas_cifsdelete	Delete NFS/CIFS shares								
4			Non-public	Mount	PAIR	PAIR	–	–		
5	enas_fsumount	Un-mount								
6			Non-public	Un-mount	PAIR	PAIR	–	–		
7	pairsplit -rw	Split a pair ***Note 1***								
8	Execute pairvolchk several times		Non-public	Un-mount	PSUS	SSUS	–	–		
9	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
10			Non-public	Un-mount	PSUS	SSUS	–	–		
11	enas_fsmount	Mount								
12			Non-public	Mount	PSUS	SSUS	–	–		
13	enas_nfscreate/ enas_cifscreate	Share								
14			Public	Mount	PSUS	SSUS	–	–		
15	Restart a program that accesses the file system at the main site.									
16			Public	Mount	PSUS	SSUS	–	–		
17									Connect the file system ***Note 2***	horc_svol _vmimport (when LVM is not used, horc_svol _import)
18			Public	Mount	PSUS	SSUS	Un-mount	Non-public		
19									Mount	enas_fsm

										ount
20			Public	Mount	PSUS	SSUS	Mount	Non-public		

21									Create a subdirectory/ change a directory attributes ***Note 3***	enas_dircreate/ enas_diredit
22			Public	Mount	PSUS	SSUS	Mount	Non-public		
23									Share	enas_nfscreate/ enas_cifscreate
24			Public	Mount	PSUS	SSUS	Mount	Public		
25						Start a program that accesses the file system at the remote site.				

Note 1*: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.*

Note 2*: Specify the same file system name as the main site for the file system name connected with NAS unit.*

Note 3*: It is not necessary for the same public directory name as the main site. When the main site and the public directory name are changed, it is necessary.*

E.3 Resynchronizing Pairs

E.3.1 Commands for Resynchronizing Pairs

Table E.7 lists commands used for resynchronizing pairs.

Table E.7 Commands for Re-synchronizing Pairs

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
2		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
3		sudo enas_fsumount file system name	Un-mount a file system.
4		sudo enas_fsdelete file system name	Delete a file system.
5		sudo horc_svol_define -d device file number	When LVM is not used, reserves device files.
6		sudo horc_svol_vmdefine -d device file number	When LVM is used, reserves device files.
7		syncumount	Un-mounts the differential-data snapshot.
8		syncstop	Releases the differential-data storage device.
9	CCI commands	sudo pairresync {-g group name | -d volume name}	Resynchronizes the split pairs.
10		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair re-synchronizing.

E.3.2 Procedure for Re-synchronizing Pairs

The following section describes the procedure for resynchronizing a pair with volume status PSUS/SSUS. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.8 Procedure for Resynchronizing Pairs

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
1						Stop a program that accesses the file system at the remote site.				
2			Public	Mount	PSUS	SSUS	Mount	Public		
3									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
4			Public	Mount	PSUS	SSUS	Mount	Non-public		
5									Un-mount	enas_fsumount
6			Public	Mount	PSUS	SSUS	Un-mount	Non-public		
7									Delete file system ***Note***	enas_fsdelete
8			Public	Mount	PSUS	SSUS	–	–		
9									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
10			Public	Mount	PSUS	SSUS	–	–		
11	pairresync	Resynchronizes the pairs								
12			Public	Mount	COPY	COPY	–	–		
13	pairvolchk	At this point a pair status is COPY								
14	Execute pairvolchk several times		Public	Mount	COPY	COPY	–	–		
15	pairvol chk	When a pair status change to PAIR, re-synchronizing pair is completed								
16			Public	Mount	PAIR	PAIR	–	–		

Note: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

E.4 Deleting Pairs

E.4.1 Commands for Deleting Pairs

Table E.9 lists commands used for deleting pairs.

Table E.9 Commands for Deleting Pairs

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
2		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
3		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
4		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
5		sudo enas_fsumount file system name	Un-mount a file system.
6		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
7		sudo enas_fsdelete file system name	Delete a file system.
8		sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
9		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
10		enas_dircreate -u owner -g group name -m mode of directory to be changed path of subdirectory	Create a subdirectory.
11		enas_diredit -u owner -g group name -m mode of directory to be changed path of subdirectory	Change directory attributes.
12		syncumount	Un-mounts the differential-data snapshot.
13		syncstop	Releases the differential-data storage device.
14		sudo horc_svol_delete	When LVM is not used, releases device files.
15		sudo horc_svol_vmdelete	When LVM is used, releases device files.
16	CCI commands	sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.
17		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.
18		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair deleting.

E.4.2 Procedure for Deleting Pairs (S-VOL Continuously Used-1)

The following section describes the procedure for deleting a pair when the S-VOL is in PAIR status and is continuously used after pair deletion. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.10 Procedure for Deleting Pairs (S-VOL Continuously Used-1)

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
1			Public	Mount	PAIR	PAIR	–	–		
2	pairsplit -S	Deletes a volume pair ***Note 1***								
3			Public	Mount	PAIR	PAIR	–	–		
4	pairsplit -R	Brings S-VOL into SMPL forcibly								
5	Execute pairvolchk several times		Public	Mount	SMPL	SMPL	–	–		
6	pairvolchk	When a pair status change to SMPL, deleting is completed								
7			Public	Mount	SMPL	SMPL	–	–		
8									Connect the file system ***Note 2***	horc_svol_vmimport (when LVM is not used, horc_svol_import)
9			Public	Mount	SMPL	SMPL	Un-mount	Non-public		
10									Mount	enas_fsmount
11			Public	Mount	SMPL	SMPL	Mount	Non-public		
12									Create a subdirectory/ change a directory attributes ***Note 3***	enas_dircreate/ enas_diredit
13			Public	Mount	SMPL	SMPL	Mount	Non-public		
14									Share	enas_nfscre

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
										ate/ enas_cifscre ate
15			Public	Mount	SMPL	SMPL	Mount	Public		
16						Start a program that accesses the file system at the remote site.				

Note 1: The time required for the response to a command varies depending on the size of the differential data in order t o respond to the host after the guarantee of the S-VOL data is completed.

Note 2: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 3: It is not necessary for the same public directory name as the main site. When the main site and the public directory name are changed, it is necessary.

E.4.3 Procedure for Deleting Pairs (S-VOL Continuously Used-2)

The following section describes the procedure for deleting a pair when the S-VOL is in PSUS/SSUS status and is continuously used after pair deletion. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.11 Procedure for Deleting Pairs (S-VOL Continuously Used-2)

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
1			Public	Mount	PSUS	SSUS	Mount	Public		
2	pairsplit -S	Deletes a volume pair ***Note***								
3			Public	Mount	SMPL	SMPL	Mount	Public		
4	pairsplit -R	Brings S-VOL into SMPL forcibly								
5			Public	Mount	SMPL	SMPL	Mount	Public		
6	pairvolchk	When a pair status change to SMPL, deleting is completed								
7			Public	Mount	SMPL	SMPL	Mount	Public		

Note: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

E.4.4 Procedure for Deleting Pairs (S-VOL Not Used-1)

The following section describes the procedure for deleting a pair when the S-VOL is in PAIR status and is not used after pair deletion. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.12 Procedure for Deleting Pairs (S-VOL Not Used-1)

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
1			Public	Mount	PAIR	PAIR	–	–		
2	pairsplit -S	Deletes a volume pair ***Note***								
3			Public	Mount	SMPL	SMPL	–	–		
4	pairsplit -R	Brings S-VOL into SMPL forcibly								
5	Execute pairvolchk several times		Public	Mount	SMPL	SMPL	–	–		
6	pairvolchk	When a pair status change to SMPL, deleting is completed								
7			Public	Mount	SMPL	SMPL	–	–		
8									Releases device files	horc_svol_vmdelete (when LVM is not used, horc_svol_delete)
9			Public	Mount	SMPL	SMPL	–	–		

Note: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

E.4.5 Procedure for Deleting Pairs (S-VOL Not Used-2)

The following section describes the procedure for deleting a pair when the S-VOL is in PSUS/SSUS status and is not used after pair deletion. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted in the following table. You should specify the appropriate options for the actual operation.

Table E.13 Procedure for Deleting Pairs (S-VOL Not Used-2)

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
1						Stop a program that accesses the file system at the remote site.				
2			Public	Mount	PSUS	SSUS	Mount	Public		
3									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
4			Public	Mount	PSUS	SSUS	Mount	Non-public		
5									Un-mount	enas_fsumount
6			Public	Mount	PSUS	SSUS	Un-mount	Non-public		
7									Delete file system ***Note1***	enas_fsdelete
8			Public	Mount	PSUS	SSUS	–	–		
9	pairsplit -S	Deletes a volume pair ***Note 2***								
10			Public	Mount	SMPL	SMPL	–	–		
11	pairsplit -R	Brings S-VOL into SMPL forcibly								
12			Public	Mount	SMPL	SMPL	–	–		
13	pairvolchk	When a pair status change to SMPL, deleting is completed								
14			Public	Mount	SMPL	SMPL	–	–		

Note 1: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

Note 2: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

Appendix F Failure Operations on NAS Modular

In this appendix, explain the five assumed failures. The following sections are included:

- Isolation when Failures Occur
- Operation when the Main Site or Subsystem Went Down
- Operation when the NAS Cluster of the Main Site Went Down
- Operation when Multiple Failures Occurred in All the Storages on the Main Site
- Operation when Multiple Failures Occurred in a Part of Storages on the Main Site
- Operation when Network Failures Occurred in the Main Site

The commands and the important options are described, but specify the appropriate options referring to the following manual.

– *NAS Manager Modular User's Guide*

– *NAS Backup Restore Modular User's Guide*

– Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage Command Control Interface (CCI) User and Reference Guide (MK-95DF701)

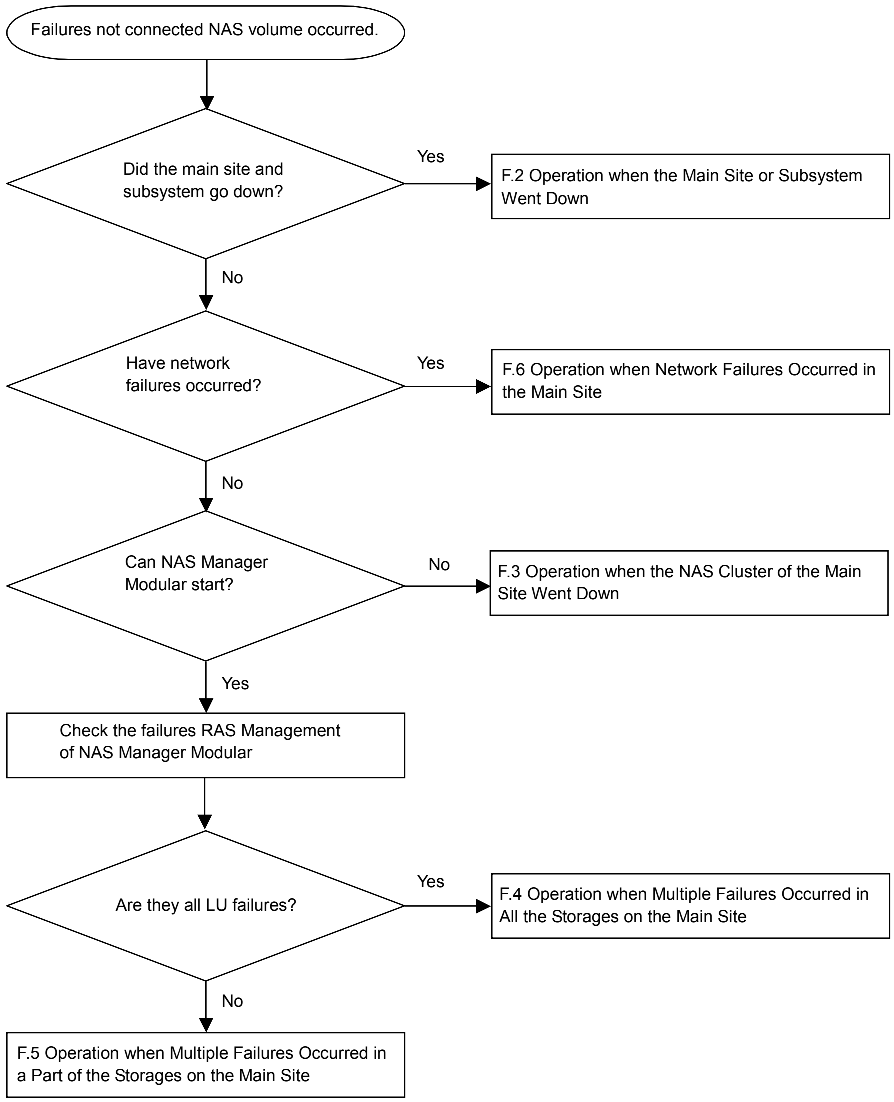
Note: During the fail-over of the NAS unit, the commands shown in Table F.1 cannot be used. When the fail-over occurs because of a failure caused by a controller blockade or an NAS unit failure, continue the pair operation procedure by executing the fail-back.

Table F.1 Commands that Cannot be Used During the Fail-over

Category	Command
NAS Backup Restore Modular commands	horc_svol_import, horc_svol_vmimport
NAS Manager Modular commands	enas_cifscreate, enas_cifsdelete
	enas_fscreate, enas_fsdelete
	enas_fsmount, enas_fsumount
	enas_nfscreate, enas_nfsde

F.1 Isolation when Failures Occur

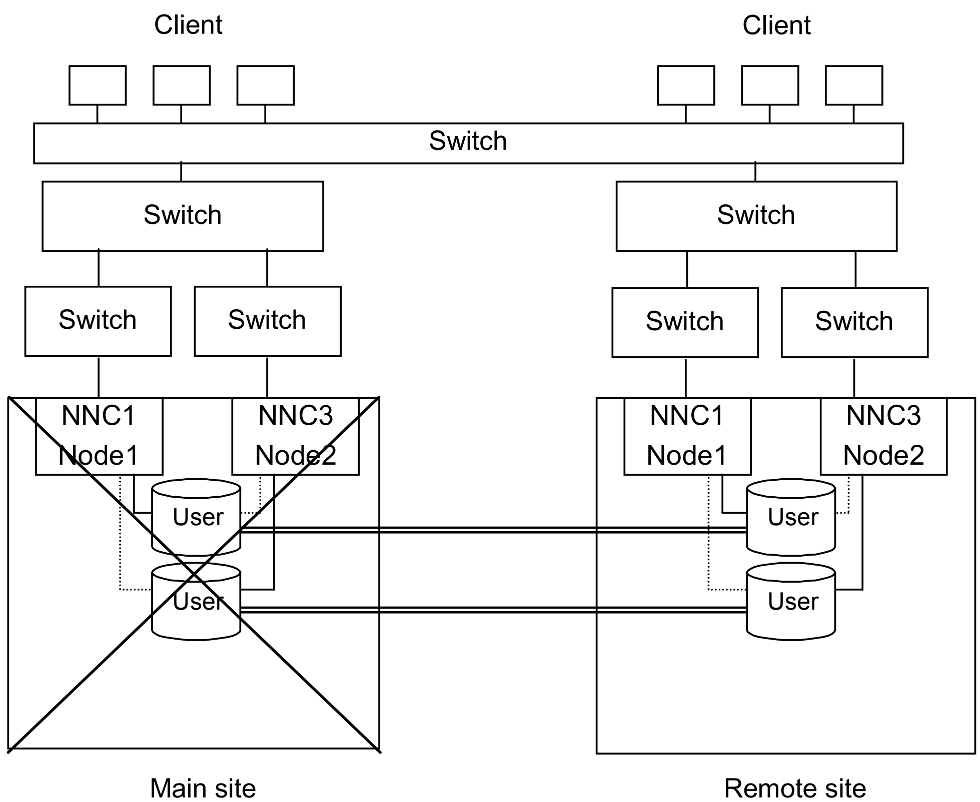
Perform the isolation based on the following flow chart when failures occurred. Moreover, failures of the case, which is not applied, to this flowchart may have occurred, so that perform the isolation contacting the service personnel. It is described by separating the cases when it can be connected to the NAS volume from the client and when it cannot be connected to the client.



F.2 Operation when the Main Site or Subsystem Went Down

F.2.1 Assumed Scenarios

The main site is struck, the business operation on the main site cannot be performed, and the business operation is continued on the remote site. The main site cannot be restored, and new subsystem was introduced.



F.2.2 Commands to be Used

The following commands are used when the main site or subsystem went down.

Table F.2 Commands for Recovery when the Main Site or Subsystem Went Down

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
2		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
3		sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
4		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
5		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
6		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS sharF.
7		sudo enas_fsumount file system name	Un-mount a file system.
8		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
9		sudo enas_fsdelete file system name	Delete a file system.
10		sudo horc_setenv HORCMINST instance number	Sets up or modifies the CCI environment variable.
11		sudo horc_unsetenv HORCC_MRCF	Deletes the CCI environment variable.
12		sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
13		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
14		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
15		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
16		sudo horcmstart.sh	Start CCI (1-instance configuration)
17		sudo horcmstart.sh 16 17	Start CCI (2-instance configuration)
18	CCI commands	sudo horctakeover {-g group name | -d volume name} [-t waittime]	Takeover the pair.
19		sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
20		sudo paircreate {-g group name | -d volume name} -f async -vl -nocopy	Creates a volume pair.
21		sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.

No.	Category	Command	Description
22		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.
23		sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
24		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair.
25		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.
26		sudo raidvchkset {-g group name | -d volume name} -vg	Releases S-VOL Disable.

F.2.3 Recovery Procedure from Failures

When the main site is struck, recover it in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation.
It is a prerequisite that S-VOL is reserved.

Table F.3 Recovery Procedure when the Main Site or Subsystem Went Down

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	Main site is down.									
3			–	–	–	PAIR	–	–		
4						The takeover execution is decided by the customer judgment.				
5									Execute takeover	horctakeover
6			–	–	–	SSWS	–	–		
7									Connect the file system ***Note 1***	horc_svol_vmimport (when LVM is not used, horc_svol_import)
8			–	–	–	SSWS	Un-mount	Non-public		
9									Mount	enas_fsmount
10			–	–	–	SSWS	Mount	Non-public		
11									Share	enas_nfscreate/ enas_cifscreate

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
12			–	–	–	SSWS	Mount	Public		
13						Start the business operation at the remote site. ***Note 2***				
14	Main site is recovered.									
15			–	–	–	SSWS	Mount	Public		
16	NAS cluster in the main site was started.									
17			–	–	–	SSWS	Mount	Public		
18	horc_setenv	Set CCI environment								
19			–	–	–	SSWS	Mount	Public		
20	horc_unsetenv	Set CCI environment								
21			–	–	–	SSWS	Mount	Public		
22									Deletes a volume pair	pairsplit -R
23			–	–	–	SMPL	Mount	Public		
24									When a pair status change to SMPL, deleting is completed	pairvolchk
25			–	–	–	SMPL	Mount	Public		
26	horcmstart.sh	Start CCI								
27			–	–	SMPL	SMPL	Mount	Public		
28	horc_svol_vmdefine (when LVM is not used, horc_svol_define)	Reserves the old P-VOL								
29			–	–	SMPL	SMPL	Mount	Public		
30									Confirm the configuration definition file before crating pair	pairdisplay
31			–	–	SMPL	SMPL	Mount	Public		
32									Creates a volume pair	paircreate
33			–	–	COPY	COPY	Mount	Public		
34									Confirm a volume pair status	pairdisplay
35			–	–	COPY	COPY	Mount	Public		
36									At this point a pair status is COPY	pairvolchk
37			–	–	COPY	COPY	Mount	Public	Execute pairvolchk several times	

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
38									When a pair status change to PAIR, crating pair is completed	pairvolchk
39			–	–	PAIR	PAIR	Mount	Public		
40						Stop the business operation at the remote site.				
41									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
42			–	–	PAIR	PAIR	Mount	Non-public		
43									Un-mount	enas_fsumount
44			–	–	PAIR	PAIR	Un-mount	Non-public		
45									Split a pair ***Note 3***	pairsplit -rw
46			–	–	SSUS	PSUS	Un-mount	Non-public	Execute pairvolchk several times	
47									When a pair status change to PSUS, splitting pair is completed	pairvolchk
48			–	–	SSUS	PSUS	Un-mount	Non-public		
49									Deletes a volume pair	pairsplit -S
50			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
51										pairsplit -R
52			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
53									When a pair status change to SMPL, deleting is completed	pairvolchk
54			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
55									Releases S-VOL Disable	raidvchkset
56			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
57	paircreate -nocopy	Creates a volume pair								
58			Non-public	Un-mount	PAIR	PAIR	Un-mount	Non-public		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
59	pairvolchk	When a pair status change to PAIR, crating pair is completed								
60			Non-public	Un-mount	PAIR	PAIR	Un-mount	Non-public		
61	pairsplit -rw	Split a pair ***Note 3***								
62			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
63	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
64			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
65									Setting S-VOL Disable	raidvchkset -vg svd
66			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
67	horc_svol_vmimport (when LVM is not used, horc_svol_import)	Connect the file system ***Note 1***								
68			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
69									Delete file system ***Note 4***	enas_fsdelete
70			Non-public	Un-mount	PSUS	SSUS	–	–		
71									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
72			Non-public	Un-mount	PSUS	SSUS	–	–		
73	pairresync	Resynchronize								
74			Non-public	Un-mount	COPY	COPY	–	–		
75	pairvolchk	At this point a pair status is COPY								
76	Execute pairvolchk several times		Non-public	Un-mount	COPY	COPY	–	–		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
77	pairvolchk	When a pair status change to PAIR, re-synchronizing pair is completed								
78			Non-public	Un-mount	PAIR	PAIR	–	–		
79	enas_fsmount	Mount								
80			Non-public	Mount	PAIR	PAIR	–	–		
81	enas_nfscreate/ enas_cifscreate	Share								
82			Public	Mount	PAIR	PAIR	–	–		
83	Start the business operation at the main site. ***Note 2***									

Note 1: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 2: When the takeover is performed, the IP address cannot be taken over from the primary site to the secondary site. When starting a job at the secondary site, un-mount the client from the primary site, change the IP address of the site on which the client is to be mounted to that of the secondary site, and then mount the client again. When the job is resumed at the primary site because the primary site has been restored, un-mount the client from the secondary site, return the IP address of the site on which the client is to be mounted to that of the primary site, and then mount the client again.

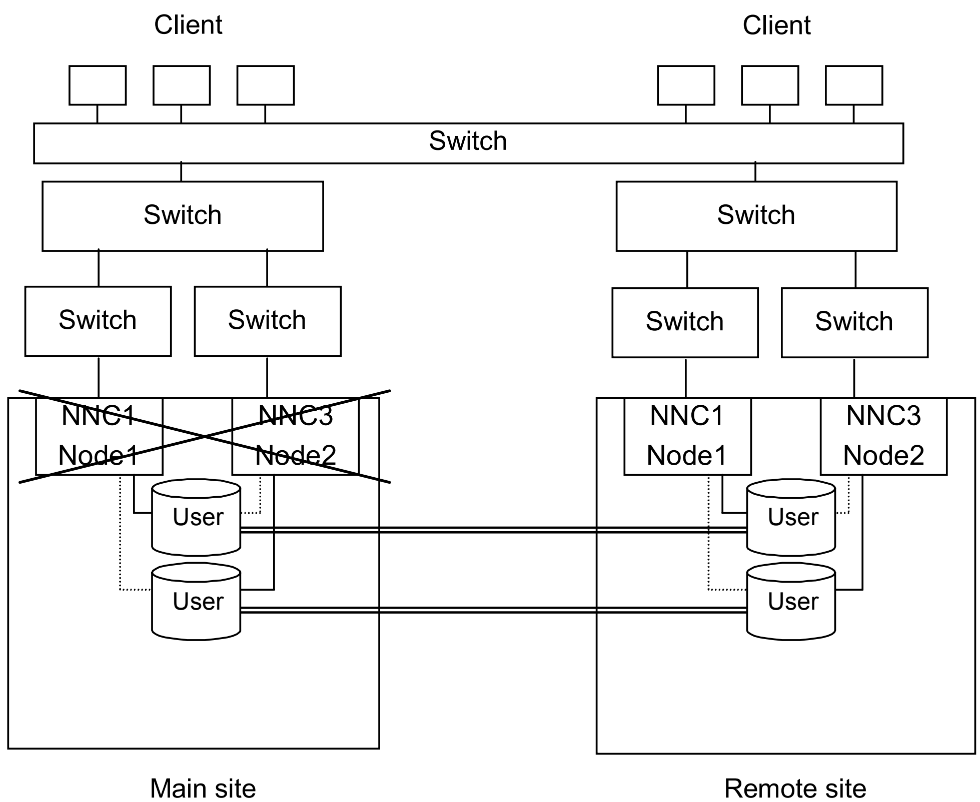
Note 3: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

Note 4: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

F.3 Operation when the NAS Cluster of the Main Site Went Down

F.3.1 Assumed Scenarios

Both two nodes of NAS OS on the main side went down. Because OS does not start, the business operation on the main site cannot be performed, and the business operation is continued on the remote site.



F.3.2 Commands to be Used

The following commands are used when the NAS cluster of the main site went down.

Table F.4 Commands for Recovery when the NAS Cluster of Main Site Went Down

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
2		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
3		sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
4		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
5		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
6		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
7		sudo enas_fsumount file system name	Un-mount a file system.
8		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
9		sudo enas_fsdelete file system name	Delete a file system.
10		sudo horc_setenv HORCMINST instance number	Sets up or modifies the CCI environment variable.
11		sudo horc_unsetenv HORCC_MRCF	Deletes the CCI environment variable.
12		sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
13		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
14		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
15		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
16		sudo horcmstart.sh	Start CCI (1-instance configuration)
17		sudo horcmstart.sh 16 17	Start CCI (2-instance configuration)
18	CCI commands	sudo horctakeover {-g group name | -d volume name} [-t waittime]	Takeover the pair.
19		sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
20		sudo paircreate {-g group name | -d volume name} -f async -vl -nocopy	Creates a volume pair.
21		sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.

No.	Category	Command	Description
22		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.
23		sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
24		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair.
25		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.
26		sudo raidvchkset {-g group name | -d volume name} -vg	Releases S-VOL Disable.

F.3.3 Recovery Procedure from Failures

When the NAS cluster of the main site is struck, recover it in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation. It is a prerequisite that S-VOL is reserved.

Table F.5 Recovery Procedure when the NAS Cluster of the Main Site Went Down

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	NAS cluster is down.									
3			–	–	–	PAIR	–	–		
4						The takeover execution is decided by the customer judgment.				
5									Execute takeover	horctakeover
6			–	–	PSUE	SSWS	Un-mount	Non-public		
7									Confirm a volume pair status to be SSWS	pairvolchk
8			–	–	PSUE	SSWS	Un-mount	Non-public		
9									Connect the file system ***Note 1***	horc_svol_vmimport (when LVM is not used, horc_svol_import)
10			–	–	PSUE	SSWS	Un-mount	Non-public		
11									Mount	enas_fsmount

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
12			–	–	PSUE	SSWS	Mount	Non-public		
13									Share	enas_nfscreate/ enas_cifscreate
14			–	–	PSUE	SSWS	Mount	Public		
15						Start the business operation at the remote site. ***Note 2***				
16	NAS cluster is recovered.									
17			Public	Mount	PSUE	SSWS	Mount	Public		
18	enas_nfsdelete/ enas_cifsdelete	Delete NFS/CIFS shares								
19			Non-public	Mount	PSUE	SSWS	Mount	Public		
20	enas_fsumount	Un-mount								
21			Non-public	Un-mount	PSUE	SSWS	Mount	Public		
22	enas_fsdelete	Delete file system ***Note 3***								
23			–	–	PSUE	SSWS	Mount	Public		
24	horc_setenv	Set CCI environment								
25			–	–	PSUE	SSWS	Mount	Public		
26	horc_unsetenv	Set CCI environment								
27			–	–	PSUE	SSWS	Mount	Public		
28	horcmstart.sh	Start CCI								
29			–	–	PSUE	SSWS	Mount	Public		
30	horc_svol_vmdefine (when LVM is not used, horc_svol_define)	Reserves the old P-VOL								
31			–	–	PSUE	SSWS	Mount	Public		
32									Deletes a volume pair	pairsplit -S
33			–	–	SMPL	SMPL	Mount	Public		
34										pairsplit -R
35			–	–	SMPL	SMPL	Mount	Public		
36									When a pair status change to SMPL, deleting is completed	pairvolchk

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
37			–	–	SMPL	SMPL	Mount	Public		
38									Creates a volume pair	paircreate
39			–	–	COPY	COPY	Mount	Public		
40									Confirm a volume pair status	pairdisplay
41			–	–	COPY	COPY	Mount	Public		
42									At this point a pair status is COPY	pairvolchk
43			–	–	COPY	COPY	Mount	Public	Execute pairvolchk several times	
44									When a pair status change to PAIR, crating pair is completed	pairvolchk
45			–	–	PAIR	PAIR	Mount	Public		
46						Stop the business operation at the remote site.				
47									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
48			–	–	PAIR	PAIR	Mount	Non-public		
49									Un-mount	enas_fsumount
50			–	–	PAIR	PAIR	Un-mount	Non-public		
51									Split a pair ***Note 4***	pairsplit -rw
52			–	–	SSUS	PSUS	Un-mount	Non-public		
53									When a pair status change to PSUS, splitting pair is completed	pairvolchk
54			–	–	SSUS	PSUS	Un-mount	Non-public		
55									Deletes a volume pair	pairsplit -S
56			–	–	SMPL	SMPL	Un-mount	Non-public		
57										pairsplit -R
58			–	–	SMPL	SMPL	Un-mount	Non-public		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
59									When a pair status change to SMPL, deleting is completed	pairvolchk
60			–	–	SMPL	SMPL	Un-mount	Non-public		
61									Releases S-VOL Disable	raidvchkset
62			–	–	SMPL	SMPL	Un-mount	Non-public		
63	paircreate -nocopy	Creates a volume pair								
64			–	–	PAIR	PAIR	Un-mount	Non-public		
65	pairvolchk	When a pair status change to PAIR, crating pair is completed								
66			–	–	PAIR	PAIR	Un-mount	Non-public		
67	pairsplit -rw	Split a pair ***Note 4***								
68			–	–	PSUS	SSUS	Un-mount	Non-public		
69	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
70			–	–	PSUS	SSUS	Un-mount	Non-public		
71									Setting S-VOL Disable	raidvchkset -vg svd
72			–	–	PSUS	SSUS	Un-mount	Non-public		
73	horc_svol_vmimport (when LVM is not used, horc_svol_import)	Connect the file system ***Note 1***								
74			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
75									Delete file system	enas_fsdelete
76			Non-public	Un-mount	PSUS	SSUS	–	–		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
77									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
78			Non-public	Un-mount	PSUS	SSUS	–	–		
79	pairresync	Resynchronize								
80			Non-public	Un-mount	COPY	COPY	–	–		
81	pairvolchk	At this point a pair status is COPY								
82	Execute pairvolchk several times		Non-public	Un-mount	COPY	COPY	–	–		
83	pairvolchk	When a pair status change to PAIR, re-synchronizing pair is completed								
84			Non-public	Un-mount	PAIR	PAIR	–	–		
85	enas_fsmount	Mount								
86			Non-public	Mount	PAIR	PAIR	–	–		
87	enas_nfscreate/ enas_cifscreate	Share								
88			Public	Mount	PAIR	PAIR	–	–		
89	Start the business operation at the main site. ***Note 2***									

Note 1: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 2: When the takeover is performed, the IP address cannot be taken over from the primary site to the secondary site. When starting a job at the secondary site, un-mount the client from the primary site, change the IP address of the site on which the client is to be mounted to that of the secondary site, and then mount the client again. When the job is resumed at the primary site because the primary site has been restored, un-mount the client from the secondary site, return the IP address of the site on which the client is to be mounted to that of the primary site, and then mount the client again.

Note 3: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

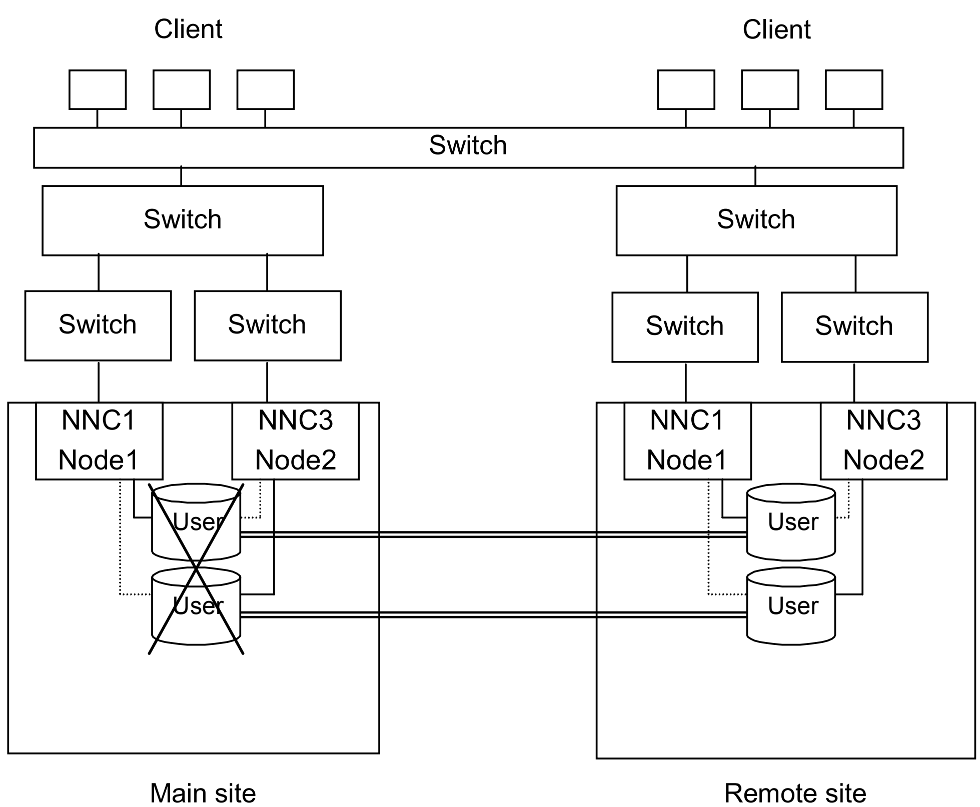
Note 4: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

F.4 Operation when Multiple Failures Occurred in All the Storages on the Main Site

F.4.1 Assumed Scenarios

When multiple failures occurred in all the storages used in the NAS user LU on the main site, the business operation cannot be performed on the main site, and the business operation is continued on the remote site. Note that ShadowImage operation is not performed.

The failed storage on the main site is recovered by the replacement (service personnel).



F.4.2 Commands to be Used

The following commands are used when multiple failures occurred in all the storages used in the NAS user LU on the main site.

Table F.6 Commands for Recovery when Multiple Failures Occurred in All the Storages Used in the NAS User LU on the Main Site

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
2		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
3		sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
4		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
5		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
6		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
7		sudo enas_fsumount file system name	Un-mount a file system.
8		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
9		sudo enas_fsdelete file system name	Delete a file system.
10		sudo horc_setenv HORCMINST instance number	Sets up or modifies the CCI environment variable.
11		sudo horc_unsetenv HORCC_MRCF	Deletes the CCI environment variable.
12		sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
13		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
14		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
15		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
16		sudo horcmstart.sh	Start CCI (1-instance configuration)
17		sudo horcmstart.sh 16 17	Start CCI (2-instance configuration)
18	CCI commands	sudo horctakeover {-g group name | -d volume name} [-t waittime]	Takeover the pair.
19		sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
20		sudo paircreate {-g group name | -d volume name} -f async -vl -nocopy	Creates a volume pair.

No.	Category	Command	Description
21		sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.
22		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.
23		sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
24		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair.
25		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.
26		sudo raidvchkset {-g group name | -d volume name} -vg	Releases S-VOL Disable.

F.4.3 Recovery Procedure from Failures

When multiple failures occurred in all the storages used in the NAS user LU on the main site, recover it in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation. It is a prerequisite that S-VOL is reserved.

Table F.7 Recovery Procedure when Multiple Failures Occurred in All the Storages Used in the NAS User LU on the Main Site

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	Multiple failures occurred in all NAS user LU.									
3			–	–	PSUE	PAIR	–	–		
4						The takeover execution is decided by the customer judgment.				
5									Execute takeover	horctakeover
6			–	–	PSUE	SSWS	–	–		
7									Deletes a volume pair	pairsplit -S
8			–	–	SMPL	SMPL	Un-mount	Non-public		
9										pairsplit -R
10			–	–	SMPL	SMPL	Un-mount	Non-public		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
11									When a pair status change to SMPL, deleting is completed	pairvolchk
12			–	–	SMPL	SMPL	Un-mount	Non-public		
13									Connect the file system ***Note 1***	horc_svol_vmimport (when LVM is not used, horc_svol_import)
14			–	–	SMPL	SMPL	Un-mount	Non-public		
15									Mount	enas_fsmount
16			–	–	SMPL	SMPL	Mount	Non-public		
17									Share	enas_nfscreate/ enas_cifscreate
18			–	–	SMPL	SMPL	Mount	Public		
19						Start the business operation at the remote site. ***Note 2***				
20	enas_nfsdelete/ enas_cifsdelete	Delete NFS/CIFS shares								
21			Non-public	Mount	SMPL	SMPL	Mount	Public		
22	enas_fsumount	Un-mount								
23			Non-public	Un-mount	SMPL	SMPL	Mount	Public		
24	enas_fsdelete	Delete file system ***Note 3***								
25			–	–	SMPL	SMPL	Mount	Public		
26	Stop the NAS cluster in the main site. ***Note 4***									
27			–	–	SMPL	SMPL	Mount	Public		
28	Shut down the NAS OS in the main site.									
29			–	–	SMPL	SMPL	Mount	Public		
30	Replace and format the storage with error in the main site.									
31			–	–	SMPL	SMPL	Mount	Public		
32	Start the NAS OS in the main site.									
33			–	–	SMPL	SMPL	Mount	Public		
34	Start the NAS cluster in the main site.									

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
35			–	–	SMPL	SMPL	Mount	Public		
36	horc_setenv	Set CCI environment								
37			–	–	SMPL	SMPL	Mount	Public		
38	horc_unsetenv	Set CCI environment								
39			–	–	SMPL	SMPL	Mount	Public		
40	horcmstart.sh	Start CCI								
41			–	–	SMPL	SMPL	Mount	Public		
42	horc_svol_vmdefine (when LVM is not used, horc_svol_define)	Reserves the old P-VOL								
43			–	–	SMPL	SMPL	Mount	Public		
44									Confirm the configuration definition file before crating pair	pairdisplay
45			–	–	SMPL	SMPL	Mount	Public		
46									Creates a volume pair	paircreate
47			–	–	COPY	COPY	Mount	Public		
48									Confirm a volume pair status	pairdisplay
49			–	–	COPY	COPY	Mount	Public		
50									At this point a pair status is COPY	pairvolchk
51			–	–	COPY	COPY	Mount	Public	Execute pairvolchk several times	
52									When a pair status change to PAIR, crating pair is completed	pairvolchk
53			–	–	PAIR	PAIR	Mount	Public		
54						Stop the business operation at the remote site.				
55									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
56			–	–	PAIR	PAIR	Mount	Non-public		
57									Un-mount	enas_fsumount
58			–	–	PAIR	PAIR	Un-mount	Non-public		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
59									Split a pair ***Note 4***	pairsplit -rw
60			–	–	SSUS	PSUS	Un-mount	Non-public	Execute pairvolchk several times	
61									When a pair status change to PSUS, splitting pair is completed	pairvolchk
62			–	–	SSUS	PSUS	Un-mount	Non-public		
63									Deletes a volume pair	pairsplit -S
64			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
65										pairsplit -R
66			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
67									When a pair status change to SMPL, deleting is completed	pairvolchk
68			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
69									Releases S-VOL Disable	raidvchkset
70			Non-public	Un-mount	SMPL	SMPL	–	–		
71	paircreate -nocopy	Creates a volume pair								
72			Non-public	Un-mount	COPY	COPY	–	–		
73	pairvolchk	When a pair status change to PAIR, crating pair is completed								
74			Non-public	Un-mount	PAIR	PAIR	–	–		
75	pairsplit -rw	Split a pair ***Note 5***								
76			Non-public	Un-mount	PSUS	SSUS	–	–		
77	pairvolchk	When a pair status change to PSUS, splitting pair is completed								

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
78			Non-public	Un-mount	PSUS	SSUS	–	–		
79									Setting S-VOL Disable	raidvchkset -vg svd
80			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
81	horc_svol_vmimport (when LVM is not used, horc_svol_import)	Connect the file system ***Note 1***								
82			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
83									Delete file system	enas_fsdelete
84			Non-public	Un-mount	PSUS	SSUS	–	–		
85									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
86			Non-public	Un-mount	PSUS	SSUS	–	–		
87	pairresync	Resynchronize								
88			Non-public	Un-mount	COPY	COPY	–	–		
89	pairvolchk	At this point a pair status is COPY								
90	Execute pairvolchk several times		Non-public	Un-mount	COPY	COPY	–	–		
91	pairvolchk	When a pair status change to PAIR, re-synchronizing pair is completed								
92			Non-public	Un-mount	PAIR	PAIR	–	–		
93	enas_fsmount	Mount								
94			Non-public	Mount	PAIR	PAIR	–	–		
95	enas_nfscreate/ enas_cifscreate	Share								
96			Public	Mount	PAIR	PAIR	–	–		
97	Start the business operation at the main site. ***Note 2***									

Note 1: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 2: When the takeover is performed, the IP address cannot be taken over from the primary site to the secondary site. When starting a job at the secondary site, un-mount the client from the primary site, change the IP address of the site on which the client is to be mounted to that of the secondary site, and then mount the client again. When the job is resumed at the primary site because the primary site has been restored, un-mount the client from the secondary site, return the IP address of the site on which the client is to be mounted to that of the primary site, and then mount the client again.

Note 3: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

Note 4: When the file system is blocked due to a drive failure etc., the NAS OS needs to be restarted so that the blockade status recognized by the NAS OS will be released.

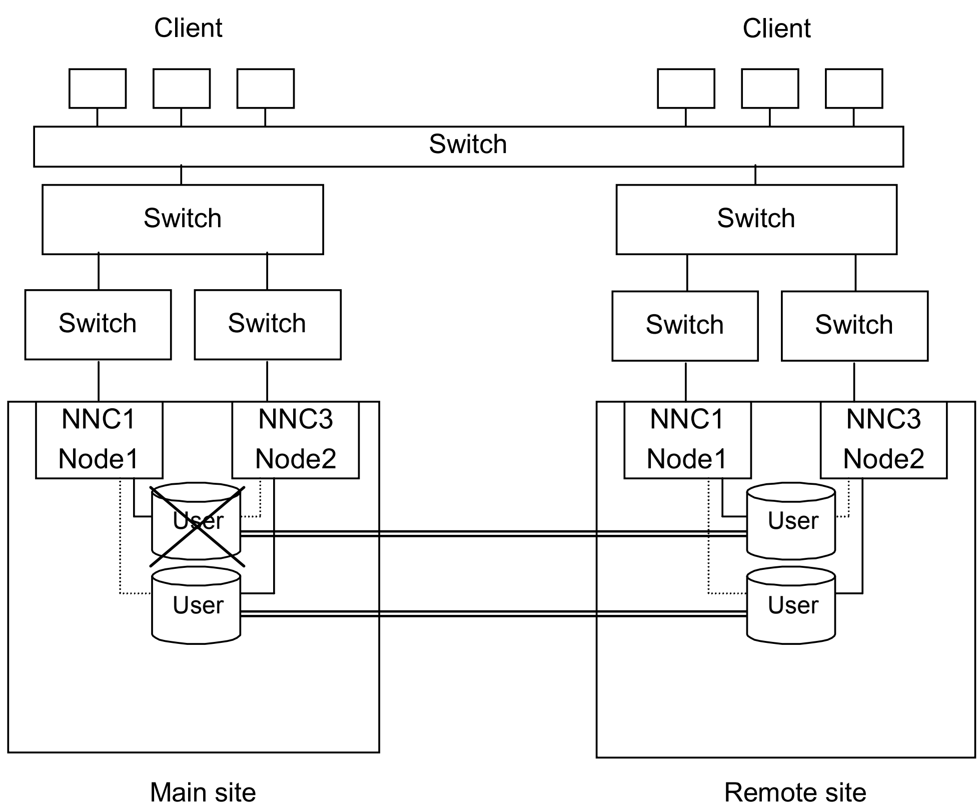
Note 5: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

F.5 Operation when Multiple Failures Occurred in a Part of Storages on the Main Site

F.5.1 Assumed Scenarios

When multiple failures occurred in a part of storages on the main site, the business operation cannot be performed on the main site, and the business operation is continued on the remote site. Note that ShadowImage operation is not performed.

The failed storage on the main site is recovered by the replacement (service personnel).



F.5.2 Commands to be Used

The following commands are used when multiple failures occurred in a part of storages used in the NAS user LU on the main site.

Table F.8 Commands for Recovery when Multiple Failures Occurred in a Part of Storages Used in the NAS User LU on the Main Site

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
2		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
3		sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
4		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
5		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
6		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
7		sudo enas_fsumount file system name	Un-mount a file system.
8		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
9		sudo enas_fsdelete file system name	Delete a file system.
10		sudo horc_setenv HORCMINST instance number	Sets up or modifies the CCI environment variable.
11		sudo horc_unsetenv HORCC_MRCF	Deletes the CCI environment variable.
12		sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
13		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
14		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
15		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
16		sudo horcmstart.sh	Start CCI (1-instance configuration)
17		sudo horcmstart.sh 16 17	Start CCI (2-instance configuration)
18	CCI commands	sudo horctakeover {-g group name | -d volume name} [-t waittime]	Takeover the pair.
19		sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
20		sudo paircreate {-g group name | -d volume name} -f async -vl -nocopy	Creates a volume pair.

No.	Category	Command	Description
21		sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.
22		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.
23		sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
24		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair.
25		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.
26		sudo raidvchkset {-g group name | -d volume name} -vg	Releases S-VOL Disable.

F.5.3 Recovery Procedure from Failures

When multiple failures occurred in a part of storages used in the NAS user LU on the main site, recover it in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation. It is a prerequisite that S-VOL is reserved.

Table F.9 Recovery Procedure when Multiple Failures Occurred in a Part of Storages Used in the NAS User LU on the Main Site

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	Multiple failures occurred in a part of NAS user LU.									
3			–	–	PSUE	PAIR	–	–		
4						The takeover execution is decided by the customer judgment.				
5									Execute takeover	horctakeover
6			–	–	PSUE	SSWS	–	–		
7									Deletes a volume pair	pairsplit -S
8			–	–	SMPL	SMPL	Un-mount	Non-public		
9										pairsplit -R
10			–	–	SMPL	SMPL	Un-mount	Non-public		

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
11									When a pair status change to SMPL, deleting is completed	pairvolchk
12			–	–	SMPL	SMPL	Un-mount	Non-public		
13									Connect the file system ***Note 1***	horc_svol_vmimport (when LVM is not used, horc_svol_import)
14			–	–	SMPL	SMPL	Un-mount	Non-public		
15									Mount	enas_fsmount
16			–	–	SMPL	SMPL	Mount	Non-public		
17									Share	enas_nfscreate/ enas_cifscreate
18			–	–	SMPL	SMPL	Mount	Public		
19						Start the business operation at the remote site. ***Note 2***				
20	Replace and format the storage with error in the main site.									
21			–	–	SMPL	SMPL	Mount	Public		
22	enas_nfsdelete/ enas_cifsdelete	Delete NFS/CIFS shares								
23			Non-public	Mount	SMPL	SMPL	Mount	Public		
24	enas_fsumount	Un-mount								
25			Non-public	Un-mount	SMPL	SMPL	Mount	Public		
26	enas_fsdelete	Delete file system ***Note 3***								
27			–	–	SMPL	SMPL	Mount	Public		
28	Failover from the node1 (blocked node) to the node2 (non-blocked node) in the main site. ***Note 4***									
29			–	–	SMPL	SMPL	Mount	Public		
30	Reboot the node1 in the main site.									
31			–	–	SMPL	SMPL	Mount	Public		
32	Failback in the main site.									
33			–	–	SMPL	SMPL	Mount	Public		
34	Failover from the node2 to the node1 in the main site.									

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
35			–	–	SMPL	SMPL	Mount	Public		
36	Reboot the node2 in the main site.									
37			–	–	SMPL	SMPL	Mount	Public		
38	Failback in the main site.									
39			–	–	SMPL	SMPL	Mount	Public		
40	horc_setenv	Set CCI environment								
41			–	–	SMPL	SMPL	Mount	Public		
42	horc_unsetenv	Set CCI environment								
43			–	–	SMPL	SMPL	Mount	Public		
44	horcmstart.sh	Start CCI								
45			–	–	SMPL	SMPL	Mount	Public		
46	horc_svol_vmdefine (when LVM is not used, horc_svol_define)	Reserves the old P-VOL								
47			–	–	SMPL	SMPL	Mount	Public		
48									Confirm the configuration definition file before crating pair	pairdisplay
49			–	–	SMPL	SMPL	Mount	Public		
50									Creates a volume pair	paircreate
51			–	–	COPY	COPY	Mount	Public		
52									Confirm a volume pair status	pairdisplay
53			–	–	COPY	COPY	Mount	Public		
54									At this point a pair status is COPY	pairvolchk
55			–	–	COPY	COPY	Mount	Public	Execute pairvolchk several times	
56									When a pair status change to PAIR, crating pair is completed	pairvolchk
57			–	–	PAIR	PAIR	Mount	Public		
58						Stop the business operation at the remote site.				
59									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
60			–	–	PAIR	PAIR	Mount	Non-public		
61									Un-mount	enas_fsumount
62			–	–	PAIR	PAIR	Un-mount	Non-public		
63									Split a pair ***Note 5***	pairsplit -rw
64			–	–	SSUS	PSUS	Un-mount	Non-public	Execute pairvolchk several times	
65									When a pair status change to PSUS, splitting pair is completed	pairvolchk
66			–	–	SSUS	PSUS	Un-mount	Non-public		
67									Deletes a volume pair	pairsplit -S
68			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
69									When a pair status change to SMPL, deleting is completed	pairvolchk
70			Non-public	Un-mount	SMPL	SMPL	Un-mount	Non-public		
71									Releases S-VOL Disable	raidvchkset
72			Non-public	Un-mount	SMPL	SMPL	–	–		
73	paircreate -nocopy	Creates a volume pair								
74			Non-public	Un-mount	COPY	COPY	–	–		
75	pairvolchk	When a pair status change to PAIR, crating pair is completed								
76			Non-public	Un-mount	PAIR	PAIR	–	–		
77	pairsplit -rw	Split a pair ***Note 5***								
78			Non-public	Un-mount	PSUS	SSUS	–	–		

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
79	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
80			Non-public	Un-mount	PSUS	SSUS	–	–		
81									Setting S-VOL Disable	raidvchkset -vg svd
82			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
83	horc_svol_vmimport (when LVM is not used, horc_svol_import)	Connect the file system ***Note 1***								
84			Non-public	Un-mount	SSUS	PSUS	Un-mount	Non-public		
85									Delete file system ***Note 3***	enas_fsdelete
86			Non-public	Un-mount	SSUS	PSUS	–	–		
87									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
88			Non-public	Un-mount	SSUS	PSUS	–	–		
89	pairresync	Resynchronize								
90			Non-public	Un-mount	COPY	COPY	–	–		
91	pairvolchk	At this point a pair status is COPY								
92	Execute pairvolchk several times		Non-public	Un-mount	COPY	COPY	–	–		
93	pairvolchk	When a pair status change to PAIR, re-synchronizing pair is completed								
94			Non-public	Un-mount	PAIR	PAIR	–	–		
95	enas_fsmount	Mount								
96			Non-public	Mount	PAIR	PAIR	–	–		

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
97	enas_nfscreate/ enas_cifscreate	Share								
98			Public	Mount	PAIR	PAIR	–	–		
99	Start the business operation at the main site. ***Note 2***									

Note 1: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 2: When the takeover is performed, the IP address cannot be taken over from the primary site to the secondary site. When starting a job at the secondary site, un-mount the client from the primary site, change the IP address of the site on which the client is to be mounted to that of the secondary site, and then mount the client again. When the job is resumed at the primary site because the primary site has been restored, un-mount the client from the secondary site, return the IP address of the site on which the client is to be mounted to that of the primary site, and then mount the client again.

Note 6: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

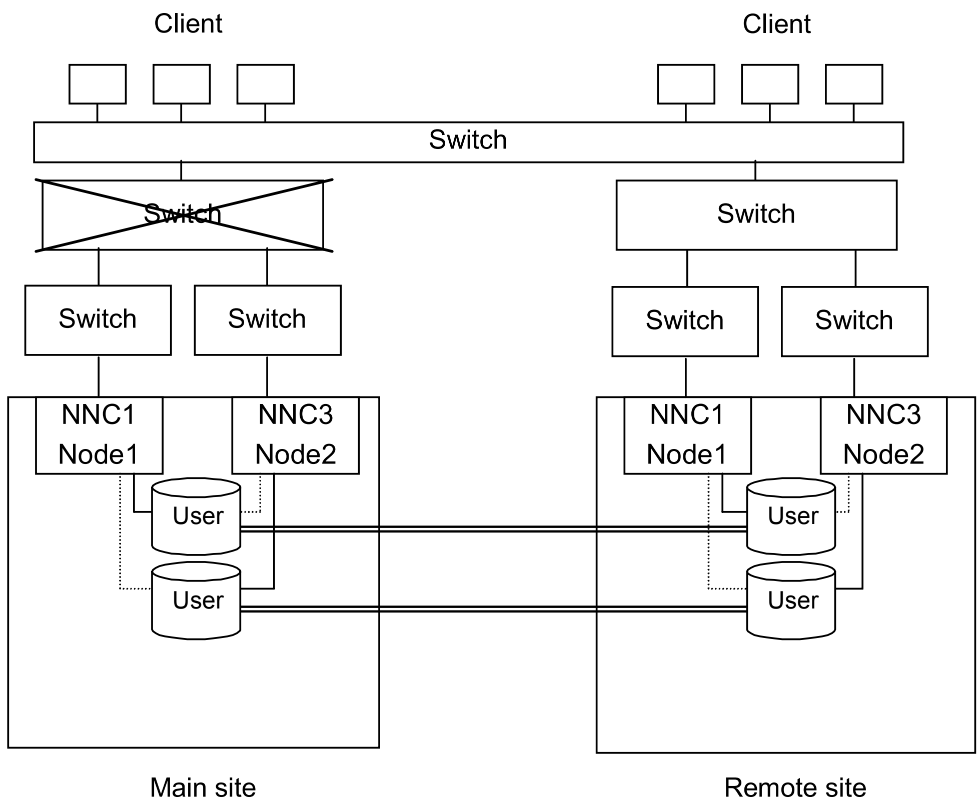
Note 4: When the file system is blocked due to a drive failure etc., the NAS OS needs to be restarted so that the blockade status recognized by the NAS OS will be released.

Note 5: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

F.6 Operation when Network Failures Occurred in the Main Site

F.6.1 Assumed Scenarios

The access from the client cannot be performed due to the switch failures in the main site.



F.6.2 Commands to be Used

Table F.10 Commands for Recovery when Network Failures Occurred in the Main Site

No.	Category	Command	Description
1	NAS Manager Modular commands	sudo horc_svol_import -f copy-destination file system name -d device file number	When LVM is not used, connects NAS unit.
2		sudo horc_svol_vmimport -f copy-destination file system name -d device file number [, device file number . . .]	When LVM is used, connects NAS unit.
3		sudo enas_nfsdelete -d shared directory {-a | -H Host}	Delete an NFS share.
4		sudo enas_nfscreate -d shared directory -H Host	Create an NFS share.
5		sudo enas_cifsdelete -x CIFS share name	Delete a CIFS share.
6		sudo enas_cifscreate -x CIFS share name -d shared directory	Create a CIFS share.
7		sudo enas_fsumount file system name	Un-mount a file system.
8		sudo enas_fsmount {-r | -w} file system name	Mount a file system.
9		sudo enas_fsdelete file system name	Delete a file system.
10		sudo horc_setenv HORCMINST instance number	Sets up or modifies the CCI environment variable.
11		sudo horc_unsetenv HORCC_MRCF	Deletes the CCI environment variable.
12		sudo horc_svol_define -d device file number	When LVM is not used, reserves device file.
13		sudo horc_svol_vmdefine -d device file number [, device file number . . .]	When LVM is used, reserves device file.
14		sudo horc_pvol_freeze -f copy-source file system name	Suppresses on the copy-source file system and holds access and data writes from clients.
15		sudo horc_pvol_unfreeze -f copy-source file system name	Restarts access and data writes on a copy-source file system from clients and cancels suppression of operations.
16		sudo horcmstart.sh	Start CCI (1-instance configuration)
17		sudo horcmstart.sh 16 17	Start CCI (2-instance configuration)
18	CCI commands	sudo horctakeover {-g group name | -d volume name} [-t waittime]	Takeover the pair.
19		sudo paircreate {-g group name | -d volume name} -f async -vl	Creates a volume pair.
20		sudo paircreate {-g group name | -d volume name} -f async -vl -nocopy	Creates a volume pair.
21		sudo pairsplit {-g group name | -d volume name} -S	Deletes a volume pair.
22		sudo pairsplit {-g group name | -d volume name} -R	Brings S-VOL into SMPL forcibly.

No.	Category	Command	Description
23		sudo pairsplit {-g group name | -d volume name} -rw	Splits a volume pair.
24		sudo pairvolchk {-g group name | -d volume name} -ss	Checks the completion of volume pair.
25		sudo pairdisplay {-g group name | -d volume name} -fc	Displays the copy operation progress.
26		sudo raidvchkset {-g group name | -d volume name} -vg	Releases S-VOL Disable.

F.6.3 Recovery Procedure from Failures

The switch failure is usually removed by replacing the switch. However, when you want to start the operation on the remote site immediately, remove it in the following procedure. The sudo command, the options of the NAS Manager Modular command, and the CCI command are omitted when it is described in the following table. Specify the appropriate options for the actual operation.

Table F.11 Recovery Procedure when Network Failures Occurred in the Main Site

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
1			Public	Mount	PAIR	PAIR	–	–		
2	Switch obstacle occurs.									
3			Public	Mount	PAIR	PAIR	–	–		
4						The takeover execution is decided by the customer judgment.				
5									Execute takeover	horctakeover
6			Public	Mount	PSUE	SSWS	Un-mount	Non-public		
7									Confirm a volume pair status to be SSWS	pairvolchk
8			Public	Mount	PSUE	SSWS	Un-mount	Non-public		
9									Connect the file system ***Note 1***	horc_svol_vmimport (when LVM is not used, horc_svol_import)

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
10			Public	Mount	PSUE	SSWS	Un-mount	Non-public		
11									Mount	enas_fsmount
12			Public	Mount	PSUE	SSWS	Mount	Non-public		
13									Share	enas_nfscreate/ enas_cifscreate
14			Public	Mount	PSUE	SSWS	Mount	Public		
15						Start the business operation at the remote site. ***Note 2***				
16	enas_nfsdelete/ enas_cifsdelete	Delete NFS/CIFS shares								
17			Non-public	Mount	PSUE	SSWS	Mount	Public		
18	enas_fsumount	Un-mount								
19			Non-public	Un-mount	PSUE	SSWS	Mount	Public		
20	enas_fsdelete	Delete file system ***Note 3***								
21			–	–	PSUE	SSWS	Mount	Public		
22	Replace the switch.									
23			–	–	PSUE	SSWS	Mount	Public		
24	horc_setenv	Set CCI environment								
25			–	–	PSUE	SSWS	Mount	Public		
26	horc_unsetenv	Set CCI environment								
27			–	–	PSUE	SSWS	Mount	Public		
28	horcmstart.sh	Start CCI								
29			–	–	PSUE	SSWS	Mount	Public		
30	horc_svol_vmdefine (when LVM is not	Reserves the old P-VOL								

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
	used, horc_svol_define)									
31			–	–	PSUE	SSWS	Mount	Public		
32									Deletes a volume pair	pairsplit -S
33			–	–	SMPL	SMPL	Mount	Public		
34										pairsplit -R
35			–	–	SMPL	SMPL	Mount	Public		
36									When a pair status change to SMPL, deleting is completed	pairvolchk
37			–	–	SMPL	SMPL	Mount	Public		
38									Creates a volume pair	paircreate
39			–	–	COPY	COPY	Mount	Public		
40									Confirm a volume pair status	pairdisplay
41			–	–	COPY	COPY	Mount	Public	Execute pairvolchk several times	
42									Confirm a volume pair status	pairvolchk
43			–	–	PAIR	PAIR	Mount	Public		
44						Stop the business operation at the remote site.				
45									Delete NFS/CIFS shares	enas_nfsdelete/ enas_cifsdelete
46			–	–	PAIR	PAIR	Mount	Non-public		
47									Un-mount	enas_fsumount
48			–	–	PAIR	PAIR	Un-mount	Non-public		
49									Split a pair ***Note 4***	pairsplit -rw

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
50			–	–	SSUS	PSUS	Un-mount	Non-public	Execute pairvolchk several times	
51									When a pair status change to PSUS, splitting pair is completed	pairvolchk
52			–	–	SSUS	PSUS	Un-mount	Non-public		
53									Deletes a volume pair	pairsplit -S
54			–	–	SMPL	SMPL	Un-mount	Non-public		
55										pairsplit -R
56			–	–	SMPL	SMPL	Un-mount	Non-public		
57									When a pair status change to SMPL, deleting is completed	pairvolchk
58			–	–	SMPL	SMPL	Un-mount	Non-public		
59									Releases S-VOL Disable	raidvchkset
60			–	–	SMPL	SMPL	–	–		
61	paircreate -nocopy	Creates a volume pair								
62			–	–	PAIR	PAIR	–	–		
63	pairvolchk	When a pair status change to PAIR, crating pair is completed								
64			–	–	PAIR	PAIR	–	–		
65	pairsplit -rw	Split a pair ***Note 4***								
66			–	–	PSUS	SSUS	–	–		

No.	Main Site					Remote Site				
	Command	Processing	NFS/ CIFS	FS	P-VOL Status	S-VOL Status	FS	NFS/ CIFS	Processing	Command
67	pairvolchk	When a pair status change to PSUS, splitting pair is completed								
68			–	–	PSUS	SSUS	–	–		
69									Setting S-VOL Disable	raidvchkset -vg svd
70			–	–	PSUS	SSUS	Un-mount	Non-public		
71	horc_svol_vmimport (when LVM is not used, horc_svol_import)	Connect the file system ***Note 1***								
72			Non-public	Un-mount	PSUS	SSUS	Un-mount	Non-public		
73									Delete file system ***Note 3***	enas_fsdelete
74			Non-public	Un-mount	PSUS	SSUS	–	–		
75									Reserves S-VOL	horc_svol_vmdefine (when LVM is not used, horc_svol_define)
76			Non-public	Un-mount	PSUS	SSUS	–	–		
77	pairresync	Resynchronize								
78			Non-public	Un-mount	COPY	COPY	–	–		
79	pairvolchk	At this point a pair status is COPY								
80	Execute pairvolchk several times		Non-public	Un-mount	COPY	COPY	–	–		
81	pairvolchk	When a pair status change to PAIR, re-synchronizing pair is completed								

No.	Main Site					Remote Site				
	Command	**Processing**	**NFS/ CIFS**	**FS**	**P-VOL Status**	**S-VOL Status**	**FS**	**NFS/ CIFS**	**Processing**	**Command**
82			Non-public	Un-mount	PAIR	PAIR	–	–		
83	enas_fsmount	Mount								
84			Non-public	Mount	PAIR	PAIR	–	–		
85	enas_nfscreate/ enas_cifscreate	Share								
86			Public	Mount	PAIR	PAIR	–	–		
87	Start the business operation at the main site. ***Note 2***									

Note 1: Specify the same file system name as the main site for the file system name connected with NAS unit.

Note 2: When the takeover is performed, the IP address cannot be taken over from the primary site to the secondary site. When starting a job at the secondary site, un-mount the client from the primary site, change the IP address of the site on which the client is to be mounted to that of the secondary site, and then mount the client again. When the job is resumed at the primary site because the primary site has been restored, un-mount the client from the secondary site, return the IP address of the site on which the client is to be mounted to that of the primary site, and then mount the client again.

Note 3: In the NAS Sync Image Modular target file system, un-mount the differential-data snapshot using the **syncumount** command before executing the **enas_fsdelete** command, and release the device which stores the differential data using the **syncstop** command.

Note 4: The time required for the response to a command varies depending on the size of the differential data in order to respond to the host after the guarantee of the S-VOL data is completed.

Acronyms and Abbreviations

Acronym	Expansion
AMS	Adaptable Modular Storage
API	application programming interface
ATM	asynchronous transfer mode
CCI	command control interface
CFW	cache fast write
CKD	count-key data
CLI	command line interface
CTG	consistency group
CTL	controller
D-CNT	default (owner) controller
DM-LU	differential management logical unit
DWDM	dense wavelength division multiplexer
FC	fibre channel
HA	high availability
HACMP	high availability cluster multi-processing
HBA	host bus adapter
HDLM	Hitachi Dynamic Link Manager
HORCM	Hitachi Open Remote Copy Manager
H-LUN	host logical unit
H-RAIN	heterogeneous redundant array of independent nodes
HSN	hierarchical star network
HWM	high water mark
IDE	integrated drive electronics; see also ATA.
IIS	Internet Information Service

IOPS	input output operations per second
IOS	internet work operating system
iSCSI	internet small computer system interface
LDEV	logical device
LDM	logical device manager
LUN	logical unit number
LUSE	LU size expansion
LVI	logical volume image
LVM	logical volume manager
MCU	main control unit
MDB	master directory block
MU	mirror unit
NAS	network attached storage
NFS	network file system
NIS	network information service
NNC	network node controller
NSC	network storage controller
NTP	network time protocol
OSI	open systems interconnection
PFUS	pool full status
PSUE	pair suspended-error status
PSUS	pair suspended-split
PSUS(N)	pair suspended – not restored status
P-VOL	primary volume
RAID	redundant array of independent disks
RPO	recovery point objective
RTO	recovery time objective

SATA	serial ATA
SIM	service information message
SM	shared memory module
SMB	server message block
SNIA	Storage Networking Industry Association
SNMP	simple network management protocol
SONET	synchronous optical network
SSL	secure socket layer
SSWS	suspend for swapping S-VOL
S-VOL	secondary volume
TID	target identifier
USP	Universal Storage Platform
VCS	Veritas Cluster Server™
VDE	Verband Deutscher Elektrotechniker
VIB	volume information block
V-VOL	virtual volume (Snapshot Image)
VxVM	Veritas Volume Manager
WDM	wavelength division multiplexing

Glossary

access control list	An access control list (ACL) is a table that defines the access rights of each user any file directory or individual file.
arbitrated loop (AL)	A system configuration that allows multiple ports to be connected serially.
archive	An archive is an entire body of data stored in an HCA cluster, including both fixed-content data and metadata. An archive does not include the HCA software.
asynchronous	The term asynchronous is used to describe data communications between computers and devices which occurs intermittently rather than in a steady stream. Communication within a computer, however, is usually synchronous and is governed by the microprocessor clock.
attribute	An attribute is one or more qualities possessed by an object.
background copy	A physical copy of all tracks from the source volume to the target volume.
bind	To bind is to assign a value to a symbolic placeholder. For example, when a program is bound, or linked, the binder replaces the symbolic addresses in the code with real machine addresses.
cache	Cache is a temporary, high-speed storage mechanism. It can be either a reserved section of main memory or an independent high-speed storage device. Two types of caching are found in computers: memory caching and disk caching. Memory caches are built into the architecture of microprocessors and often computers have external cache memory. Disk caching works like memory caching; however, it uses slower, conventional main memory that on some devices is called a memory buffer.
cache fast write (CFW)	CFW is an attribute of record caching in which the cache fast write access function (either Simplex Write or Duplex Write) enables the specified record ID to be placed in the volatile control unit cache when a file-type macro is issued and the cache is available. If the cache is not available, the record is written directly to the DASD surface. A single write is issued to the prime module only or a duplexed write is issued to both the prime and the duplicate

modules.

capacity	Capacity is the amount of information (in bytes) that can be stored on a disk drive. The capacity of a hard disk drive is usually expressed in megabytes. Capacity is the measure of the potential contents of a device; the volume it can contain or hold. In communications, capacity refers to the maximum possible data transfer rate of a communications channel under ideal conditions.
cascade configuration	A cascade configuration is a connection configuration of volume pairs in which a P-VOL or an S-VOL from a pair belonging to one copy function is used as a P-VOL or S-VOL of the other copy function. TCE supports cascading with ShadowImage and SnapShot.
channel	A channel is the path data communication follows between two nodes of a network. It is the link between the central processor and the peripherals. A channel can be the physical cabling that connects the nodes on a network, an electronic signal traveling over a pathway, or a sub-channel in a carrier frequency.
channel adapter (CHA)	Provides the channel interface control functions and intercache data transfer functions. It is used to convert the data format between CKD and FBA. The CHA contains an internal processor and 128 bytes of edit buffer memory.
channel extender	A channel extender is a device used to increase the communication distances between channel-connected mainframe computers or between a computer and peripheral devices such as workstations, printers, and storage devices. Optical fiber channel connections are part of the system.
CIFS	Common Internet File System is a protocol used to expose the contents of an archive. CIFS allows clients to access files on a remote Windows computer as if they were part of the local file system.
cluster	A cluster is group of disk sectors. The operating system assigns a unique number to each cluster and then keeps track of files according to which clusters they use.
cluster capacity	Cluster capacity is the total amount of disk space in a cluster, excluding the space required for system overhead and operating system. Cluster capacity is the amount of space available for all archive data, including original file data, metadata, and redundant data.
command devices	Command devices are dedicated logical volumes that are

used only by management software such as CCI, to interface with the storage subsystems. Command devices are not used by ordinary applications. Command devices can be shared between several hosts. Up to two command devices can be configured per TCE subsystem.

concurrency of S-VOL	A state resulting when an S-VOL is synchronized, by simultaneously updating S-VOL with P-VOL data and data cached in the primary host memory. There may be discrepancies in S-VOL data, if data is cached in the primary host memory between two write operations. This data which is not available on the P-VOL, is not reflected on to the S-VOL. To ensure concurrency of S-VOL, cached data is written onto the P-VOL before subsequent remote copy operations.
concurrent copy	Concurrent copy is a combined hardware, license, and software systems management solution that creates data dumps or copies while other applications are updating that data, allowing end-user processing to continue. Concurrent copy allows you to update the data in the files being copied, but the copy or dump of the data it secures does not contain any of the intervening updates.
configuration	Configuration for hardware involves setting various switches and jumpers. For software it means defining the values of parameters. For hardware and software respectively, configuration is the arrangement of the components that make up the system or the set up and set values of the software.
configuration definition file	The configuration definition file describes the system configuration for making CCI operational in a TCE environment. The configuration definition file is a text file created and/or edited using any standard text editor, and can be defined from the PC where the CCI software is installed. The configuration definition file describes configuration of new TCE pairs on the primary or remote subsystem.
consistency group	A consistency group is a group of two or more logical units in a file system or logical volume. When a file system or a logical volume which stores application data, is configured from two or more logical units, these multiple logical units are managed as a consistency group (CTG) and treated as a single entity. A set of volume pairs can also be managed and operated as a consistency group.
consistency of S-VOL	A state in which a reliable copy of S-VOL data from a

	previous update cycle, is available at all times on the remote subsystem. A consistent copy of S-VOL data is internally pre-determined during each update cycle and maintained in the remote data pool. When remote takeover operations are performed, this reliable copy is restored to the S-VOL, eliminating any data discrepancies. Data consistency at the remote site enables quicker restart of operations upon disaster recovery.
console (administrative)	The cluster-specific web application that allows monitoring and managing of a HCA cluster and its individual nodes.
control unit (CU)	The control unit is a CPU component that implements microprocessor instructions.
count-key data (CKD)	Format for encoding data on hard drives, typically used in the mainframe environment. It is a physical disc format (Count, Key, Data) introduced by IBM with Series/360 2311 disks in 1964. Count-key-data (CKD) disks format each track as a new file is written on that track (all files have at least one track). CKD disks like SCSI and IDE have sector (count) ID fields and data fields. CKD disks can also have a third kind of field between ID and data called a key field.
cycle time	Cycle time is a user specified time interval used to execute recurring data updates for remote copying. Cycle times are set for each subsystem and are calculated based on the number of CTGs.
cycle update	Cycle update processing involves periodically transferring differential data updates from the P-VOL to the S-VOL. TCE remote replication processes are implemented as recurring cycle update operations executed in specific time periods (cycles).
dark fiber	A dark fiber is an optical fiber cable that has been physically laid but is still unactivated.
data pools	A data pool is a group of one or more disk volumes, designated to temporarily store untransferred differential data (in the local subsystem) or snapshots of backup data (in the remote subsystem). The saved snapshots are useful for accurate data restoration (of the P-VOL) and faster remote takeover processing (using the S-VOL).
data volume	A data volume is a volume that stores database information, whereas other files, such as index files and data dictionaries, store administrative information, known as metadata.
dataset	A dataset is a named collection of data in an IBM

	mainframe operating system. A dataset in an IBM mainframe is the equivalent of a file in other operating systems, such as Mac OS, Windows and UNIX, for PCs.
device emulation	See logical volume image (LVI).
Differential Management Logical Unit (DM-LUs)	Differential management logical units are logical units used to manage differential data in a storage subsystem. In a TCE system, there may be up to two DM-LUs configured per subsystem.
differential-data	Differential-data is a set of snapshot data that consists of only the data that changed since a previous snapshot was taken. A backup of differential-data is called an incremental backup.
direct access storage device (DASD) fast write (DFW)	DFW is an attribute of record caching (while DASD Fast Write Access is a function of record caching) in which a specified record ID is placed in the cache and nonvolatile storage when a file-type macro is issued. If the cache is not available or the nonvolatile storage is not available, the record is written directly to the DASD surface.
disaster recovery	Disaster recovery implies procedures executed to recover critical application data and processing after a disaster. Disaster recovery processes include failover and failback procedures.
disk array device	In this manual, a disk array subsystem is sometimes referred to as a disk array device.
disk controller unit (DKC)	A disk controller unit consisting of CHA, CHF, DKA, Cache and other components except DKU.
DNS manager	The Domain Name System (manager) provides host-name resolution services to clients. It also balances requests across all nodes to ensure maximum cluster throughput and availability.
dual copy	Dual copy is simultaneously updating of a P-VOL and S-VOL using a single write operation.
emulation	Emulation is the ability of a program or device to imitate another program or device. Emulation causes a software package to accept that a device it is talking to is really another device. At the system level, emulation is a package of hardware, firmware, and software able to recreate a target machine environment on a new system.
extender	An extender is a converter used to change signals when data is transmitted over long distances. For instance,

	changing a fiber channel signal to a signal for dark fiber or an Ethernet (IP).
extent	An extent is a contiguous area of storage in a computer file system that is reserved for writing a file.
fabric	The entity which interconnects various N-Ports attached to it and is capable of routing frames.
failover	A failover operation involves takeover of critical application processing by an alternate host in the event of a failure at the primary site. Automatic switching of hosts is possible using High Availability software.
fallback	Fallback refers to the process of restarting business operations at a local site using the P-VOL, after the storage subsystems have been recovered.
fibre channel (FC)	Input/output channel using optical fiber cables. It is the physical media that forms the lowest layer of fibre channel transport.
firmware	Microcode complementing the hardware used to implement the architecture of a system.
fixed block architecture	A model of disks in which storage space is organized as linear, dense address spaces of blocks of a fixed size. Abbreviated FBA. Fixed block architecture is the disk model on which SCSI is predicated. cf. count-key-data.
fixed-content data	Fixed-content data is an exact digital reproduction of a data file as it existed before the file was archived. Fixed-content data cannot be modified or deleted before its retention period expires.
gateway	A gateway is a protocol that provides users and applications access to data in a cluster.
granularity of differential data	The granularity of differential data refers to the size or amount of data transferred to the S-VOL during an update cycle. Since only the differential data in the P-VOL is transferred to the S-VOL, the size of data sent to S-VOL is often the same as that of data written to P-VOL. The amount of differential data that can be managed per write command is limited by the difference between the number of incoming host write operations (inflow) and outgoing data transfers (outflow).
HCA cluster	An HCA cluster is an implementation of Hitachi Content Archiver. An HCA cluster is both a repository that stores terabytes of data and a gateway that enables access to

	that data.
High Availability (HA) software	High Availability software is used for automatically switching to a stand-by host, in the event of primary host or disk failure. High availability software has to be installed on the primary and secondary hosts.
H-RAIN	Heterogeneous Redundant Array of Independent Nodes. A collection of networked servers that can differ by vendor or model.
HTTP	Hypertext Transfer Protocol is one of the protocols used to expose the contents of an archive. Using HTTP, archived files and directories can be viewed on a web page.
HTTPS	HTTPS is HTTP with SSL security.
ICKDSF	A DSF command used to perform media maintenance.
initial copy	An initial copy operation involves copying all data in the primary volume to the secondary volume prior to any update processing. Initial copy is performed when a volume pair is created.
initiator ports	A port-type used for MCU port of Fibre Remote Copy function.
iSCSI	iSCSI (Internet-Small Computer Systems Interface) is used as an IP-based standard for carrying SCSI commands over IP networks which link data storage devices and allows the transfer of data.
Java applet	A Java applet is a program written in the JavaTM programming language that can be included in an HTML page. When you use a Java technology-enabled browser to view a page that contains an applet, the applet's code is transferred to your system and executed by the browser's Java Virtual Machine (JVM).
journal volume	A journal volume is a volume in which the hard disk maintains data integrity in the event of a system crash. The journal volume maintains a log or journal of the activities that have taken place in the volume. This journal allows any lost data to be recreated, because updates to the metadata in directories and bit maps have been written to a serial log.
logical	Logical is used to describe a user's view of the way data or systems are organized. The opposite of logical is physical, which refers to the real organization of a system. A logical description of a file is that it is a quantity of data collected

	together in one place. The file appears this way to users. Physically, the elements of the file could live in segments across a disk.
logical device (LDEV)	A logical device is a group of hardware items that the operating system treats as a single unit.
logical unit number (LUN)	LUN is a three-bit code identifier for a logical unit. LUN0-7 can be assigned.
logical volume	An area on a disk consisting of device files that are logically integrated using a volume manager.
logical volume image (LVI)	LVI (also called device emulation) is a feature used to create virtual LUs that are up to 36 times larger than the standard OPEN-x LUs.
mainframe	A mainframe is a large and expensive computer capable of simultaneously supporting thousands of users. In this document the term mainframe is used for IBM computers (zSeries® and S/390®-based systems). This term also marks a distinction between Unix or Windows server computers and the larger more, powerful mainframe. A mainframe is also commonly referred to as the *host*, even though any computer host having a unique IP address can equally be referred to as a *host*.
metadata	Metadata is information about an archived object.
microcode	Microcode is the lowest-level instructions directly controlling a microprocessor. Microcode is generally hardwired and cannot be modified.
Microsoft cluster server	Microsoft Cluster Server is a clustering technology built that supports clustering of two NT servers to provide a single fault-tolerant server.
middleware	Middleware is software that connects two otherwise separate applications. For example, a middleware product can be used to link a database system to a Web server. Using forms, users request data from the database; then, based on the user's requests and profile, the Web server returns dynamic Web pages to the user.
mount	To mount a device or a system means to make a storage device available to a host or platform.
mount point	Mount points are the location in your system where you mount your file systems or devices. For a volume that is attached to an empty folder on an NTFS file system volume, the empty folder is a mount point. In some

	systems a mount point is simply a directory.
multiple allegiance support	The multiple allegiance support means that the storage unit can accept concurrent I/O requests for a volume from multiple channel paths. Therefore the storage unit can process requests from separate FICON hosts in parallel, improving throughput and performance.
multiple virtual storage	MVS (including MVS/370, MVS/ESA, MVS/XA) is an operating system that runs on IBM or compatible mainframe computers. The host component works on MVS/ESA (Enterprise Systems Architecture) and MVS/XA (Extended Architecture).
NDMP	Network Data Management Protocol is a protocol used to backup and restore archived objects.
NFS	Network File System is a protocol used to expose the contents of an archive. NFS allows clients to access files on a remote computer as if they were part of a local file system.
node	In networks, a node is a processing location. A node can be a computer or some other device, such as a printer. Every node has a unique network address. In HCA clusters, nodes are Linux-based servers running HCA software and networked to form an HCA cluster.
open device	Collectively refers to the host computer, peripheral control units, and intelligent peripherals that are connected to fibre channel.
pair splitting	Pair splitting refers to the termination of a volume pair relationship to temporarily stop update copy processing for the specified volume pair. Pairs may also be split before system reduction tasks.
pair status	A pair status is an internal status assigned to a volume pair before or after pair operations. Pair status transitions occur when pair operations are performed or as a result of failures. Pair statuses are used to monitor copy operations and detect system failures.
paired volumes	Paired volumes are primary and secondary volumes comprising a volume pair.
panel	In this document a *panel* is equivalent to a window.
parity groups	RAID groups can contain single or multiple parity groups where the parity group acts as a partition of that container.

path blockade watch	The path blockade watch setting specifies the time for monitoring blockade in the Fibre Channel paths on the MCU side. The path blockade watch value must be from 0 to 45 seconds. This setting is available for Fibre Channel interface only.
pattern file	A pattern is a table that contains the access attributes of all logical volumes. Pattern files enable administrators to change the access attributes of all logical volumes quickly and easily.
permission	Using SMB/CIFS, HTTP, WebDAV, and SMTO gateways, permissions are granted from the owner to members of a group or other users to allow access to data files or directories in an archive. Read, write, and execute permissions can be granted for files, directories, or symbolic links.
point-in-time logical copy	A logical copy or snapshot of a volume at a point in time. This allows a backup or mirroring application to run concurrently with the system.
point-to-point	A configuration that allows two ports to be connected serially.
policy	A policy is a process that performs a specific function to aid maintenance of the overall health of a cluster. For example, authentication, cluster balance, garbage collection, protection, retention, and scavenging.
pool volume	A pool volume is used to store backup versions of files, archive copies of files, and files migrated from other storage.
primary or local site	A site where the production applications run.
primary volume (P-VOL)	A primary volume is the storage volume in a volume pair, used as the source of a copy operation. In copy operations a copy source volume is called the "P-VOL" while the copy destination volume is called "S-VOL" (secondary volume).
quota values	Quota values are set for users with write access to volumes providing data storage limits for that user/volume. Quota values are applied to a snapshot and set for the target file system when the snapshot is taken.
RCU target port	A port-type used for RCU port of Fibre Remote Copy function. This port allows LOGIN of host computers and MCUs.

recovery point objective	Recovery point objective is the maximum desired time period prior to a disaster, during which changes to data may be lost as a result of recovery. This measure determines up to what point in time data should be recovered in the event of a disaster. Data changes preceding the disaster by at least this time period are preserved by recovery.
recovery time objective	Recovery time objective is the maximum desired time period required to bring one or more applications and associated data back to a correct operational state. It defines the time frame within which specific business operations or data must be restored to avoid any business disruption.
remote or target site	A site that has the mirrored data of the primary site.
remote path	A remote path is a route connecting identical ports on the local subsystem and the remote subsystem. Two remote paths must be setup for each subsystem, (one path for each of the two controllers built in the subsystem).
remote volume (R-VOL)	In TrueCopy operations, the remote volume (R-VOL) is a volume located in a different subsystem from the primary host subsystem.
repeater	A repeater is a network device used to regenerate or replicate a signal. A repeater relays messages between sub-networks that use different protocols or cable types. A repeater cannot do the intelligent routing performed by bridges and routers.
resynchronization	Resynchronization is a copy operation performed to make data in the secondary volume consistent with data in the primary volume. This operation involves copying only untransferred differential data to the target secondary volume.
secondary volume (S-VOL)	A secondary volume (S-VOL) is a replica of the primary data volume (P-VOL), maintained on the standby subsystem. Recurring differential data updates are performed to keep the data in the S-VOL consistent with data in the P-VOL.
sequential data striping	Sequential data striping refers to writing to multiple disk drives in a pre-planned sequence. Because the processor writes faster than the disk can accept, it has the left-over capacity to locate the next segment of the logically sequential data and prepare to write to it.

server set identifier	The SSID is an alphanumeric name that is 1-32 bytes. The purpose of an SSID is to help hardware clients find and connect to an access point (AP) on the correct network.
service	A service is the set of functions that one of the seven (7) Open Systems Interconnection (OSI) model layers delivers to the layer above it. For example, the TCP layer provides a reliable byte-stream service to the application layer above it.
session	A session is a series of communications or exchanges of data between two end points that occurs during the span of a single connection. The session begins when the connection is established at both ends and terminates when the connection is ended. For some applications each session is related to a particular port. In this document a session the exchange of data between groups of primary and secondary volumes.
ShadowImage	ShadowImage is a software program that replicates user data on TagmaStore AMS/WMS disks, bypassing the host system.
shared memory module (SM)	Stores the shared information about the system and the cache control information (director names). This type of information is used for the exclusive control of the system. Like CACHE, shared memory is controlled as two areas of memory and fully non-volatile (sustained for approximately 7 days).
sidefile	A sidefile is an area of a controller's storage that occupies about 1MB of the controller storage. The sidefile is used to hold data that is changed or updated while being backed up or copied. The sidefile holds it for later integration into the copied data.
SMB	Server Message Block is a protocol used to expose the contents of an archive. SMB allows clients to access files on a remote computer as if they were part of a local file system.
SMTP	Simple Mail Transfer protocol is a protocol used to receive and store email data directly from email servers.
snapshot	A term used to denote a copy of the data and data-file organization on a node in a disk file system. A snapshot is a replica of the data as it existed at a particular point in time.
SNMP	Simple Network Management Protocol is a protocol used to

	facilitate monitoring and management of clusters through an external interface. SNMP sends notifications to IP addresses whenever certain types of events occur.
source copy	Source copy is the place from which data is taken. The place from which the data is moved is called the source. The source can also indicate the node on a network from which data is sent to its destination.
SSL	Secure Sockets Layer is a key-based Internet protocol for transmitting documents through an encrypted link.
SSL certificate	An SSL certificate is a file containing the cryptographic keys and signatures used with an SSL protocol to verify the authenticity of web sites to protect data sent to or from that site.
S-VOL determination	S-VOL determination is an internal process on the remote subsystem, which replicates the S-VOL independent of the update operations. This process occurs at the end of each update cycle. This process allows a pre-determined copy of S-VOL data consistent with P-VOL data (as of the previous cycle) to be maintained on the remote site at all times.
sysplex	Sysplex denotes *system complex*. This is a processor complex formed by connecting a number of processors together into a single unit through channel-to-channel adapters or ESCON/FICON fiber optic links. The processors are synchronized using a Sysplex Timer and are managed as a single system image (SSI1). The Sysplex Timer is an invaluable component when systems on multiple CPCs share access to the same data.
system reduction	System reduction refers to maintenance tasks performed to improve system performance. These tasks may include pair deletion, deletion of command devices and data pools.
takeover processing	Takeover processing involves transferring of critical application processing to the S-VOL on the remote standby subsystem. The remote S-VOL is immediately enabled to process subsequent host I/O operations.
target copy	Target copy is a file, device or any type of location to which data is moved or copied.
target port	A port-type which is different from "Initiator Port" and "RCU Target Port". This port is a normal target port which is used without configuration of Fibre Remote Copy. This "Target port" allows LOGIN of host computers. It does not allow LOGIN of MCUs.

tier architecture	Tier 1 is fully supported computing expected to be production quality. Tier 2 platforms are not supported by the security officer and release engineering teams. Tier 2 systems are targeted at Tier 1 support, but are still under development. Tier 3 platforms are architectures for which hardware is not or will not be available or which are considered legacy systems unlikely to see broad future use. Tier 4 systems are not supported in any way. Tier 1: Static content, Tier 2: Application logic, Tier 3: Database
track	A track is a ring on a disk where data can be written. For hard disks, tracks aggregate into platters and a single track location that cuts through all platters is termed a cylinder. Each track can be subdivided into a number of sectors. The operating system and disk drive find stored information using its track and sector numbers.
truck size	Truck size represents a fixed sector size for each volume type.
TrueCopy	TrueCopy is a software program that replicates data between two TagmaStore disks, independent of the host system. TrueCopy versions are available for TagmaStore AMS/WMS and USP/NSC subsystems. TrueCopy for z/OS is a mainframe version.
user logical unit (LU)	A user logical unit is a term used to describe any device file located on an external disk subsystem connected to the TagmaStore subsystem by a fibre channel.
virtual LVI/ LUN (VLL)	Virtual LVI/LUN is an option that enables the configuration of custom-size logical device images and logical units, which are smaller than standard size devices.
volume	A volume is the basic unit of storage that includes recovery logs and storage pools. A volume can be a logical volume management (LVM) logical volume, a standard file system file, a tape cartridge, or an optical cartridge. The various types of defined volumes include: external, internal, copy source, copy destination, reserve, data, journal, virtual, pool, system, LUSE, copy pair, and USP.
volume pair	A volume pair is formed by pairing two logical data volumes. It typically consists of one primary volume (P-VOL) on the local storage subsystem and one secondary volume (S-VOL) on the remote storage subsystems. TCE remote copy operations are performed using logical volume pairs.

volume signature	A volume signature is an integer that is an element in the master directory block (MDB) (volume information block (VIB)). For example, for HFS volumes, this field (drSigWord) contains the number $4244.
WDM/DWDM	WDM denotes Wave Division Multiplexing technology used to multiplex optical signals from multiple channels. DWDM denotes Dense Wave Division Multiplexing and is used to multiplex optical signals of several dozen channels.
World Wide Name (WWN)	A unique identifier for an open systems host. It consists of a 64-bit physical address (the IEEE 48-bit format with a 12-bit extension and a 4-bit prefix). The WWN is essential for defining the SANtinel™ parameters because it determines whether the open systems host is to be allowed or denied access to a specified LU or a group of LUs.
WORM	Write once, read many is a data storage technique in which files are protected from being modified, overwritten, or deleted.
write order guarantee	The write order guarantee feature ensures that data is updated in S-VOL in the same order that the host was updated data in the P-VOL, particularly when there are multiple write operations in one update cycle. This feature is critical to maintain data consistency in the remote S-VOL and is implemented by inserting sequence numbers in each update record. Update records are then sorted in the cache within the remote system, to assure write sequencing.

Index