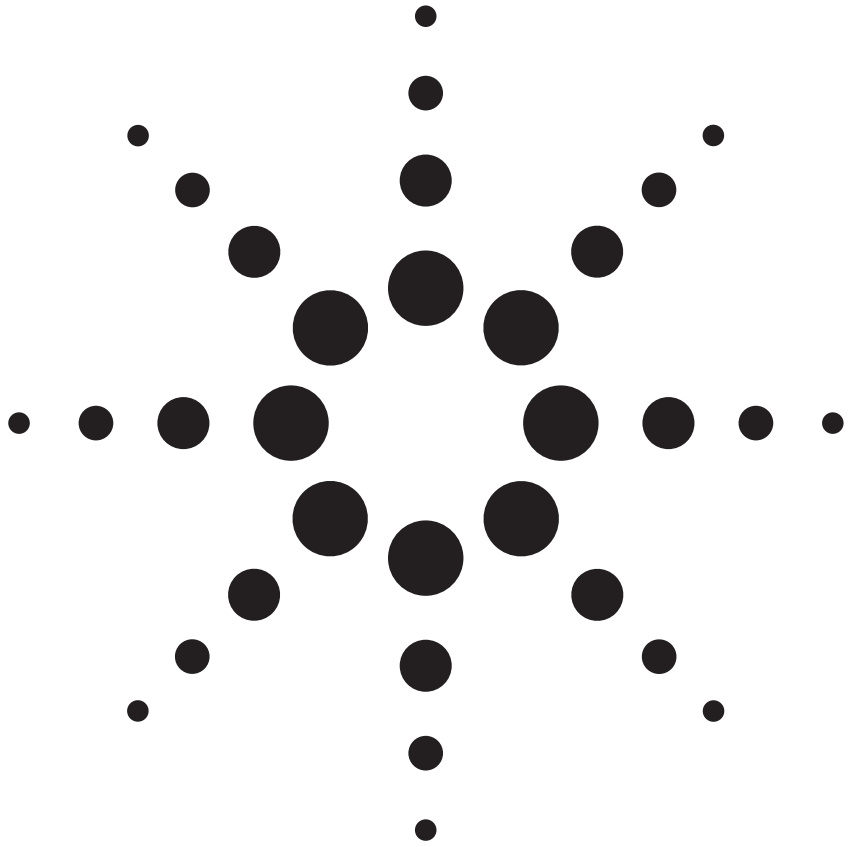
DIF (Data Integrity Field):
Provides End-to-End Protection of User Data

White Paper



By: Narayan Ayalasomayajula
Jim Shupenis
John Connolly

Agilent Technologies
Semiconductor Products Group
Storage Networking Division

Agilent Technologies

Introduction:

Error detection is critical in storage system implementations. Unfortunately, today's SCSI architectures lack a reliable mechanism for detecting corruption of the data generated by the host/server and written to the target and corruption of data retrieved from the target by the host. Additional risks of data corruption are introduced by intermediate data buffers and translations in the data path.

Current techniques for detecting data corruption are based on proprietary schemes. Implementing effective, universal SCSI error detection and correction would be greatly simplified if there were a standardized method of providing end-to-end data integrity.

Figure 1 shows the basic components involved in a Host-Controller-Storage model.

Within this storage model, the existence of non-overlapping protection domains can lead to undetected data corruption occurring during the check and regenerate process. Corruption can be caused by changes to the data or by errors that cause data to be written to or read from the wrong location. Corrupted data being passed through intermediate storage devices creates a situation where recovery and correction of any unseen error becomes very difficult.

This situation can be greatly improved with the adoption of a standardized protection algorithm. A standardized approach would allow any element in the data path to take part in the data integrity process, allowing early detection of and recovery from data corruption.

Agilent's End-To-End Data Integrity Proposal Using DIF (Data Integrity Field) :

This paper describes Agilent's Data Integrity Field (DIF) used for end-to-end data protection. DIF is the new End-to-End data protection standard described in document T10/03-111 of the T10 technical committee of the InterNational Committee for Information Technology Standards (INCITS)[1]

Agilent's data protection mechanism is based on the SCSI command structure and SCSI architecture model. Using DIF, application level protection is accomplished while supporting legacy implementations.

The primary benefits of using DIF within the Host-Controller-Storage Model are:

- 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 read/write operations
- Isolation/correction of bad data occurs as early as possible, with minimum impact to system integrity and performance
- A standards-based common methodology between vendors simplifies management and maintenance for customers
- The flexibility to allow application-specific data and options for data integrity checks

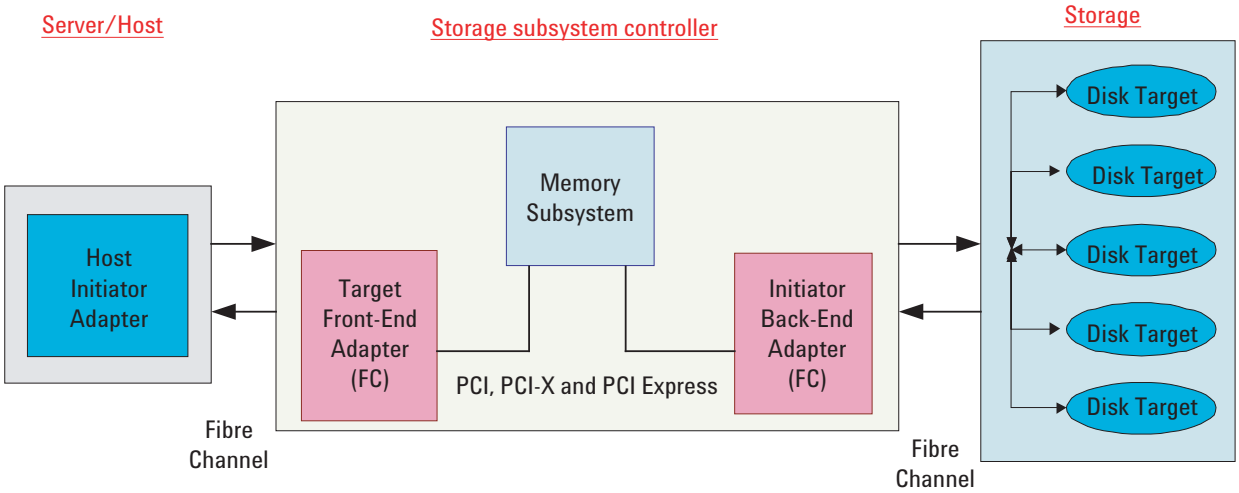


Figure 1. Host Controller - Storage Model

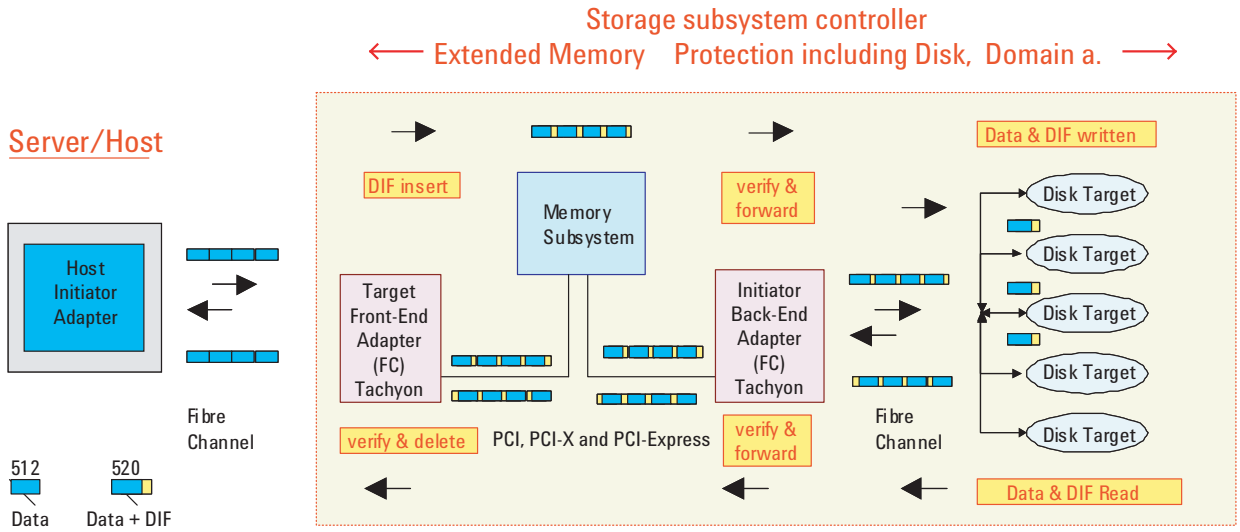


Figure 2. Host Controller - Storage with DIF

Introduction to DIF (Data Integrity Field):

DIF provides 8 bytes of error detecting code for each block of 512 bytes of data.

Immediately following the data are 2 bytes of cyclical redundancy code (CRC) that is computed over the 512 bytes of data. Following the CRC are 6 bytes of DIF referred to as the User Defined Tag (UDT), which can be used by the application. A typical use of the UDT is to store the Logical Block Address (LBA) information for the block of data with which it is associated.

Note: Fibre Channel CRC is computed over the frame header plus the entire payload in that frame. The DIF CRC is computed solely over the payload for 512 bytes of contiguous data, and is reset at every 512-byte block boundary. If a Fibre Channel frame has the 8 bytes of DIF embedded in the frame, the DIF bytes are treated as payload and are included in the FC-2 CRC calculation, providing overlapped protection.

DIF protection scope:

1. Data errors due to software bugs
2. Double bit bus parity errors for data transfers (in single bit parity protected buses)
3. Disk errors

Data Integrity using DIF

Agilent's DIF implementation provides three basic actions that can be performed on a 512-byte data block. The DIF action determines whether the DIF bytes are *Inserted*, *Verified and Forwarded*, or *Verified and Deleted* for every 512 bytes of payload.

Insert Action

The *Insert* action causes the origination of the protection domain. The ingress data stream does not have any DIF information embedded in the payload. When disabled, DIF provides transparent usage in normal legacy contexts. When enabled, each block on the device is expected to have DIF information attached.

As a result of the *Insert* action, 8 bytes of DIF are inserted into the egress data stream after every 512 bytes of data.

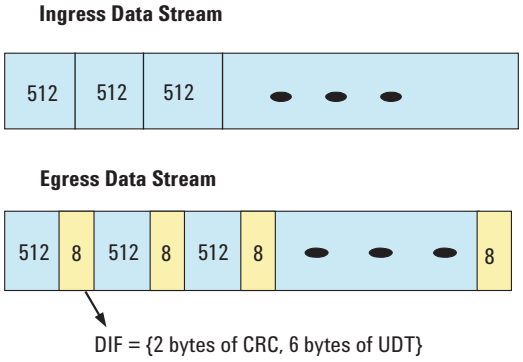


Figure 3. Insert Action

Verify and Forward Action

The *Verify and Forward* action causes the propagation of the protection domain. The ingress and egress data streams both have DIF information embedded after every 512 bytes of data. The 2 bytes of CRC and 6 bytes of UDT are verified for each 512 bytes of data.

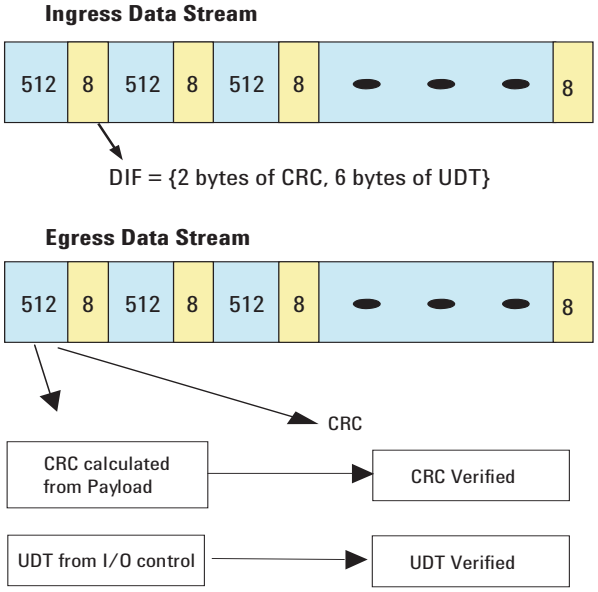


Figure 4. Verify and Forward Action

Verify and Delete Action

The *Verify and Delete* action causes the termination of the protection domain. The ingress data stream has DIF information embedded after every 512 bytes of data. The 2 bytes of CRC and 6 bytes of UDT are verified for each block of data. The egress data stream does not have any DIF information.

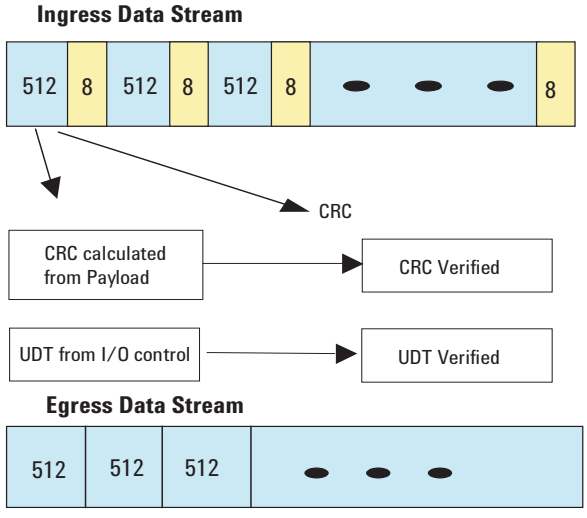


Figure 5. Verify and Delete Action

Implementation of protection domains utilizing Agilent's PCI-X based 2G and 4G Fibre Channel Interface Controller IC's or PCI-Express based 4G Fibre Channel Controller IC's (either dual and quad channel) when compared to previous-generation FC interface controler ICs. On page 5 are listed Fibre Channel Interface Controller IC's which support DIF.

With DIF features incorporated within an Agilent FC controller IC, improved performance and reduction of overall solution cost can be realized in PCI-X based 2 or 4G Fibre Channel Interface Controller ICs, or dual or quad channel PCI-Express based 4G Fibre Channel ICs.

Storage Subsystem

The following operations take place between the FC Target, memory subsystem, FC Initiator and disk (protection 'domain a' figure-2):

- Target Write Operation

  In this protection mode, the frames received off the Fibre Channel link do not contain DIF information. The Fibre Channel Interface Controller IC can be programmed to originate the protection domain using the *Insert* action. This embeds DIF within the data that is written to the memory system.

- Initiator Write Operation

  As an initiator, data that is fetched from the memory subsystem has embedded DIF at 512 byte block boundaries. The Fibre Channel Interface Controller IC performs a *Verified and Forward* operation to check data integrity. Since the data transmitted on the FC link has DIF at the block boundaries, the protection domain is extended to the target disk drive.

- Initiator Read Operation

  As an initiator, the frames received off the Fibre Channel link will have DIF information embedded at the block boundaries. The Fibre Channel Interface Controller IC can be programmed to propagate the protection domain using the *Verify and Forward* action. This causes the protection domain to be extended to the memory subsystem. Any data corruption will be flagged as closely as possible to the block where the corruption is detected. The Fibre Channel Interface Controller IC will inform the driver about the corruption. Depending on the way that a driver is programmed regarding what constitutes an DIF error, it could treat it as fatal or attempt to retrieve the data.

• Target Read Operation

As a target, data that is fetched from the Memory Subsystem has DIF embedded within. The Agilent Fibre Channel Interface Controller IC performs a *Verify and Delete* operation to check for data integrity. Data transmitted back on the FC link does not contain any DIF information.

If the target disk drive is configured for a 512-byte sector size, it may not be able to handle the DIF information that is appended by the Fibre Channel Interface Controller IC in the operations described above. In that case, the Initiator Fibre Channel Interface Controller IC could be programmed to terminate the protection domain during the Initiator Write operation by performing a *Verified and Delete* action when writing data to the target disk drive. During Initiator Read operations, the Fibre Channel Interface Controller IC may be programmed with the *Insert* action to start the protection domain. This causes the data that is written to the memory subsystem to contain DIF information.

Conclusion

The need for a standardized end-to-end data integrity protection solution can now be realized using DIF provided by Agilent's Fibre Channel Interface Controller IC. As can be seen from the various examples, the standardized approach allows for the expansion of the protection domain to the maximum extent possible. It is therefore conceivable to have complete end-to-end protection from the host server system to the disk drives. By standardization of the protection mechanism, early detection of data corruption is facilitated, leading to more effective fault isolation. Agilent's Fibre Channel Interface Controller IC has implemented this protection scheme in hardware at 2 Gb/s, leading the way for a high performance solution that is critical as the Fibre Channel link scales from 2 Gb/s to 4 Gb/s and 10 Gb/s.

Note:

The T10 technical committee of the InterNational Committee for Information Technology Standards (INCITS) develops standards and technical reports on I/O interfaces, particularly the series of SCSI standards.

Document T10/03-111 is available at:

http://www.t10.org/ftp/t10/document.03/03-224r0.pdf

Other related documents:

http://www.t10.org/ftp/t10/document.03/03-111r0.pdf

http://www.t10.org/ftp/t10/document.03/03-290r0.pdf

Following Fibre Channel Controller IC's support DIF

HPFC5700A	4G 2 port PCI-X Fibre Channel Controller IC
HPFC6200A	2G 4 port PCI-Express x8 Fibre Channel Controller IC
HPFC6400A	4G 4 port PCI-Express x8 Fibre Channel Controller IC

www.agilent.com/
semiconductors

For product information and a complete list of distributors, please go to our web site.

For technical assistance call:

Americas/Canada: +1 (800) 235-0312
or (916) 788-6763

Europe: +49 (0) 6441 92460

China: 10800 650 0017

Hong Kong: (+65) 6756 2394

India, Australia, New Zealand: (+65) 6755 1939

Japan: (+81 3) 3335-8152 (Domestic/International), or 0120-61-1280 (Domestic Only)

Korea: (+65) 6755 1989

Singapore, Malaysia, Vietnam, Thailand, Philippines, Indonesia: (+65) 6755 2044

Taiwan: (+65) 6755 1843

Data subject to change.

August 31, 2004
5989-0892EN

Agilent Technologies