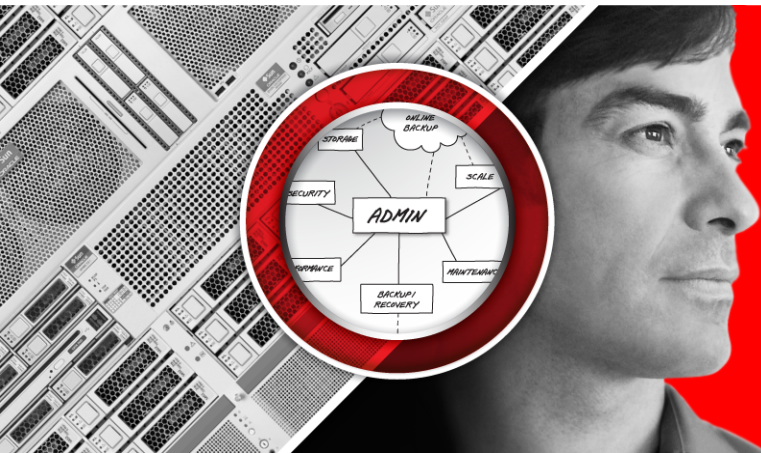


ORACLE®



**ENGINEERED
FOR INNOVATION**

**ORACLE
OPEN
WORLD**

ORACLE®

An Integrated End-to-End Data Integrity Solution for Oracle Products

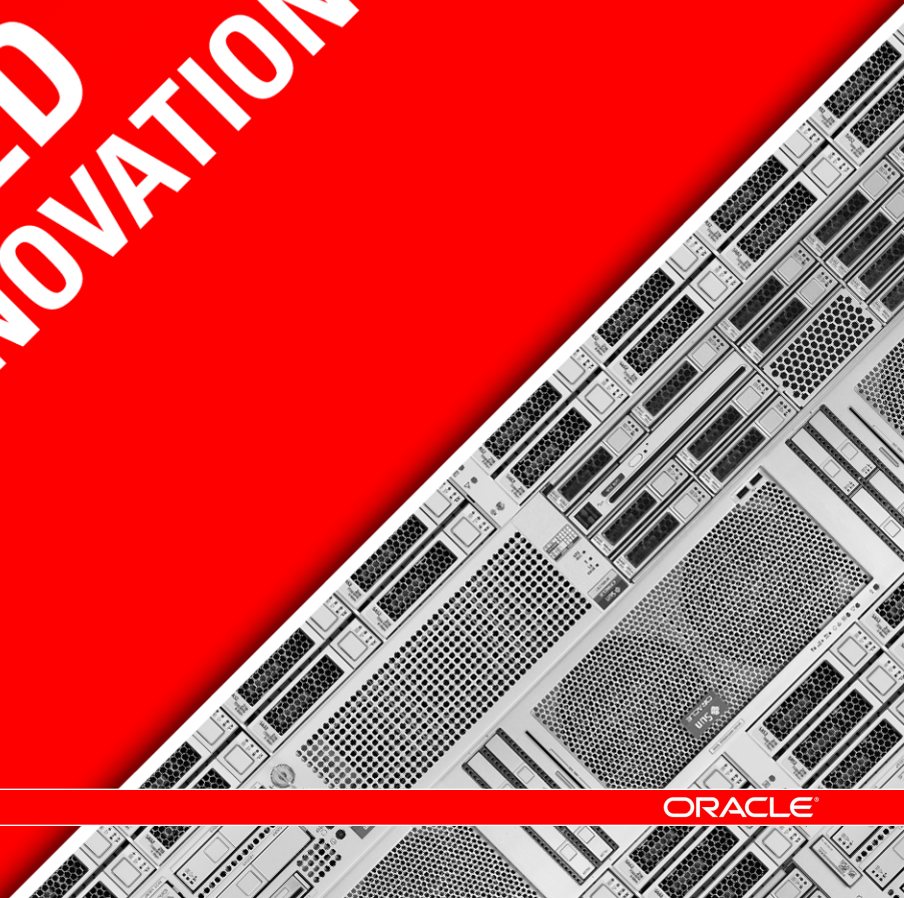
Martin K. Petersen, Oracle
Ken Taylor, EMC

Presenting with

EMC²



**ENGINEERED
FOR INNOVATION**





Latin America 2011

December 6–8, 2011

Tokyo 2012

April 4–6, 2012

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.



Program Agenda

- Data Corruption
- T10 PI & DIX
- EMC Presentation
- Technology Demonstration
- Q&A

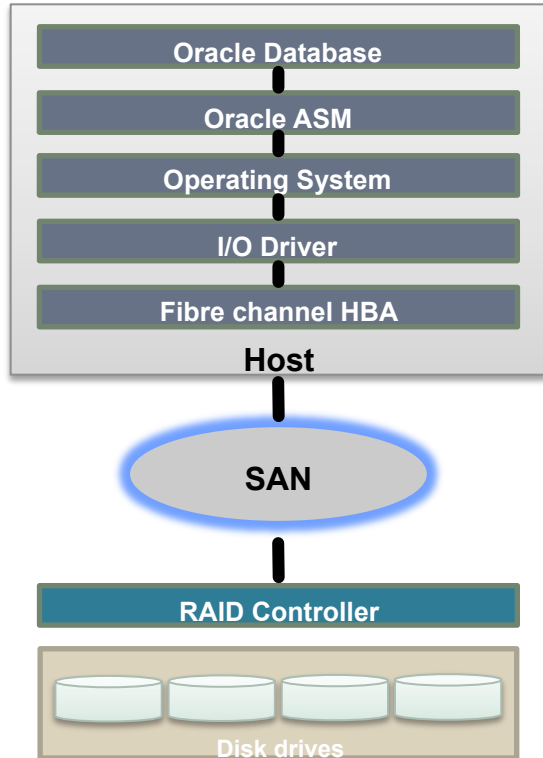


Data Corruption

Data Corruption

- Unintentional loss of data
- Data loss prevention
 - RAID, remote replication
 - Backups
- Data corruption detection
 - Logical block checksums: Oracle DB, Linux btrfs, Solaris ZFS
 - Scrubbing
- Not all data corruption scenarios are handled by the existing technologies

Data Corruption



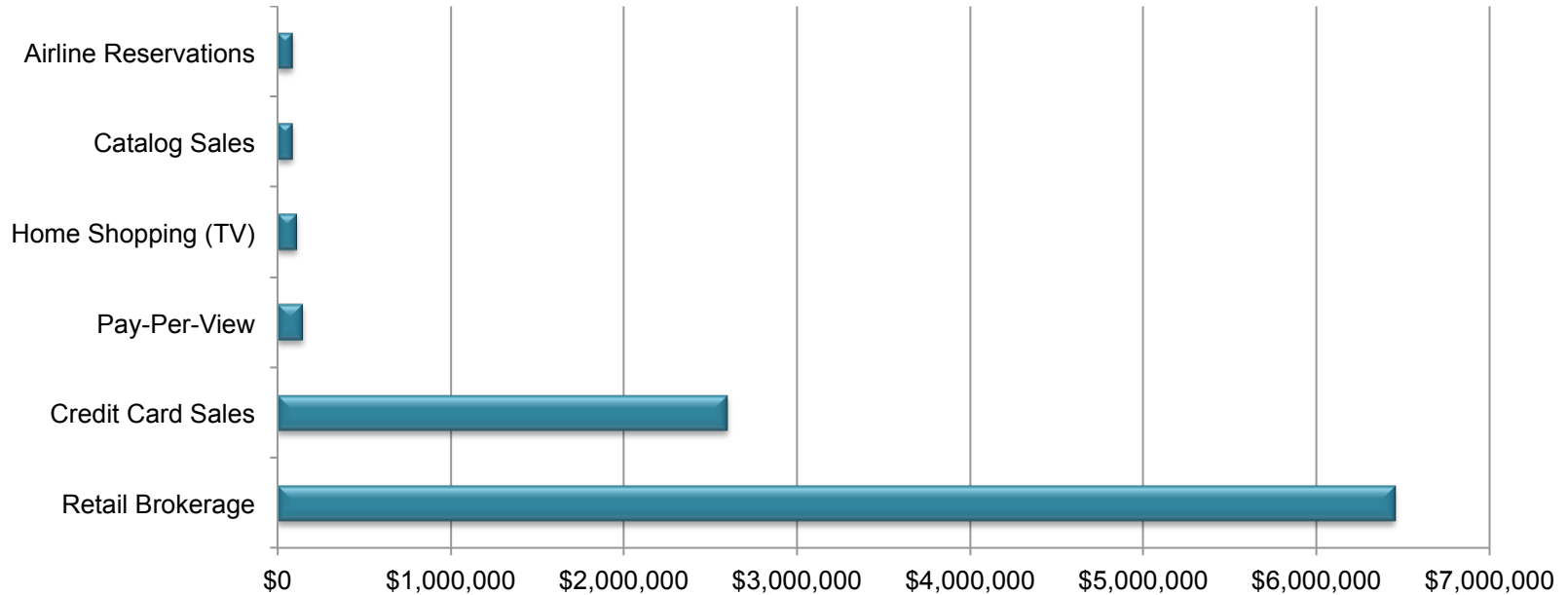
- Many potential sources of error in the I/O stack
- While many components employ data protection measures internally, handoffs between adjacent layers are typically unprotected
- Hardware
 - Memory, CPU, disk
- Firmware
 - HBA, RAID controller, disk
- Software
 - Library, OS kernel, hypervisor, device driver

Silent Data Corruption

- Data corruption that goes unnoticed
 - No errors or warning
 - Often discovered long time after the fact
 - Hard to root cause
- Logical block checksums are not an effective measure in these scenarios
 - Deployed at read time, original data buffers have been erased from host memory

Silent Data Corruption

Average financial impact per hour downtime by industry



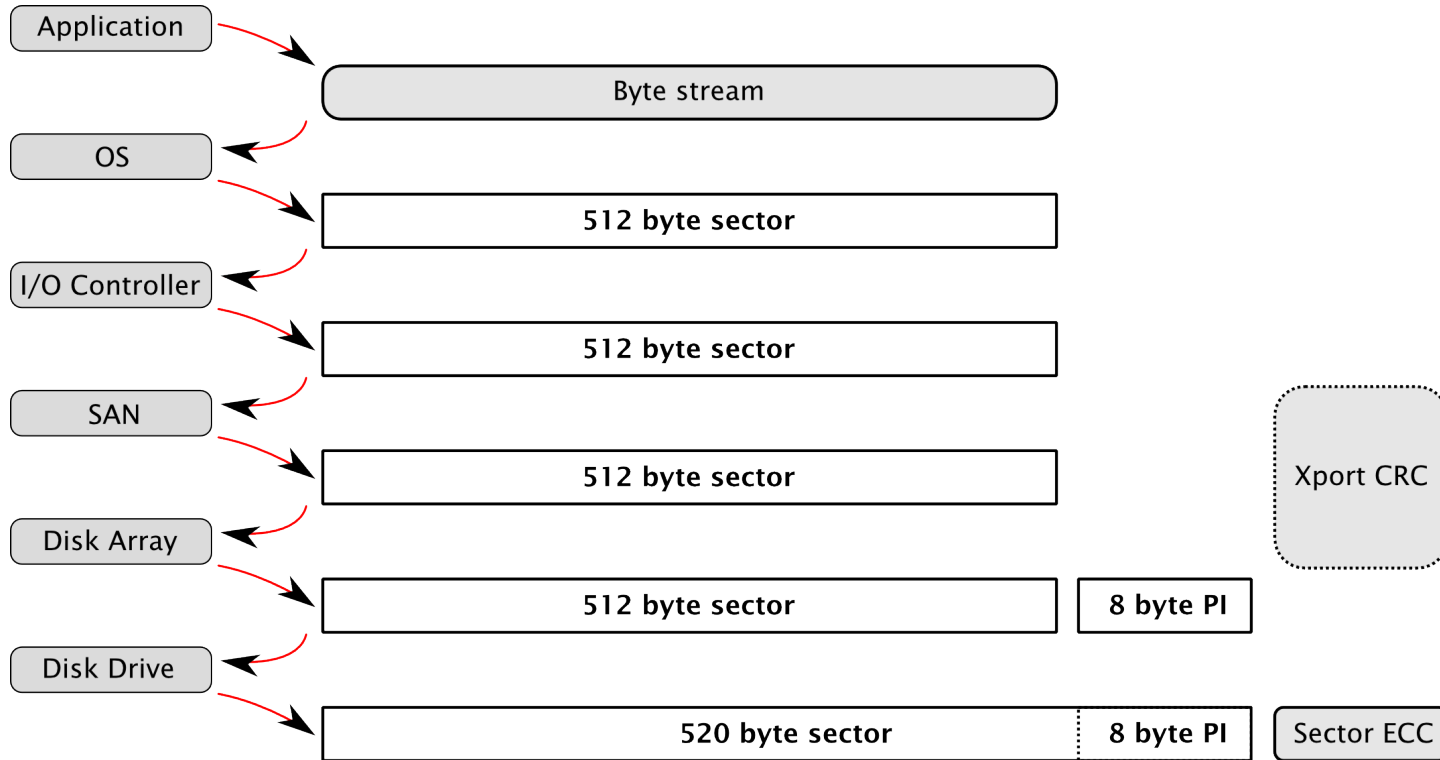
Source: Gartner Group & Contingency Planning Research, Inc.

Silent Data Corruption

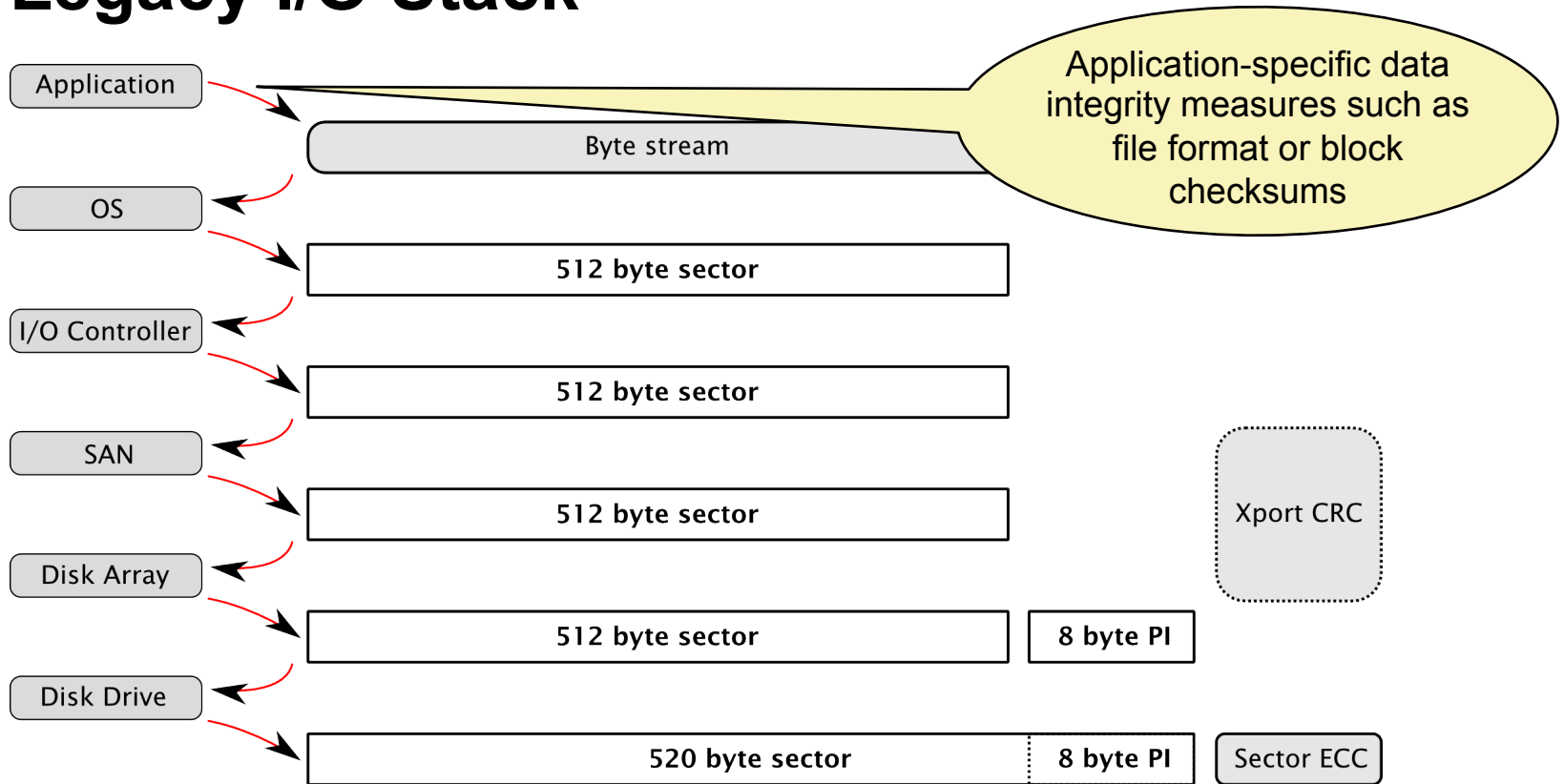
- NetApp, Univ. of Wisconsin, and Univ. of Toronto study¹
 - 41 month period
 - More than 1.5 million SATA and enterprise class fibre channel drives
- Silent data corruption detected:
 - 3,078 SATA drives
 - 760 fibre channel drives
- CERN study, 2007²
 - Write known data patterns to more than 3,000 nodes
 - 5 week period
 - 22 out of 33,700 files (8.7TB) corrupt
 - Nearly 1 in 1500 files

T10 PI & DIX

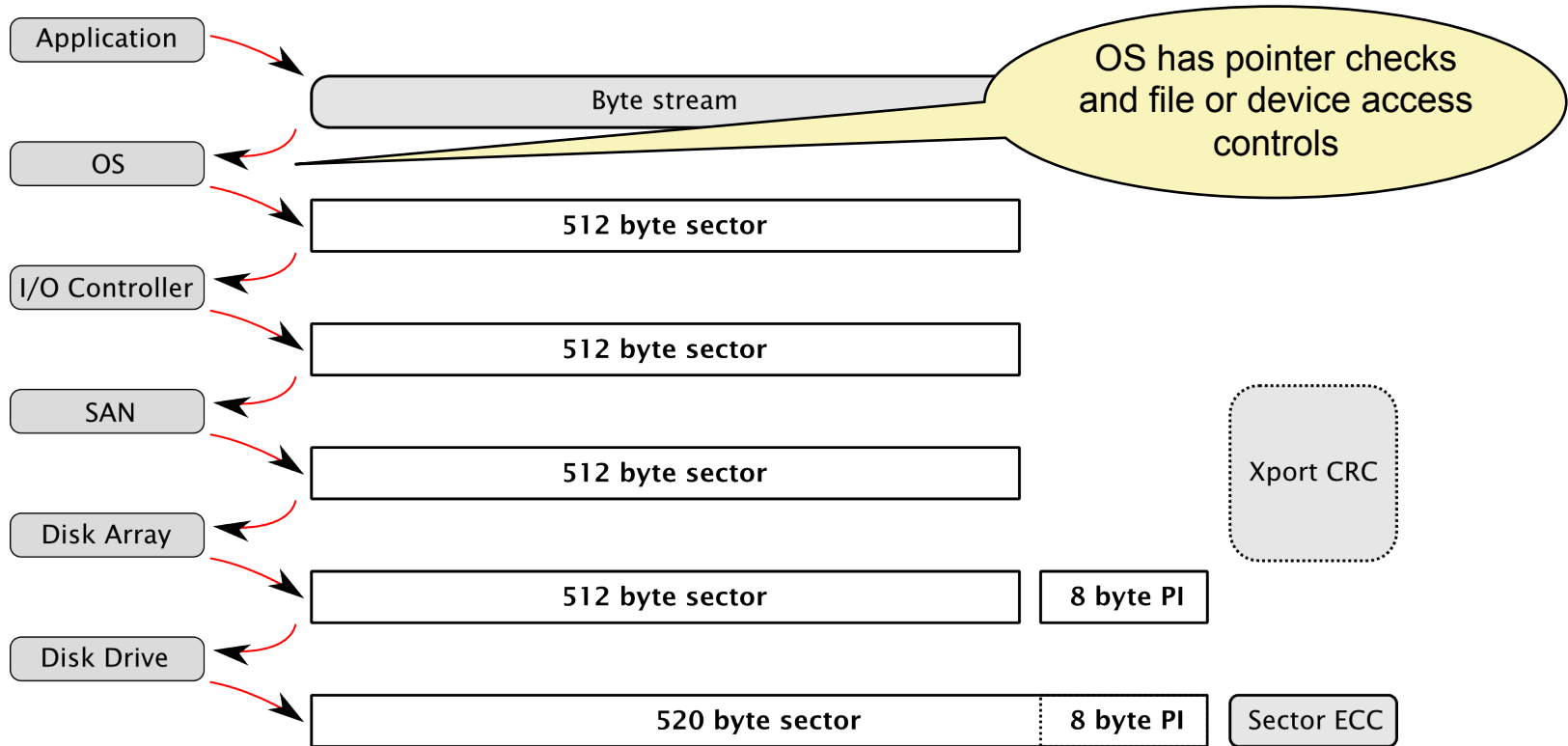
Legacy I/O Stack



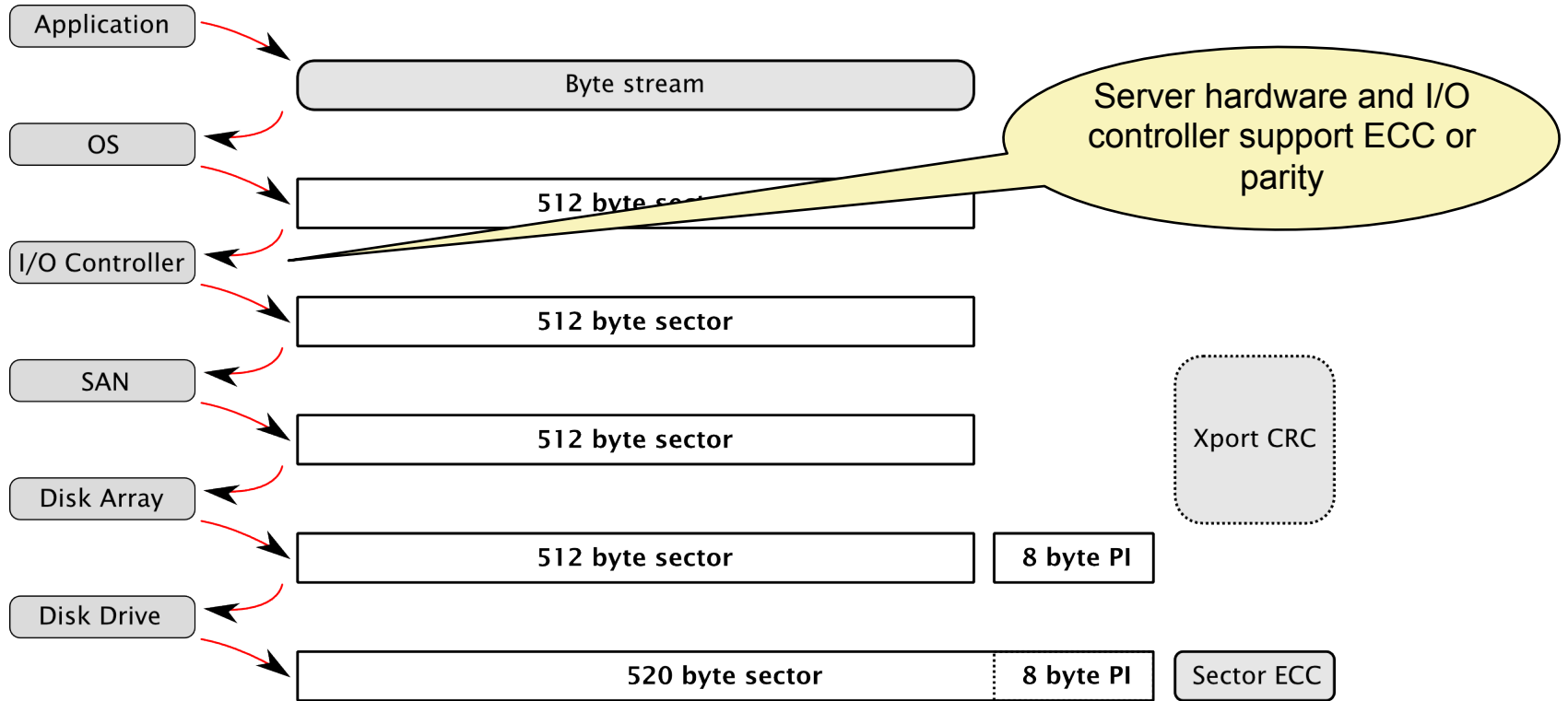
Legacy I/O Stack



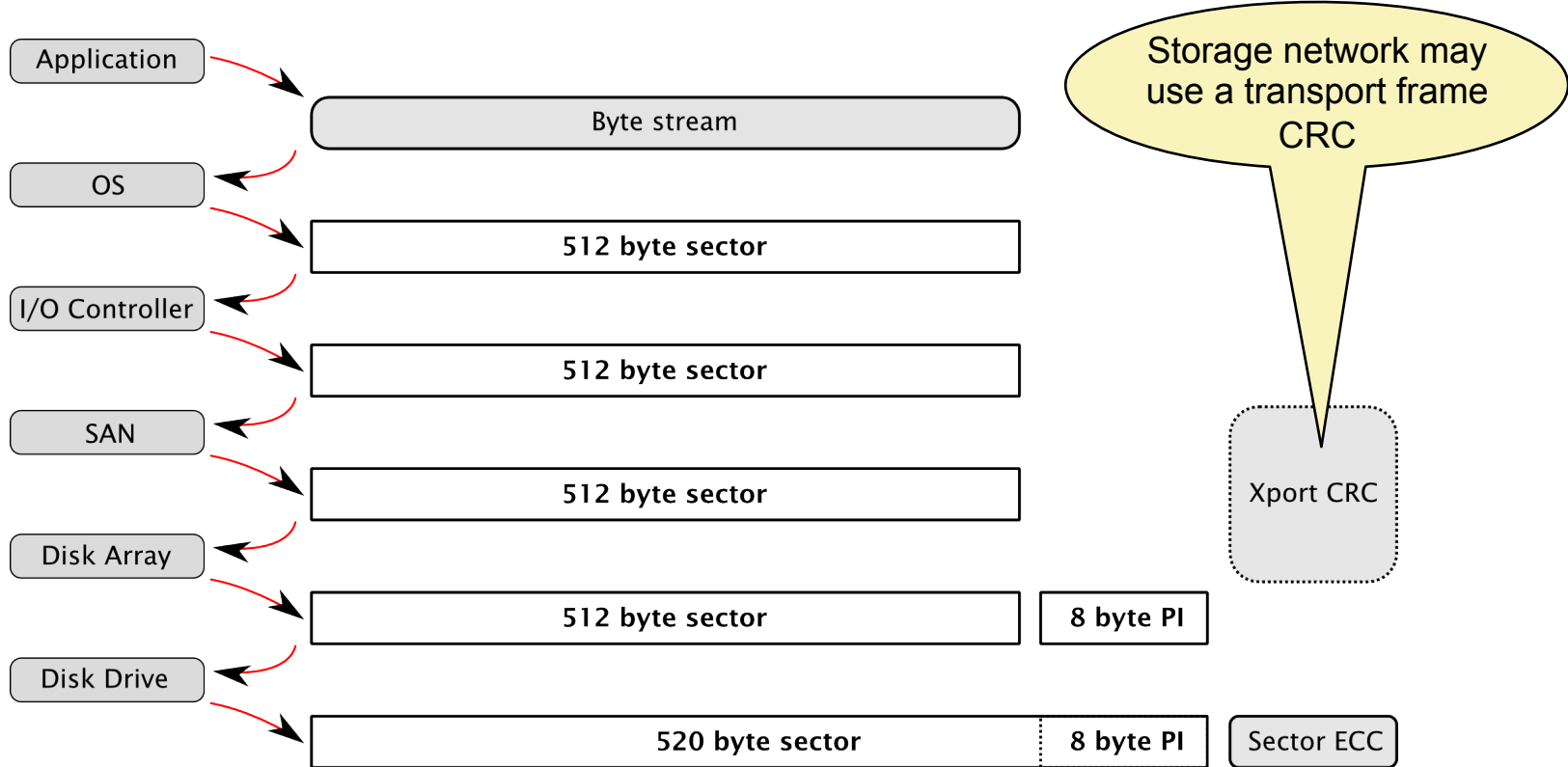
Legacy I/O Stack



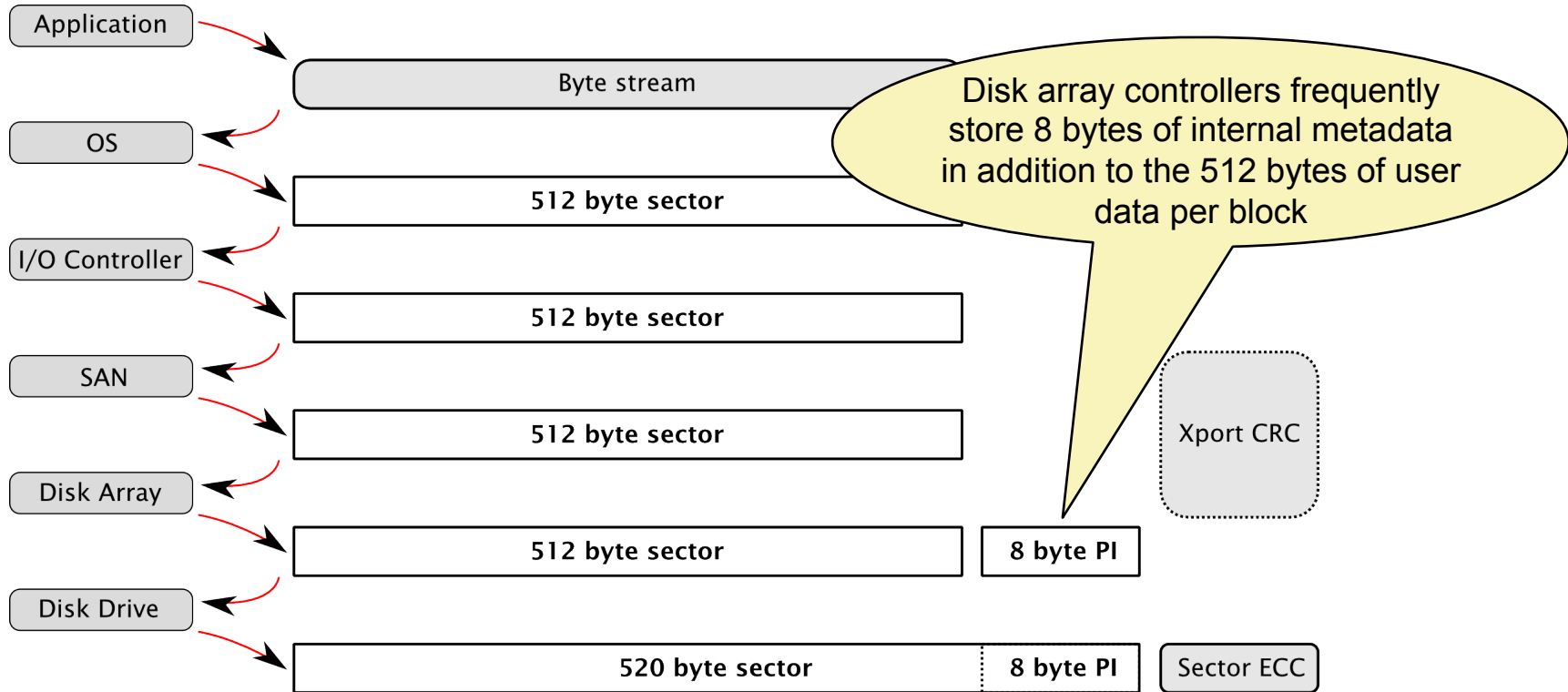
Legacy I/O Stack



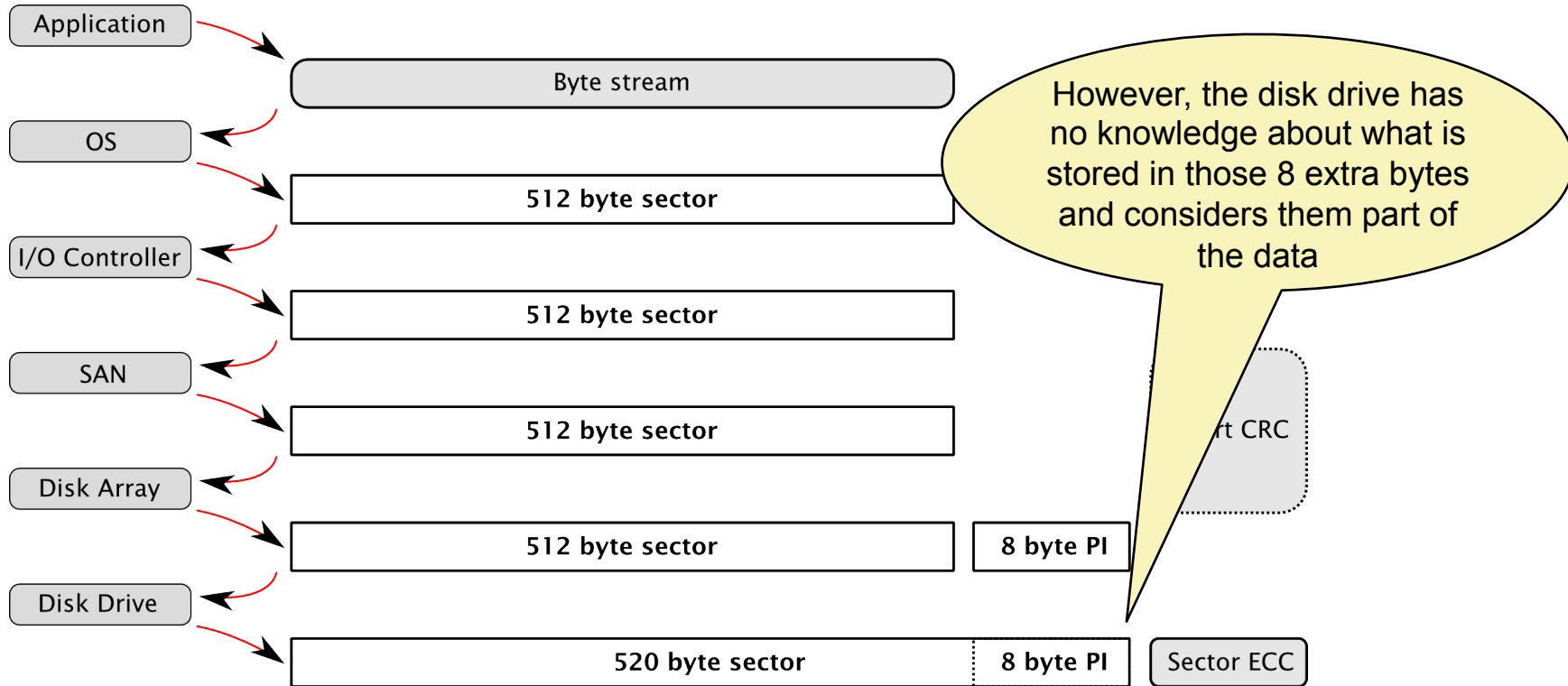
Legacy I/O Stack



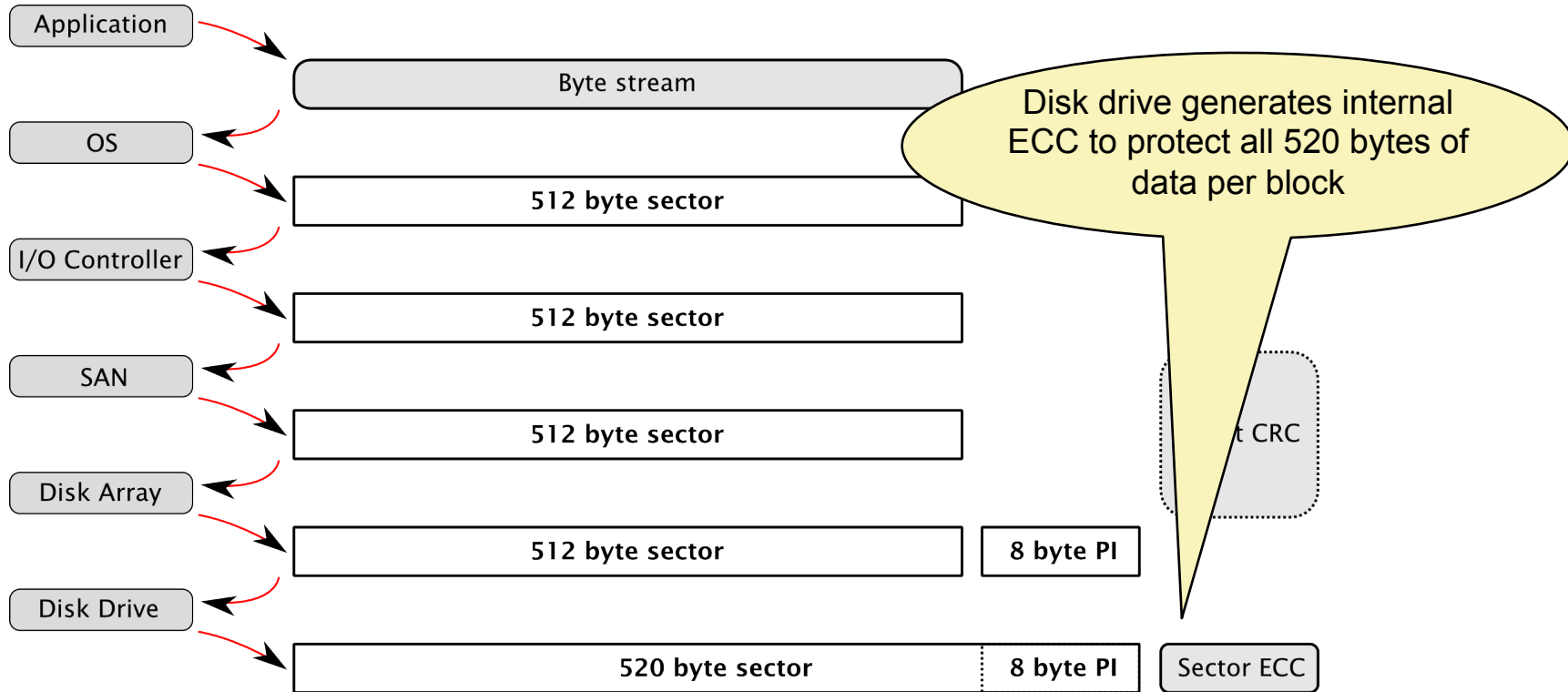
Legacy I/O Stack



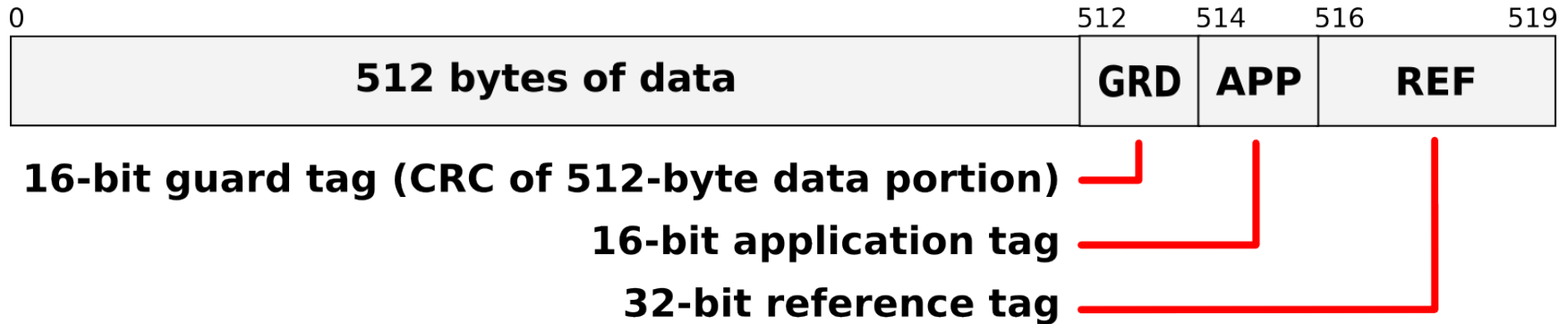
Legacy I/O Stack



Legacy I/O Stack

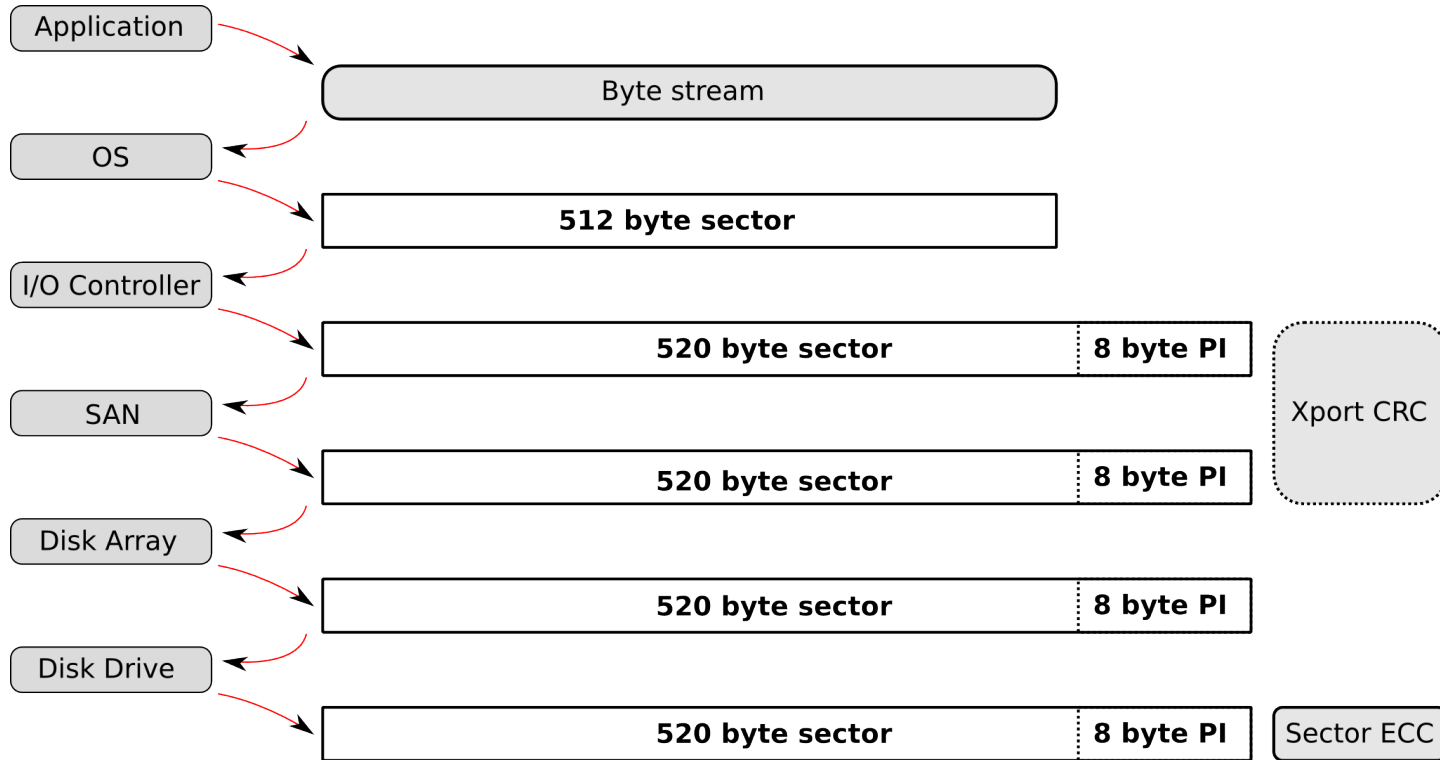


T10 Protection Information Model

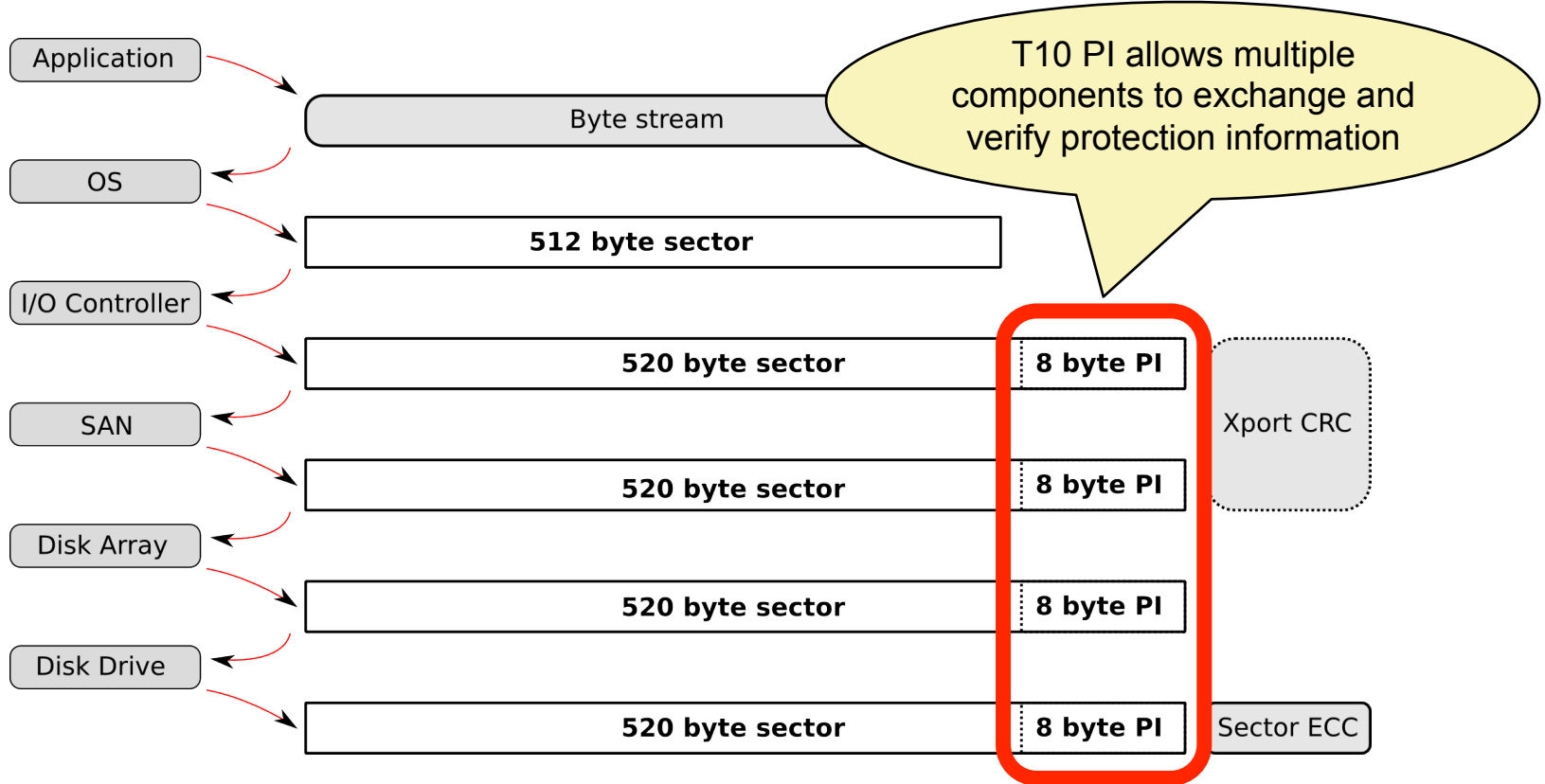


- Standardizes those extra 8 bytes
- Prevents content corruption and misplacement errors
- Protects path between HBA and storage device
- Protection information is interleaved with data on the wire, i.e. effectively 520-byte logical blocks

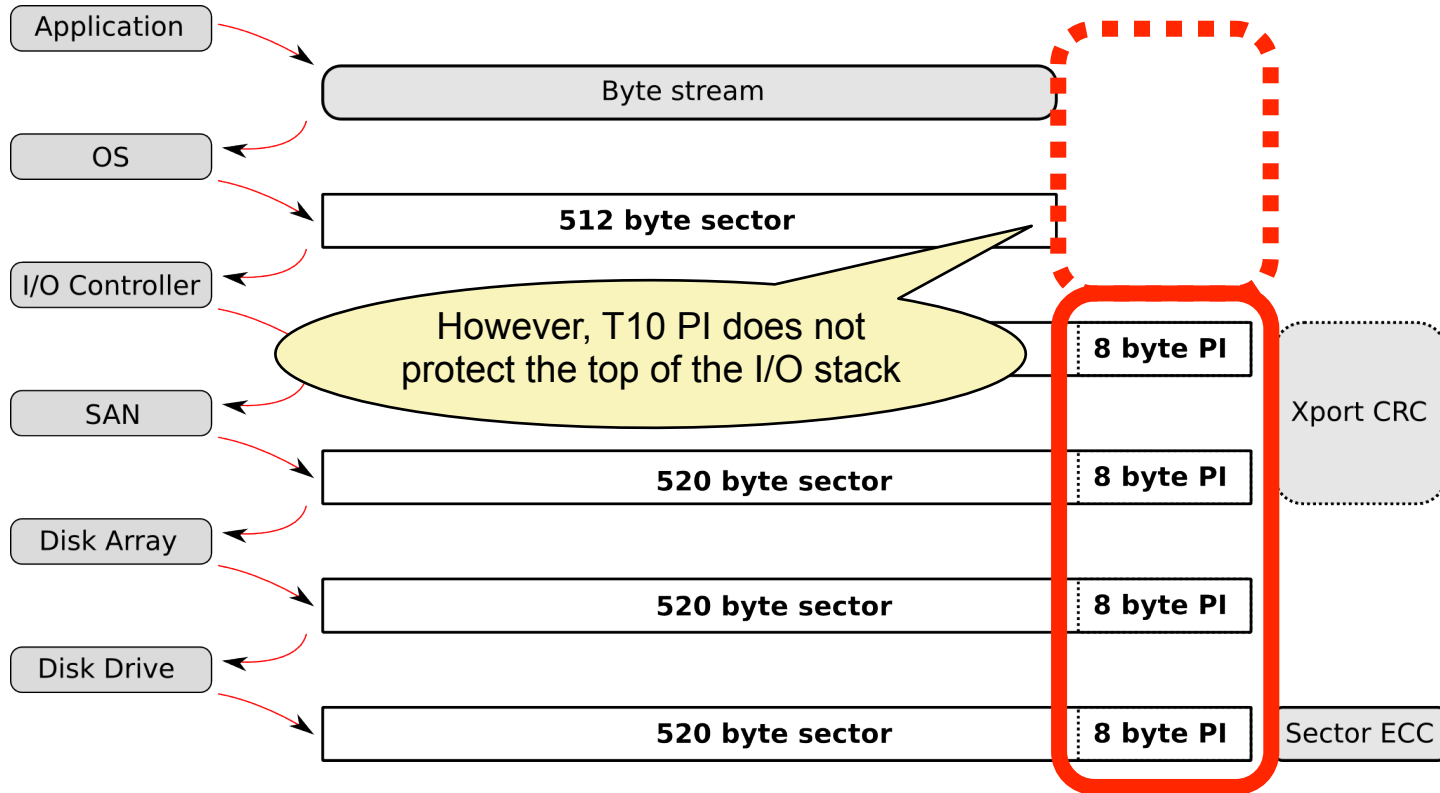
T10 Protection Information Model



T10 Protection Information Model



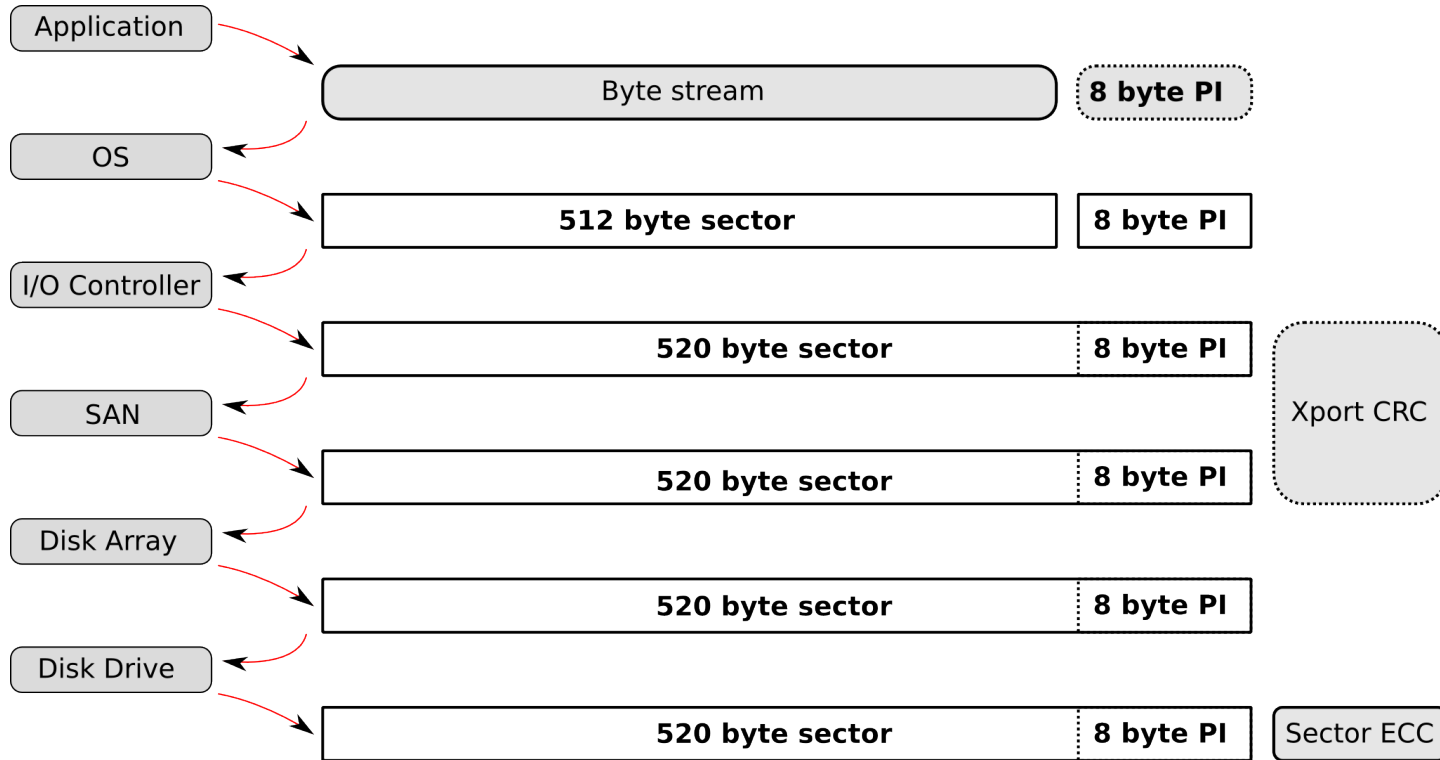
T10 Protection Information Model



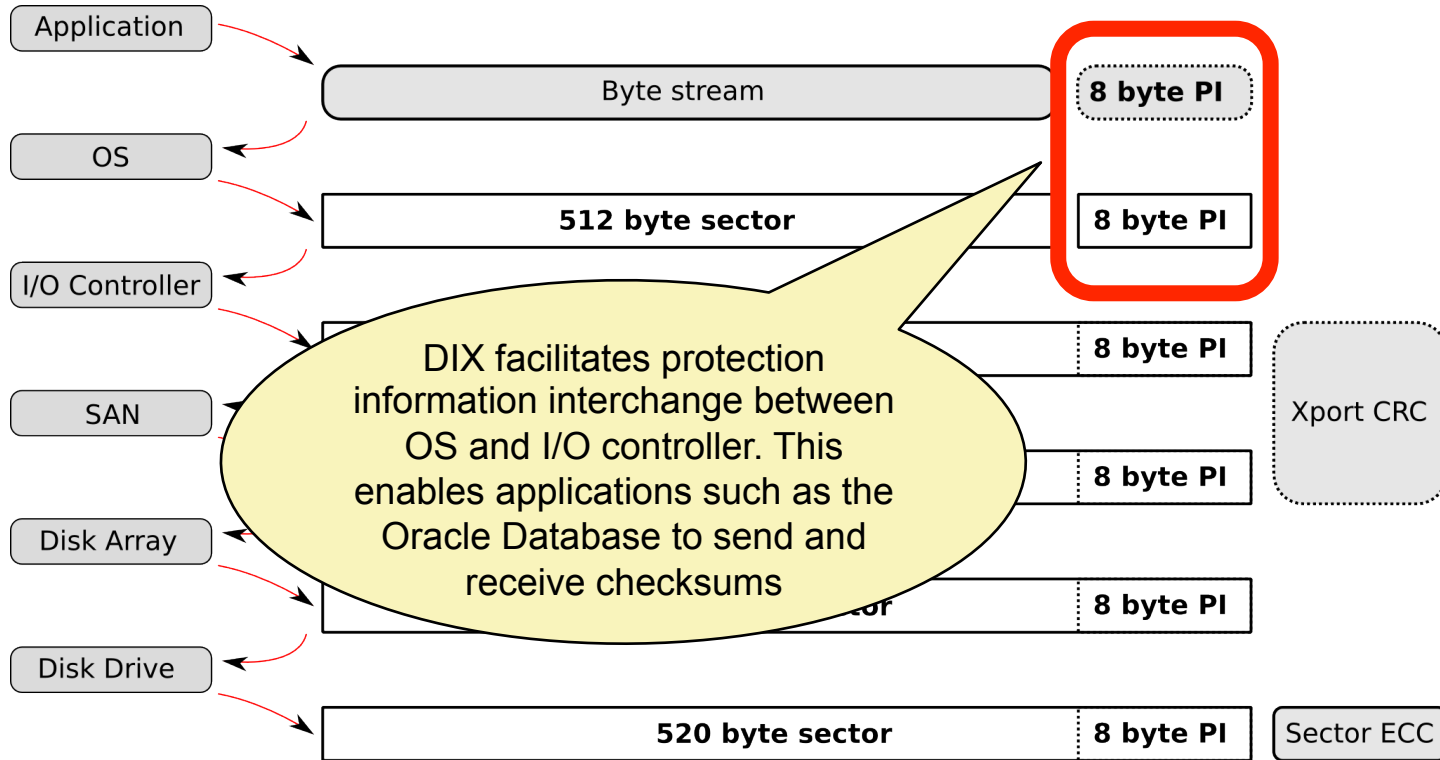
Data Integrity Extensions

- Extends T10 PI all the way up to the application, enabling true end-to-end data integrity protection
- The Data Integrity Extensions (DIX)
 - Enable DMA transfer of protection information to and from host memory
 - Separate data and protection information buffers to avoid inefficient 512+8+512+8+512+8 scatter-gather lists
 - Provide a set of commands that tell HBA how to handle the I/O: *Generate, Strip, Forward, Verify, etc.*

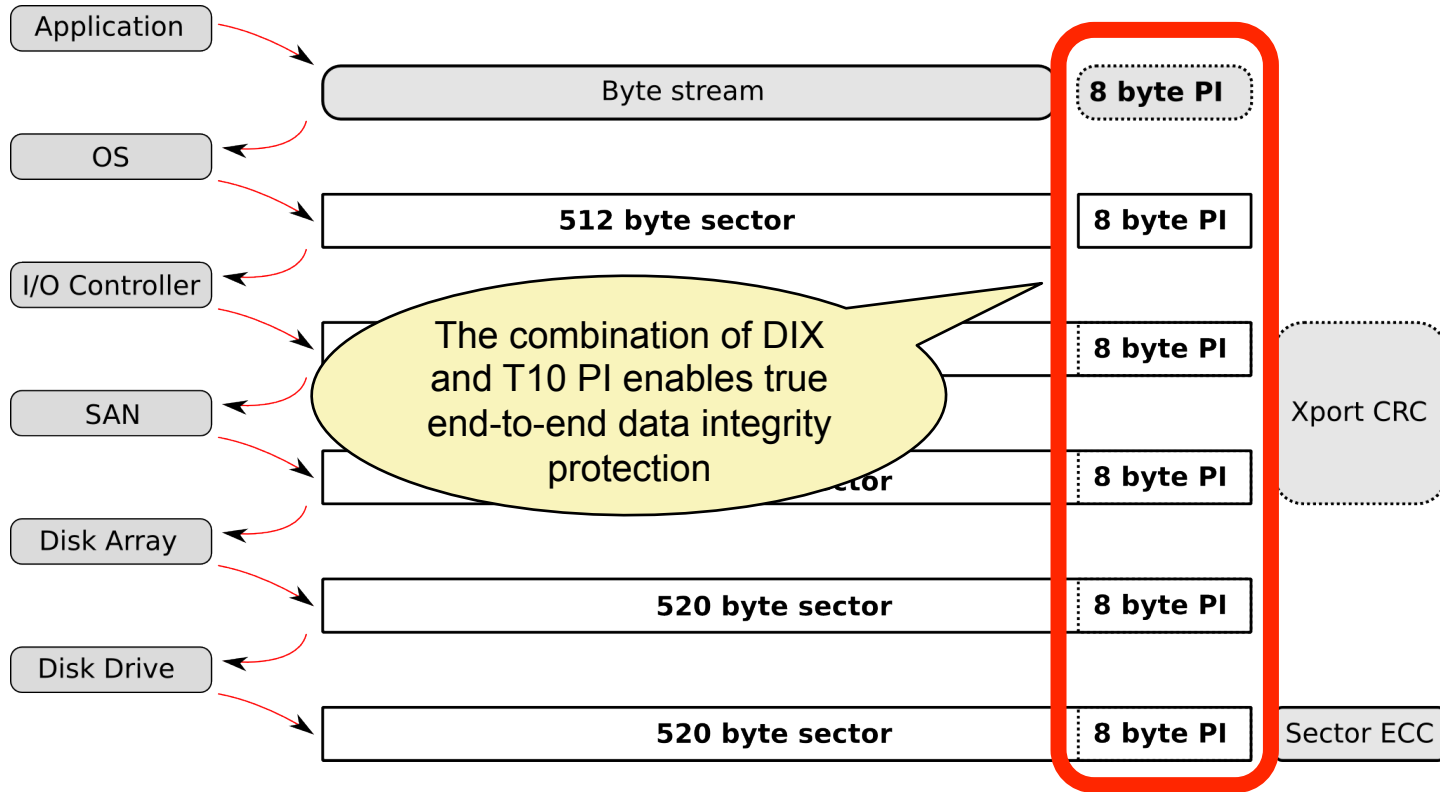
Data Integrity Extensions + T10 PI



Data Integrity Extensions + T10 PI



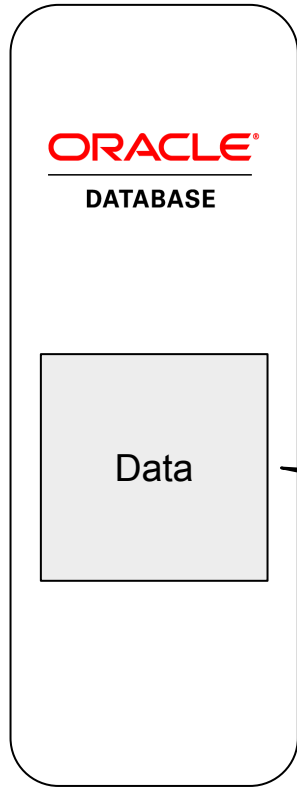
Data Integrity Extensions + T10 PI



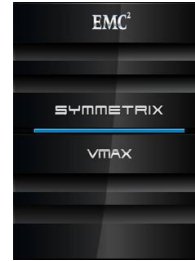
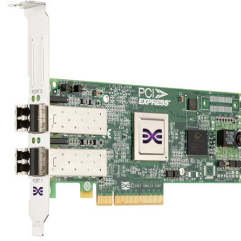
EMC Presentation

Technology Demonstration

Demo

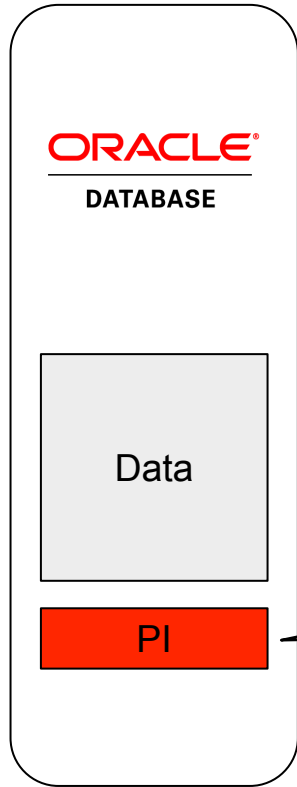


ORACLE
LINUX



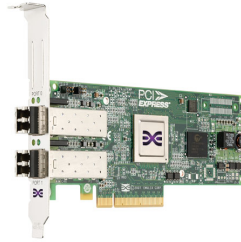
Database application creates data in system memory

Demo



ORACLE®
DATABASE

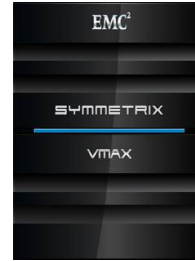
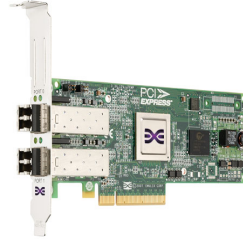
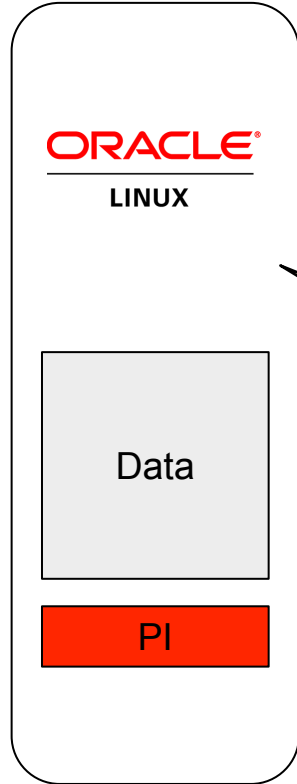
ORACLE®
LINUX



ASM generates
matching protection
information

Demo

ORACLE®
DATABASE

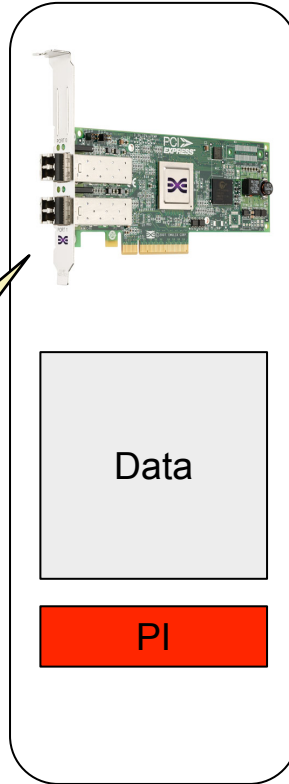


ASM submits request to the Linux kernel. Data and protection information reside in separate buffers

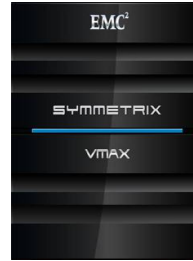
Demo

ORACLE[®]
DATABASE

ORACLE[®]
LINUX



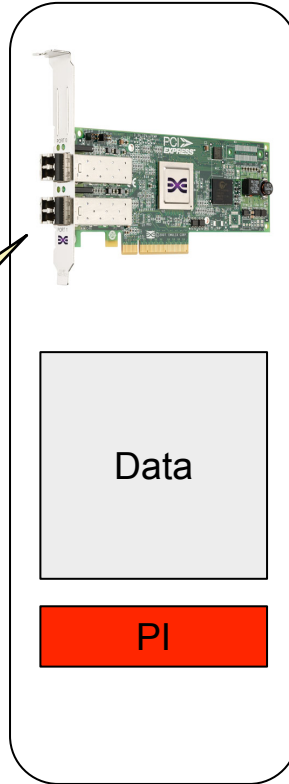
Linux kernel forwards data and protection information to HBA using DIX



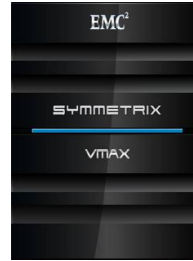
Demo

ORACLE®
DATABASE

ORACLE®
LINUX



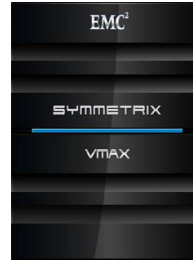
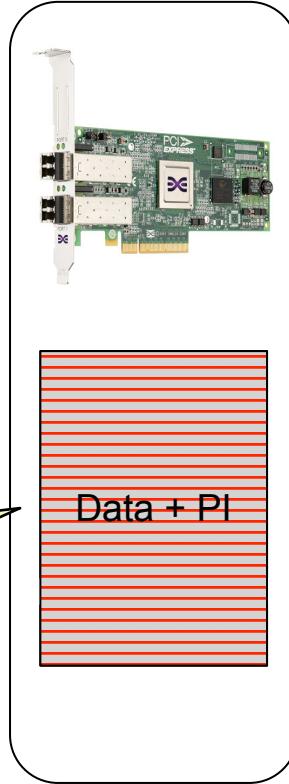
HBA verifies data, protection information, and request parameters



Demo

ORACLE®
DATABASE

ORACLE®
LINUX

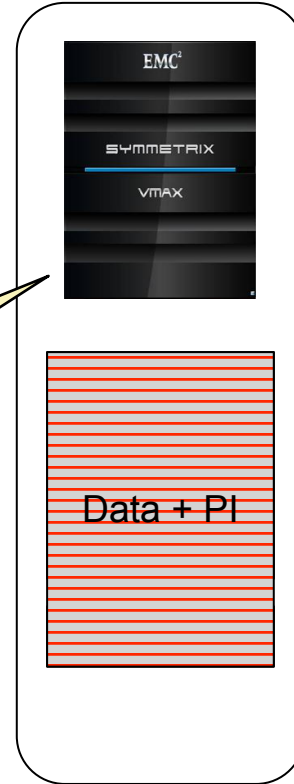
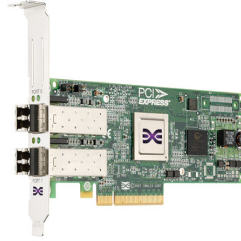


HBA interleaves data and protection information while transmitting 520-byte blocks to storage

Demo

ORACLE®
DATABASE

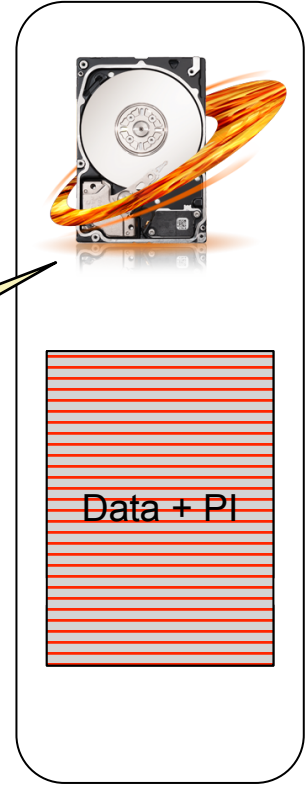
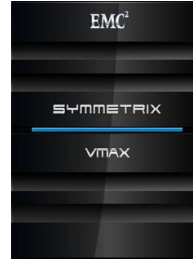
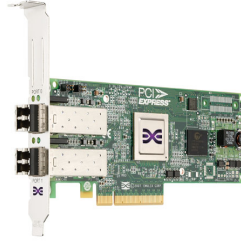
ORACLE®
LINUX



Demo

ORACLE
DATABASE

ORACLE
LINUX

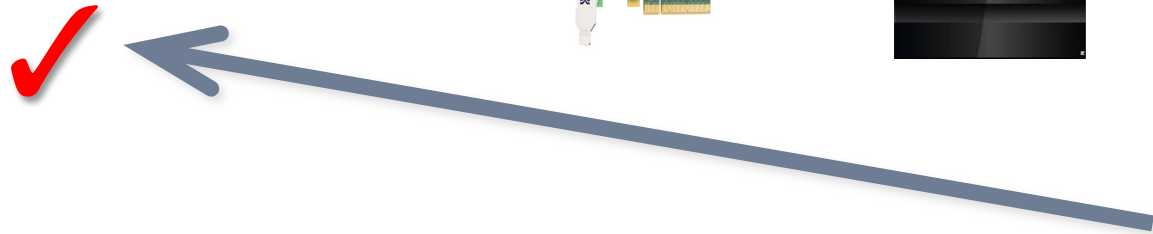
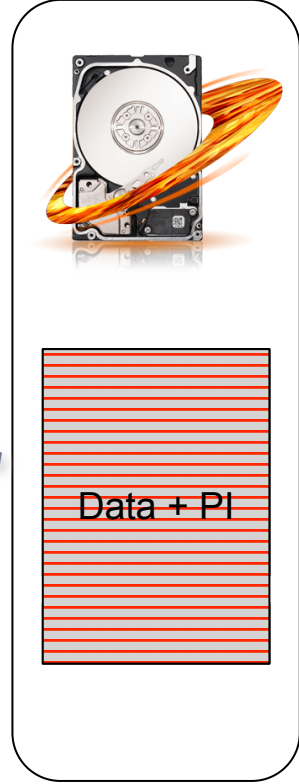
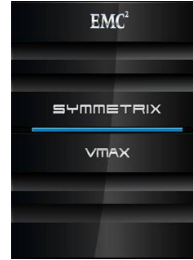
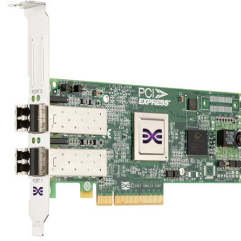


Disk drive firmware verifies that data, protection information, and request parameters match

Demo

ORACLE[®]
DATABASE

ORACLE[®]
LINUX

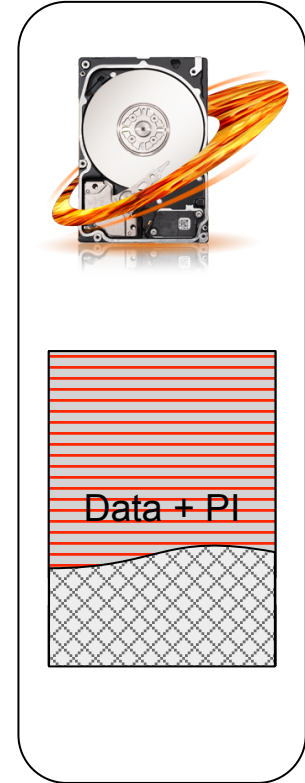
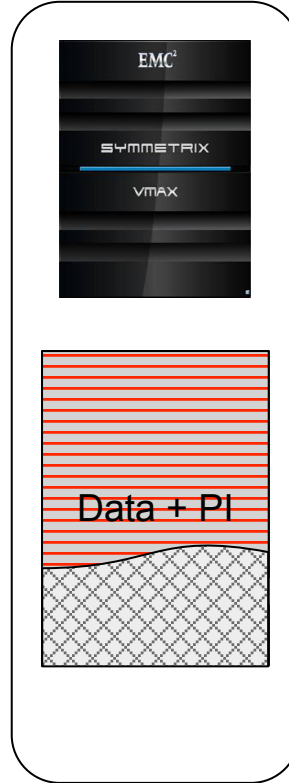
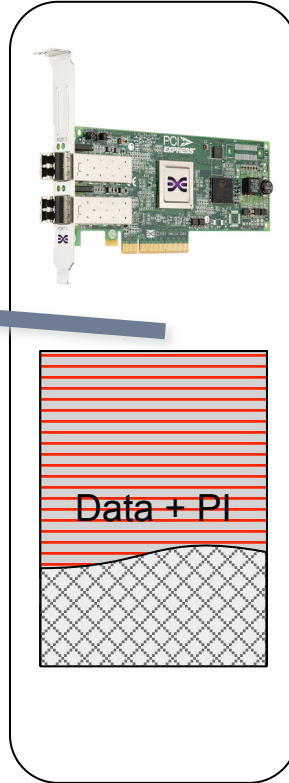


Successful I/O completion is reported to the application

Demo

ORACLE®
DATABASE

ORACLE®
LINUX



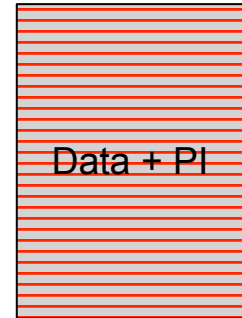
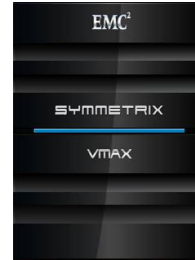
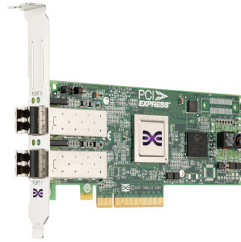
Mismatches detected by HBA, storage array, or disk drive will cause I/O to be aborted. The error is passed back to the application

Demo

ORACLE[®]
DATABASE



ORACLE[®]
LINUX

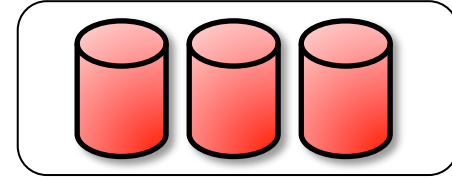
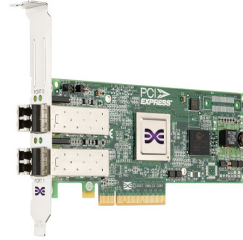


Protection information is also transmitted for read requests and ASM will verify data before signaling completion to the application

Demo

ORACLE[®]
DATABASE

ORACLE[®]
LINUX



Protection information exchange is dynamically negotiated. Protection is automatically enabled between ASM and HBA when using legacy storage

End-to-End Data Integrity

- Storage devices that support the T10 Protection Information Model
- Data Integrity Extensions implemented in Oracle 8Gbps Fibre Channel Host Adapters
- Oracle Linux with Unbreakable Enterprise Kernel
- Oracle Database 11g with ASM



Q&A

Oracle Linux Pavilion

Exhibitors	Location
Visit Oracle Linux Pavilion to get insight into partners that implement Oracle Linux and Oracle VM in their products.	Moscone South Booth 2241



- Open Exhibition Hours on : Monday, Tuesday, Wednesday
- Giveaways and prizes
- Tuesday, 2:15-6pm “Chill with Tux over Smoothies” Reception

Oracle DemoPods

Demo	Location
Oracle Linux	Moscone South Exhibition Hall Booth 1133 Pod: S-155
Oracle Linux with Unbreakable Enterprise Kernel	Moscone South Exhibition Hall Booth 1133 Pod: S-154

Thursday, October 6

Oracle OpenWorld Sessions

Time	Title	Location
9:00am - 10:00am	Debugging and Configuration Best Practices for Oracle Linux 6 Greg Marsden, Software Development Director, ORACLE	Moscone South Room 301

Oracle OpenWorld Hands-On Labs

Time	Title	Location
1:30pm - 2:30pm	Oracle Linux Package Management: Configuring and Enabling Services Avi Miller, Principal Product Manager, ORACLE Lenz Grimmer, Sr. Product Manager, Oracle Linux, ORACLE	Marriott Marquis Salon 5/6
3:00pm - 4:00pm	Oracle Linux Storage Management with LVM and Device Mapper Avi Miller, Principal Product Manager, ORACLE Lenz Grimmer, Sr. Product Manager, Oracle Linux, ORACLE	Marriott Marquis Salon 5/6

Call to Action

- **Linux Home Page**
oracle.com/linux
- **Follow us on Twitter**
twitter.com/ORCL_Linux
- **Read our blog**
[Blogs.oracle.com/Linux](https://blogs.oracle.com/Linux)
- **Become a fan on Facebook**
Facebook.com/OracleLinux
- **Subscribe to our channel YouTube**
YouTube.com/OracleLinuxChannel
- **Free Download: Oracle Linux**
edelivery.oracle.com/linux



ORACLE®
LINUX

Hardware and Software

ORACLE®

Engineered to Work Together

ORACLE®

ORACLE®