

Usage Guide

iStorage Series Disk Array System
with
Data Trust for Oracle

iS-TI-02-002
Rev-1.10 April 2003
NEC Corporation.

Copyright (C) 2003 NEC Corporation. All rights reserved.

THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND WHATSOEVER, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE WARRANTIES REGARDING MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT OR ACCURACY OF DATA.

iStorage, Data Trusty, iStorageManager and WebSAM are trademarks or registered trademarks of NEC Corporation in Japan.

Oracle, Oracle9i and other product names of Oracle are trademarks or registered trademarks of Oracle Corporation in the U.S.A. and other countries.

CONTENTS

1	INTRODUCTION	4
2	OVERVIEW OF DATA TRUST FOR ORACLE	4
2.1	OVERVIEW	4
2.2	DATA CORRUPTION	4
2.3	ORACLE CHECKSUM	4
2.4	DATA TRUST FOR ORACLE.....	5
2.5	ORACLE H.A.R.D. INITIATIVE	7
3	SYSTEM CONFIGURATION OF DATA TRUST FOR ORACLE	7
3.1	HARDWARE REQUIREMENTS.....	7
3.2	SOFTWARE REQUIREMENTS.....	7
3.3	ENVIRONMENT CONFIGURATION	7
3.4	COMBINED USE WITH VOLUME MANAGER	8
4	OPERATIONS OF DATA TRUST FOR ORACLE	8
4.1	OVERVIEW	8
4.2	ISMCHKSUM COMMAND.....	9
4.3	AT DATABASE SETUP	9
4.4	ACTIONS UPON ILLEGAL WRITE	9
4.5	SETTING CONFIRMATION.....	9
4.6	ANALYSIS UPON DETECTING ILLEGAL WRITE.....	9
4.7	CORRUPTED DATA.....	10
4.8	TEMPORARILY DISABLING CHECK TARGET.....	10
4.9	PERMANENTLY DISABLING CHECK TARGET	10
5	CONSIDERATIONS IN SYSTEM OPERATION	11
5.1	ACTIONS UPON OCCURRENCE OF ERROR.....	11
5.2	BACKUP AND RESTORATION.....	11
5.3	DELETION OF ORACLE DATABASE	12
5.4	AFFECT ON PERFORMANCE.....	12
6	SUMMARY.....	12
7	REFERENCES	13
	ACKNOWLEDGEMENTS.....	13

1 Introduction

The importance of the data accumulated by business every day is increasing, and if the consistency of the data accumulated on the disk array is lost, it will have a serious effect on business. Data Trust for Oracle is a function of the iStorage series disk array, and improves the reliability of the data stored in the Oracle database by performing consistency verification between the Oracle database and the disk array. Data Trust for Oracle is provided as a software product named “Data Trusty for Oracle” by NEC Corporation.

2 Overview of Data Trust for Oracle

2.1 Overview

Data Trust for Oracle protects the database from data damage (*) caused by hardware or software faults or human error, by checking the checksum within the Oracle database block in the iStorage disk array system before data is actually written to the disk.

(*) The state in which the data of the Oracle database is corrupted. Although each data transfer mechanism is generally equipped with its own data checking mechanism, it only checks for data damage in the transmission unit, etc., independently of the data structure of the higher layer applications.

2.2 Data Corruption

The following are the causes of corrupted data being written to the disk array:

- 1) Damage in the software layer between the Oracle database and the data in the disk array;
- 2) Damage in the hardware layer between the Oracle database and the data in the disk array;
- 3) Application error or human error (e.g. incorrect operation).

In order to prevent corrupted data from being written to the disk array, it is necessary to prevent data corruption from occurring in all of the hardware/software layers from the Oracle database to the disk array, and by human operations.

However, most of the data in each hardware/software layer currently available is only transmitted/recorded from the adjacent layer as is, and consequently the corrupted data in the middle of the layers is written on the disk array in its corrupted state.

2.3 Oracle Checksum

The Oracle database offers a function to verify data consistency by adding a checksum to the Oracle database blocks when writing and by verifying the checksum when reading. If this function detects data corruption from the checksum when reading, damage to business operations or the spread of corrupted data can be prevented by making the tablespace with the corrupted data reside partially offline or by shutting down the entire database. However, when data corruption occurs during writing or corrupted data is written by an application or human operator, the data corruption is not detected until the corrupted data is read, so it

takes a long time to recover the data and check its consistency, meaning that database service suspension will have a huge impact on business operations.

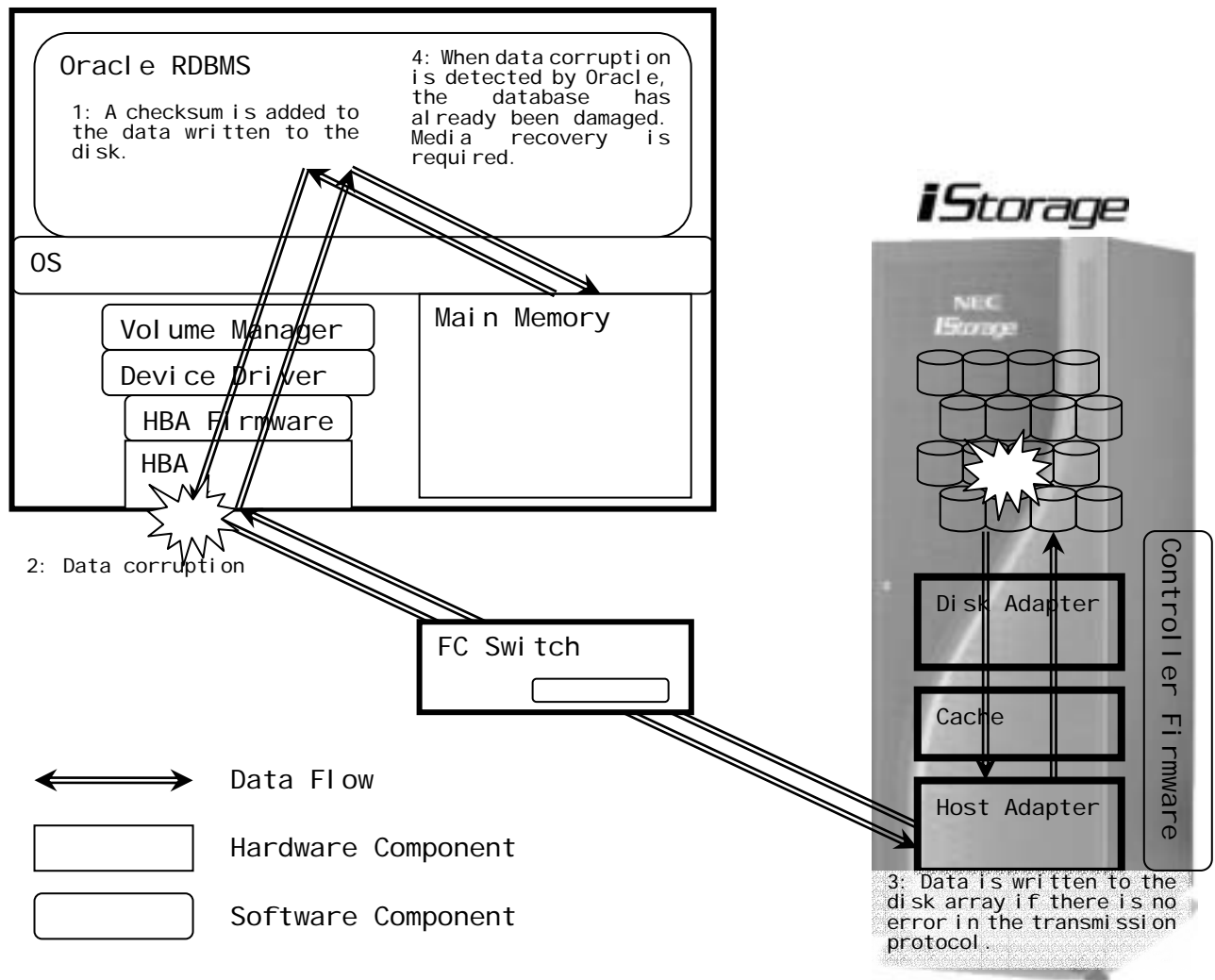


Figure 1: Conventional Oracle database system configuration with storage system

2.4 Data Trust for Oracle

Data Trust for Oracle blocks writing to the disk array when corrupted data has been detected by checking that the Oracle database checksum is the same as that on the disk array side when the disk array is written, using the same technique as the Oracle database checksum check during reading. This function makes it possible to guarantee the consistency of the Oracle database without damaging the data on the disk array.

Moreover, since the check of the Oracle checksum on the disk array side operates independently of the operation of the host computer, even if the host computer running the Oracle database stops, Data Trust for Oracle can detect and prevent corrupted data from

being written from another host computer or the writing of corrupted data due to trouble in related equipment.

To use Data Trust for Oracle, it is necessary to preset the disk space used by the Oracle database to the disk array. Moreover, the performance of the write operations to the disk space that is specified as the Oracle checksum target might decrease due to the overheads required to verify the checksum inside the disk array. When writing of corrupted data is detected by Data Trust for Oracle, it is necessary to check the fault status in each hardware/software layer, replace faulty parts and correct the applicable software. Since Data Trust for Oracle cannot clarify the cause of corrupted data, the following two actions are also required to achieve early solutions to the causes of corrupted data: implementation of a damage detection mechanism for all the products that constitute the database system, and recording of the administration procedures and operations of the database system. The iStorage series disk array has an automatic self-detection and recording mechanism for errors and abnormalities, which can be very helpful for early damage and cause detection.

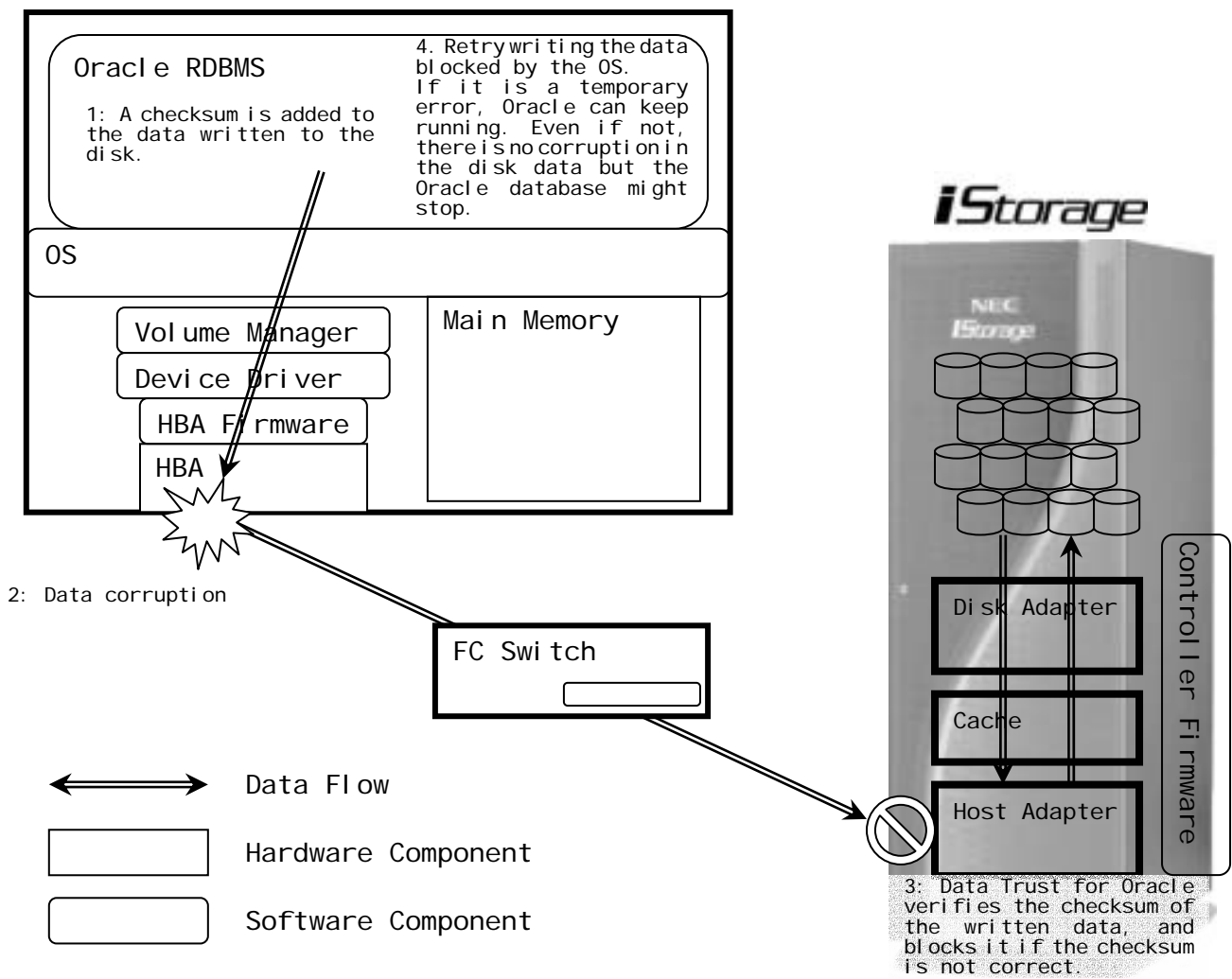


Figure 2: Oracle database system configuration with Data Trust for Oracle

2.5 Oracle H.A.R.D. Initiative

The checksum function of the Oracle database used by Data Trust for Oracle is fully compliant with the specifications of the Oracle H.A.R.D. (*) Initiative provided by Oracle Corporation, and has passed the test suite of Oracle H.A.R.D.

(*) H.A.R.D. = **H**ardware **A**ssisted **R**esilient **D**ata

3 System Configuration of Data Trust for Oracle

3.1 Hardware Requirements

In order to adopt and utilize Data Trust for Oracle, your system must meet the minimum hardware requirements listed below.

- 1) Disk array: S4300 for high-end or S3300 for mid-range disk array system of NEC iStorage series disk array.
- 2) Database server: Server computer running Oracle database.

3.2 Software Requirements

The software requirements for the iStorage series disk array and Data Trust for Oracle are listed below.

- 1) Data Trust for Oracle: Software to control the Oracle checksum function with the iStorage series disk array, and that is installed in the database server.
- 2) Operating system: Operating system of the database server.
- 3) Oracle database: Oracle database software.
- 4) WebSAM iStorage Manager: Storage management software that provides functions to monitor the system configuration and status of the iStorage series disk array.

Please refer to the product manuals of Data Trust for Oracle regarding detailed platform information (server hardware and operating system) and the software version.

3.3 Environment Configuration

The configuration of the Oracle database is as follows:

- 1) Oracle checksum enable

In order to enable Data Trust for Oracle, the Oracle data block checksum function should be enabled by specifying the following initialization parameter.

In the init<SID>.ora file:

```
DB_BLOCK_CHECKSUM = TRUE
```

After restarting the Oracle database with the above parameter, the Oracle checksum function becomes valid. This parameter is true by default from Oracle 9i.

- 2) Types of Oracle files to be checked

The types of Oracle files for Data Trust for Oracle are the same as the ones of the Oracle checksum function. When applying this function on Oracle 9iR2, the following three types

of Oracle files are available.

- 1 Data file
- 2 Control file
- 3 REDO log file

3) Oracle block size

Multiple Oracle block sizes within one Oracle database are available from Oracle9i. Data Trust for Oracle can verify the multiple Oracle block sizes listed below.

2048 bytes
4096 bytes
8192 bytes
16384 bytes
32768 bytes

Please refer to the product manual of the Oracle database for specifying multiple Oracle block sizes within one Oracle database.

3.4 Combined Use with Volume Manager

Data Trust for Oracle is only available with the logical raw device files created by using Volume Manager. Please refer to each product manual for the version and usage requirements of the available Volume Manager products.

1) Notes for the combined use with Volume Manager

If Volume Manager cannot handle the checksum error detected by Data Trust for Oracle, an error for the logical volume that contains the raw devices used by Oracle may remain on the monitor console in spite of checksum error dissolution. But this error does not affect the application and database operations. Please perform the required operations, such as compulsive release of error output on the display, by using the Volume Manager functions. Refer to the product manuals of Volume Manager for function details.

2) Use with file system

Data Trust for Oracle cannot be applied to files created by the file system. Please use the raw device files created by Volume Manager.

3) Use with OS-level raw device files

Do not use OS-level raw device files with Data Trust for Oracle. Some software continues to retry write operations infinitely, even if an error occurs that should stop operations during the file I/O procedure. If Data Trust for Oracle detects corrupted data and blocks the write operation of such software, the corresponding software cannot continue to work due to infinite retries. This is why the above condition exists.

4 Operations of Data Trust for Oracle

4.1 Overview

Data Trust for Oracle is an optional command of the WebSAM iStorageManager storage management software.

Data Trust for Oracle is managed by the iSMChksum command from the host computer

that is running the Oracle database or from the management server of the iStorage series disk array. Since the iSMChksum command can manage all devices (raw device files) used by the targeted Oracle database instances, it can prevent operational mistakes such as the omission of target device information when the Oracle database uses a lot of devices.

4.2 iSMChksum Command

This is the main command of Data Trust for Oracle and is invoked from the command line of the host computer running the Oracle database. This command should be executed under super user (root) privileges. For more details, refer to the manual of Data Trust for Oracle. The typical command format of iSMChksum is as follows. For details of sub commands (subcmd) or options, refer to the manual.

```
# iSMChksum subcmd -option arg -option ...
```

Examples of basic operations using iSMChksum are described below.

4.3 At Database Setup

By using the 'set' sub command, you can register a newly installed or existing Oracle database as the target of Data Trust for Oracle.

```
# iSMChksum set -sid orclSID -user system/manager -all
```

orclSID is the SID of the Oracle database instance which is registered to Data Trust for Oracle. The username and password with SYSDBA privileges in the Oracle database instance are specified as the arguments of the '-user' option. By utilizing the '-append' option, you can append a new target to a disk array for which targets have already been set.

4.4 Actions upon Illegal Write

With the 'change' sub command, you can specify the action when an illegal write to a target device is detected.

```
# iSMChksum change -sid orclSID -user system/manager -illegalwrite reject
```

The following two actions can be specified by using the '-illegalwrite' option:

reject: Block the writing of corrupted data

accept: Perform the write even if the data is corrupted. This option can be applied only for testing purposes, and must not be used in normal operations.

4.5 Setting Confirmation

The 'query' sub command is used to confirm the settings.

```
# iSMChksum query -sid orclSID -user system/manager -all
```

4.6 Analysis upon Detecting Illegal Write

When an illegal write is detected, the 'queryerrlog' sub command is used to identify the Oracle database and the physical file to which the illegal write was performed. In addition,

error information is stored in the syslog file and displayed on the console of the iSM management client.

```
# iSMchksum queryerrlog -sid orclSID -user system/manager -all
```

Identification of the database instance

The contents of error information stored in the syslog file or displayed on the iSM management client is the device unit of the disk array. Therefore, when the iStorage series disk array is shared by multiple Oracle instances, you can identify in which Oracle instance the data error occurred by executing the 'queryerrlog' sub command from the iStorage Manager using a specific device name.

```
# iSMchksum queryerrlog -noconnect -arrayname S4300/01 -ldn 0h
```

The above example displays the checksum error detected by the specific disk array named S4300/01. The argument of the '-ldn' option, 0h, indicates the logical device number (LDN) of S4300/01 where the Oracle database file is located. Refer to the manuals of the iStorage disk array and WebSAM iStorageManager for details of the management procedures, such as the retrieval method of LDN. In addition, if the iStorageManager manages two or more servers running Oracle databases, you need to clarify the server name to identify the place where the error occurred. You can obtain the server name from the error information stored in the syslog file.

4.7 Corrupted Data

Since writing of corrupted data is blocked by Data Trust for Oracle, the Oracle database data on the disk array is never damaged. ('reject' is the default action of the 'change' sub command.)

4.8 Temporarily Disabling Check Target

The functions of Data Trust for Oracle need to be temporarily disabled when executing database backup and/or restoration using a backup tool. Use the 'disable' and 'enable' sub commands to suspend and resume the functions of Data Trust for Oracle.

```
# iSMchksum disable -sid orclSID -user system/manager -all
```

After completion of backup and/or restore operations, you can resume checking by using the 'enable' sub command.

```
# iSMchksum enable -sid orclSID -user system/manager -all
```

4.9 Permanently Disabling Check Target

When you no longer need to use the iStorage series disk array for Oracle database operations, you have to disable the use of Data Trust for Oracle for the disk array. The 'cancel' sub command is used to do this.

```
# iSMchksum cancel -sid orclSID -user system/manager -all
```

By specifying the Oracle instance SID, all the devices used by the Oracle database can be

anceled collectively, which helps to prevent operational mistakes such as the omission of targets to be disabled. After the 'cancel' sub command is applied, all the settings for targets are removed.

5 Considerations in System Operation

5.1 Actions upon Occurrence of Error

The occurrence of a corrupted data write to the target devices can be checked from the syslog file or the display on the iSM management client.

Because an illegal write caused by human error is blocked by Data Trust for Oracle and has no effect on the database data on the disk array, you can continue database operations without concern. If the corrupted data is caused by a temporary hardware error but the corrupted data is corrected by retry processing at the OS level, the correct data is written to the disk array in the end, so database operations can be continued.

If the corrupted data write is caused by faults such as a permanent hardware error, the Oracle database receives the write error from underlying layers and makes the corresponding tablespace offline or shuts down the database automatically. Since the database data in the disk array itself is protected from corruption even in such a situation, you can make the tablespace online or restart the database after replacing and repairing the equipment and parts that caused the hardware error. However, in some cases, you may have to perform a database recovery operation when starting up the database.

WebSAM iStorage Manager, which is the integrated management software of the iStorage series disk array, also provides functions to report abnormal information from the disk array to the system administrator by e-mail, and to invoke recovery programs (shell script etc.). For details of functions such as e-mail notification, refer to the manuals of WebSAM iStorage Manager.

Moreover, the iStorage series disk array has powerful data backup functions, `DynamicDataReplication` and `RemoteDataReplication`. For more details, refer to the manuals of the iStorage series disk array, `DynamicDataReplication` and `RemoteDataReplication`.

5.2 Backup and Restoration

While backing up and restoring the Oracle database, Data Trust for Oracle may affect the backup tool's behavior and results. In particular, if the backup tool does not make allowances for the Oracle block size when writing to the files, Data Trust for Oracle sometimes judges correct data to be erroneous. In this case, it is necessary to suspend the functions of Data Trust for Oracle temporarily. However, backing up or restoring the Oracle database can be done with Data Trust for Oracle enabled using the following methods. (Each command and tool can handle the Oracle checksum function.)

- Export/import utility
- RMAN (only Oracle9iR2)(*)

- `DynamicDataReplication` and `RemoteDataReplication` provided by the iStorage disk array.

When using other methods (for example, dump/restore, dd command, etc.), it is necessary to disable the functions of Data Trust for Oracle temporarily using the 'disable' sub command.

(*) Data Trust for Oracle does not check the backup files created by RMAN.

5.3 Deletion of Oracle Database

In order to delete the Oracle database from the system, it is also necessary to remove all the targets of Data Trust for Oracle. This is because write operations to the registered storage devices from an application that is not the Oracle database are blocked by Data Trust for Oracle because the write data does not have the Oracle checksum.

The targets should be deleted before deleting the Oracle database. The `iSMChksum` command can retrieve the information of devices used from the Oracle database instances, so you can remove all the registered devices from the targets of Data Trust for Oracle together in an easy operation before Oracle database deletion.

The sequence of operations is as follows:

- 1) Shut down the Oracle database
- 2) Remove the targets of Data Trust for Oracle
- 3) Delete the Oracle database

5.4 Affect on Performance

Data Trust for Oracle causes a slight increase in the I/O overhead when writing to the disk due to its operating principle. When corrupted data is written, the response (notification of an error) to the Oracle database is delayed until the write to the device (raw device file) reaches the TIMEOUT of the OS processing. Therefore, in cases when data is temporarily damaged but then corrected by the retry processing of the OS and the write finally succeeds, the response performance of the corresponding processing is degraded slightly. When the retry succeeds at the OS level, the subsequent processes are executed normally.

If the data damage is permanent, the tablespace is placed offline or the Oracle database instance stops, so the damaged data needs to be restored and the Oracle database restarted.

6 Summary

Data Trust for Oracle and the iStorage disk array combined with Oracle checksum protects the Oracle database from data corruption in the disk array by detecting and blocking corrupted data before it is written. This is a sophisticated approach to data protection that was not achieved by conventional methods. Moreover, this new approach can protect the Oracle database from human error. These functions increase the reliability and availability of the Oracle database stated by the NEC OMCS (Open Mission Critical System) solution model.

Data Trust for Oracle makes system management of disk arrays in the Oracle database

system easy because Data Trust for Oracle works closely with the Oracle database and manages the device information (raw device files) that the Oracle database uses.

7 References

- [1] Wei Hu, J. Bill Lee and Juan Loaiza, "End-to-End Validation of Oracle Database Blocks," Oracle Corporation, November 2001

Acknowledgements

We would like to thank Mr. Paul Tsien, Mr. Mathew Phillips and Mr. James Viscusi in Oracle Corporation for their helpful reviews and support of our H.A.R.D. project. We also would like to thank colleagues in our company associated with this project for their continuous support and advice.