



Education

## **Why Data Integrity is important to you**

Richard Vanderbilt, LSI

- The material contained in this tutorial is copyrighted by the SNIA.
  - Member companies and individuals may use this material in presentations and literature under the following conditions:
    - ◆ Any slide or slides used must be reproduced without modification
    - ◆ The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
  - This presentation is a project of the SNIA Education Committee.
  - Neither the Author nor the Presenter is an attorney and nothing in this presentation is intended to be nor should be construed as legal advice or opinion. If you need legal advice or legal opinion please contact an attorney.
  - The information presented herein represents the Author's personal opinion and current understanding of the issues involved. The Author, the Presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.
- NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

## ➤ Why Data Integrity is important to you

- ◆ This session will appeal to Data Center Managers, Development Managers, and those that are seeking a fundamental understanding how and why silent data corruptions occur, the affects it has on a data center and how the industry is moving to minimize its occurrence.

# Data Integrity VS Corruption

- At the Storage Level, There Are Two Kinds of Data Corruption
  - ◆ **Latent sector errors**
  - ◆ **Silent data corruption**
- For a Storage Device, it is Usually the Case That...

**it is Better to Not Return Any Data,  
Than Return the Wrong Data!**

# Typical Causes of Data Integrity Problems

- **Operating System Bugs**
  - ◆ **Core O/S**
  - ◆ **Device drivers**
- **Storage Hardware and Firmware Bugs**
  - ◆ **HBA's**
  - ◆ **Arrays**
  - ◆ **Disks**
- **Administrative Errors- “The human factor”**
  - ◆ **System administrators**
  - ◆ **Database administrators**
  - ◆ **Storage administrators**

Why Data Integrity is important to you!

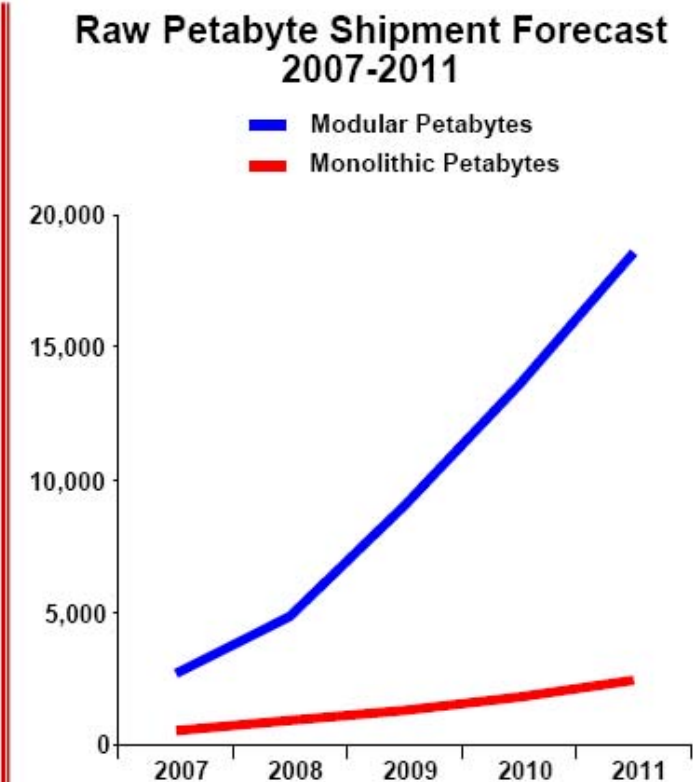
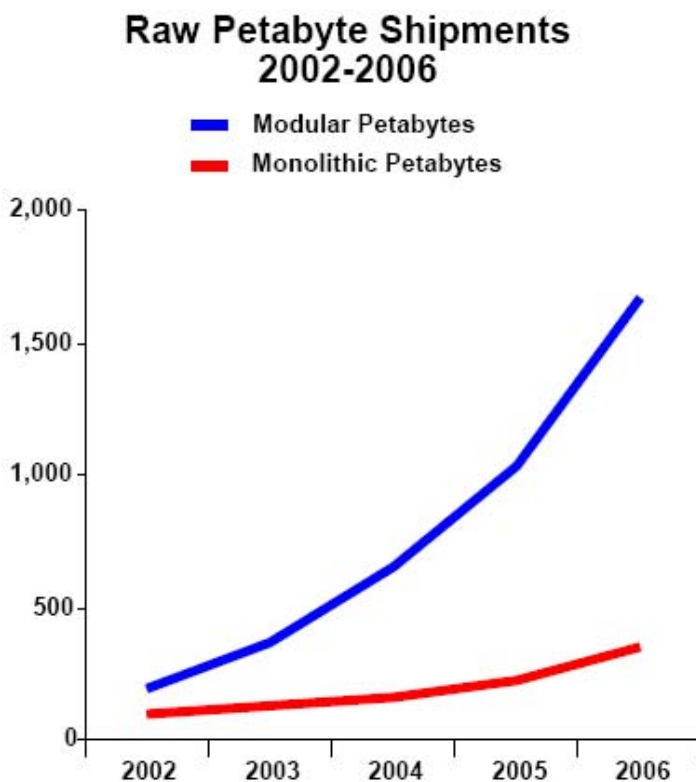
© 2008 Storage Networking Industry Association. All Rights Reserved.

# Technology transitions can add risk! SNIA

- Server and Storage ecosystem is changing
  - ◆ Multi Core processors 2, 4, 8, 16....
  - ◆ Ever expanding memory 2GB, 4GB, 8GB, 32GB...
  - ◆ Doubling wire speeds 8GFC, 10GbE, FCoE, Infiniband...
  - ◆ Higher density drives, doubling every six months
  - ◆ Geographically dispersed processing, virtual data centers, Grids, Clouds ...
  - ◆ Commodity SATA drives entering the enterprise
- Technology sectors progress independently of each other
  - ◆ Plugfests focus on their particular sector, not the entire solution

# Growth can add risk!

## Petabyte Shipment Trends — External Controller-Based Disk Storage



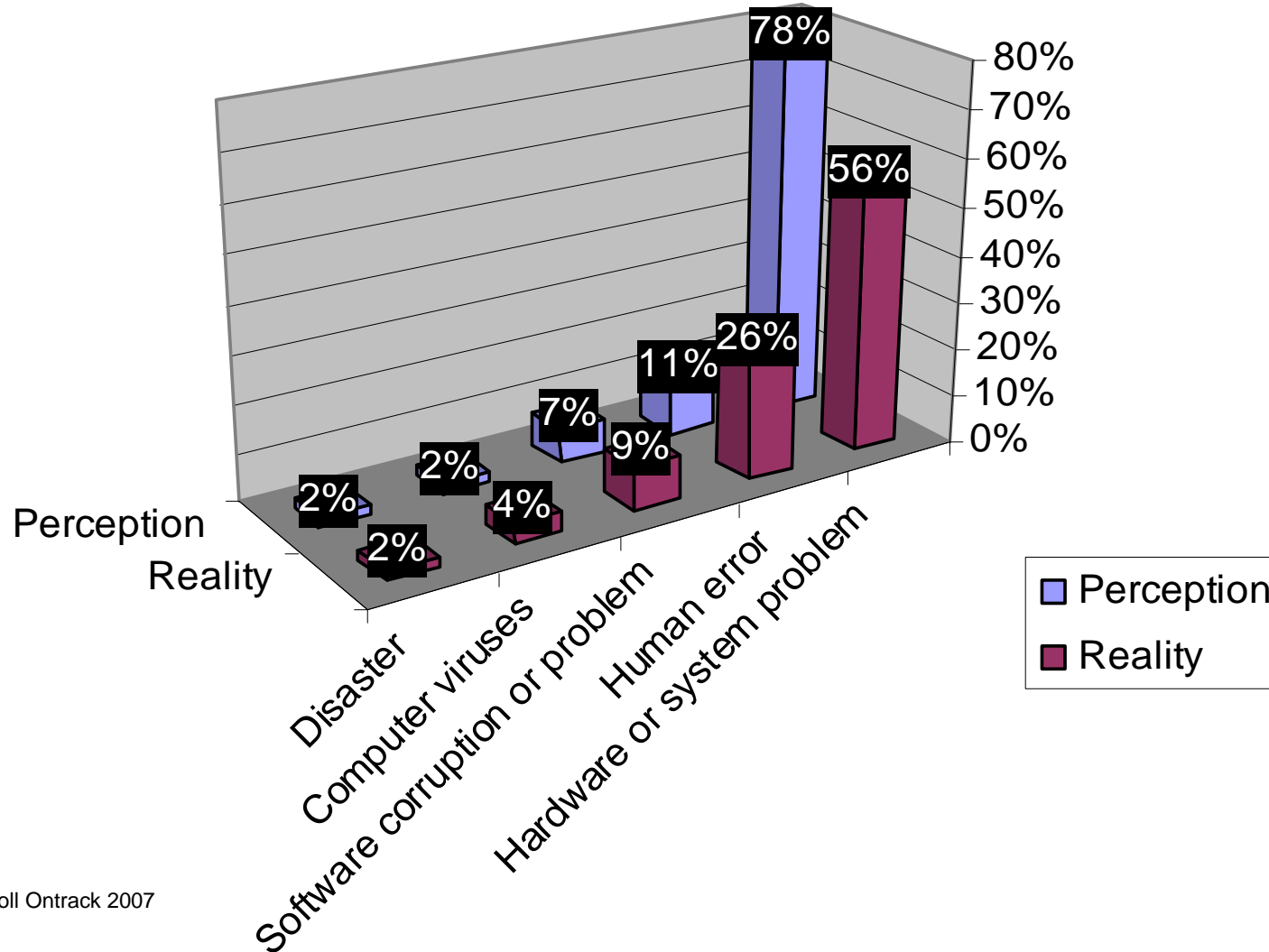
Ref 1 Gartner '07

Gartner

Why Data Integrity is important to you!

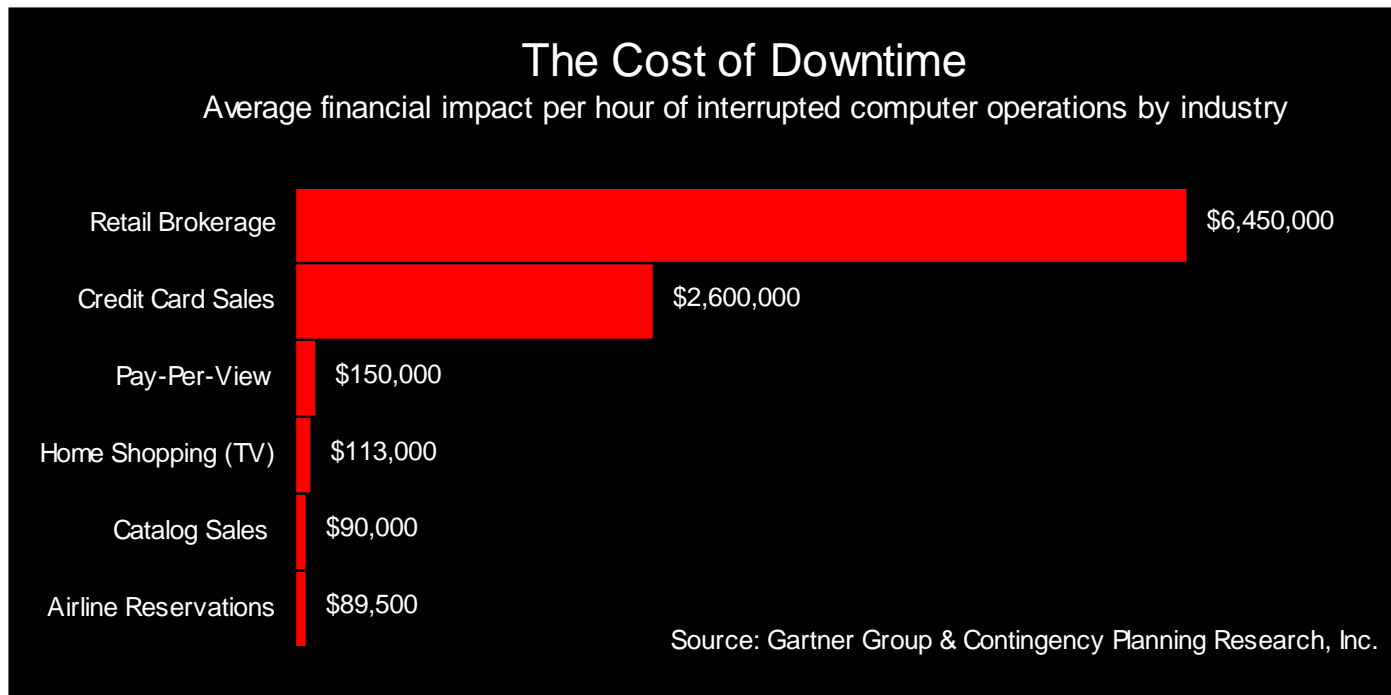
© 2008 Storage Networking Industry Association. All Rights Reserved.

# The Cause!



# The Effect!

- A well known e-commerce company was forced to shut down for days when a bug in the file manager caused bad data to be written onto their database
- A leading financial services company experienced repeated corruptions when a problem in the virtual memory system caused the wrong data to be written onto the database



Why Data Integrity is important to you!

© 2008 Storage Networking Industry Association. All Rights Reserved.

# File Systems Today

## ➤ Observations:

- ◆ **Bugs are common**
- ◆ **Numerous bugs across the file systems tested, some of which are serious, and many of which are not found by other sophisticated techniques**

## ➤ Detection:

- ◆ **Sanity checking is of limited utility**
- ◆ **Many of the file systems use sanity checking**
- ◆ **Modern disk failure modes such as misdirected and phantom writes lead to cases where:**
  - › **Bad block thus passes sanity checks, is used and can corrupt the file system**
  - › **Indeed, all file systems tested exhibit this behavior**

## ➤ Recovery:

- ◆ **Automatic repair is used rarely by the file systems**
- ◆ **Most of the file systems require manual intervention . . . (i.e., running fsck)**

# Data Integrity Study 1

## ➤ Disk Errors

- ◆ **A special 2 GB file was written to more than 3,000 nodes every 2 hours and read it back**
- ◆ **Errors were checked for 5 weeks**
- ◆ **500 errors were found on 100 nodes**
  - Single bit errors: 10% of disk errors
  - Sector (512 bytes) sized errors: 10% of disk errors
  - 64 KB regions: 80% of disk errors

## ➤ RAID Errors

- ◆ **Verify command was run on 492 RAID systems each week for 4 weeks**
- ◆ **The disks are spec'd at a Bit Error Rate of  $10^{14}$  read/written**
- ◆ **The good news**
  - That the observed BER was only about a 3rd of the spec'd rate
- ◆ **The bad news**
  - That in reading/writing 2.4 Petabytes of data there were some 300 errors

# Data Integrity Study 1 Cont..

## ➤ Memory Errors

- ◆ **Good news:**

- > 600,000 single bit errors were detected (1/3 of the expected rate at  $10^{12}$ )
- > only 3 double-bit errors in 3 months on 1300 nodes

- ◆ **Bad news:**

- > according to the spec there shouldn't have been any double bit errors AND these can't be corrected

## ➤ All of These Errors Will Corrupt User Data

- ◆ **8.7 TB of user data was checked for corruption, some 33,700 files...**
- ◆ **They found 22 corrupted files**
- ◆ **Which is 1 in every 1500 files**

# Data Integrity Study 2

- **1.53M Drives Monitored Over 41 Months**
  - ◆ **1.17M FC drives, 358K SATA drives**
- **Three Classes of Corruption Were Identified**
  - ◆ **Random**
  - ◆ **Lost writes (365 disks)**
  - ◆ **Parity inconsistencies (data scrubbing)**
- **Random Corruptions Detected With Data Scrubbing**
  - ◆ **49% for SATA**
  - ◆ **73% for FC**
- **Silent Data Corruption**
  - ◆ **Any media error not detected by the drives internal protection**
  - ◆ **400,000 instances were recorded**
  - ◆ **8% (on average) were found during RAID reconstructions**
  - ◆ **3088 SATA drives developed silent data corruption (0.86%)**
  - ◆ **767 FC drives developed silent data corruption (0.065%)**

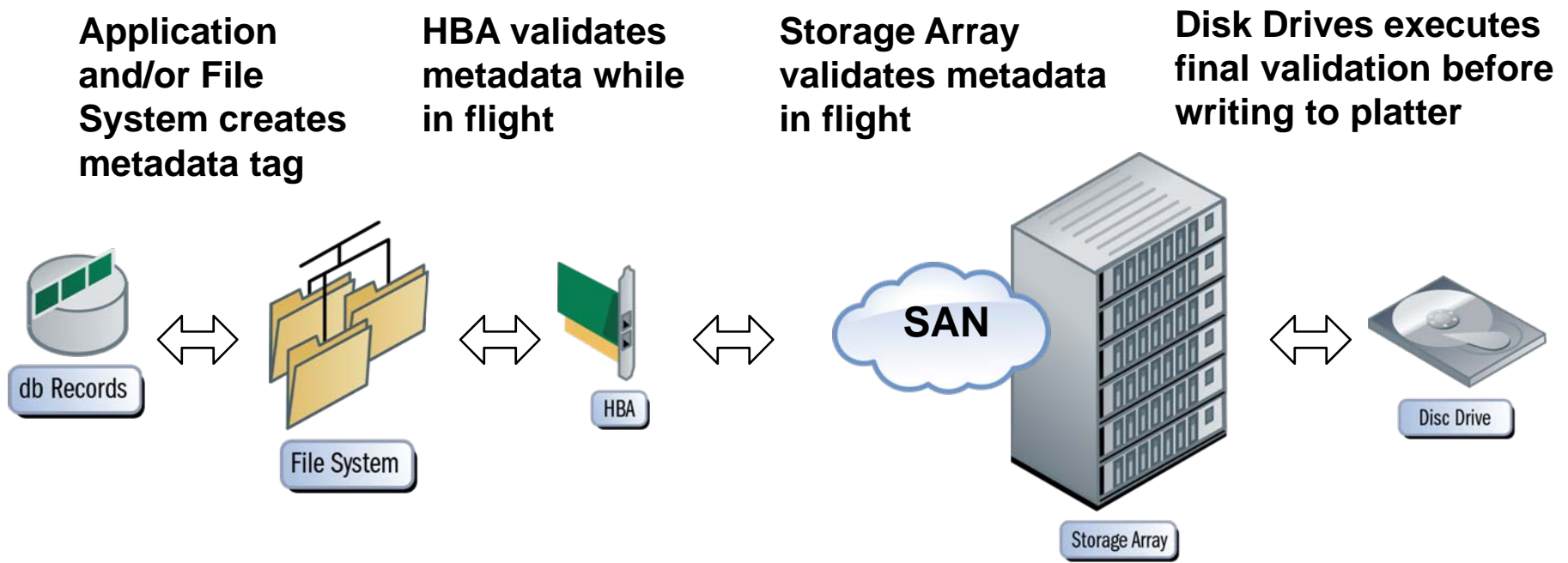
# True E2E ... 'IT' Takes a Village

- Individual companies strive to build better protection mechanisms within their own products
  - ◆ This addresses component failures only
  - ◆ Many disparate and proprietary technologies addressing device failures
  
- True end to end data integrity solutions require cooperation of everyone in the stack
  - ◆ Exchanging protection metadata between application and spindle
  - ◆ Application and or file system creates protection metadata
  - ◆ Metadata is verified in-flight and at rest
  - ◆ Corruption detection can occur before data is written to disk

# Application-to-Disk Data Integrity

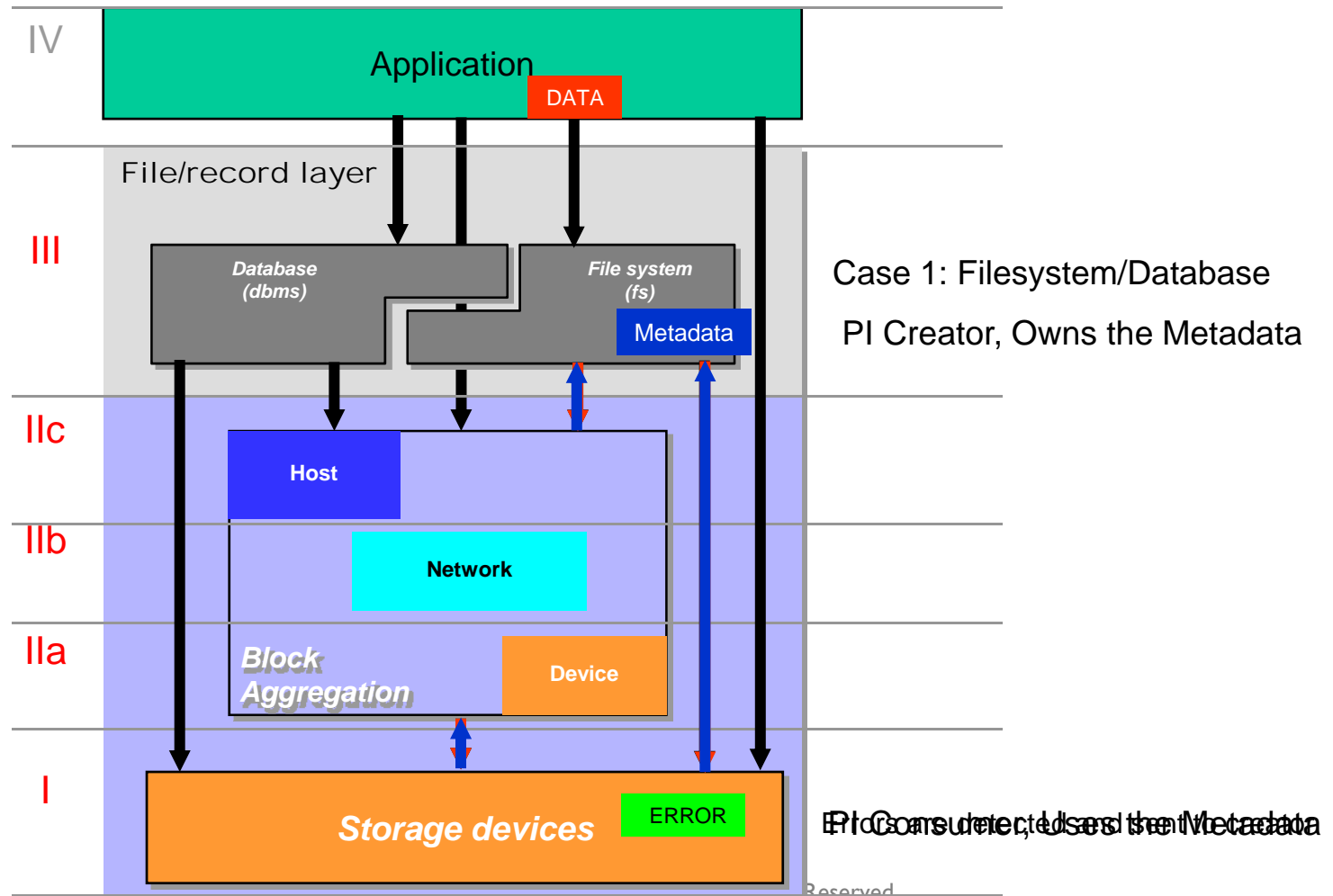
## ➤ Data Integrity Solution Overview

- ◆ Create integrity metadata coincident with data creation
- ◆ Validate the metadata throughout the data path
- ◆ Errors are directed to the application for remediation



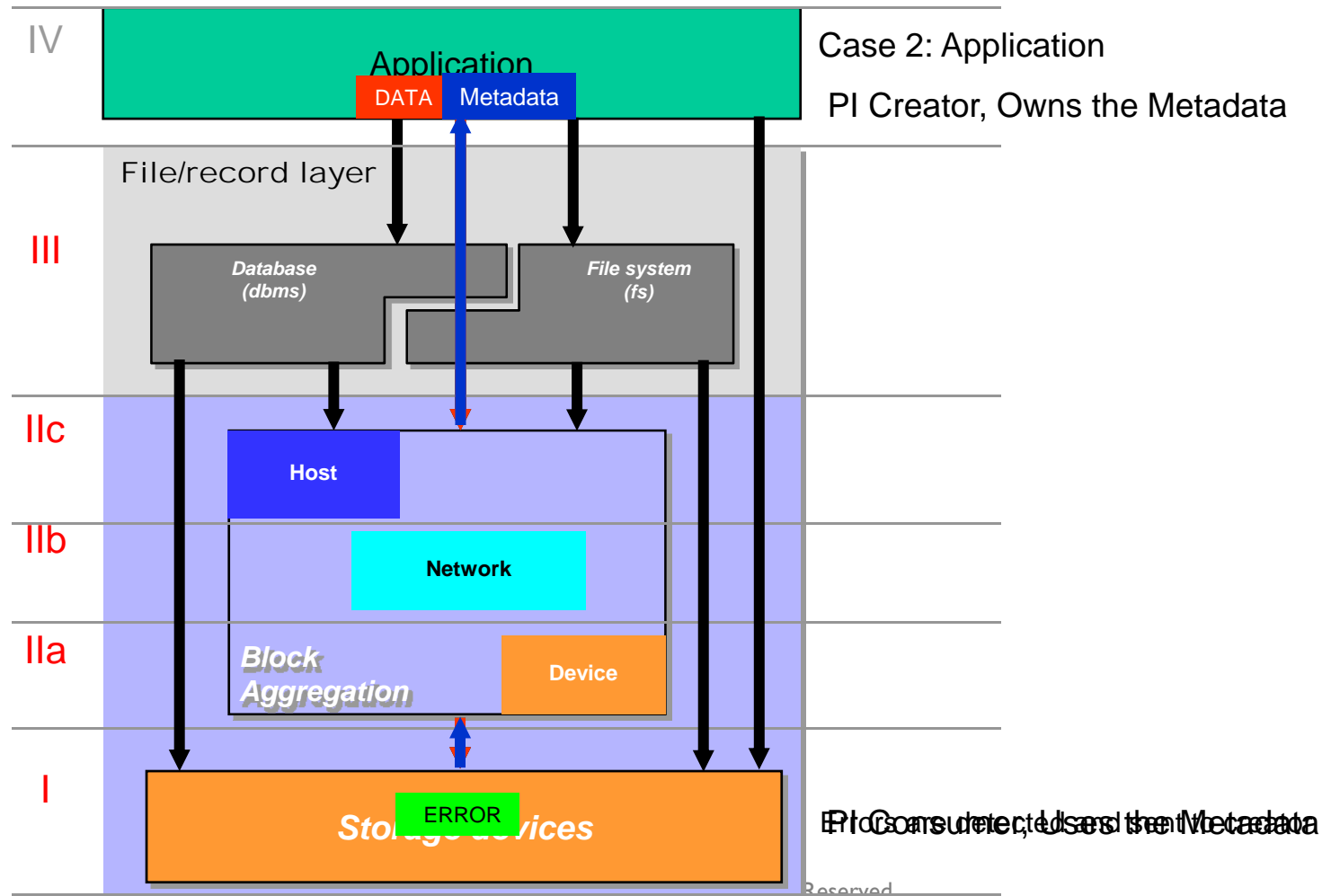
# Data Integrity Model

## SNIA Storage Model



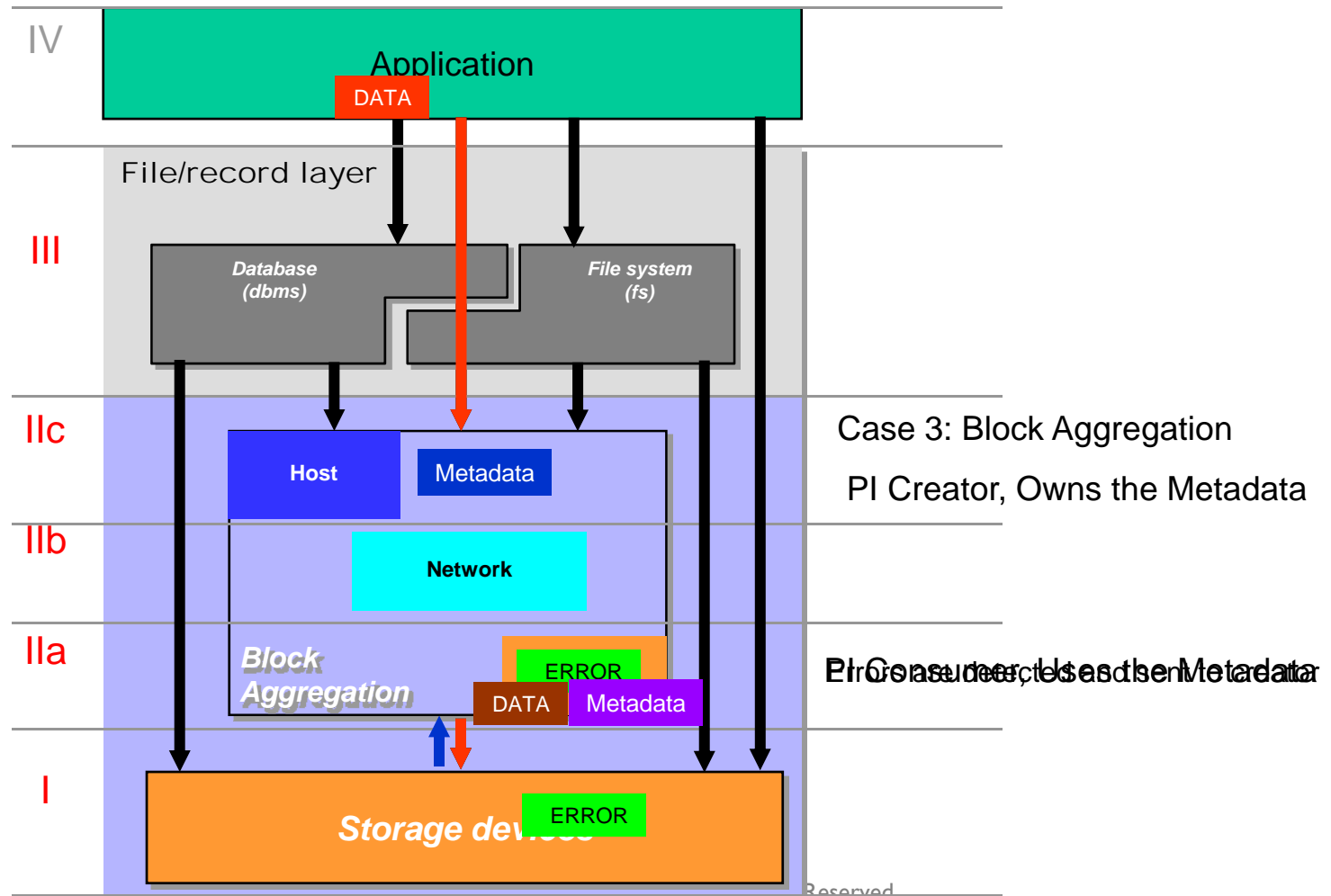
# Data Integrity Model

## SNIA Storage Model



# Data Integrity Model

## SNIA Storage Model

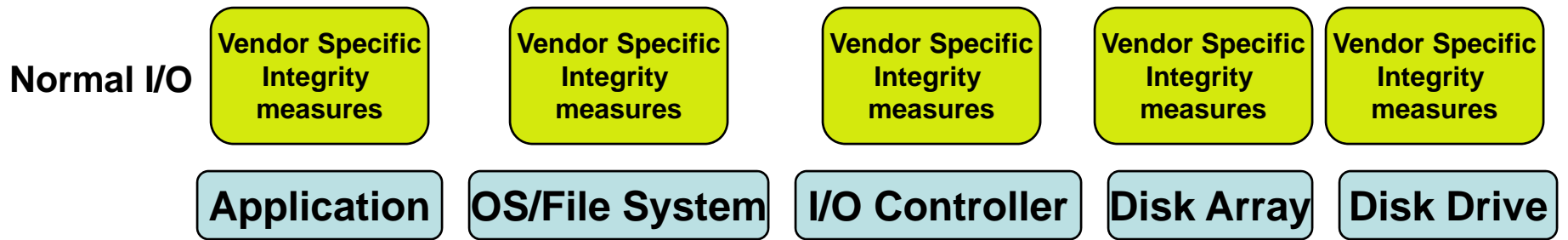


# Data Integrity Schemes

**T10-DIF + DIX(Data integrity plus extensions)** **Validates Data From End to End (True E2E)**

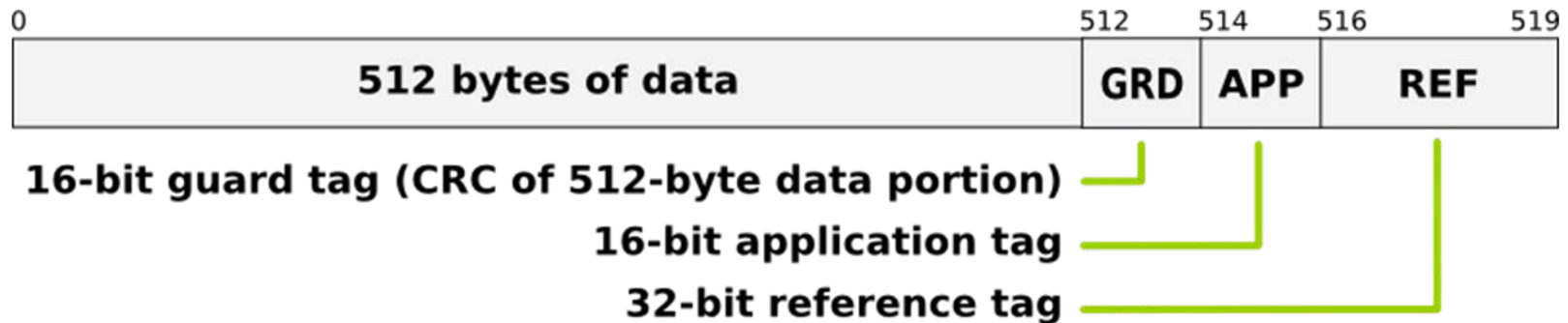
**T10-DIF(Data Integrity Field)** **Validates Data from Driver to Spindle**

**Other industry solutions** **Validates Data in flight to the Array**



From SNIA DITWG 080131

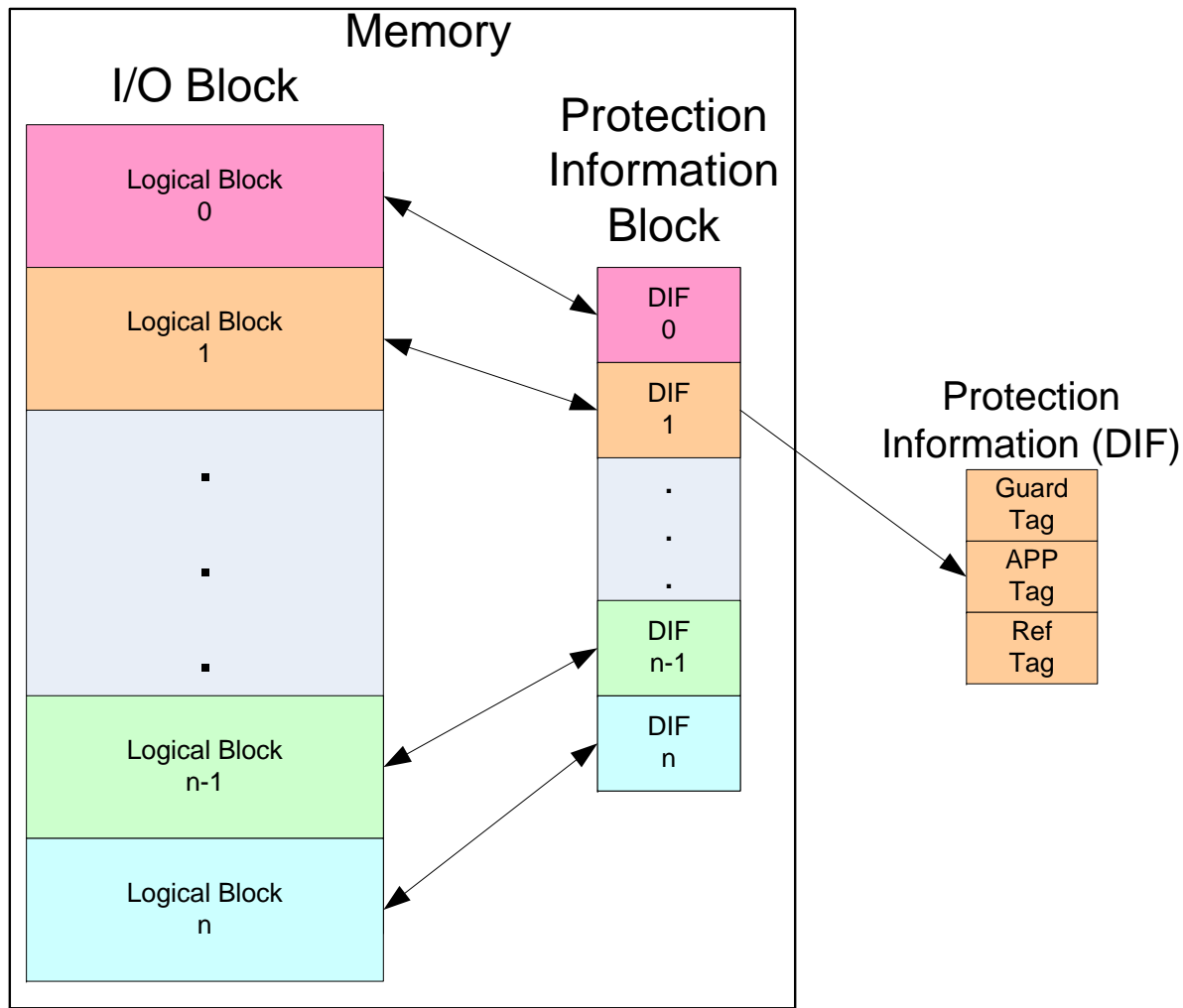
# T10 Data Integrity Field (DIF)



- Interleaved With Data Sectors on the Wire
- Three Protection Schemes: Type 1, 2 & 3
  - ◆ All have guard tag defined
  - ◆ Type 1 reference tag is lower 32 bits of target Logical Block Address
  - ◆ Type 2 reference tag is seeded in 32-byte Command (CDB)
  - ◆ Type 3 reference tag and application tag are combined

# T10 Data Integrity Field (DIF)

## + Data Integrity Extensions (DIX)



# Why Data Integrity is Important to you

- Silent data corruptions can have irreparable damage
- Backups may have bad data, and can't be trusted
- A data corruption could cause your DBA to take the wrong corrective action
- Downtime or corruptions will be costly, and may be fatal to a company
- **What Should You Do?**
  - ◆ Promote Standards - SNIA has recently approved Data Integrity (DI) as a formal Technical Working Group(TWG)
  - ◆ Promote the Technology – Research with your server/storage supplier their plans to enhance data integrity

- Please send any questions or comments on this presentation to SNIA: [trackstorage@snia.org](mailto:trackstorage@snia.org)

**Many thanks to the following individuals  
for their contributions to this tutorial.**

**- SNIA Education Committee**

**Dave Crespi  
Tom Hammond-Doel  
Brian McKean**

# Supporting Studies

- ◆ Ref 1: Gartner Study 2007
- ◆ Ref 2: Kroll On track 2007
- ◆ Ref 3: Gartner Group & Contingency Planning Research Inc.
- ◆ Ref 4: IRON File Systems
  - › V. Prabhakaran, L. Bairavasundaram, N. Agrawal, H. Gunawi – University of Wisconsin, Madison
- ◆ Ref 5: Data Integrity
  - › Bernd Panzer-Steindel CERN/IT, April '07
- ◆ Ref 6: Disk Failures in the Real World
  - › L. Bairavasundaram, G. Goodson, B. Schroeder, A. Arpaci-Dusseau, R. Arpaci-Dusseau, FAST'08