

# EMC Double Checksum for Oracle

Date 11/1/2001

---

Copyright © 2001 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

EMC, EMC<sup>2</sup>, and Symmetrix are registered trademarks and Double Checksum, Enginuity, and where information lives are trademarks of EMC Corporation.

All other brand names are trademarks or registered trademarks of their respective owners.

Part Number H445 Rev B

Printed 11/1/2001

---

## Table of Contents

<b>Introduction .....</b>	<b>3</b>
Related Documentation .....	3
<b>Business Consequences of Data Corruption .....</b>	<b>3</b>
<b>EMC Double Checksum .....</b>	<b>4</b>
The Data Block Corruption Problem .....	4
Enabling the Oracle Checksum Feature .....	5
The EMC Double Checksum Feature .....	5
Enabling the EMC Double Checksum Feature .....	5
Responding to the Detection of Errors .....	6
Reject I/O Option Selected .....	6
Reject I/O Option Not Selected .....	6
Recovery Actions in Cases of Corruption .....	6
Items to be Aware of With Version 1.0 of EMC Double Checksum .....	7
Performance Considerations .....	7
<b>Examples .....</b>	<b>8</b>

## Introduction

The “simple” task of writing a data block to disk actually involves complex interactions between multiple layers of hardware and software from different vendors. Of these layers, only the database, storage systems, and network protocols generate and verify data integrity checks. The database and storage systems are at the extreme ends of the I/O path. With conventional technology, there is no way of recognizing or responding to data corruption introduced in one of the intermediate layers. Although this type of corruption problem is not common, it can remain undetected for an indefinite period of time. The recovery procedures can require significant downtime and resources and become more demanding with the passage of time.

The Oracle database and the EMC<sup>®</sup> Symmetrix<sup>®</sup> storage system are two industry-proven highly reliable and robust IT infrastructure platforms. Both products have long possessed the ability to reliably protect data within their respective environments. EMC Double Checksum<sup>™</sup> is an innovative solution that extends this protection from the database through all the layers to the Symmetrix. If a corruption is detected, notification occurs quickly.

This white paper outlines the features of EMC Double Checksum and how it works in conjunction with Oracle’s checksum functionality to improve the detection and notification of corruption.

## ***Related Documentation***

The following Symmetrix reference manuals describe the topics discussed in this paper:

- For a list of supported hosts, refer to the *EMC Solutions Enabler Installation Guide* (P/N 300-000-047).
- For more information on managing Symmetrix units and devices in an enterprise storage environment, refer to the *EMC Solutions Enabler SYMCLI Base Component Product Guide* (P/N 300-000-048).
- For more information on the host system mapping of relational databases and how the databases and their structures can be examined, refer to the *EMC Solutions Enabler SYMCLI Mapping Component Product Guide* (P/N 300-000-051).
- For more information on the EMC Double Checksum feature, refer to the *EMC Solutions Enabler SYMCLI V4.3 Checksum Release Notes* (P/N 300-000-046).

## Business Consequences of Data Corruption

Information is becoming increasingly important for today’s businesses. A company that is not able to deliver accurate information effectively and quickly will find itself falling behind the competition. One of the major factors that can prevent a company from delivering valid information in a timely manner is data corruption. This can have severe penalties for any business; it can cause system downtime and require significant resources (time and personnel) to remedy, and lead to millions of dollars in lost revenue.

The impact on customers can be just as dramatic. Customers will remember that their stock trade did not go through, or that they were unable to order products online, or that their request was not serviced in a timely manner, and they will take their business elsewhere.

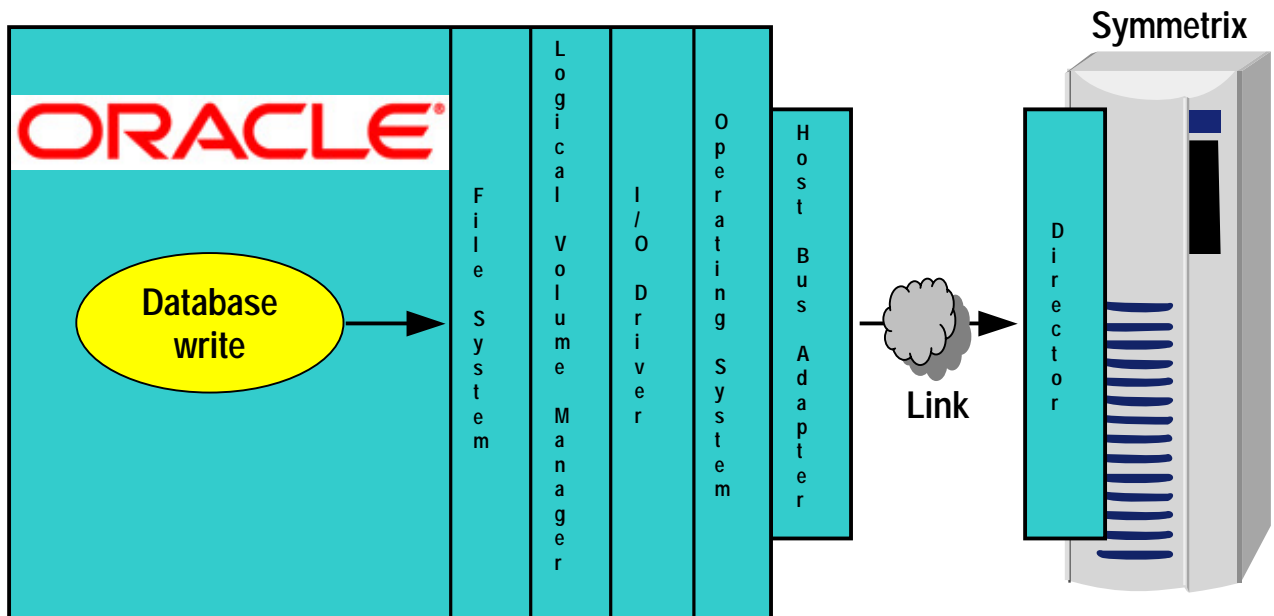
## EMC Double Checksum

It is because of the risk of data corruption and the subsequent high cost of remedial action that EMC and Oracle have combined efforts to deliver EMC Double Checksum. EMC Double Checksum, when used with Oracle checksum, provides an end-to-end data integrity solution, quickly detecting and reporting data corruption when it occurs.

### *The Data Block Corruption Problem*

How is it architecturally possible for data used with an Oracle database to become corrupt?

When an Oracle process writes a data block to disk, the data block in the host memory is intact and consistent. However, as shown in Figure 1, there are layers of software and hardware (for example, File System (FS), Logical Volume Manager (LVM), I/O Driver, Operating System (OS), Host Bus Adapter (HBA), Link, and Director) between the point when Oracle writes an I/O to the OS and when the Symmetrix receives the data. It is in these layers that unwanted erroneous corruption could be introduced. For example, when a faulty HBA begins to generate bad data blocks, the database may incur corruption in hundreds of locations. Other possible causes of corruption are bugs in the various software layers or faulty components.



**Figure 1. Database Transfers**

Without turning on the Oracle checksum and EMC Double Checksum features, or if using storage systems other than EMC Symmetrix, the storage system would not be aware of bad data blocks and they would be written to disk as though no problem existed. Then, at some time in the future when the bad block is read, the Oracle session will receive a checksum error, indicating that recovery from a good backup of this block must be performed. Since RMAN (the Oracle Recovery Manager utility) automatically does a checksum during backups of the database, the corruption, in the worst case, will be detected at the next database backup. However, by this time the corruption has already been written to the current database files, and corrective action must now be taken. If a site deploys Oracle on EMC Symmetrix with both the Oracle checksum and EMC Double Checksum features enabled, errors detected by the checksum calculations are quickly reported and can be rejected, triggering retry of the faulty I/O in real time. Since many corruptions are transient in nature, retries are often able to correct the corruption. In this case, the data is corrected in real time, minimizing the impact on the business. By providing an early warning when a corruption is detected, recovery procedures can be avoided all together or at least initiated quickly.

## ***Enabling the Oracle Checksum Feature***

Starting with Oracle 9i, the Oracle checksum feature is automatically enabled by default. For releases prior to Oracle 9i, the Oracle checksum is enabled by setting the initialization parameter `db_block_checksum` in the `init.ora` file to `TRUE` and restarting the database instance, or by using an `ALTER SYSTEM` command. From that point on, Oracle will calculate the data block checksum and store it in every data block when it is written to disk. This checksum is verified by Oracle when data blocks are read back at a later time. Checksums are also automatically calculated and stored by RMAN when backups and restores are performed. Oracle also writes checksum data for redo log blocks and control file blocks; however, redo log blocks and control file blocks are not checked by the initial release of EMC Double Checksum.

Prior to Oracle 9i, if `db_block_checksum` is enabled using the `ALTER SYSTEM` command, users must also ensure to set the parameter in the corresponding parameter file. Starting with Oracle 9i, the `ALTER SYSTEM` command is the preferred method for changing Oracle parameters since this is automatically propagated into the corresponding `spfile`.

The Oracle checksum feature and EMC Double Checksum can be enabled on a live system at any time. If desired, RMAN can be used to do a `BACKUP VALIDATE` of the database files prior to enabling the checksum features.

## ***The EMC Double Checksum Feature***

EMC Double Checksum, a feature of the Symmetrix Engenuity™ software, is a solution that ensures end-to-end data integrity in the handling of Oracle data blocks from the Oracle instance to the Symmetrix volume. When the EMC Double Checksum feature is enabled, the checksum information written by Oracle is used to verify the write at the time the write is *received* by the Symmetrix from the host. This verification can detect erroneous modifications to the data block in the layers of software and hardware between the point when Oracle releases the data block and when the Symmetrix receives the data block.

## ***Enabling the EMC Double Checksum Feature***

The EMC Double Checksum feature is enabled by issuing a `symchksum` command from the Symmetrix Command Line Interface (SYMCLI). When enabled, if the Symmetrix detects an Oracle data block checksum error, the error is logged in the error log facilities internal to the Symmetrix unit. The Symmetrix can also be configured to respond by returning an I/O error to the Oracle instance and/or calling EMC Customer Service (“phone home”), letting the Oracle instance continue with the bad block in the database.

EMC recommends using only the “error log” and/or “phone home” options initially until confirming that the configuration is operating as expected. Once the system has been fully exercised and operating for a period of time, the “reject I/O” option can be enabled. Traditional change management procedures are useful, including initial use in test systems before deployment.

For example, to enable EMC Double Checksum on the extents of all devices that define the current database instance, and to call EMC Customer Service and reject the I/O on error, enter:

```
symchksum enable -type Oracle -phone_home -reject_io
```

Please refer to the *EMC Solutions Enabler SYMCLI V4.3 Checksum Release Notes* (P/N 300-000-046) for more details on the `symchksum` command.

## ***Responding to the Detection of Errors***

Before deploying EMC Double Checksum, it is important to understand the expected behavior of the product when it is enabled. The product is designed to detect bad Oracle data blocks and take suitable action when this detection occurs. When a checksum error is detected, an error is automatically logged in the error log facilities internal to the Symmetrix. The user can specify that checksum errors cause the I/O to be rejected and/or have the Symmetrix phone home. The following two sections describe the different responses based on whether or not the “reject I/O” option is specified. Oracle recovery is a complex subject; please refer to Oracle documentation for further details.

### **Reject I/O Option Selected**

If the “reject I/O” option is specified, the corrupted I/O with incorrect checksum is rejected in real time. This option causes the I/O to be retried. Because many corruptions are transient in nature, the corruption is often corrected upon retry. If the retry successfully corrects the error, Oracle continues to run. If the error is long-lasting and not corrected by a retry, an error is logged by Oracle.

Depending on what the Oracle instance was attempting to write, the result ranges from the data file going offline to the instance stopping. Once the source of the corruption is fixed, the data file can be put back online. If the instance has stopped, restarting the instance should cause Oracle to go into automatic crash recovery.

This option is useful for maximum protection of the database. It is important to note that this option may cause the entire database to be unavailable until the source of the corruption is repaired. Although this option may cause the instance to stop, total resolution time may be shorter.

### **Reject I/O Option Not Selected**

If the “reject I/O” option is *not* specified, the corrupted I/O is accepted in real time. The database does not receive an error, but the problem is logged in the Symmetrix and handled using normal Symmetrix error reporting facilities. At the earliest opportunity, the corrupted block(s) in the data file(s) must be repaired with Oracle media recovery. Media recovery requires a backup that predates the corruption, and the application of archived redo logs from that point.

This option is useful for sites that prefer to not have an immediate outage. However, if the database attempts to read a corrupted block, it will notice the problem and may still cause an outage.

### **Recovery Actions in Cases of Corruption**

If data block corruption is detected when the EMC Double Checksum feature is enabled with the “reject I/O” option, the blocks that have been rejected are logged and may not be propagated to the Symmetrix device. It is important to investigate the problem and verify that the database integrity has not been compromised. The RMAN BACKUP VALIDATE option can be used to verify that the current database structure is clear of any corruption. If errors are detected, appropriate actions will be required, such as restoring from a backup and/or performing media recovery. In this case, please consult EMC Customer Service and/or Oracle Support if further help is required. Most importantly, further investigation as to what is causing the “reject I/O” error is required. If this is due to potential hardware failures or software bugs, the appropriate onsite administrators and vendors will need to be contacted to repair the problem as quickly as possible.

## ***Items to be Aware of With Version 1.0 of EMC Double Checksum***

- The initial release of EMC Double Checksum is limited to 31 EMC Symmetrix contiguous extents per host physical device. Symmetrix extents are internal to the Symmetrix and are used to map Oracle tablespace locations. Oracle extents are a completely different concept from Symmetrix extents and are not limited by this constraint. An Oracle extent is a logical construct and is completely contained within a given Oracle tablespace. Since EMC Double Checksum is managed at the tablespace level, Oracle extents are not even visible.

Certain factors affect the number of Symmetrix extents consumed when raw volumes are deployed. At least one Symmetrix extent is used for each raw volume (Oracle data file) residing on a given host physical device. A single host physical device would have to be partitioned with an LVM into more than 31 separate raw volumes to hit the maximum. This scenario is probably not very likely since each raw volume would then be so small.

However, raw volumes are often striped across multiple host physical devices in a volume group using an LVM. Regardless of the number of host physical devices in the volume group, it can be divided into no more than 31 striped raw volumes. Again, this will not affect the size of the overall database, but must be taken into account when planning the database layout.

- File system-based Oracle files are not supported in the initial release of EMC Double Checksum. All Oracle tablespace files must be placed on raw volumes (including LVM raw volumes). If host-based striping is used on the raw volumes, then the stripe size must be an exact multiple of the Oracle block size. Please note that the use of host-based striping will increase the number of Symmetrix extents that are utilized.
- Checksums are not currently checked on Oracle redo log file blocks or control file blocks.
- A raw file may not be extended using an LVM. The object should be re-enabled to cover the extension. Validate can be used periodically to check the setup.
- Before restoring Oracle data files from a backup, the EMC Double Checksum feature may need to be temporarily turned off for those data files that were backed up prior to the Oracle checksum being enabled. Old blocks may exist on disk without checksum information in them if the database was running without checksum enabled in the past. Also, an online backup may contain Oracle “split” data blocks that do not have the correct checksum — a normal situation corrected during recovery. Performing a host restore with such blocks can inadvertently trigger the feature.

This limitation does not apply if RMAN is being used. In this case, there is no need to disable the EMC Double Checksum feature upon restore, since RMAN calculates checksums by default for backups and restores. (Note: this does not apply to RMAN when proxy copies are being used, since checksums are not written in this case, so the same restrictions apply.) Please refer to the Oracle Recovery Manager documentation for further details on using RMAN.

## ***Performance Considerations***

One tradeoff to using the checksum feature is that a small performance overhead may be incurred. Benchmark testing so far indicates that this overhead is minimal. However, the impact of enabling the checksum feature will ultimately depend upon the application profile and usage, so this will vary for each customer. It is therefore important to always test and validate that any impact is acceptable to the customer service-level requirements before introducing this feature (or any new feature) into a production system. In addition, it is also important to note that any overhead incurred will be significantly less than having a data block corruption go undetected for a long period of time. The longer the duration a corruption goes undetected, the more time it will potentially take to recover the system, and can, in turn, cause more significant outages. Specifically for this reason, the Oracle High Availability solutions, such as Oracle Parallel Fail Safe, highly recommend that all checksumming features be enabled.

## Examples

To list the devices on Symmetrix 3890 that have extents being checked for checksum errors, enter:

```
# symchksum list -sid 3890
```

To show all the extents of Symmetrix device 0A1 on Symmetrix 3890 that are being checked for checksum errors, enter:

```
# symchksum show dev 0A1 -sid 3890
```

To enable EMC Double Checksum on the extents of all the devices that define the current database instance and then to phone home on error, enter:

```
# symchksum enable -type Oracle -phone_home
```

To enable EMC Double Checksum on the extents of all the devices that define the tablespace and then to log on error, enter:

```
# symchksum enable -type Oracle -tbs SYSTEM
```

To verify that Oracle tablespace USER01 has EMC Double Checksum enabled on all the devices that have defined it, enter:

```
# symchksum verify -type Oracle -tbs USER01
```

To disable EMC Double Checksum on the current database instance, enter:

```
# symchksum disable -type Oracle
```

Disable by device should only be used under special circumstances. For example, this option can be used to remove extents if a database or a tablespace has been dropped without first doing a normal disable. In this case, disable by device can be used to remove the extents. To disable (with force) EMC Double Checksum for all checksum extents on Symmetrix device 0A1 on Symmetrix 3890, enter:

```
# symchksum disable dev 0A1 -sid 3890 -force
```