End to End Data Protection

T10 - 05/07/03

LSI LOGIC STORAGE SYSTEMS

Agilent Technologies

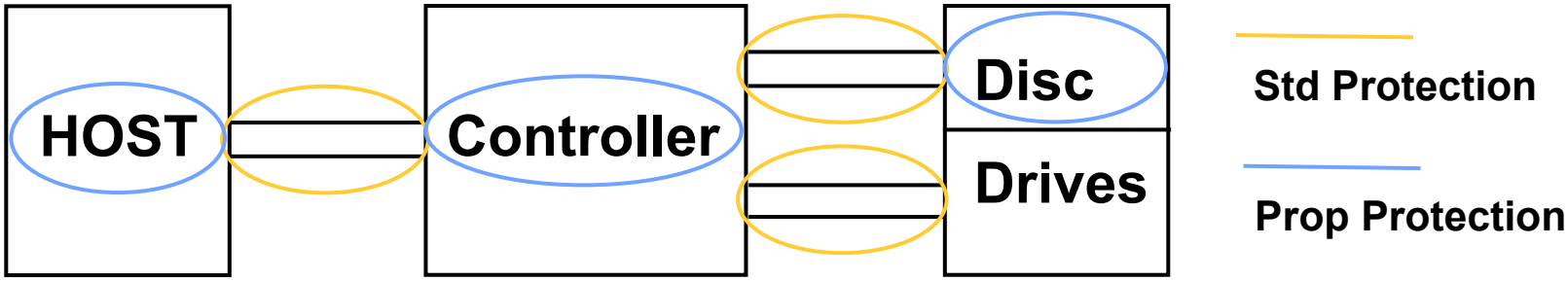
intel

Seagate We turn on ideas

Proposal Development Process

- Basic framework of Data Integrity scheme provided to four (4) companies for their review and comment.
- Companies sought out customers for input and incorporated that knowledge into the scheme.
- Overall scheme grew complex, but was simplified through cooperative efforts based on commitment of all to support standardized approach.
- Now it's time to bring the approach to T10 for wider review, optimization and final standards approval.

Current Industry Practice - 1



Host – Controller – Storage Model

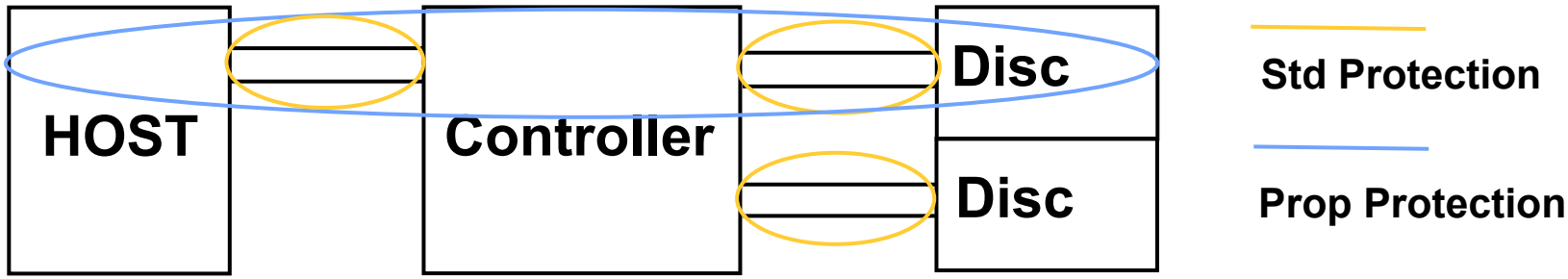
Standardized protection defined for Interfaces

Each component has proprietary protection schemes

System philosophy is “check and regenerate” protection for each region.

Disadvantage: Opportunities for undetected corruption can occur during check and regenerate processes.

Current Industry Practice - 2



Host – Controller – Storage Model

Host can add proprietary protection bytes (end-to-end protection) to each block written to disc and recheck on read.

Disadvantage: Intermediate devices cannot check, just pass through. If error is seen, recovery is difficult or impossible. Also difficult to determine source of error. Timeframe may be long between error occurrence and discovery.

Solution: End-to-End Data Protection from Host to Drive

Implement a Standard Approach for End-to-End Data Protection

Support Data Protection Checks in Intermediate Controllers and Drives

Allow Application-Specific Data or Functions

Terminology

- Data Integrity Field (DIF) – Combination of Tag Data and Guard Fields
- (Incrementing) Reference Tag – May be used to validate that data is sent in correct order, with no duplicates, and that all data is transferred. The initial value of the Tag is specified in the CDB. Use of the ID Tag is specified in CDB and may be turned ON or OFF from command to command.
- (Fixed) Meta Tag – Fixed value is specified in CDB (e.g. Constant could be a logical unit number in a RAID system).
- Guard Field – Method of verifying (by calculating and comparing data checking algorithm) that correct data is transferred from interface to disk media, or from memory (could cross page boundaries) to the interface.

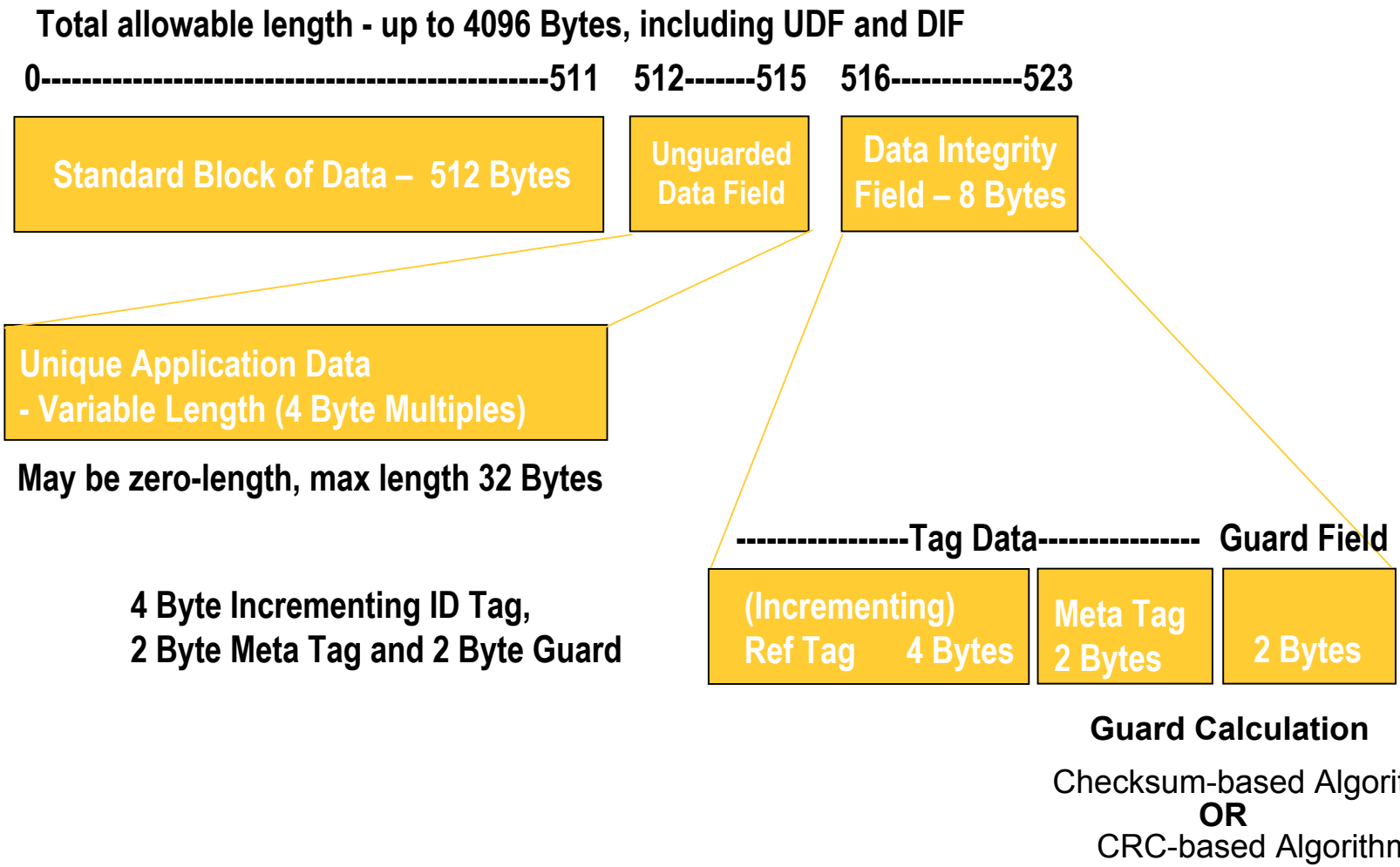
Data Integrity - Approach

- Generic Host-to-Drive Data Format
  - Standard block of data, length defined as Logical Block Size (LBS)
  - Unguarded Data Field, where field length is included in LBS
  - 8-Byte Data Integrity Field (in addition to LBS)

- Unguarded Data Field (UDF)
  - e.g. Software Stamp to indicate creator, data type, creation and/or last modification date, data path, or other identification information.
- Data Integrity Field (DIF)
  - Data Integrity Field includes Tag Data and Guard field.

- New CDBs for this new Data type
  - Read / Write / Verify / Write & Verify / Write Same
  - New Op Codes (16-Byte and 32-Byte structures)

Generic Data Format – Host to HDD



Guard Calculation

Checksum-based Algorithm
OR
CRC-based Algorithm

Data Integrity Configuration Settings

The Data Integrity Mode Page is used to specify configuration settings and data that will not change from command to command.

The Command Descriptor Block (CDB) is used to specify configuration settings and data that may change from command to command.

Data Integrity Mode Page

Bit	7	6	5	4	3	2	1	0
Byte 0	PS	Reserved	PAGE CODE (TBDh)					
1	PAGE LENGTH (0Eh)							
2	STOR_DIF	Vendor Specific		RESV	EXCL_Bytes (# 4B incr.)			
3	META ECHO	REF METHOD		GUARD METHOD			RESERVED	
4	{Primary Verification Controls:}		{Alternate Verification Controls:}		{Legacy Verification Controls:}		{Legacy Write Controls:}	
	REF_ CK_p	GRD_ CK_p	REF_ CK_a	GRD_ CK_a	REF_ CK_l	GRD_ CK_l	ZAP_ REF	MRK_ GRD
5	STK_ META	STK_ REF	STK_ GRD	DI_ AVAIL	Reserved			
6	(MSB)	META TAG MASK Primary						
7								(LSB)
8	(MSB)	META TAG MASK Alternate						
9								(LSB)
10	(MSB)	META TAG MASK Legacy						
11								(LSB)
12	(MSB)	META TAG DEFAULT (legacy)						
13								(LSB)
14	Reserved							
15	Reserved							

16 Byte CDB w Tag Data

	Bit							
Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (TBD)							
1		CHECK_OPTION		Command Specific Flags				
2								
3								
4								
5								
6								
7								
8								
9								
10	(MSB)			META TAG				
11				(2 Bytes)				(LSB)
12	(MSB)			TRANSFER LENGTH				
13				(2 Bytes)				(LSB)
14				RESERVED				(LSB)
15				CONTROL				

CMD_FMT – Specifies 8 Byte LBA, where 4 Byte Tag serves as part of LBA.

CHK_OPTN – Allows disabling of Tag Data and Guard field verification.

(Mode Page specifies type of Guard).

REF TAG FIELD – Contains 4-Byte Incrementing Tag Data value.

META TAG FIELD – Contains 2-Byte Fixed Tag Data value.

DIF Write Operation Sequence

- Host generates Guard value at the Application, Driver, or HBA levels. Data Stream includes DIF Field for each Logical Block.
- Hosts communicates format (Fixed and/or Incr) and value of Tag Fields in the CDB.
- For each block, verification is done at Intermediate Controller and Target Device.
  - Guard value calculated and compared to Host value
  - Verify Tag Fields in CDB compare to Data Stream
  - Write operation halts when error is detected

DIF Read Operation Sequence

- For each block, verification will be done at Target, Intermediate Controller and Host Devices.
  - Guard value calculated and compared to value in Data Stream
  - Tag Fields in CDB are compared to Data Stream
  - Read operation halts when error is detected

Legacy Write Operation Sequence

- Host generates Data Stream at the Application level. Block of Data with no DIF Field for each Logical Block.
- When STOR_DIF flag set, DIF Field is inserted at Intermediate Controller or Target.
- For each Block of Data, DIF verification is done at points after DIF insertion point.
  - Guard value calculated and compared to Data Stream
  - Verify Tag Fields in CDB compare to Data Stream
  - Write operation halts when error is detected

Legacy Read Operation Sequence

- For each Block of Data, verification is done at Target and Intermediate Controller (prior to DIF removal point).
  - Guard value calculated and compared to value in Data Stream
  - Tag Fields in CDB are compared to Data Stream
  - Read operation halts when error is detected
- When STOR_DIF flag set, DIF Field is removed by Target or Intermediate Controller.

Error Handling

- **Specific Error cases and Methods of Handling those cases will need to be defined prior to release of a standard.**
- **This feature creates a new type of error not previously considered in the T-10 standard. A new subclass of error code (e.g. 03/11/xx) is required to indicate a system error condition as opposed to a drive or interface error.**

Benefits

Primary benefits of a cooperative scheme:

- Data Integrity information is written to the drive media to provide end to end assurance of data integrity.
- Detection of data failures at the drive level is enabled during write operations and read operations.
- Isolation / Correction of bad data occurs as early as possible, with minimum impact to system integrity and performance.
- Standard approach results in common methodology between vendors to ease management and maintenance for customers.
- Flexibility to allow Unguarded (Application-Specific) Data, and options for Data Integrity checks, provide varied implementation possibilities for end-users.

Next Steps

- A paper is provided for T10 review that represents the combined position of 4 companies, with input from their customers, on the proposed framework for Data Integrity Extensions.
- It is hoped that T10 will adopt this effort and continue to develop the proposal into a final standard with full industry support.
- The members of the 4 companies that have worked on the current version of the proposal will continue to work with T10 as the standard is brought to final form and adopted.