End-to-end Data Protection

Dan Colegrove
Senior Technical Staff Member

Introduction

End-to-end Data Protection is a new feature in SAS and Fibre Channel hard drives that extends error detection to cover the entire path from the computer system to the hard drive media and back. Data protection information is appended to the data in the computer system. It stays with the data from the computer, through Fibre Channel or SAS connections, through RAID controllers, and through drive electronics to the drive media. When read, the same data protection information returns with the data to the computer system. The protection information is used to verify the correctness of the data at multiple points in the path.

Error Detection at the Hard Drive

For the first time, hard drives can detect an error created anywhere in the path before it is written to the drive media and before sending erroneous data back to the computer system.

Some storage systems without End-to-end Data Protection attach proprietary error-checking information to user data. The disadvantage of error checking that takes place only in the computer system is that hard drives can not understand the proprietary checking information, so they can not detect an error that occurred between the computer and the drive or check the proprietary information when data is read, before sending it to the computer system.

End-to-end, EDP and DIF are the Same

End-to-end Data Protection is also known as EDP or the Data Protection Model. During development it was also referred to as Data Integrity Field or DIF.

Error Isolation

Errors can be isolated to the host adapter, RAID controller or hard drive. As a standard feature, each device in the path from system to drive can check and report a data error (Figure 1). This simplifies isolation of the component that had the error. This allows for multiple points of failure isolation:

1. As it passes through intermediate devices (RAID Controllers)
2. Before it is written to the hard drive media
3. When it is read from the media
4. As it passes back through intermediate devices (RAID Controllers)
5. When it arrives at the computer system

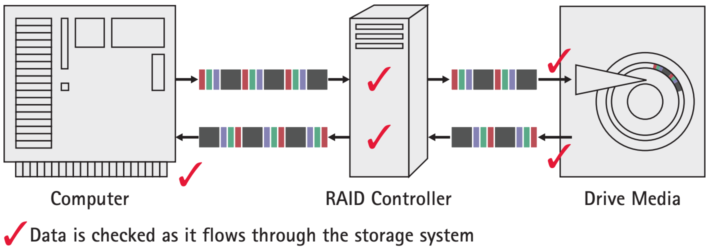

Figure 1. End-to-end means data protection information stays with the data from the computer to the hard drive media and back so errors can be isolated to a component when they occur.

Integration

EDP allows integration of hardware and software from multiple vendors. As a standard feature, all devices use the same rules for interpreting the protection information. There is no need to customize controllers, drives or software to support proprietary checking information.

Protection Three Ways

The protection is provided by eight bytes of data appended to each user data block. The eight bytes are divided into three fields (Figure 2), the Guard, the Reference Tag and the Application Tag. The protection data is created by computer systems, transmitted with the user data block, and written to the drive media.

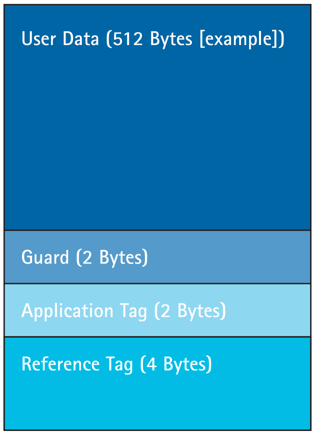

Figure 2. Three fields are added to each block of user data.

Guard

The Guard field protects against errors in the data. The two-byte Guard is a Cyclic Redundancy Check (CRC) on the data in the data block. This allows each device along the path from the computer system to the drive to check that the data in the block is still correct. Computer systems use the CRC algorithm specified by the protection specification to fill the Guard when sending a write data command. Because a standard CRC algorithm is used, hard drives can check that the data is correct before writing it to the media. On a read command, the drive reads the protection information from the media, checks it, and then sends the protection information along with the data. When a block of data is received by the computer system, the user data block is checked against Guard to verify that the data was received correctly.

Reference Tag

The four-byte Reference Tag contains block address information. As data is moved from computer to drive media and back, perhaps through intermediate devices, such as a RAID controllers, there is a possibility of an addressing error. An addressing error causes the wrong blocks of correct data to be sent, or blocks to be sent in the wrong order. The Reference Tag numbers the blocks so hard drives and computer systems can check to see that the correct data blocks are received in the proper order.

Application Tag

The two-byte Application Tag may be used by storage applications to add additional checking information to each data block. A typical use for the Application Tag is to associate RAID configuration data with data blocks. The Application Tag is created by the application and checked by the application, and may be checked by hard drives.

Because the Guard, Reference Tag and Application Tag are created with a standard algorithm, hard drives can check the received data and report errors when data is being written and read back. Proprietary systems can only detect the error when the data is read back.

Implementing End-to-end Data Protection

The EDP system has a wide range of controls to allow checking of some or all of the protection fields. Hard drives indicate support for EDP and which of the fields they can check. Drives may be formatted for one of four protection types, which define how the protection data is used. In the read, write and verify commands used with protection data there are controls to indicate which fields should be checked for that command. New 32-byte commands are available, if the drive is formatted to use them, which carry additional initial Reference Tag and Application Tag data so that drives can check the Reference tag at a custom starting point and check the Application tag.

Because the protection model was developed in several stages, the fields used to control EDP operations are spread over multiple locations. Tables in the SCSI Block Commands-3 (SBC-3)[1] standard describe the relationship of the fields.

Determining Support for Protection

The PROTECT bit in the INQUIRY data indicates if the drive supports protection. The SPT field in the Extended INQUIRY Data VPD page indicates the highest type supported. Additional bits in Extended INQUIRY indicate if the drive can check the Guard, Reference Tag and Application Tag. The type of protection is selected with a Format Unit command. The protection type selected may be read with the Read Capacity(16) command.

The type of protection determines which read, write and verify commands are enabled and how the Reference Tag is used. The read, write and verify commands can be roughly broken into two groups: the 32-byte commands and the non-32-byte commands. Only one of the two groups may be used depending on the protection type.

Four Types of Protection

There are currently four types of protection:

Type 0: No protection
Type 1: Protection is enabled and the 32-byte commands are not valid.
Type 2: Protection is enabled and only the 32-byte commands are valid.
Type 3: Protection is enabled and the 32-byte commands are not valid. The Reference Tag is not defined and may be used as an extension of the Application Tag. The drive will not check the Reference Tag.

The non-32-byte commands are the family of commonly used read, write and verify commands (10-, 12-, 16-byte commands, XOR 32-byte commands). See the SBC-3 standard for the full list. In the non-32-byte commands, the Reference Tag is the low-order bytes of the Logical Block Address. Because the non-32-byte commands do not contain an expected Application Tag value, the Application Tag is not checked by the drive.

The 32-byte commands, Read(32), Verify(32), Write(32), Write and Verify(32) and Write Same(32) have additional fields that allow finer control over protection information checking. The 32-byte commands specify the starting value for checking the Reference Tag. The 32-byte commands also specify the value to use for checking the Application Tag.

The 32-byte commands are particularly useful in systems that aggregate hard drives into a single logical unit, some RAID systems, for example. The host system sees these logical units as one large storage device. The 32-byte command Reference Tag starting value allows the individual hard drives to check the Reference Tag, even though the block address at the controller level and at the drive are different.

All of the commands used with protection information contain a protect field (RDPROTECT, VRPROTECT or WRPROTECT) that specifies which protection fields are checked by the drive for that command.

Summary

The EDP feature allows detection of errors through the entire path from computer system to hard drive media and back by appending additional fields to user data blocks. The Guard field protects the data block content. The Reference Tag is used to verify that data blocks are the ones requested in the proper order. The Application Tag provides a feature that RAID systems or other applications can use to add additional checking information designed specifically for the application.

[1] *SCSI Block Commands—3 (ANSI/INCITS 306-1998) clause 4.16 Protection Information Model describes the standard implementation of EDP.*

HGST
a Western Digital company

References in this publication to HGST products, programs, or services do not imply that HGST intends to make these available in all countries in which it operates. Some countries have laws that may affect encryption-enabled devices. HGST assumes no responsibility for compliance with laws that apply to our products after they are sold or outside of our control. Accordingly, we suggest that you review all laws of the countries where you conduct business.

Product specifications provided are sample specifications and do not constitute a warranty. Information is true as of the date of publication and is subject to change. Actual specifications for unique part numbers may vary. Please visit the Support section of our website, www.hgst.com/support, for additional information on product specifications. Photographs may show design models.

HGST, a Western Digital company
3403 Yerba Buena Road
San Jose, CA 95135 USA

Produced in the United States 04/08.
All rights reserved.