Seagate®

Technology Paper

Safeguarding Data From Corruption

Introduction

As data has become the very lifeblood of virtually every business, it's no surprise that safeguarding such data is an increasingly urgent priority. Over the years, a variety of technologies have been developed to combat data corruption. Tailored to discrete stages in the I/O stack (application, operating system, I/O controller, SAN, disk array, hard disks), these technologies have each played key roles in the drive to ensure data integrity.

But the disparate nature of this approach leads to inconsistent protection, with solutions such as Error Correcting Code (ECC), Cyclic Redundancy Check (CRC), checksums, etc. only addressing data corruption at specific, isolated points in the I/O path. The risk of corrupted data falling through the inherent cracks in this protection methodology is significant, and the havoc such undetected, silent data corruption could wreak (lost or inaccurate data, significant downtime) is substantial.

Seagate Spearheads Battle Against Data Corruption

Recognizing the threat that such silent data corruption poses to businesses both large and small, Seagate has implemented the new T10 Protection Information (T10-PI) standard throughout its enterprise hard drive product lines. As one of the founders of the Data Integrity Initiative (DII), Seagate, along with Oracle and Emulex, was instrumental in defining the T10-PI standard, which, in concert with the Data Integrity Extensions (DIX) standard, delivers full end-to-end data integrity from application to disk drive.

The T10-PI standard is an extension of the existing T10 SCSI Block Commands specification; covering communication between SCSI controllers and storage devices, Protection Information (PI) adds an extra eight bytes of information to the 512-byte sectors typical of enterprise hard drives. Increasing sector size to 520 bytes, these eight bytes of metadata consist of guard (GRD), application (APP) and reference (REF) tags that are used to verify the 512 bytes of data in the sector (see Figure 1).

520-Byte Sector on Protection Information Disk Drive			
512 Bytes	8 Bytes (metadata)		
User Data	End-to-End Check		
	2 bytes (GRD tag)	2 bytes (APP tag)	4 bytes (REF tag)

Figure 1. 520-Byte Sector on Protection Information Disk Drive

Complementing PI, DIX specifies how controllers can exchange metadata with a host operating system. Thus the combination of DIX (data integrity between application and controller) and PI (data integrity between controller and disk drive) delivers end-to-end protection against silent corruption of data in flight between a sender and a receiver.

Using Seagate® Protection Information Disk Drives

Formatting PI Drives

Drives formatted with PI information provide the same sector count (and thus the same capacity) as non-PI formatted drives. To achieve this, PI-formatted drives are physically formatted to 520-byte sectors that store 512 bytes of user data with eight bytes of Protection Information appended to it. (Note: Protection Information is valid with any supported sector size; 512-byte sectors is an example.)

The advantage of PI is that the Protection Information bytes can be managed at the host bus adapter (HBA) and HBA driver level, enabling systems that typically don't support 520-byte sector formats to integrate this higher level of protection.

Important: When formatting a PI drive, the type of Protection Information to be employed must first be selected.

Types of Protection Information

There are four distinct types of PI, as shown in Figure 2.

Type 0	Describes a drive that is not formatted with PI information bytes. This allows for legacy support in non-PI systems.
Type 1	Provides support of PI protection using 10- and 16-byte commands. The RDPROTECT and WRTPROTECT bits allow for checking control through the CDB. Eight bytes of Protection Information are transmitted at sector boundaries across the interface if RDPROTECT and WRTPROTECT bits are non-zero values. Type I does not allow the use of 32-byte commands.
Type 2	Provides checking control and additional expected fields within the 32-byte CDBs. Eight bytes of Protection Information are transmitted at sector boundaries across the interface if RDPROTECT and WRTPROTECT bits are non-zero values. Type II does allow the use of 10- and 16-byte commands with zero values in the RDPROTECT and WRTPROTECT fields. The drive will generate a dummy (for example, 0xFFFF) eight bytes of Protection Information in the media, but these eight bytes will not be transferred to the host during read.
Type 3	Seagate® products do not support Type III.

Figure 2. Types of Protection Information

Setting and Determining the PI Type

A drive is initialized to a type of PI by using the format command and setting the appropriate FMTPINFO and RTO_REQ bits on a PI-capable drive. Once a drive is formatted to a specific PI type, it may be queried by a Read Capacity (16) command to report which PI type it is currently formatted to.

Multiple PI types cannot coexist on a single drive; a drive can only be formatted to a single PI type. A drive may later be changed to a new PI type, but this requires a low-level format, destroying all existing data on the drive. No other vehicle for changing the PI type is provided by the T10 SBC-3 specification.

Identifying a Protection Information Drive

The Standard Inquiry provides a bit to indicate if PI is supported by the drive. The Vital Product Data (VPD) page 0x86 provides bits to indicate the PI types supported and which PI fields the drive supports checking. Mode page 0x0A provides information on the ATO bit (Application Tag Owner) for the Application Tag fields.

How PI Fields and Information Are Transmitted

During write commands, PI-aware hosts will provide two sets of Protection Information data to the drive through the Command Descriptor Block (CDB) and by appending eight bytes of PI information to the transmitted data at sector boundaries. The bits within the CDB indicate the information provided and the appropriate action to be taken for each Protection Information field. The 32-byte command set provides fields to be used as masks or as compare data for the appropriate PI fields.

Similarly, on read commands, the host indicates to the drive which fields to check and transmit back to the host. In cases where the Application tag is saved as 0xFFFF, checking of the PI fields is disabled but the PI fields will be transmitted back to the host as requested.

Calculating Drive Transfer Rate

Due to the extra overhead of transferring PI metadata from the media (which is not calculated as part of the user’s data transferred to the host), sequential performance of a PI drive will be reduced slightly (approximately 1.56%). Thus, to determine the full transfer rate of a PI-equipped drive, transfers should be calculated by adding the eight extra bytes of PI to the transferred sector length (for example, 512 + 8 = 520).

Conclusion

Protection Information disk drives from Seagate represent a watershed in the evolution of enterprise data protection. Until now, business owners had to rely on a confusing jumble of disparate error detecting and error correcting technologies to safeguard their data as it traveled its long I/O path (application, operating system, I/O controller, SAN, disk array, hard disks). When data corruption did occur, it could go undetected; left unchecked, such silent corruption can make data recovery costly and difficult—even impossible—to perform.

Seagate understands how important data integrity is to the success of every business, whether large or small, and has proactively implemented Protection Information technology to make true end-to-end data protection a reality. Now businesses can be assured that the integrity of their valuable data is protected from the application level all the way to where the data lives—on the drive.

www.seagate.com

AMERICAS Seagate Technology LLC 920 Disc Drive, Scotts Valley, California 95066, United States, 831-438-6550
ASIA/PACIFIC Seagate Singapore International Headquarters Pte. Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, 65-6485-3888
EUROPE, MIDDLE EAST AND AFRICA Seagate Technology SAS 16–18, rue du Dôme, 92100 Boulogne-Billancourt, France, 33 1-4186 10 00

TP621.1-1104US, April 2011