End-to-End Data Protection

Ted Pang
Director, Tech Research Dept.
Infortrend Technology, Inc.
Agenda

- Data Integrity Requirement
- How to Detect Data Errors
- RAID Subsystem with DIF
- Security for Storage
- Encryption for Data on Rest
Data Integrity Requirement

- Data integrity is a fundamental aspect of storage security and reliability.
- Errors could happen at the point from host to disks.
- Errors could be caused by hardware or firmware/software
- Requirement for End-to-End Data Protection
End-to-End Data Protection

- Driver
- Application server
- Host
- Application server
- HBA
- Cable
- Switch
- Hard disks
- RAID Controller
  Hardware/Firmware
Current Mechanism for Data Integrity

- **Hard Disk Drive**
  - ECC for each data block (512 bytes) to recover error
  - More bytes for SCSI/FC/SAS drives (configurable)

- **RAID Controller**
  - RAID Parity Checking
  - Regenerate for bad sectors (media error)
  - Rebuild for disk failure

- **Data Link**
  - Checking mechanism (CRC, ECC)
Undetected Error in Storage System

- Silent Data Corruption
  - errors unreported and undetected in storage systems
  - corrupt data will be returned to applications without any warning
  - Even single bit error will affect lots of data

- Corruption could caused by
  - On-the-wire corruption
    - Data path corruption
  - Incorrect writes
    - Writing incorrect data
    - Unfinished writing
    - Misdirected writes (displacement)
  - Unknown Reason
How to Detect Data Errors

- Media scan and rebuild
  - background scrubbing and verification
- Media error and mark bad for the data block
  - Return the error to host for handling
- RAID Parity check on read
  - Parity check for the whole RAID stripe when read.
- We need more checking mechanism for data blocks
**Protection for Data Integrity**

- **Goal:**
  - Protect each 512 byte block of data with an additional 8-bytes
  - Protection information written to the disk along with the original Block of data
  - This extra parity data is checked as the data is read back from the disk

- **Data Integrity Field defined by T10**
  - EEDP (End-to-End Data Protection)
  - Blockguard

- **External Path Protection defined by T13**
  - proposed by WD
What is DIF

- **8 bytes DIF for each standard 512 block of data**
  - Logical Block Guard 2bytes - CRC (Check bit error)
  - Logical Block Application Tag 2bytes - User Defined
  - Logical Block Reference Tag 4 bytes - LBA (Check displacement)

- SATA T13/ EPP (External Path Protection) uses same format
What does DIF do

- Allow application client to mark data with integrity information during writes.
- Allow storage device to examine integrity information received during writes for consistency.
- Require storage device to return integrity information unchanged.
- Allow application client to verify integrity of returned data.
- The corrupt data is not sent to the application or written to disk.
RAID Subsystem with DIF

- Host with DIF

Check DIF before write to disk

Return fail if detect corruption
RAID Subsystem with DIF(2)

- Host without DIF
  - Insert DIF and pass to next stage
  - Remove DIF and return to host
T10 DIF Support

- **Host Initiator (HBA)**
  - Generate and check DIF
  - FC/SAS Initiator

- **Storage System**
  - Enabled/Disabled checking by user
  - Generate DIF inside the storage box

- **Disk Drives**
  - FC/SAS support 520 bytes with DIF
  - SATA is still working for T13 EPP
Impact for RAID Subsystem with DIF

- Performance could be little degraded since 8 extra bytes need to be processed.

- 1.5% extra capacity (8/512) tradeoff

- For hard disk without DIF, it could need special treatment.
Existing Products for Data Integrity

- Xanadu of RAID Inc (OEM from NEC D-series)
  - implements DIF mechanisms,
- LeftHand - Proactive Self-Healing (background scrubbing)
- DataDirect - SATAssure
  - implement continuous auto-scrubbing.
- Infortrend
  - We have supported background-scrubbing
  - We plan to support DIF in the future, even for SATA drives.
Security Requirement for Storage

- **Authentication**
  - FC-SP for FC
  - CHAP for iSCSI

- **Authorization**
  - FC WWNs
  - iSCSI IQN

- **Encryption**
  - IPSec for iSCSI
  - IEEE P1619 for Data at Rest
Security for Data at Rest

- Used to protect against unauthorized disclosure of sensitive data while they are resident in storage media (disk, tape, optical)

- Device-based
  - Full Disk Encryption
  - Encryption for Tape
  - Encryption for Disk Array

- Network-based

- Host-based
Encryption for Hard Disk

- FDE (Full Disk Encryption)
- Trusted Computing Group (TCG) based FDE
  - Automatic, always-on, hardware-based encryption
  - Transparent to end user, operating system, applications and databases
  - No performance degradation
  - IEEE1619.3 (key management) will be involved in this topic to be a standard key mgmt method in the future.

- Seagate
- HGST
IEEE P1619

- Security for storage data at rest (disk and tape)
- IEEE P1619 is an IEEE standardization project for encryption of stored data.
- Standard security transform that provides confidentiality and pseudo-integrity
  - Applied to 512-byte blocks
  - Without data expansion (no additional integrity tag)
  - Resistant to copy-and-paste attacks
  - Parallelizable for high speed HW
- Standard common format for key backup
  - Allows for decryption of a disk encrypted by any other vendor
EME-32-AES (Wide-block)

where:

\[ L = \text{AES}_k(0) \]

\[ SP = \text{MP} \oplus \text{MC} \]

\[ SP = \text{PPP}_2 \oplus \ldots \oplus \text{PPP}_{32} \]

\[ SC = \text{CCC}_2 \oplus \ldots \oplus \text{CCC}_{32} \]
IEEE P1619 Standard

- P1619 Standard Architecture for Encrypted Shared Storage Media
- P1619.1 Standard for Authenticated Encryption with Length Expansion for Storage Devices (tape)
- P1619.2 Standard for Wide-Block Encryption for Shared Storage Media (disk)
- P1619.3 Standard for Key Management Infrastructure for Cryptographic Protection of Stored Data
IEEE P1619 for RAID Subsystem

- Hardware Computation Chip
- IO Chip Support
  - PMC-Seria PMC-8002 SAS6G IOC
- RAID Controller ASIC/Firmware
Introduction for Infortrend

Basic Infortrend

- Founded in 1993, one of the most important RAID manufacturers and marketing companies worldwide.
- A listed company in Taiwan Stock Exchange (TPE: 2495)
- A global company with branches in States, UK, Germany, China, Japan...

High-Growth Infortrend

[Bar chart showing Infortrend Revenue from 2002 to 2007]
What we achieve

- **Reliable Infortrend**
  Till Now, we have delivered:
  - 80,000 RAID Subsystems
  - Over 160,000 controllers and subsystems worldwide

- **Innovative Infortrend**
  - World’s 1st external iSCSI-to-SAS/SATA RAID array w/8 x1GbE ports (2008)
  - World’s 1st external RAID array that houses 2.5” SAS drives (2007)
  - World’s 1st external RAID array that houses SAS drives (2005)
  - World’s 1st external RAID array that houses SATA drives (2003)
New Strategies for Infortrend

- Performance improvements
- Data integrity improvements
- Backup/Recovery Solution
- Virtualization Architecture