



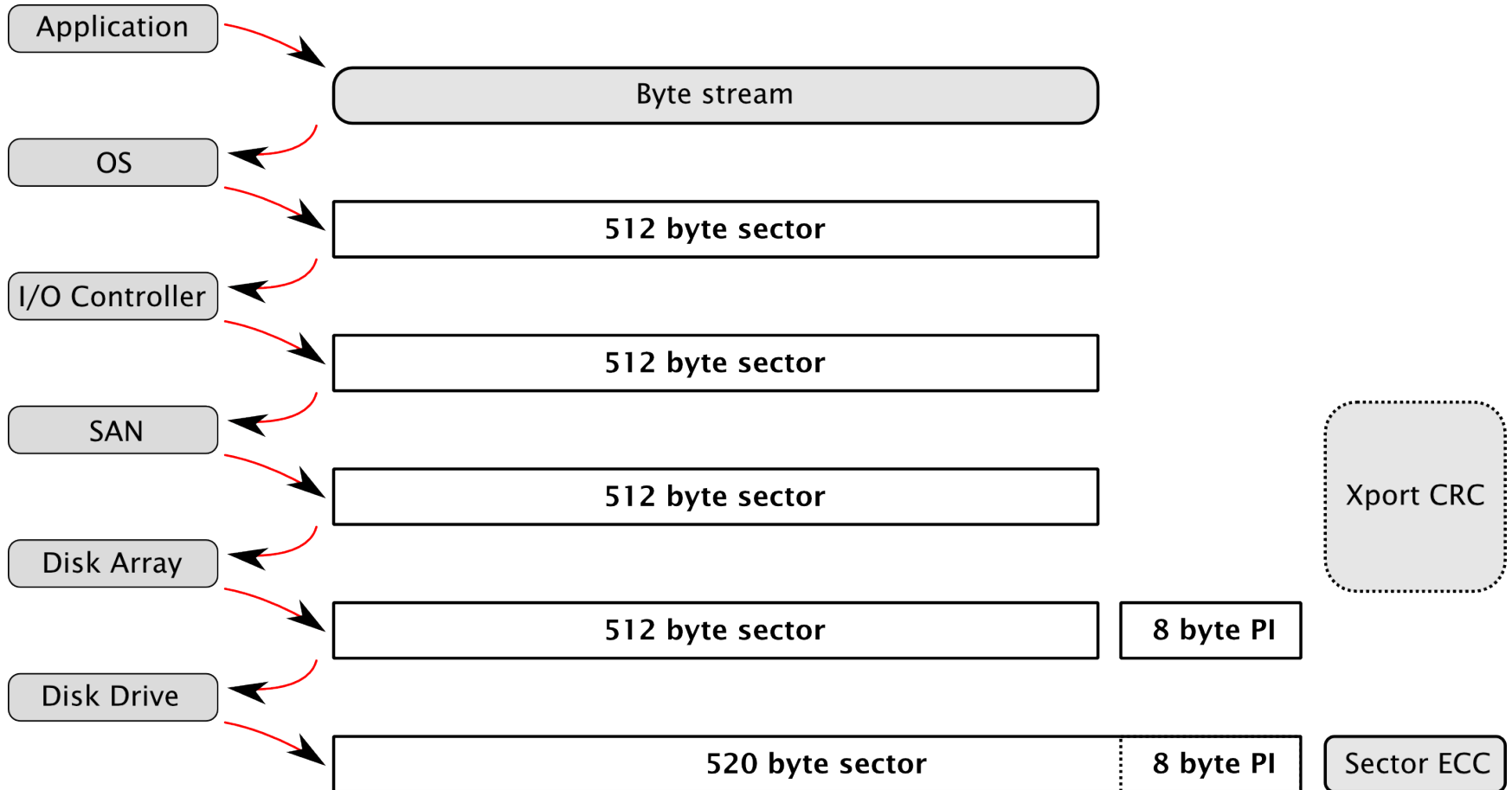
**ORACLE<sup>®</sup>**

## **Eliminating Silent Data Corruption with Oracle Linux**

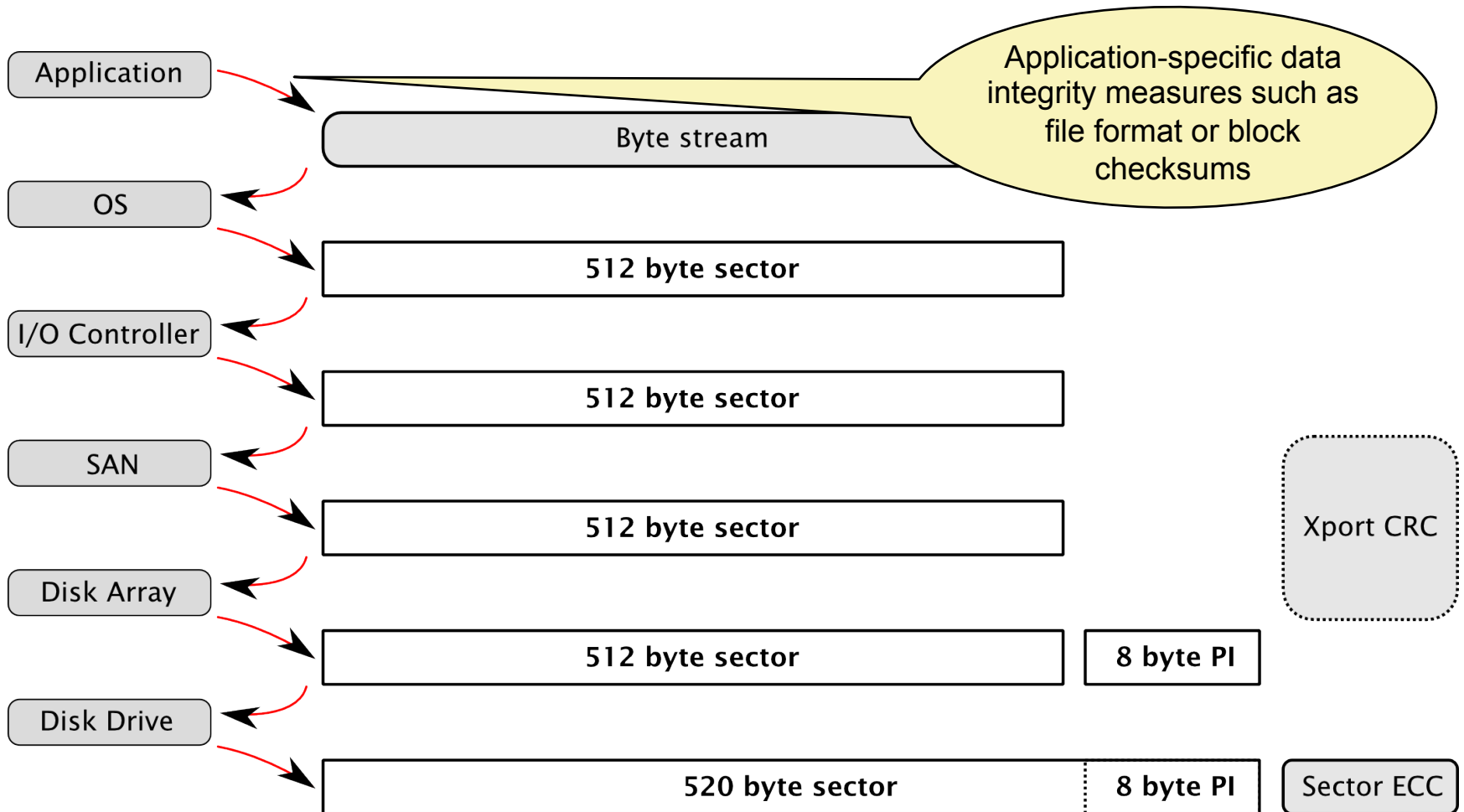
Martin K. Petersen, Architect, Linux Engineering

# T10 PI & DIX

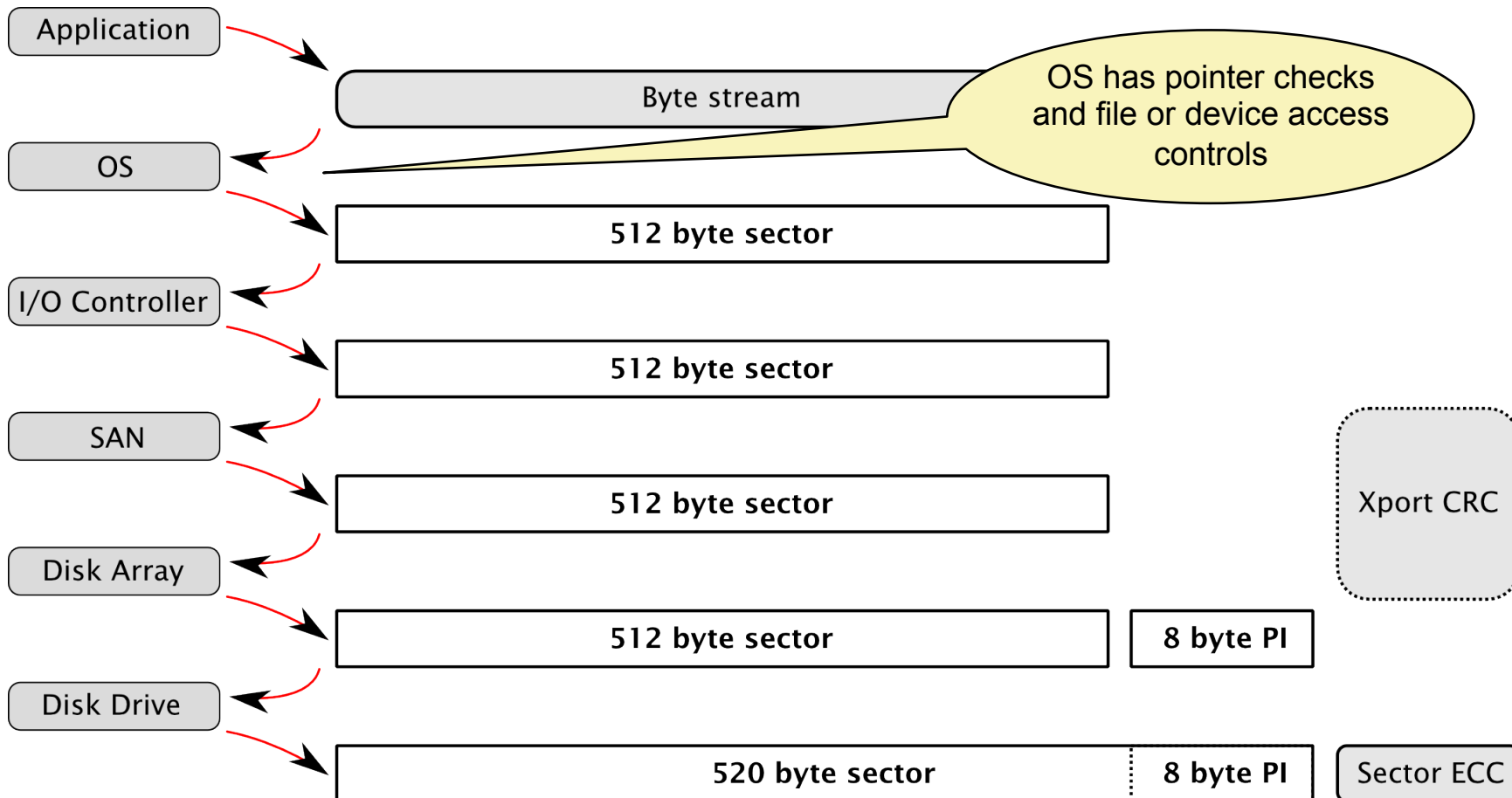
# Legacy I/O Stack



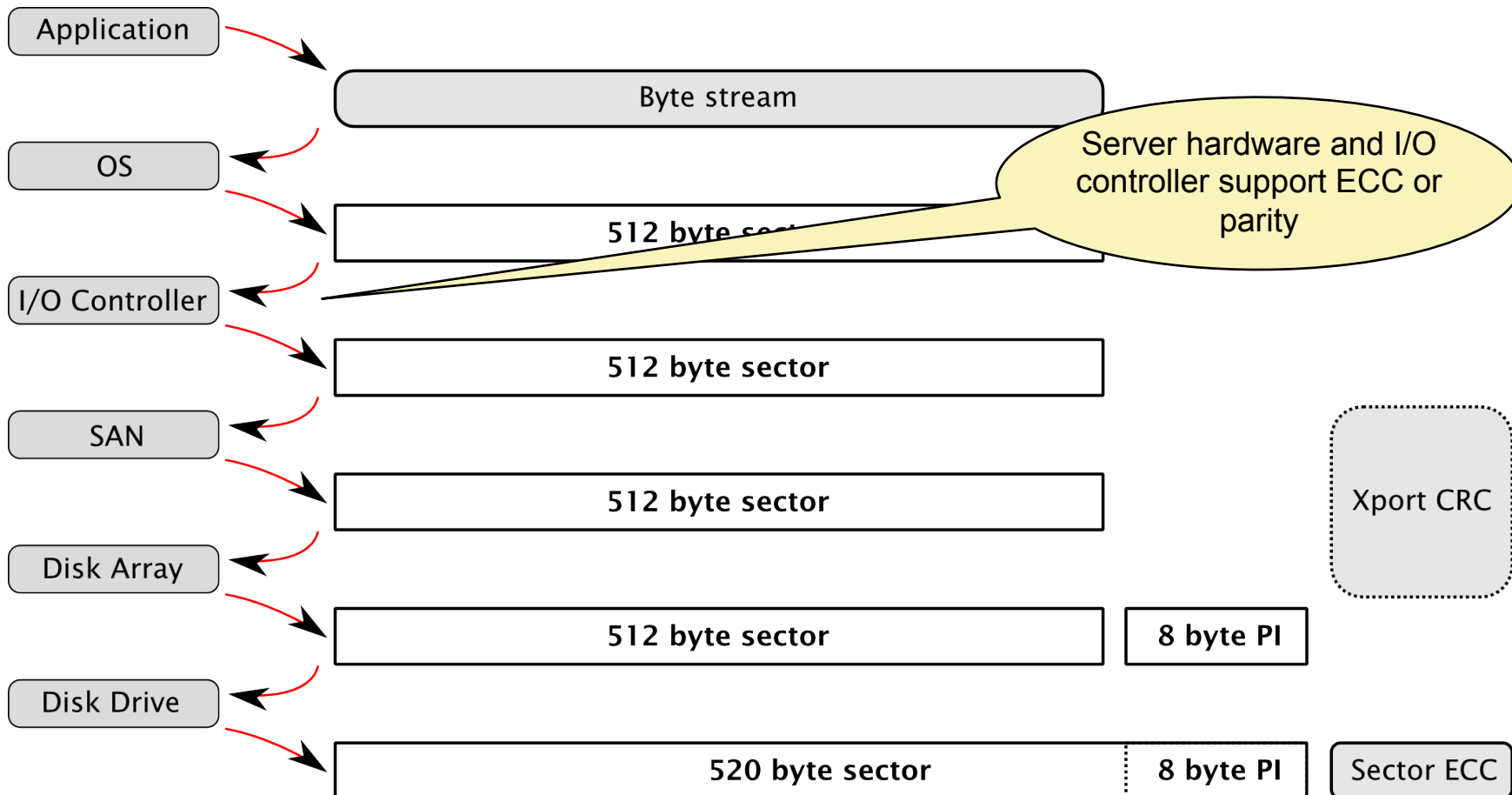
# Legacy I/O Stack



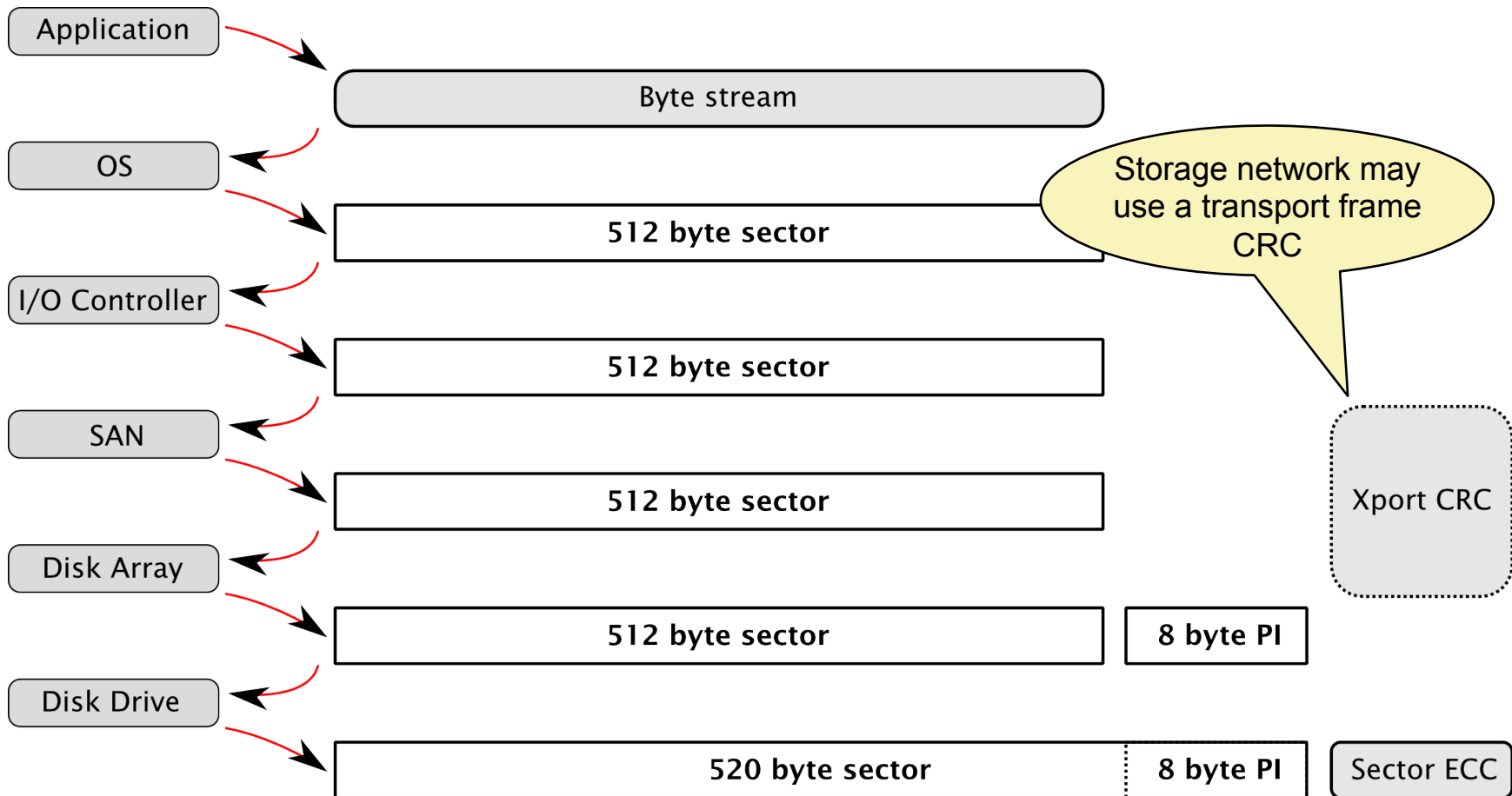
# Legacy I/O Stack



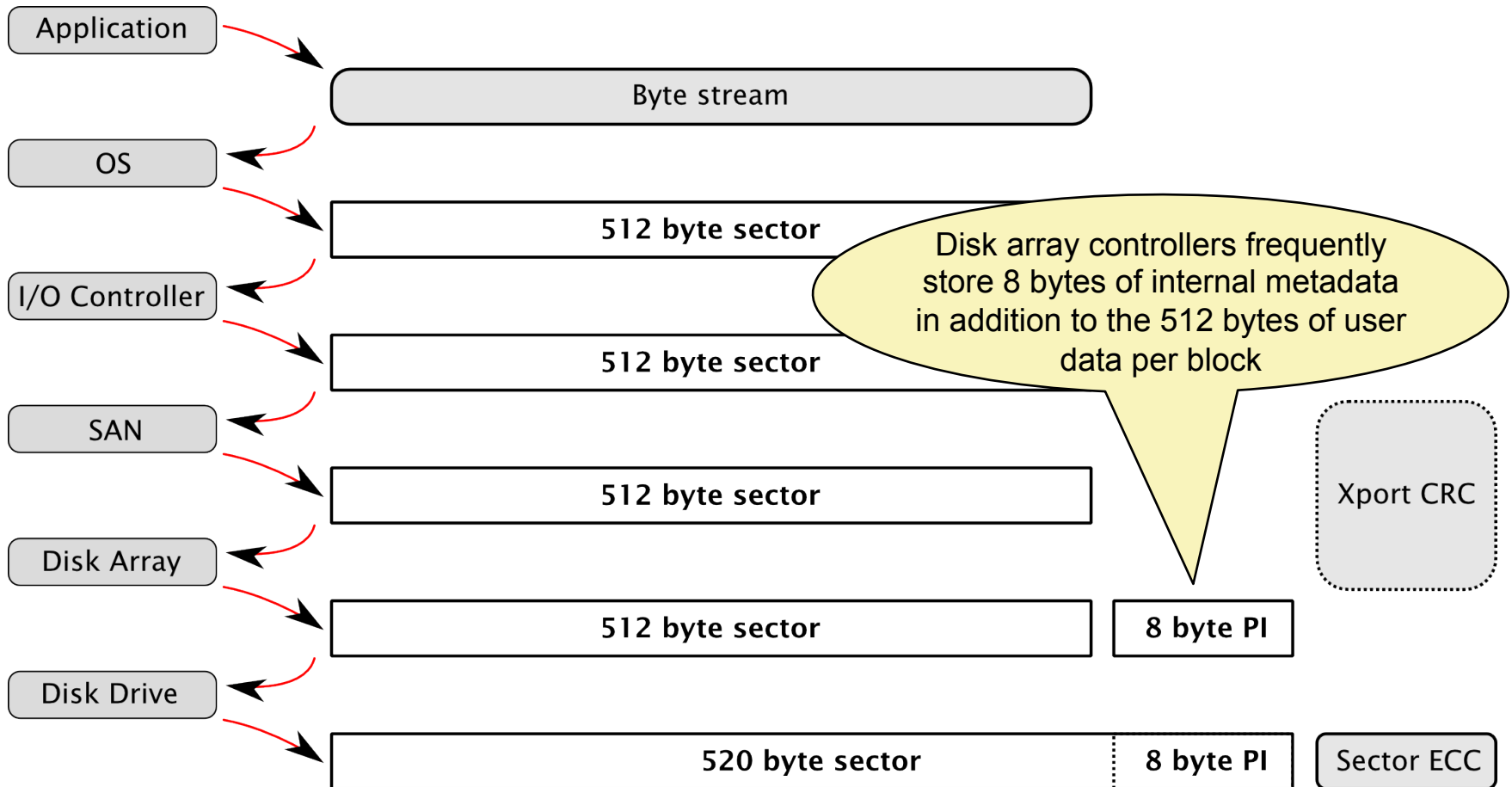
# Legacy I/O Stack



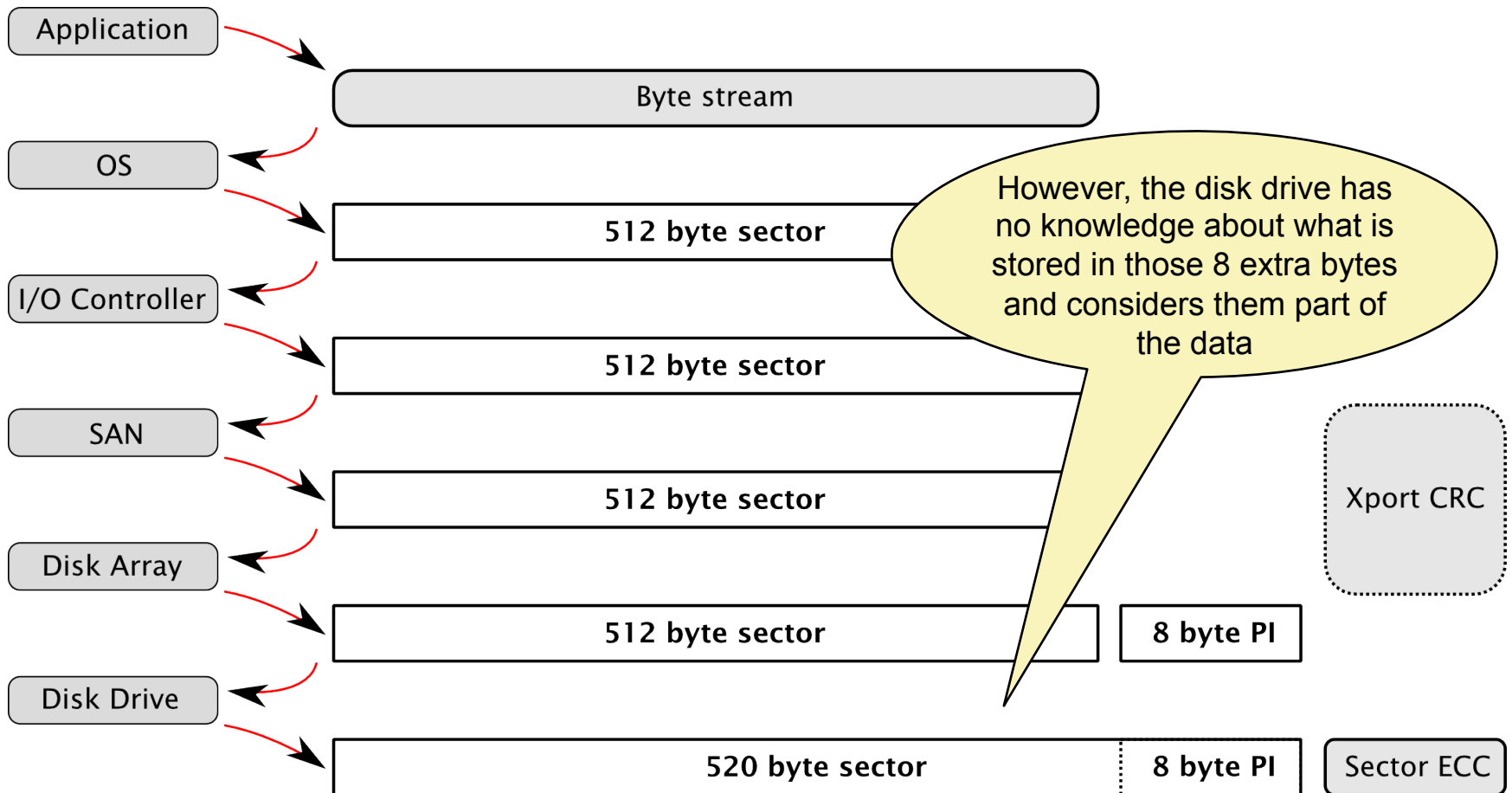
# Legacy I/O Stack



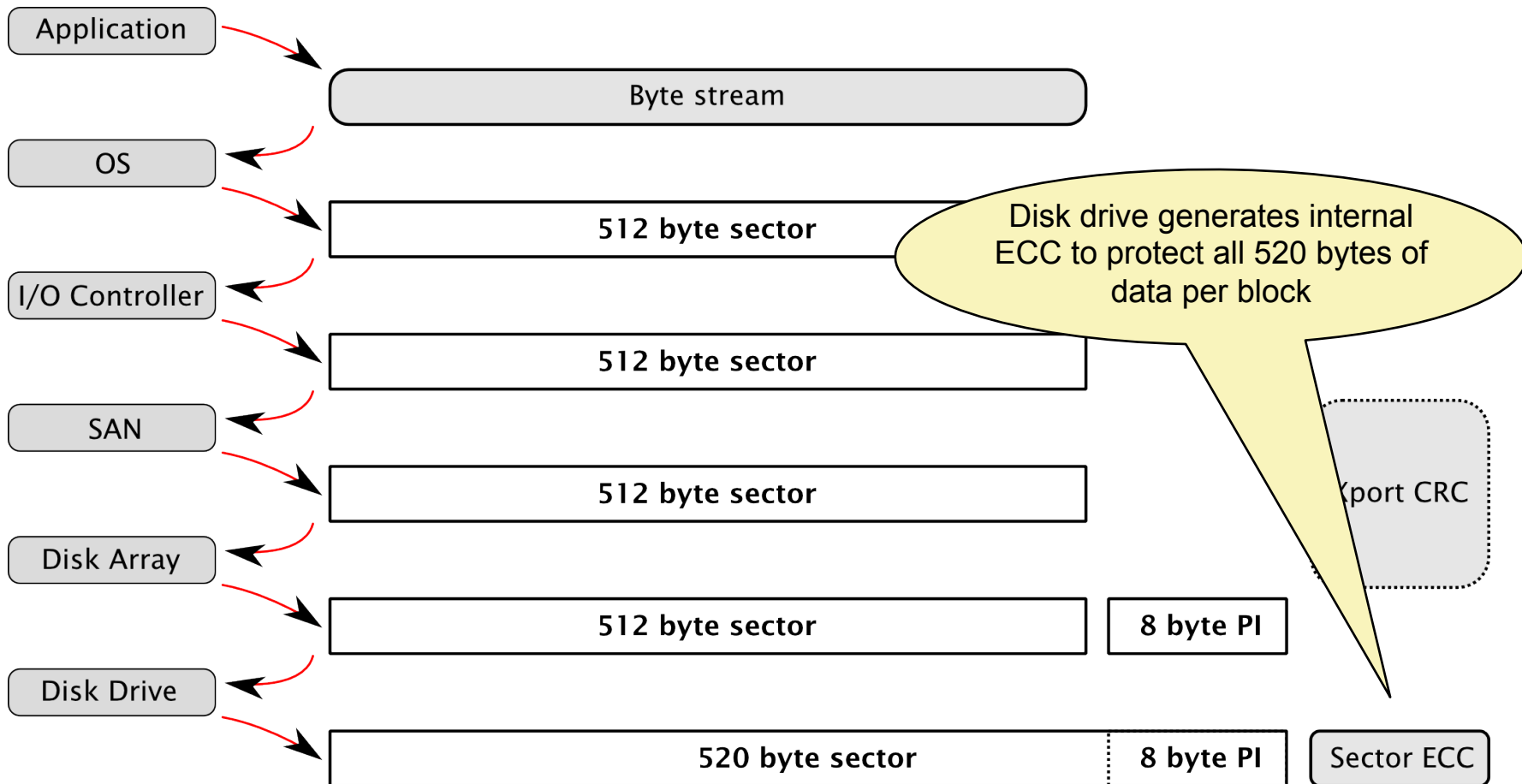
# Legacy I/O Stack



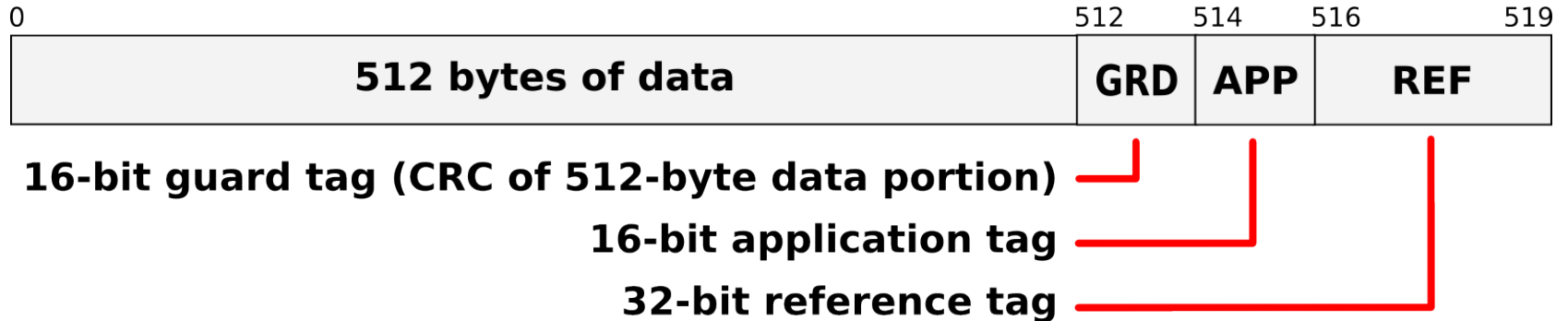
# Legacy I/O Stack



# Legacy I/O Stack

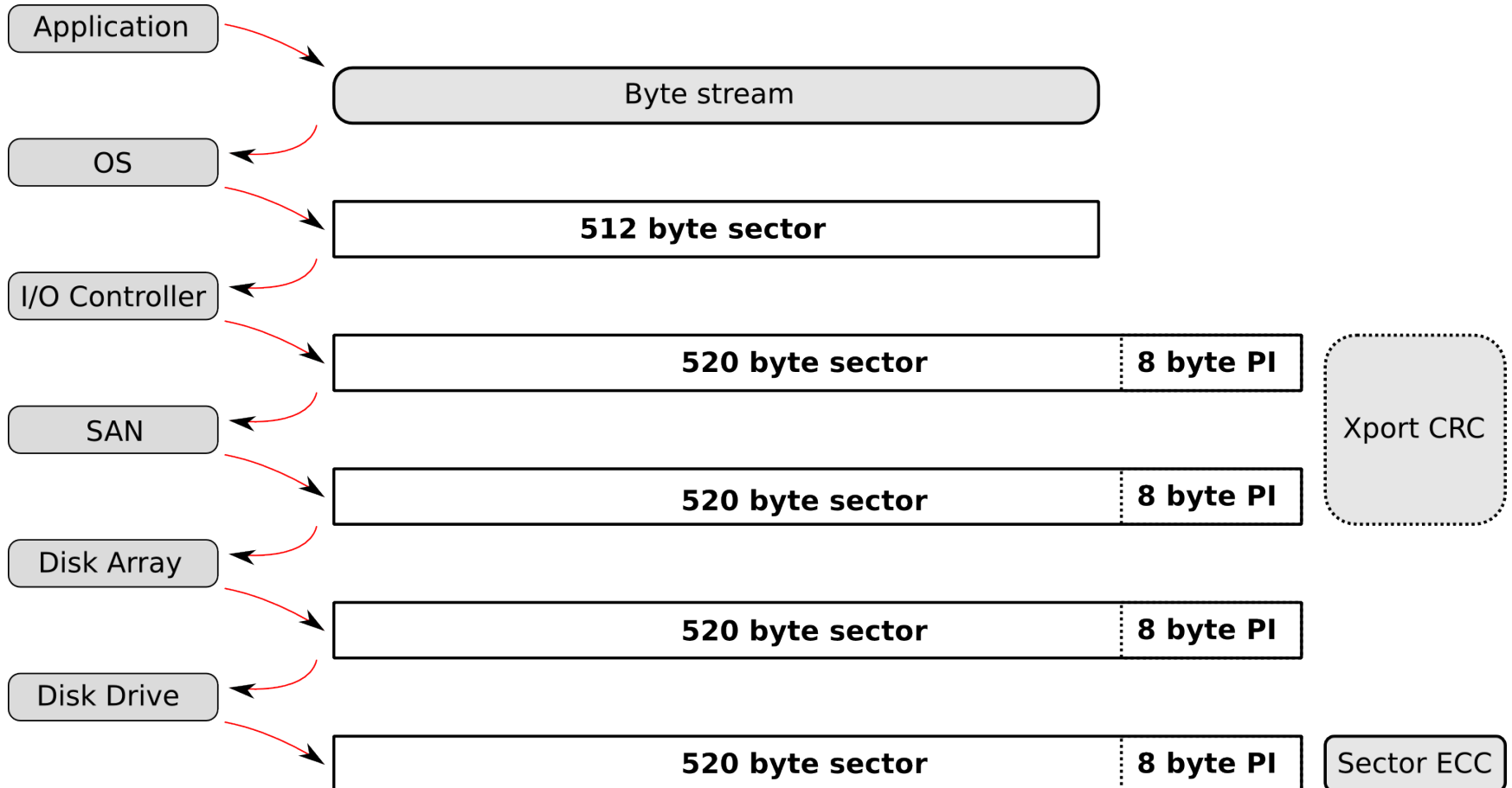


# T10 Protection Information Model

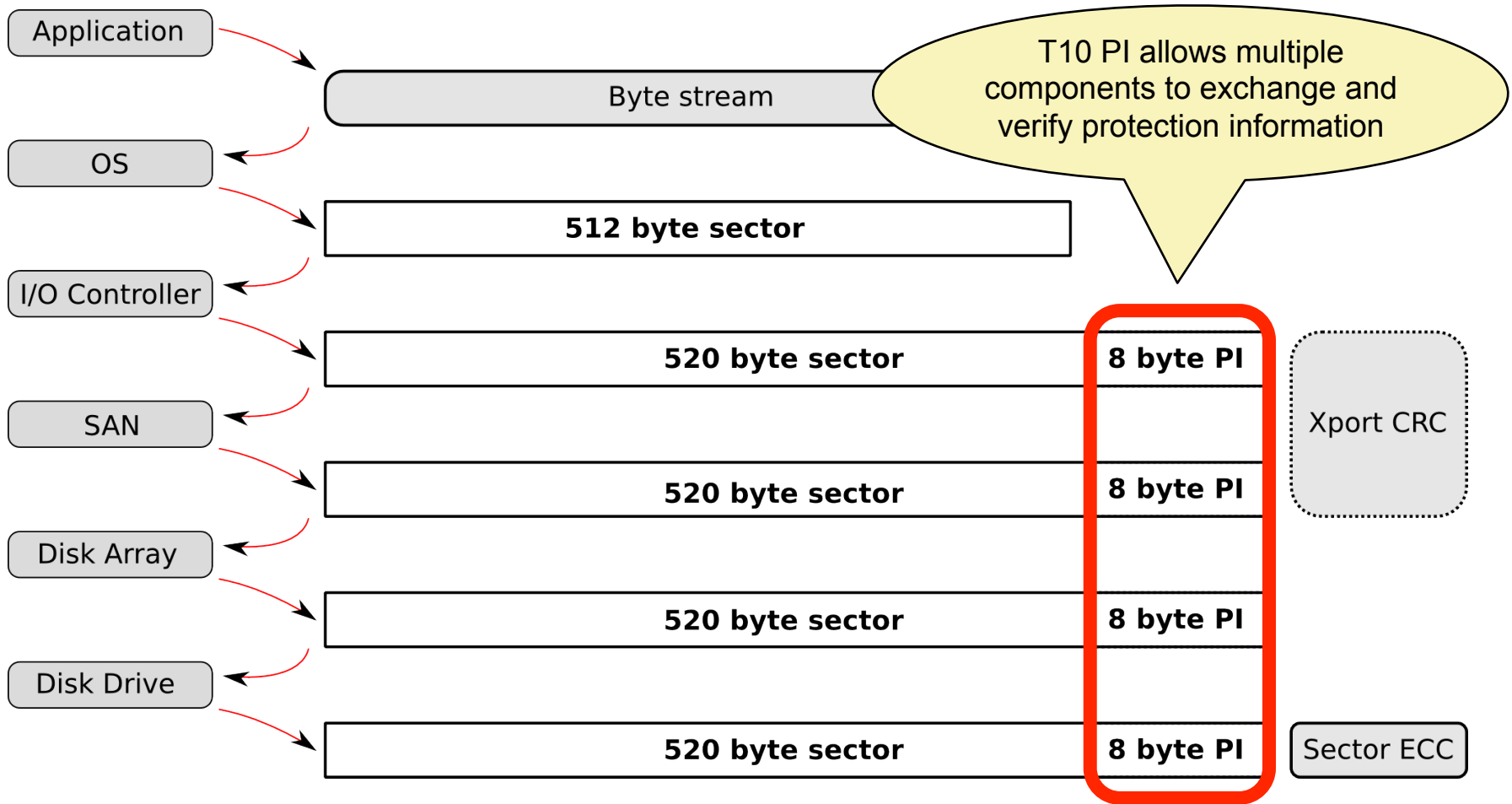


- Standardizes those extra 8 bytes
- Prevents content corruption and misplacement errors
- Protects path between HBA and storage device
- Protection information is interleaved with data on the wire, i.e. effectively 520-byte logical blocks

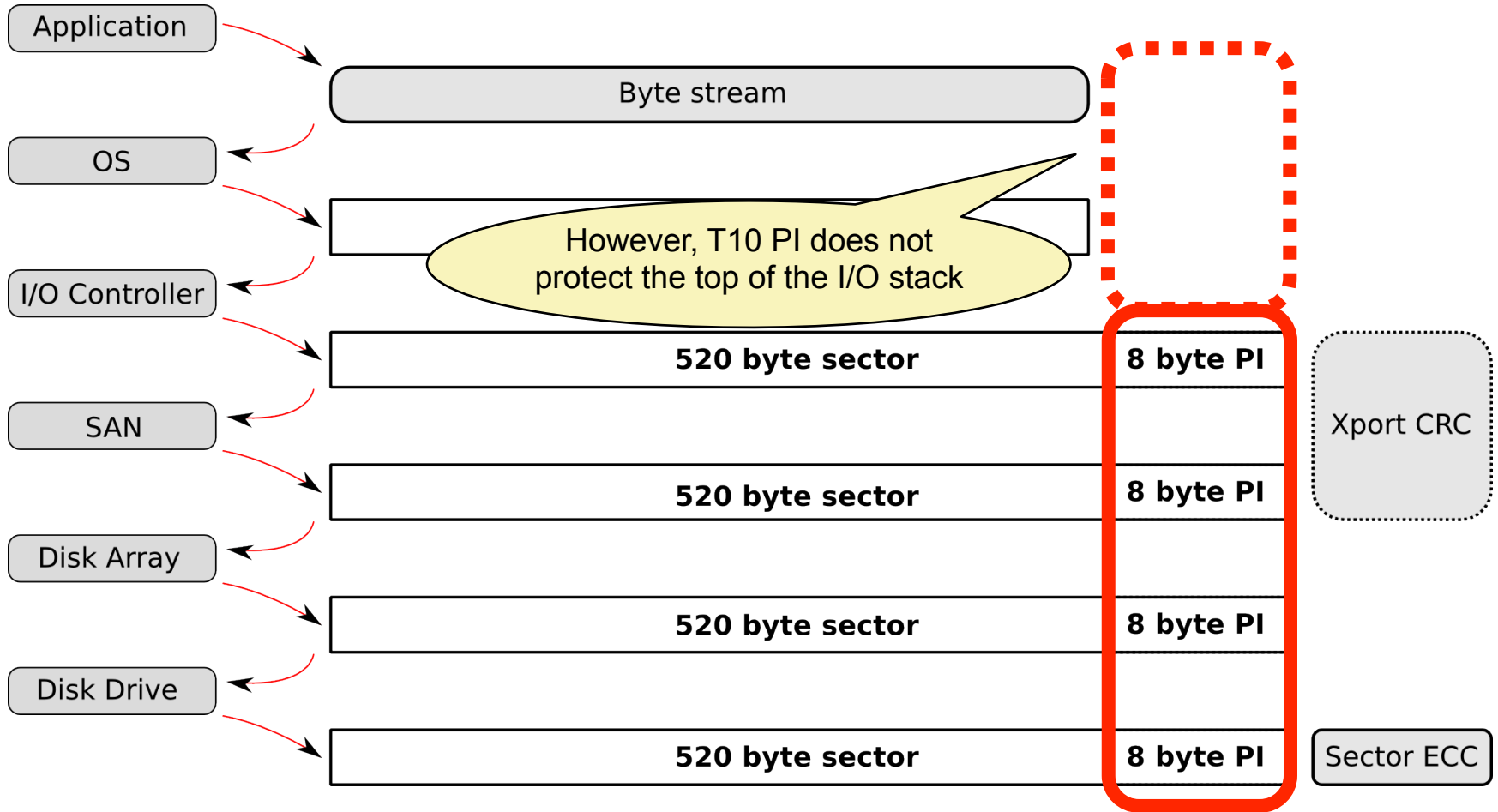
# T10 Protection Information Model



# T10 Protection Information Model



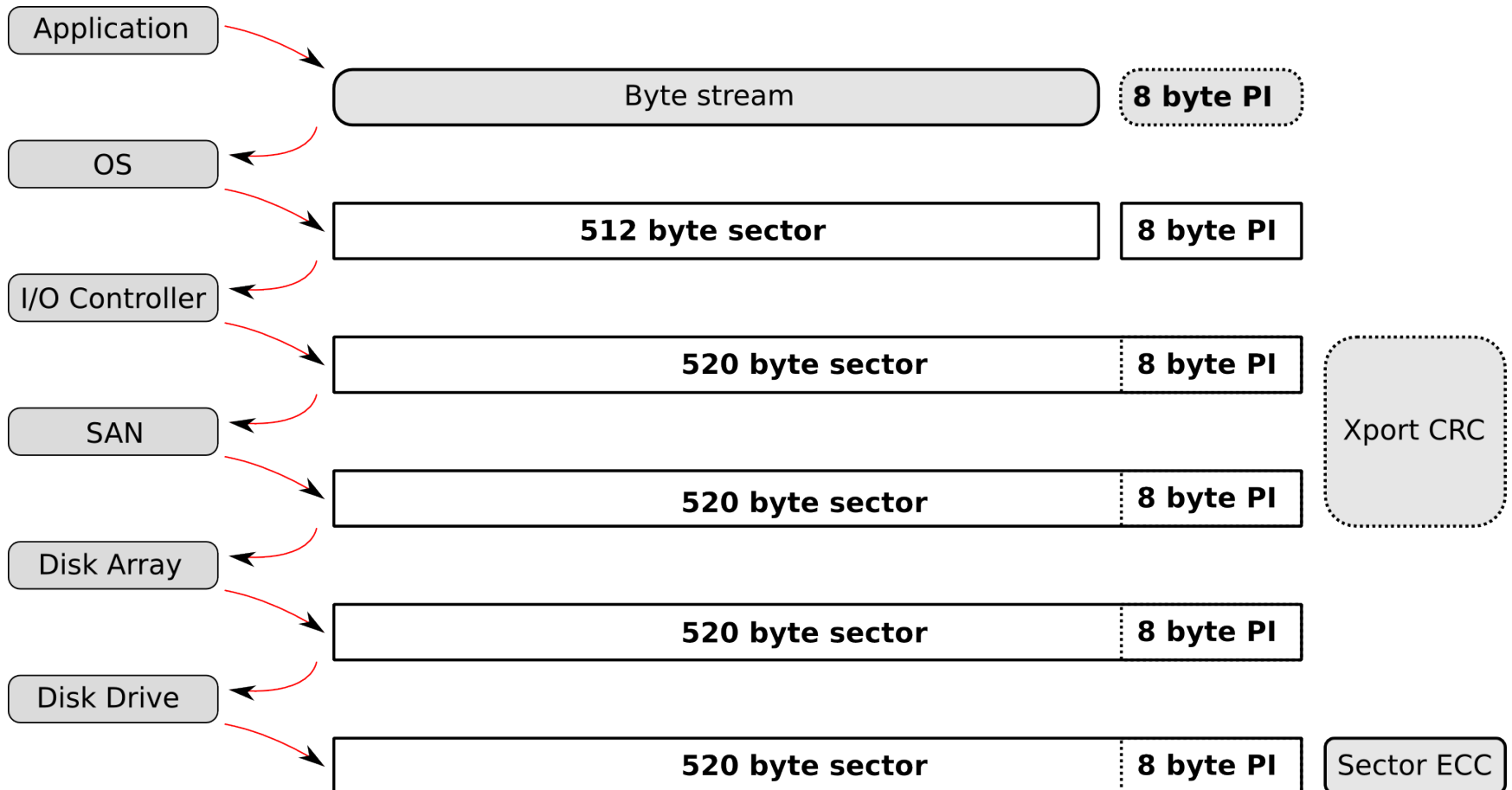
# T10 Protection Information Model



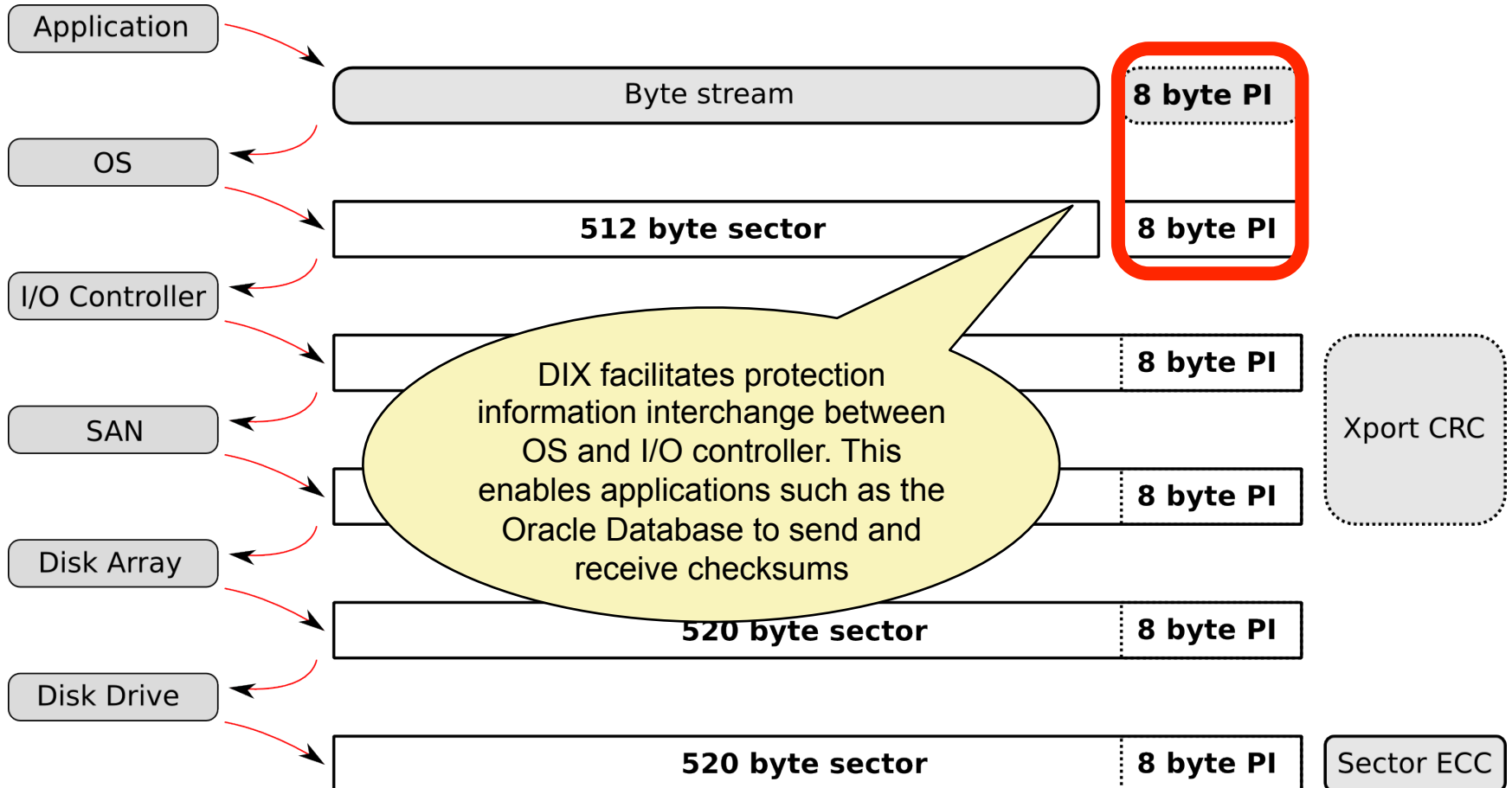
# Data Integrity Extensions

- 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.*

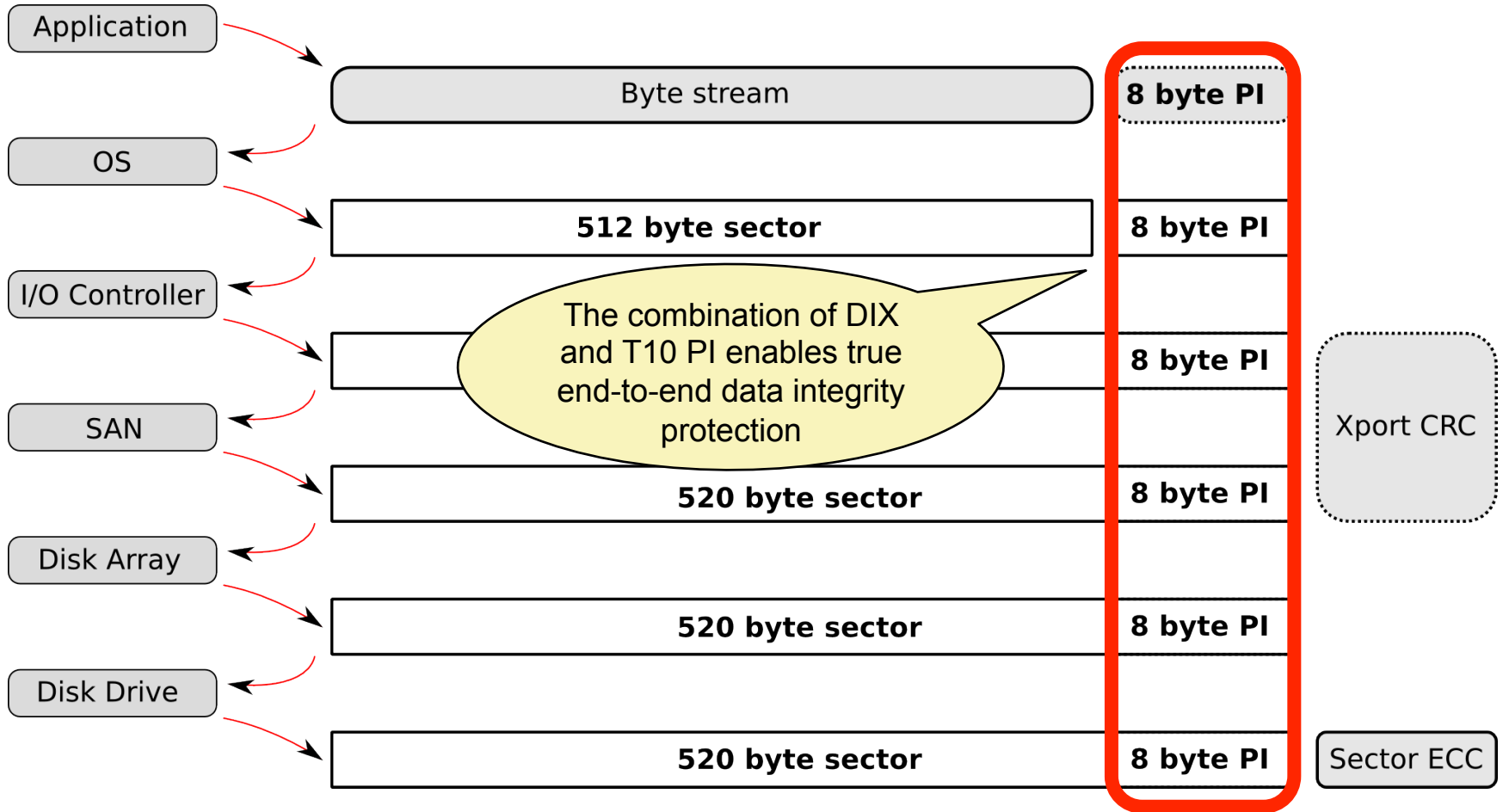
# Data Integrity Extensions + T10 PI



# Data Integrity Extensions + T10 PI

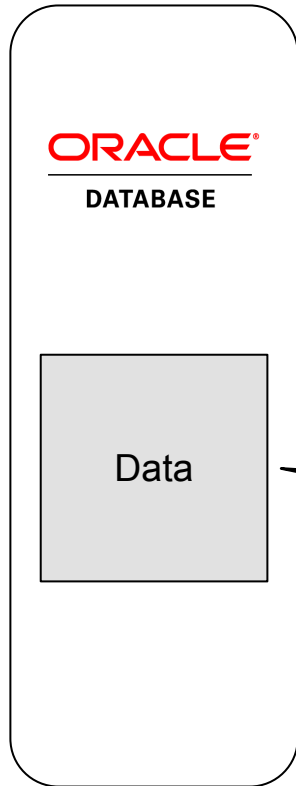


# Data Integrity Extensions + T10 PI



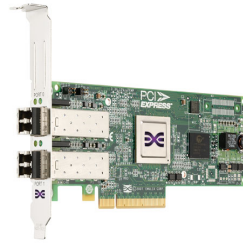
# Technology Demonstration

# Demo



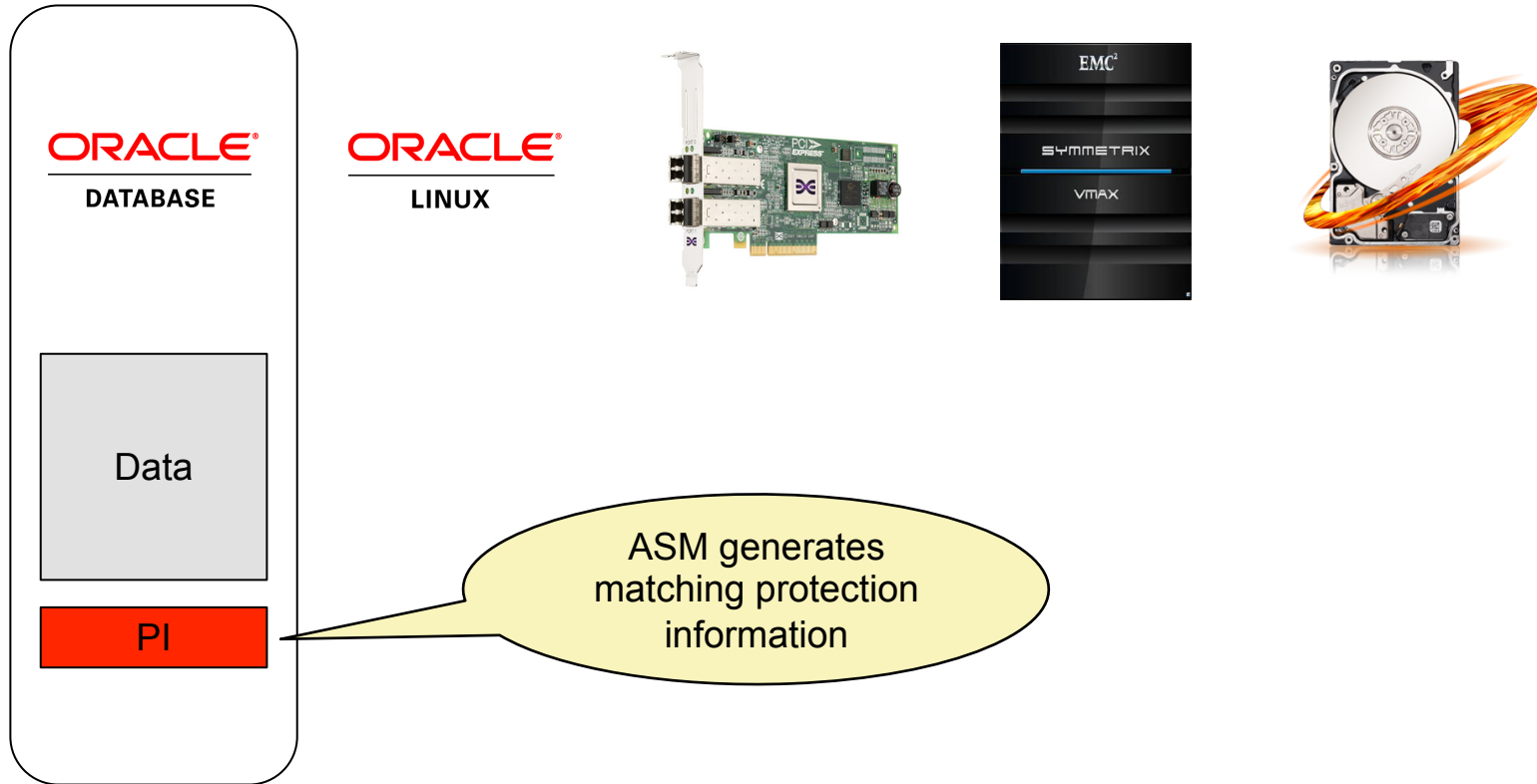
**ORACLE**  
DATABASE

**ORACLE**  
LINUX



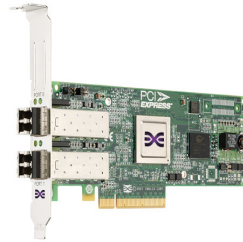
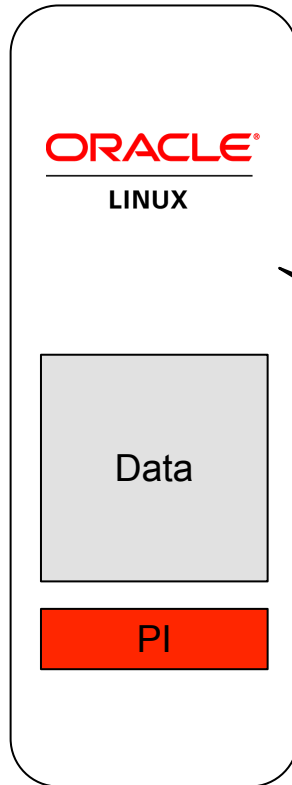
Database application creates data in system memory

# Demo



# Demo

**ORACLE**  
DATABASE



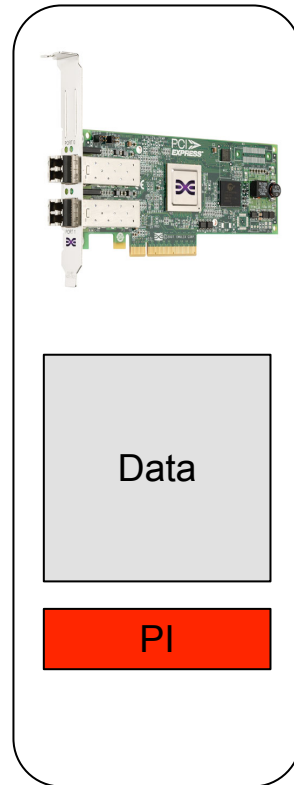
ASM submits request to the Linux kernel. Data and protection information reside in separate buffers

# Demo

**ORACLE**  
DATABASE

**ORACLE**  
LINUX

Linux kernel forwards data and protection information to HBA using DIX

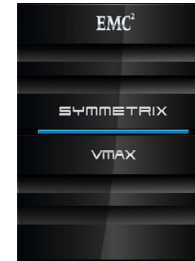
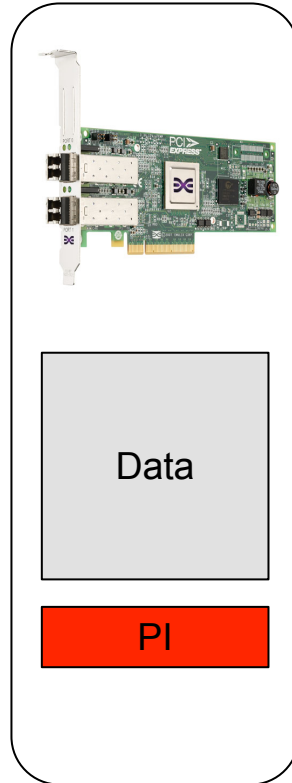


# Demo

**ORACLE**  
DATABASE

**ORACLE**  
LINUX

HBA verifies data,  
protection information, and  
request parameters

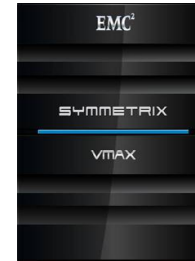
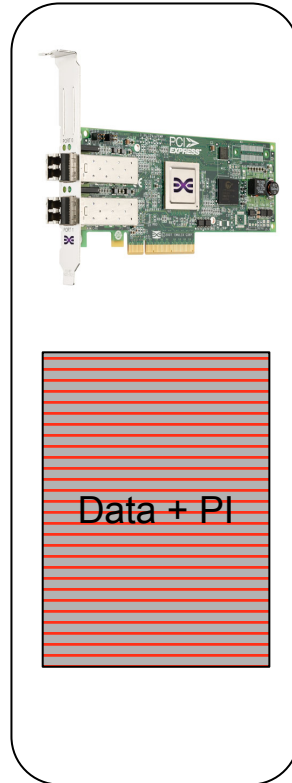


# Demo

**ORACLE**  
DATABASE

**ORACLE**  
LINUX

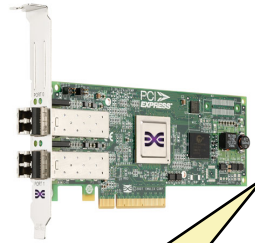
HBA interleaves data and protection information while transmitting 520-byte blocks to storage



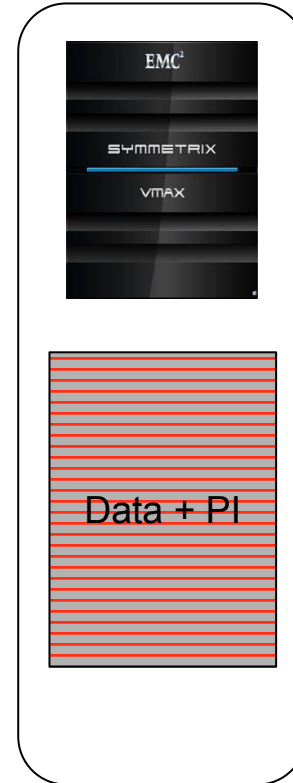
# Demo

**ORACLE**  
DATABASE

**ORACLE**  
LINUX



Storage array controller  
verifies that data,  
protection information, and  
request parameters match

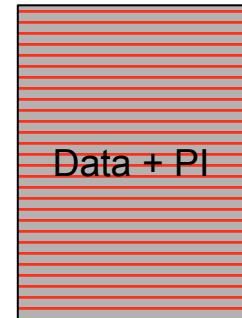
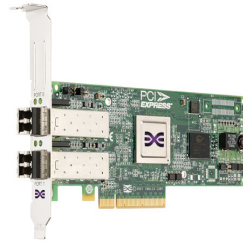


# Demo

**ORACLE**  
DATABASE



**ORACLE**  
LINUX

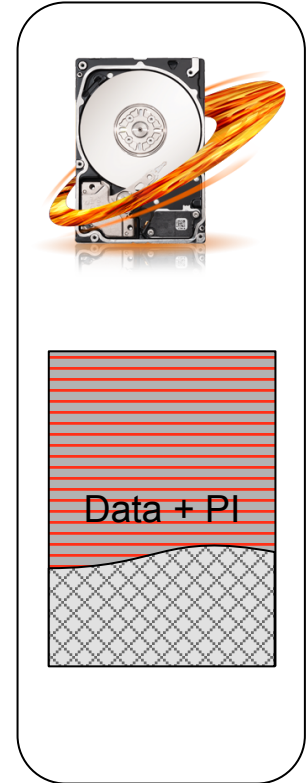
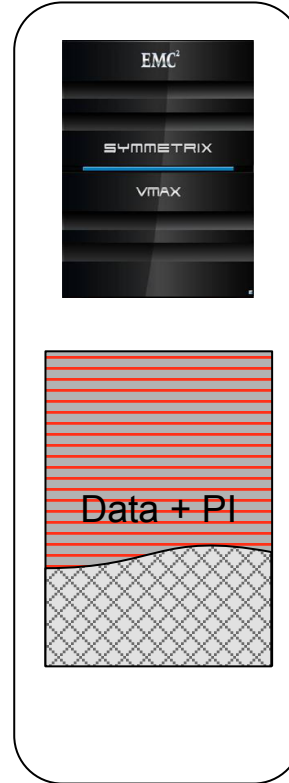
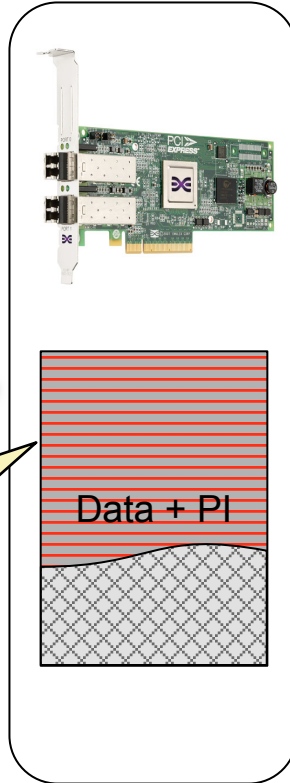


Successful I/O completion is reported to the application

# Demo

ORACLE<sup>®</sup>  
DATABASE

ORACLE<sup>®</sup>  
LINUX



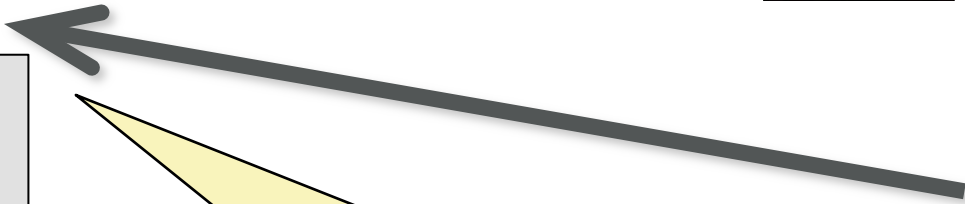
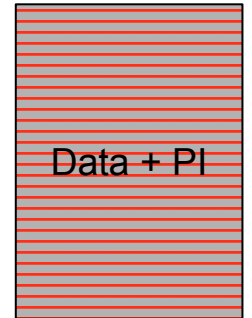
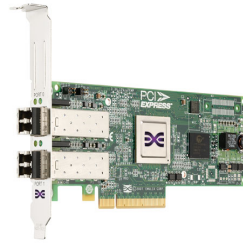
Mismatches detected by HBA, storage array, or disk drive will cause I/O to be aborted. The error is passed back to the application

# Demo

ORACLE<sup>®</sup>  
DATABASE



ORACLE<sup>®</sup>  
LINUX



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

# Resources

## Linux Home Page

[oracle.com/linux](http://oracle.com/linux)

## Free Download: Oracle Linux

[edelivery.oracle.com/linux](http://edelivery.oracle.com/linux)

## DIX Specification and Presentations

[oss.oracle.com/~mkp/](http://oss.oracle.com/~mkp/)



**ORACLE**<sup>®</sup>  
—  
**LINUX**

**Hardware and Software**

**ORACLE®**

**Engineered to Work Together**

ORACLE®