



Education

The Storage Evolution: From Blocks, Files and Objects to Object Storage Systems

Christian Bandulet, Principal Engineer, Sun Microsystems

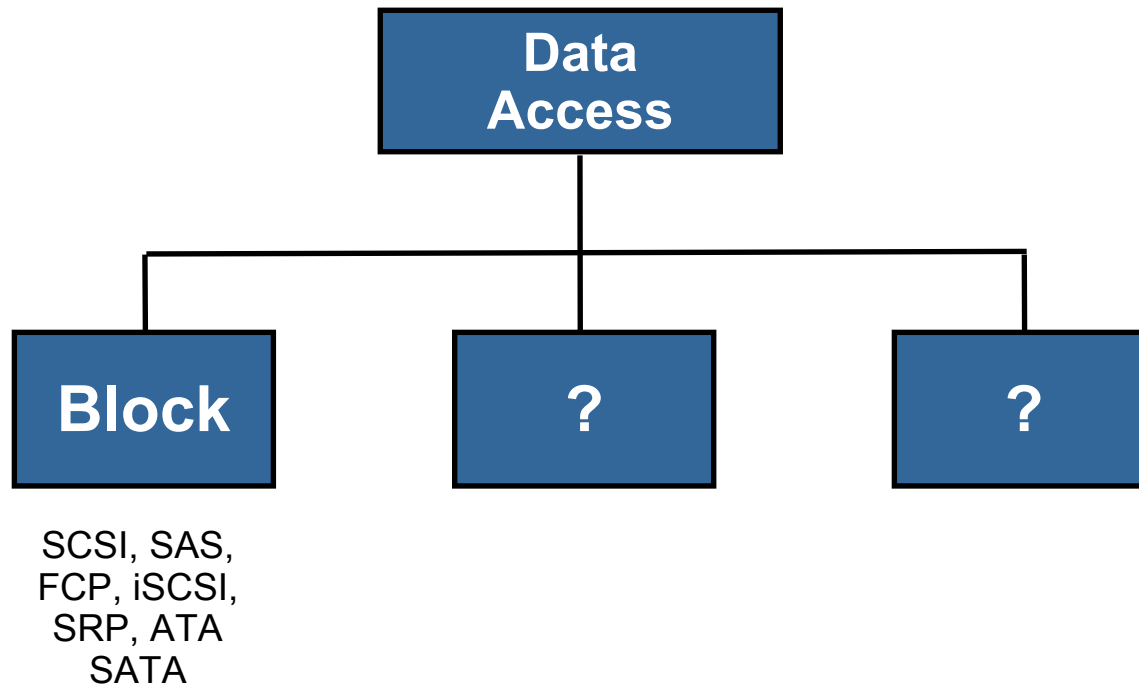
- The material contained in this tutorial is copyrighted by the SNIA.
- Member companies and individuals may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced without modification
 - ◆ The SNIA must be acknowledged as source 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.

- ▶ **The Storage Evolution: From Blocks, Files and Objects to Object Storage Systems**
 - ◆ This session will appeal to CIOs, CTOs, Consultants, System Architects and Technologists, and those that are seeking a fundamental understanding of the emerging object-based storage technologies. The audience will gain insight into the basic differences of block-, file- and object-based data access methods. The session will delve into the benefits of object storage and its value and also outline how this technology might impact future directions of storage system architectures.

- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

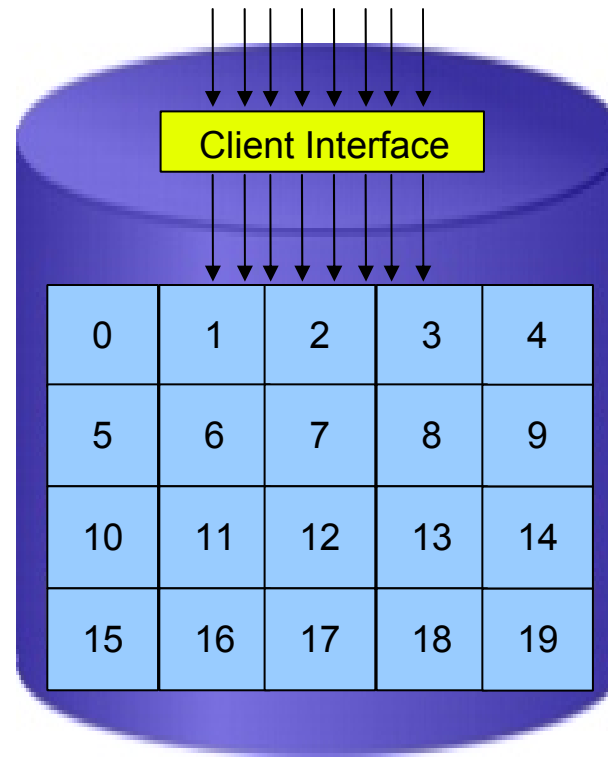
- **Block-Based Data Access**
- **File-Based Data Access**
- **Object-Based Data Access**
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- **Intelligent Storage Nodes (ISN)**

The Data Access Taxonomy



The Block Paradigm

SCSI, SAS, FCP, SRP, iSCSI, ATA, SATA



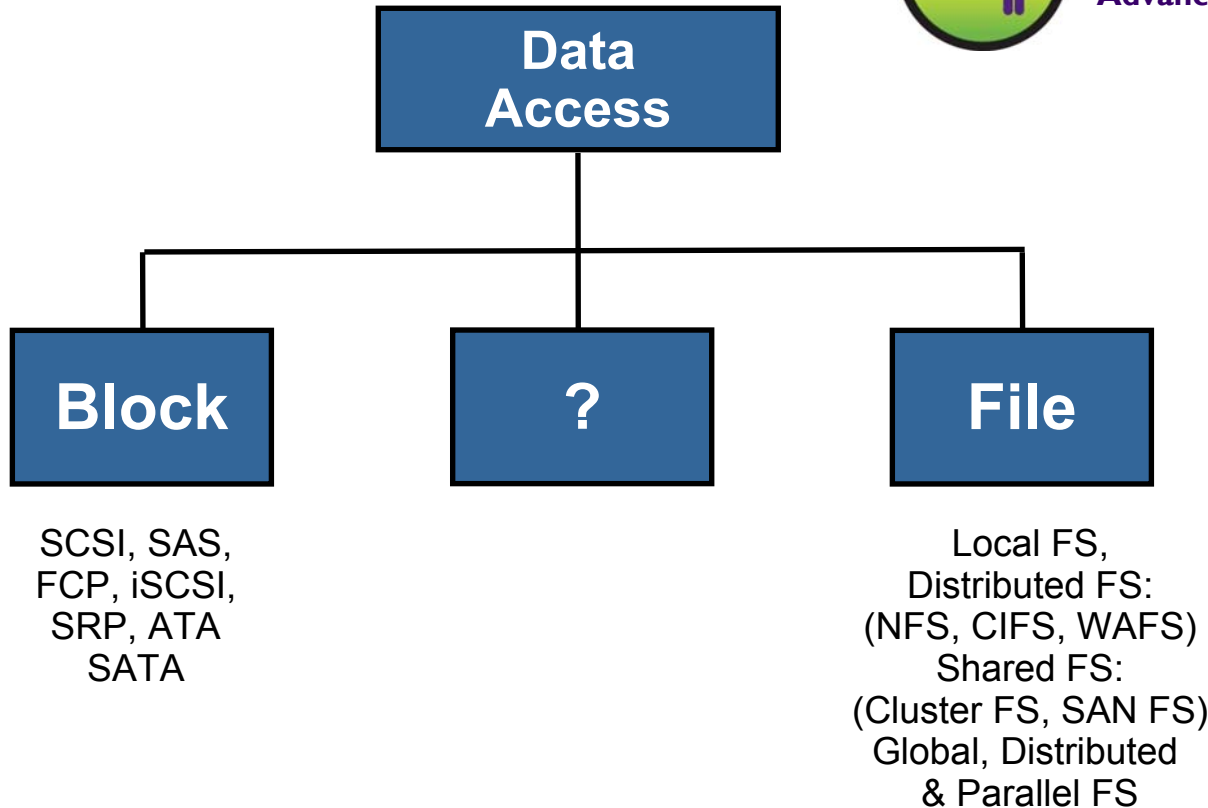
Physical Blocks:
e.g. 512 bytes

- Block-Based Data Access
- **File-Based Data Access**
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

The Data Access Taxonomy

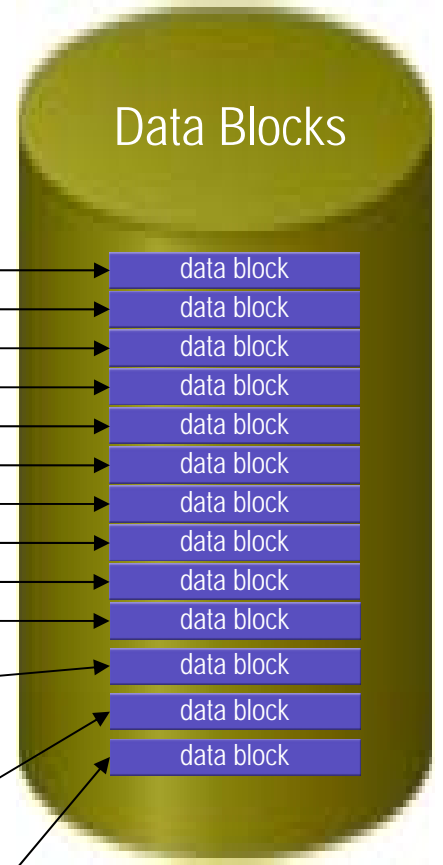
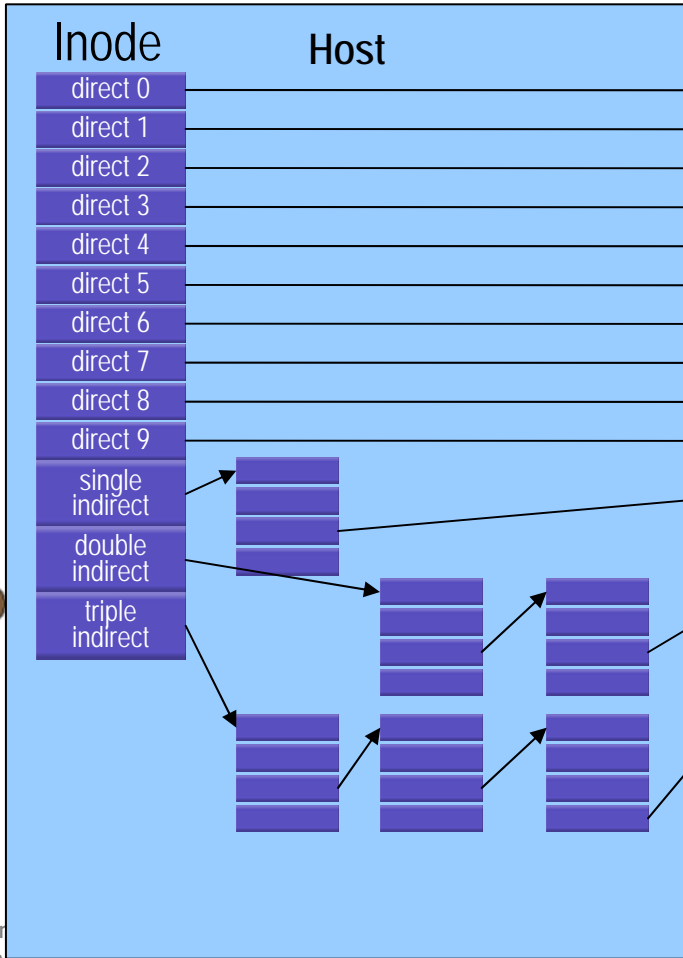
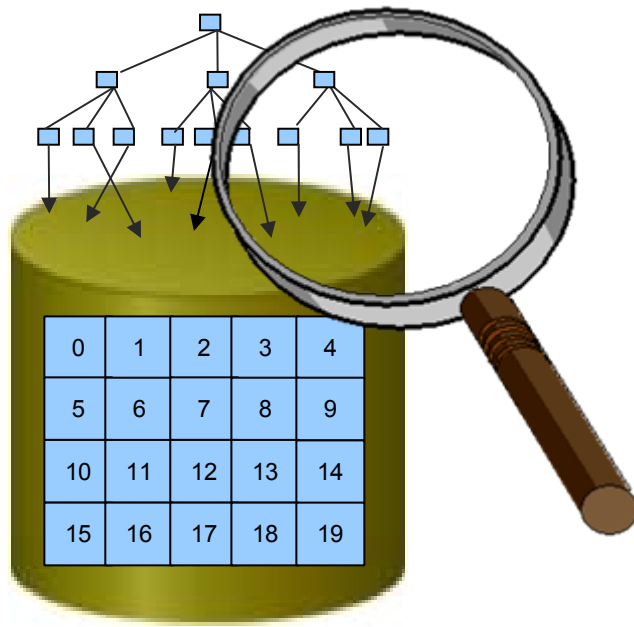


Check out SNIA Tutorial:
Advanced Data Sharing

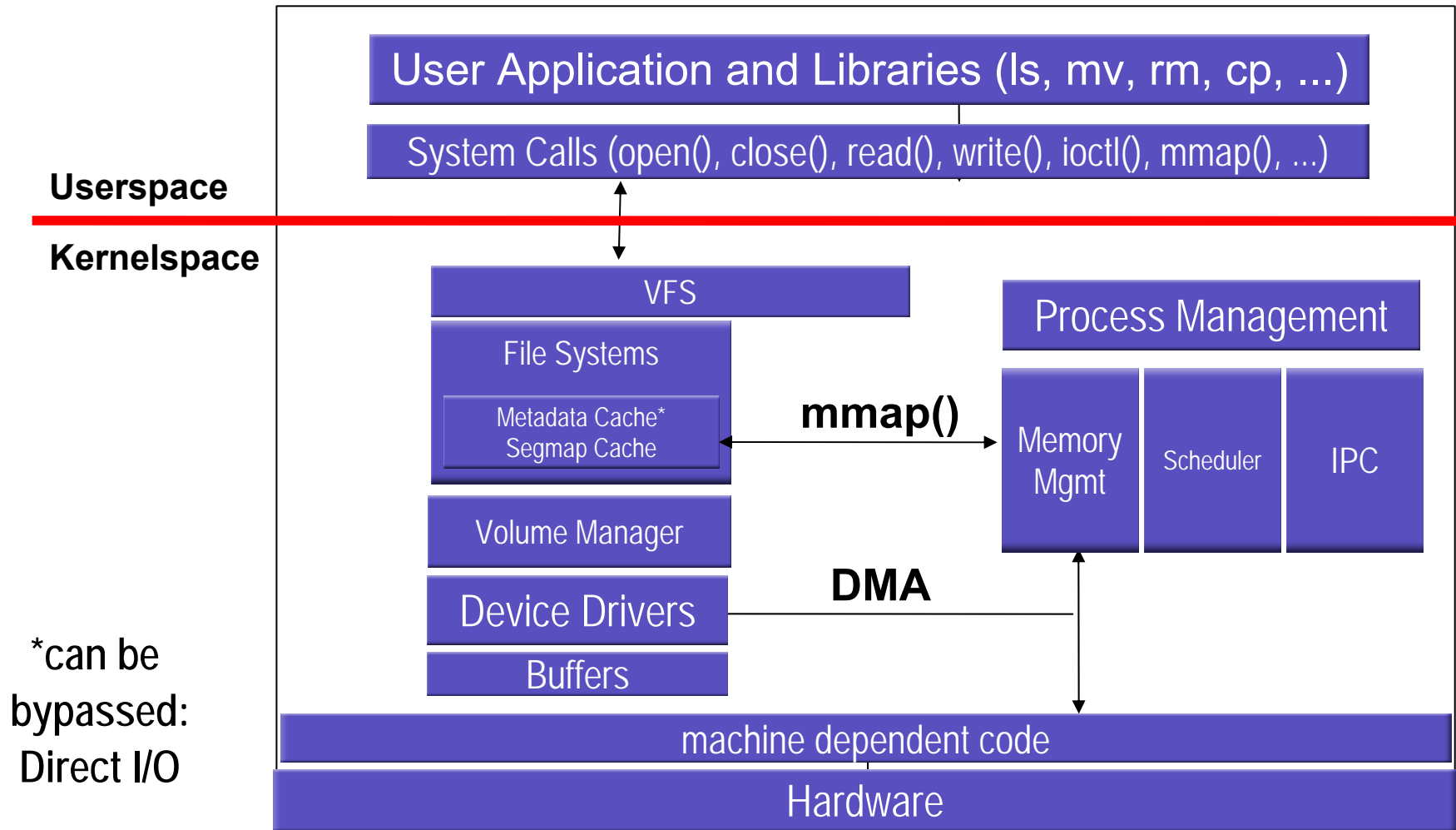


File Access - Inodes

➤ The inode contains a few block numbers to ensure efficient access to small files. Access to larger files is provided via indirect blocks that contain block numbers



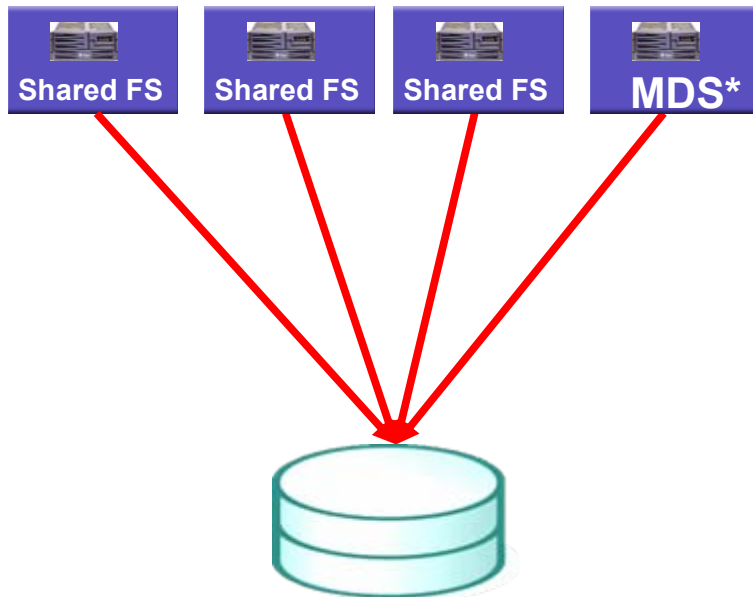
File System & Operating System



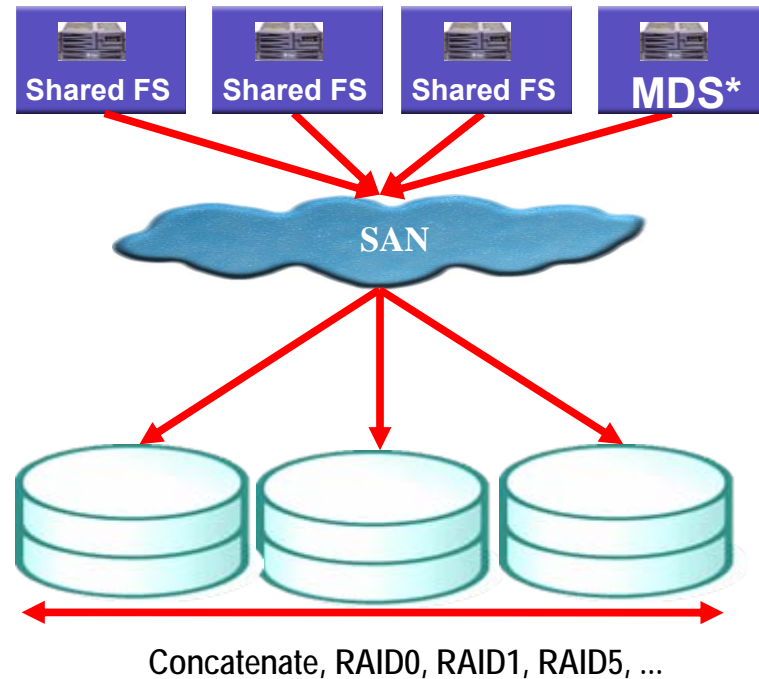
Shared File Systems – Shared Disk

- ▶ e.g. SAN FS or Cluster FS
- ▶ Dedicated or distributed Meta Data Server (MDS)

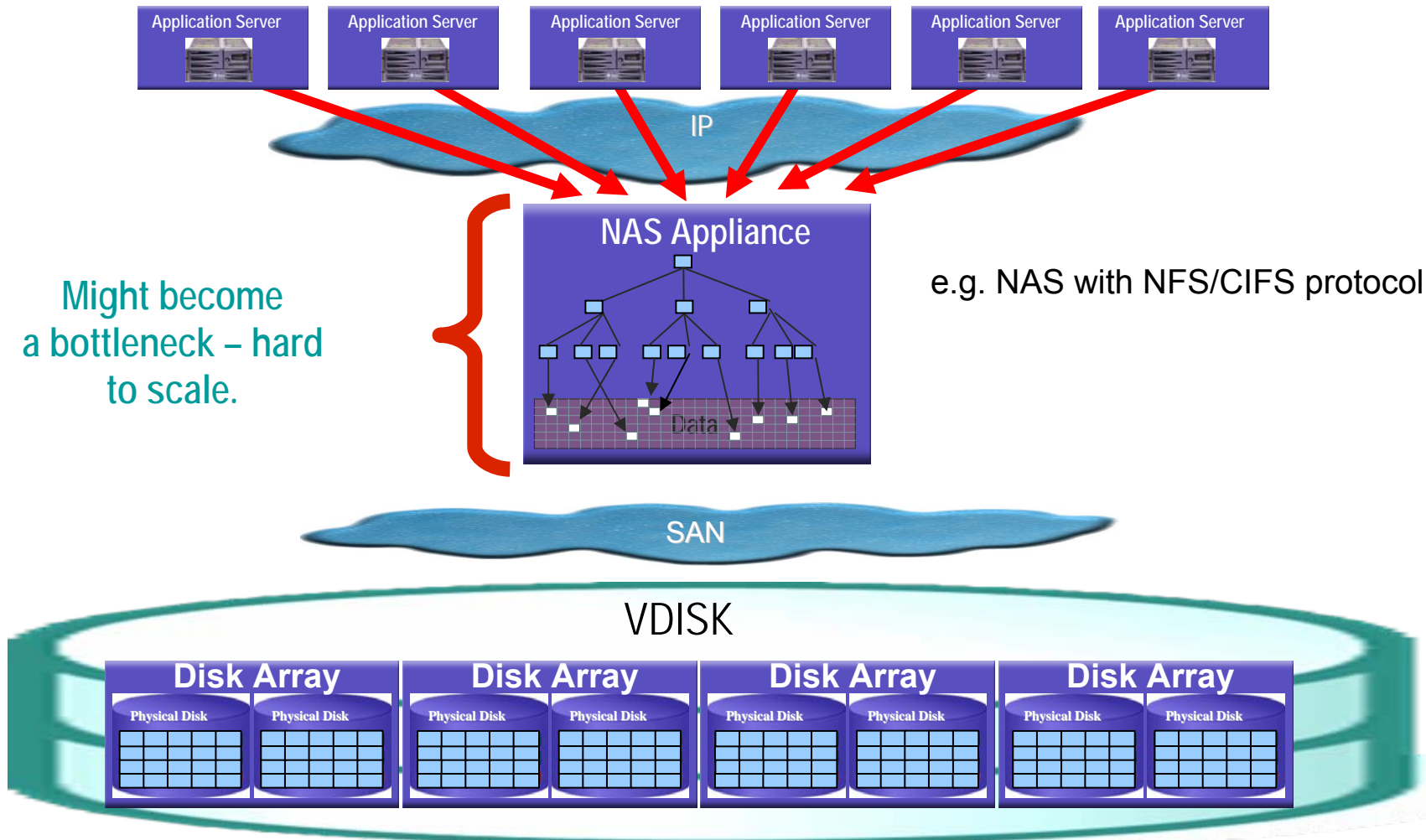
Shared FS with Shared Device



Shared FS with integrated volume management



Distributed File Systems



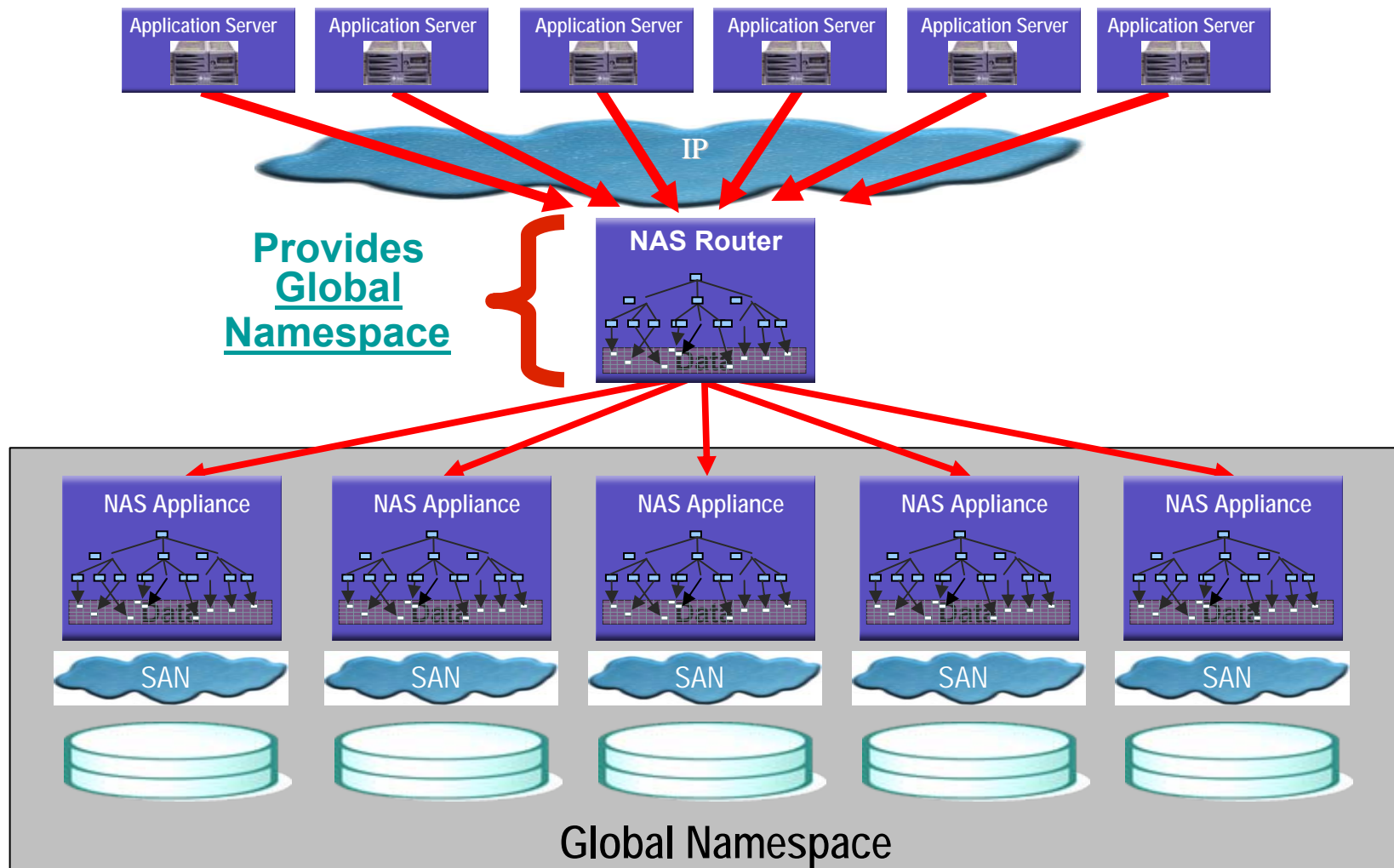
NAS Aggregation/Virtualization

Global Namespace



FS Virtualization - NAS Aggregation

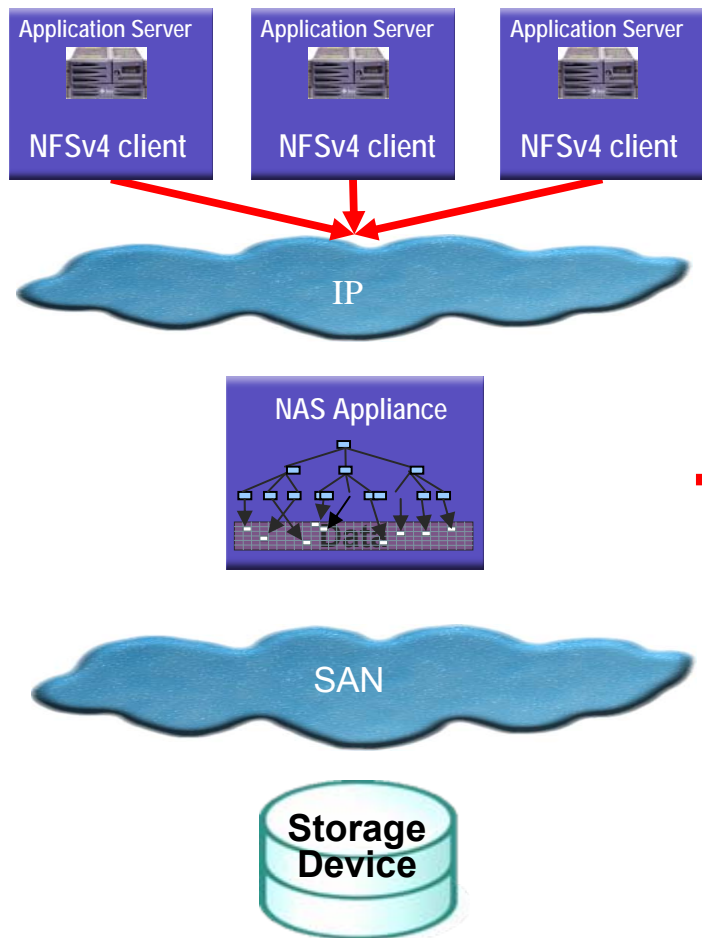
In-Band Solution



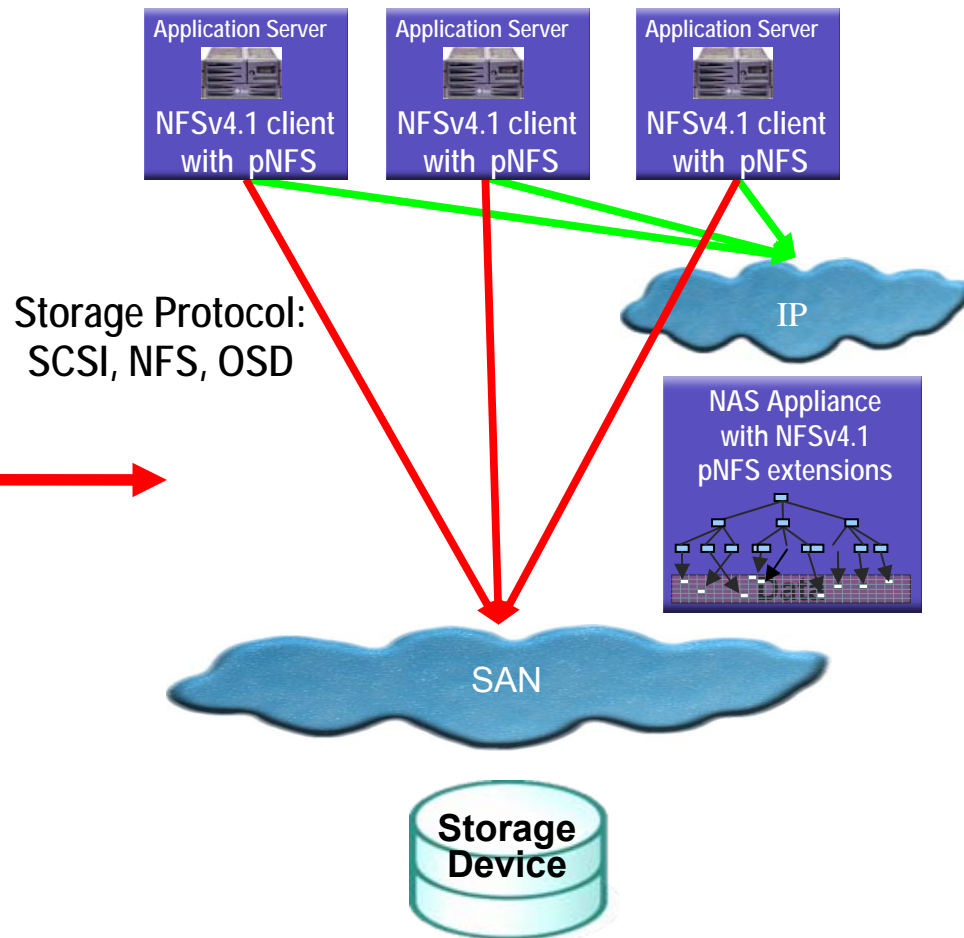
FS Virtualization – NFS4.1 pNFS

Out-of-Band Solution

In-Band NAS:



Out-of-Band NAS:

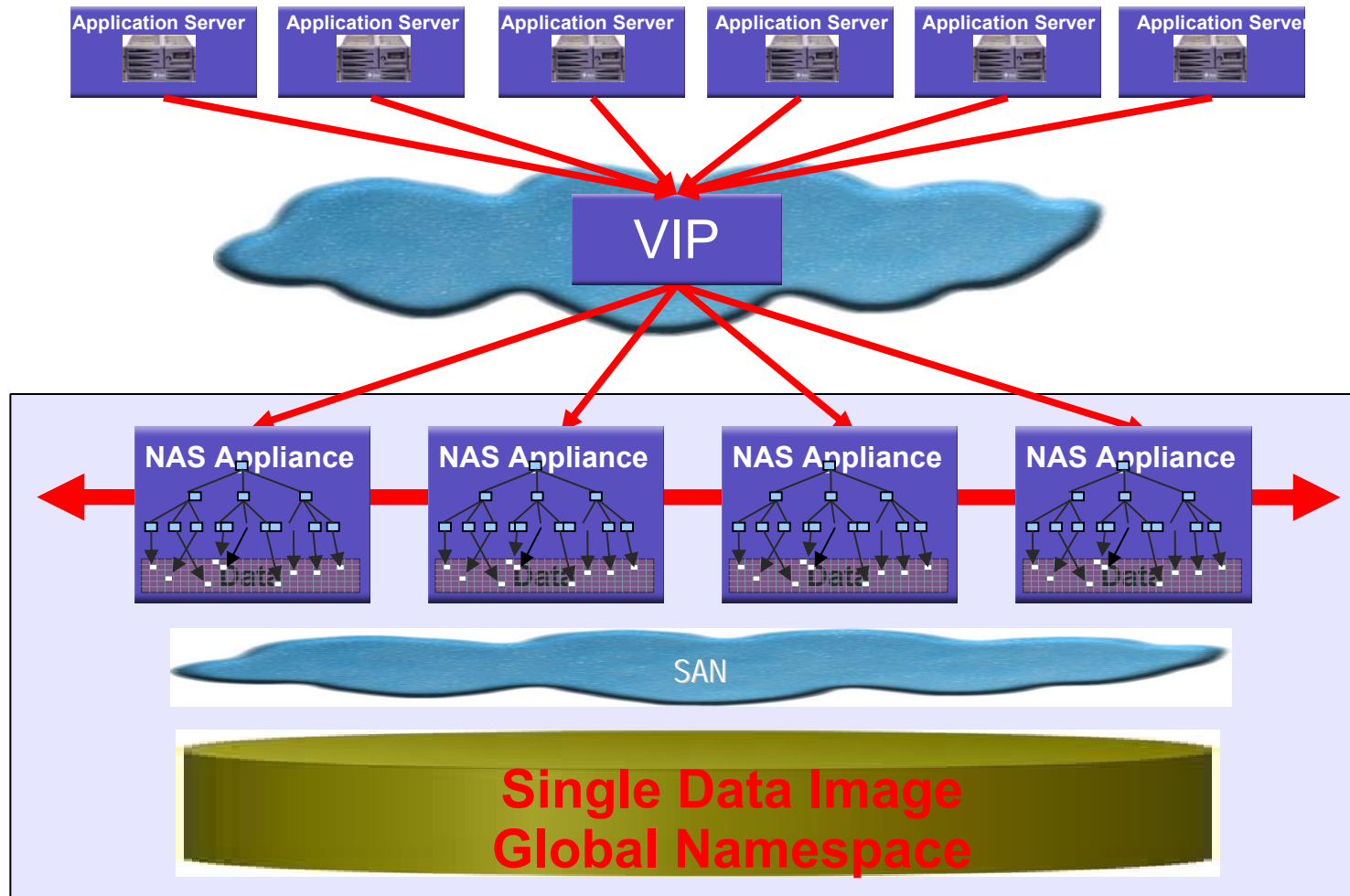


Storage Protocol:
SCSI, NFS, OSD

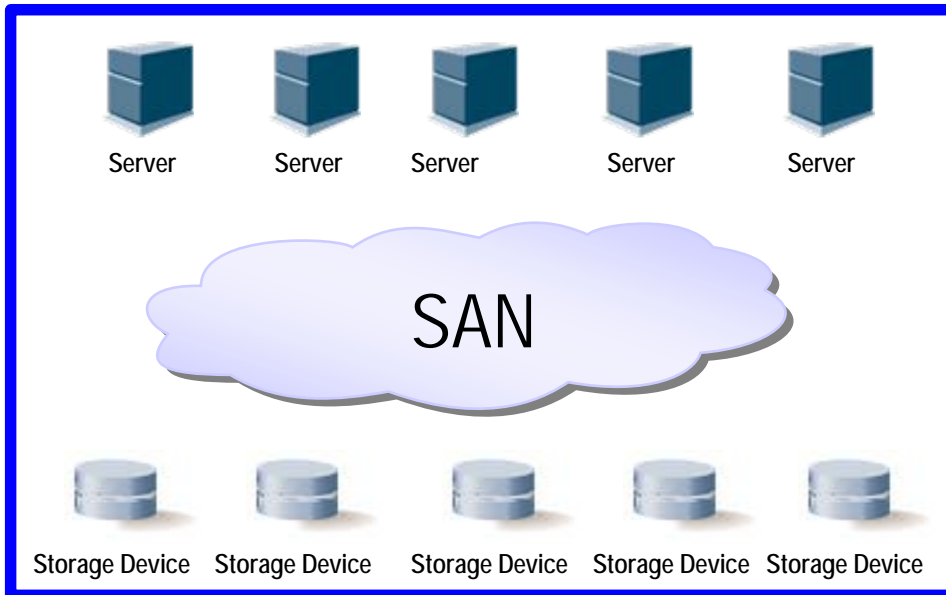
Scalable NAS - Loosely Coupled NAS Cluster

NAS Cluster – NAS Grid

aka Tightly Coupled NAS

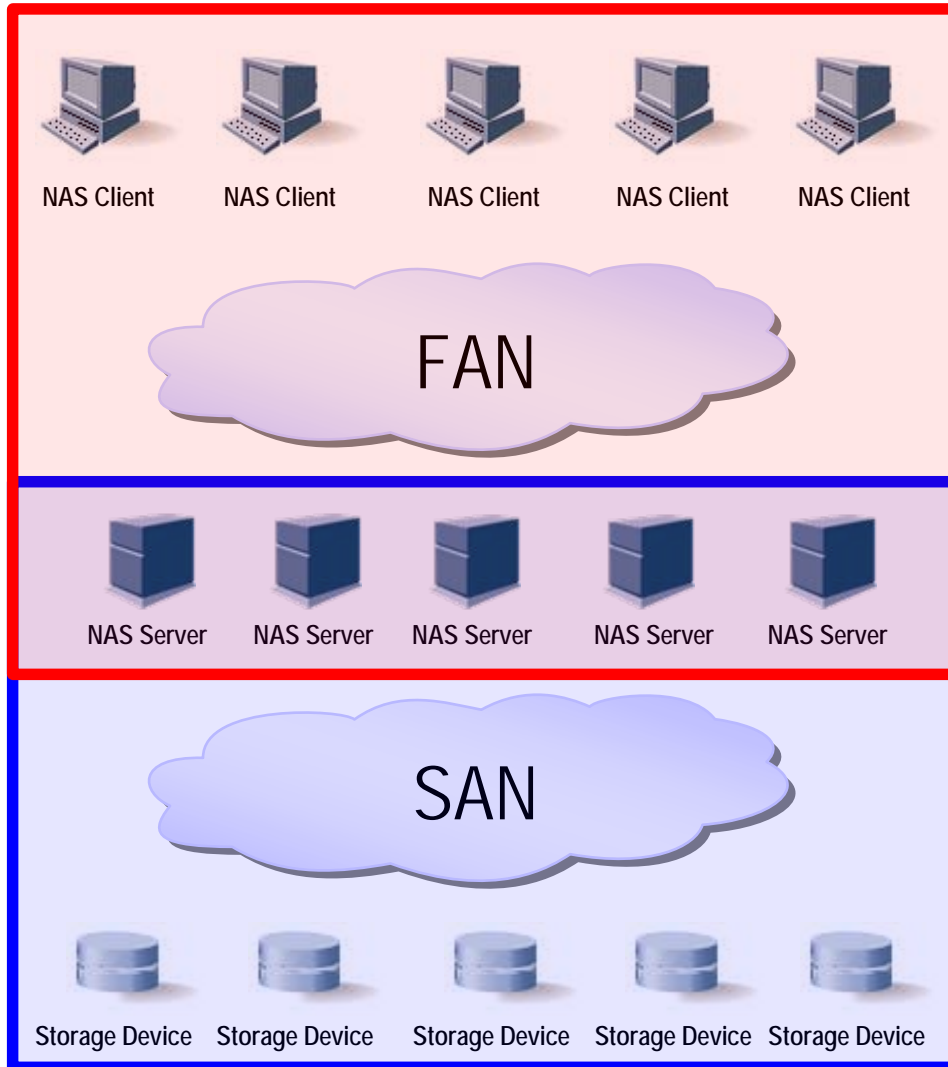


Storage Area Network - SAN



Global Address Space
(WWN, 24-bit fabric address,
nameserver),
zoning, routing, ...

File Area Network - FAN



Check out SNIA Tutorial:

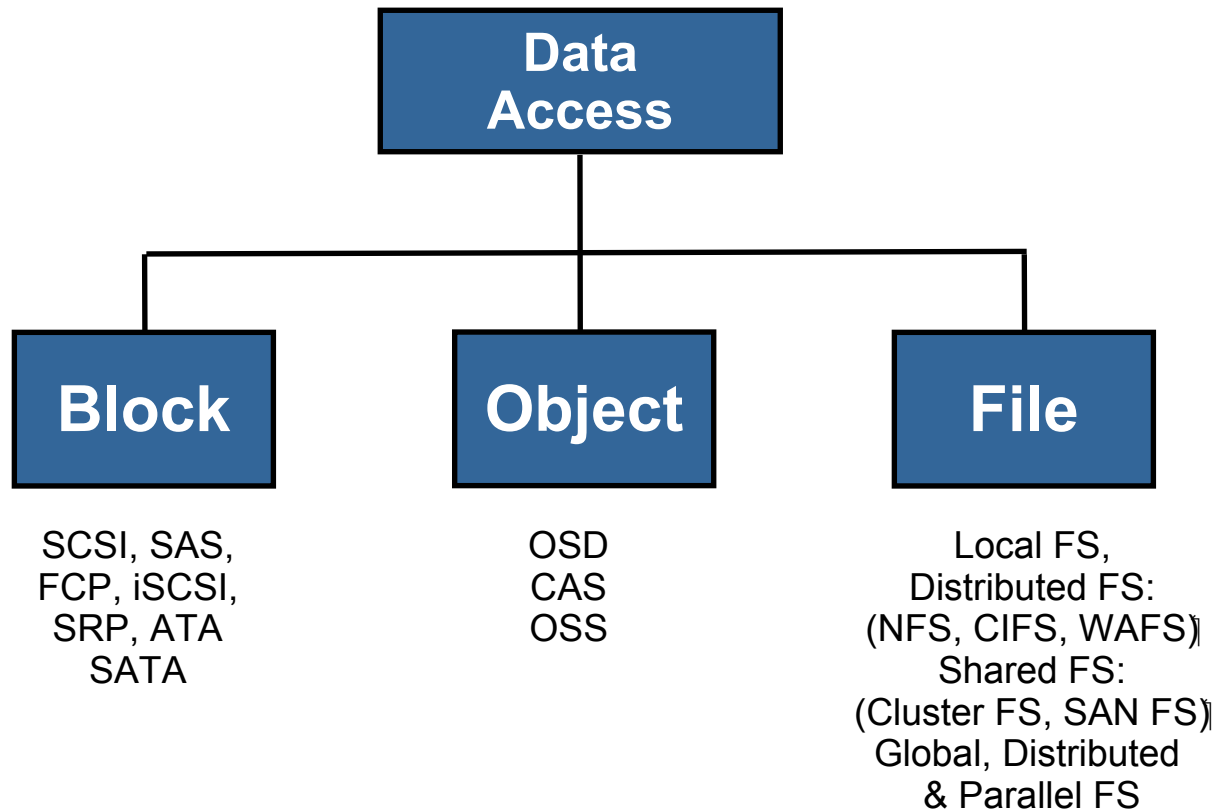
FAN: An Architecture for Scalable, Service-Oriented Data Management

Global Namespace,
load-balancing, network compression (WAFS), data protection (security, replication, ...), SLA/ILM (migration, retention), ...

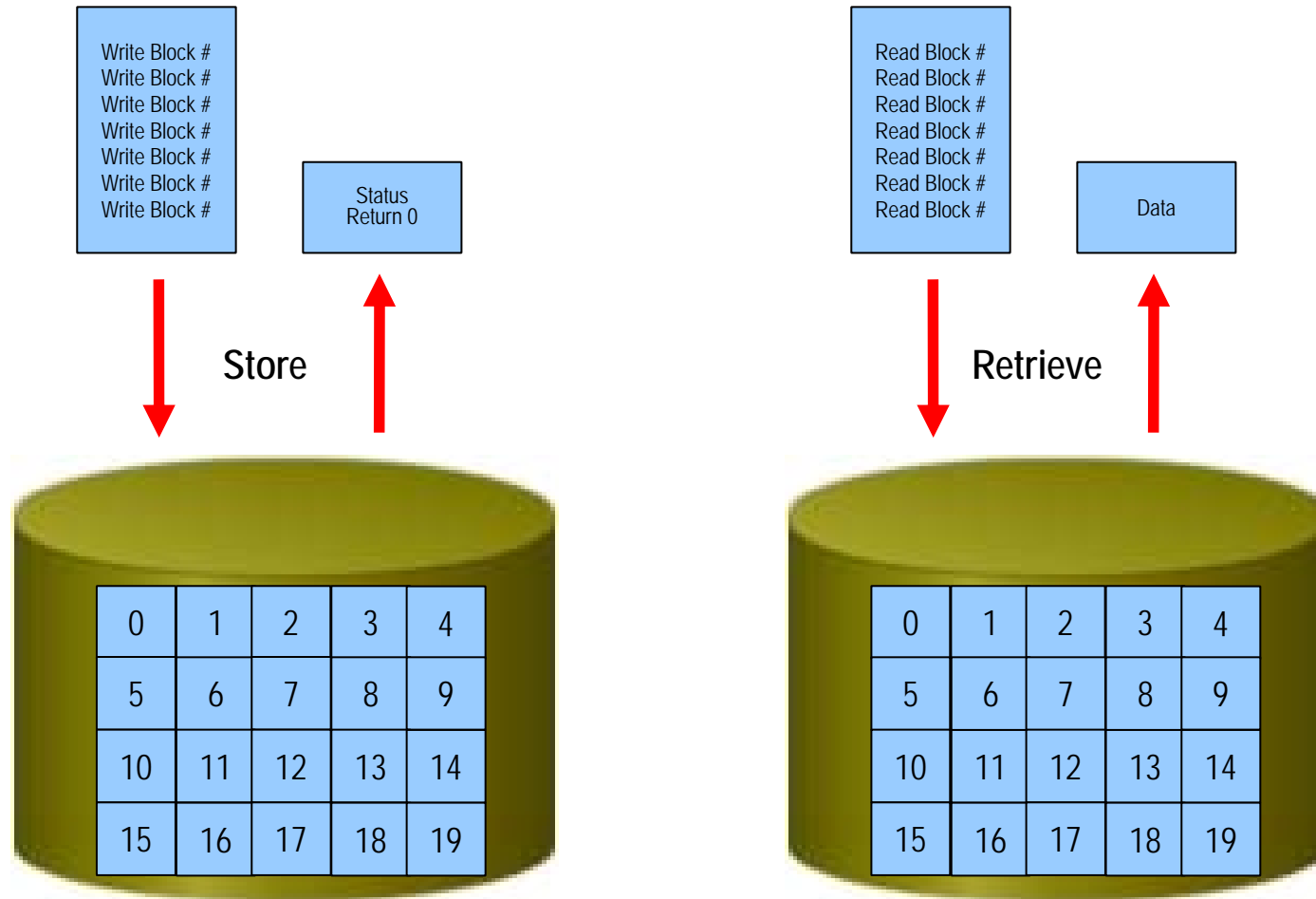
Global Address Space
(WWN, 24-bit fabric address, nameserver),
zoning, routing, ...

- Block-Based Data Access
- File-Based Data Access
- **Object-Based Data Access**
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

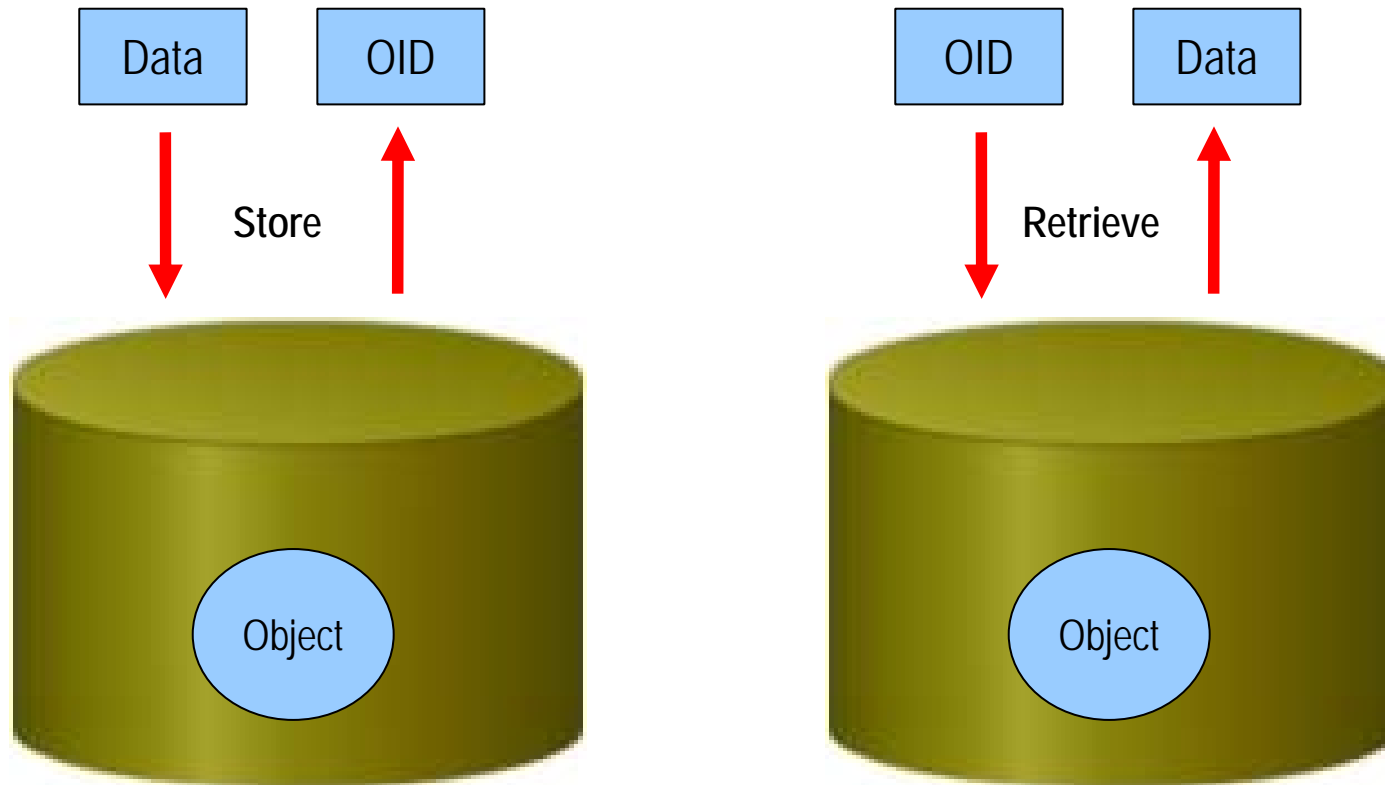
The Data Access Taxonomy



The Old Block Paradigm



The New Object Paradigm



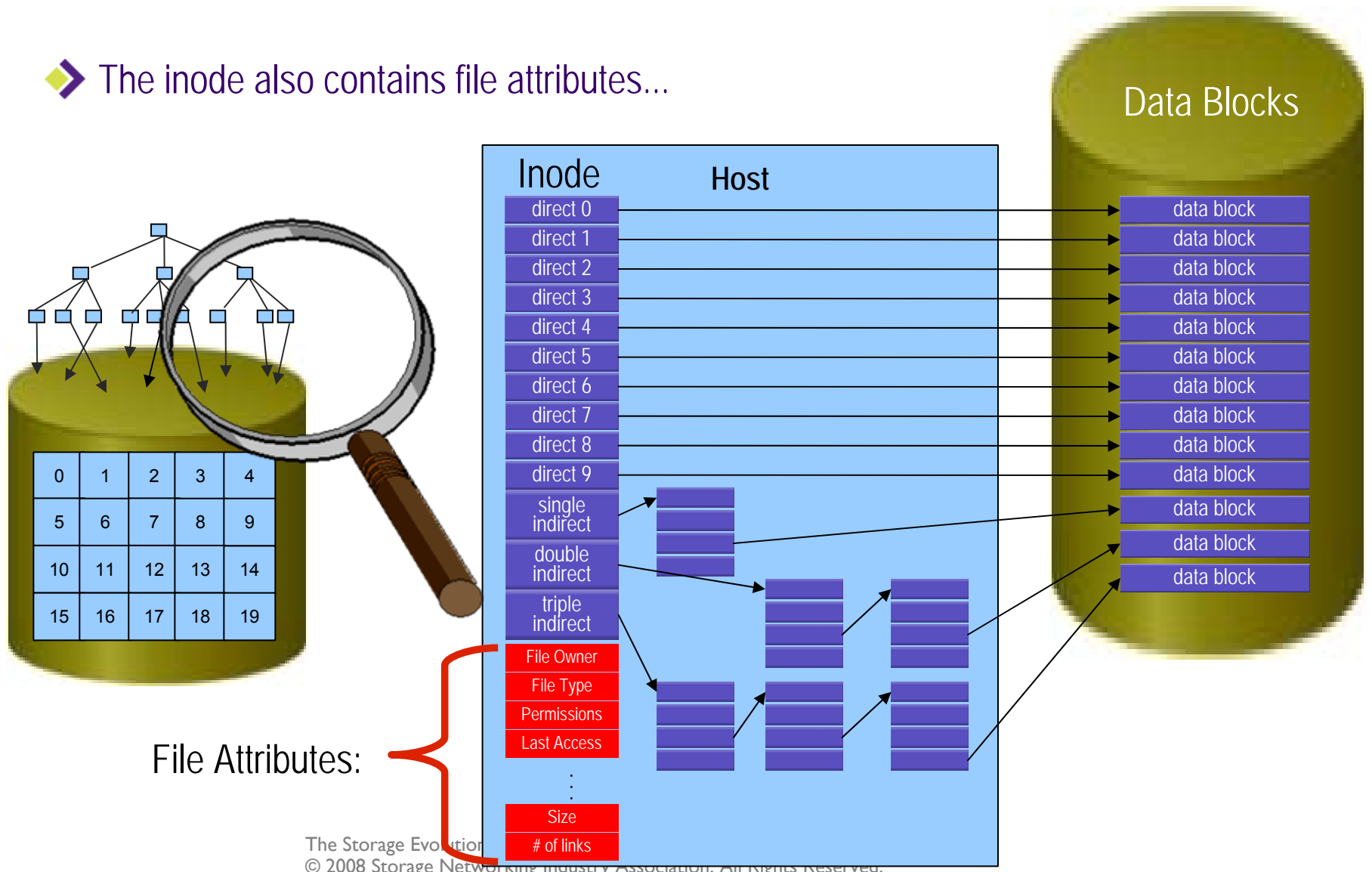
The New Challenge: Unstructured Data

- ▶ How to find one thing out of 100 million things?



File Access – Inodes

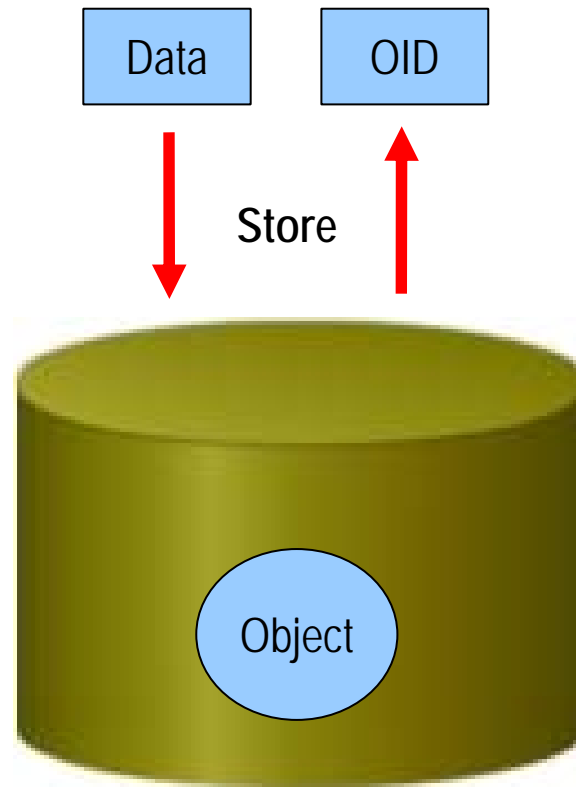
➤ The inode also contains file attributes...



File Attributes:

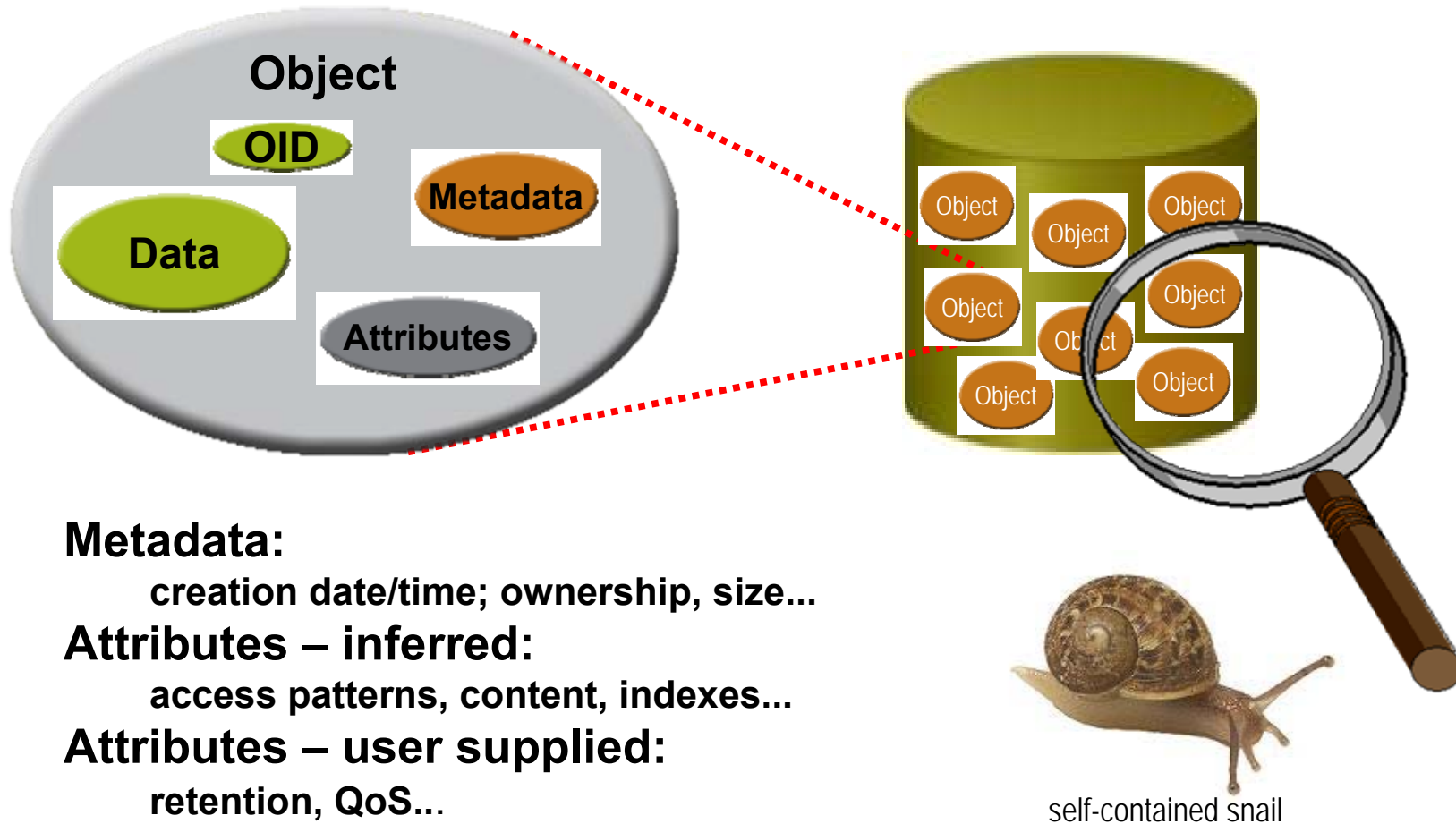
The New Object Paradigm (cont'd)

- WRITE 26,763 Bytes
- QoS = High
- Description = "X-Ray"
- Retention = 50 years
- Access Key = *&^%#
- Data Payload.....



- **Object Storage Responsibilities:**
 - Space Management
 - Access Control (Identity Mgmt)
 - QoS Management
 - Cache, Backup
 - Policy Migration, Retention

Self-Contained Objects



Metadata:

creation date/time; ownership, size...

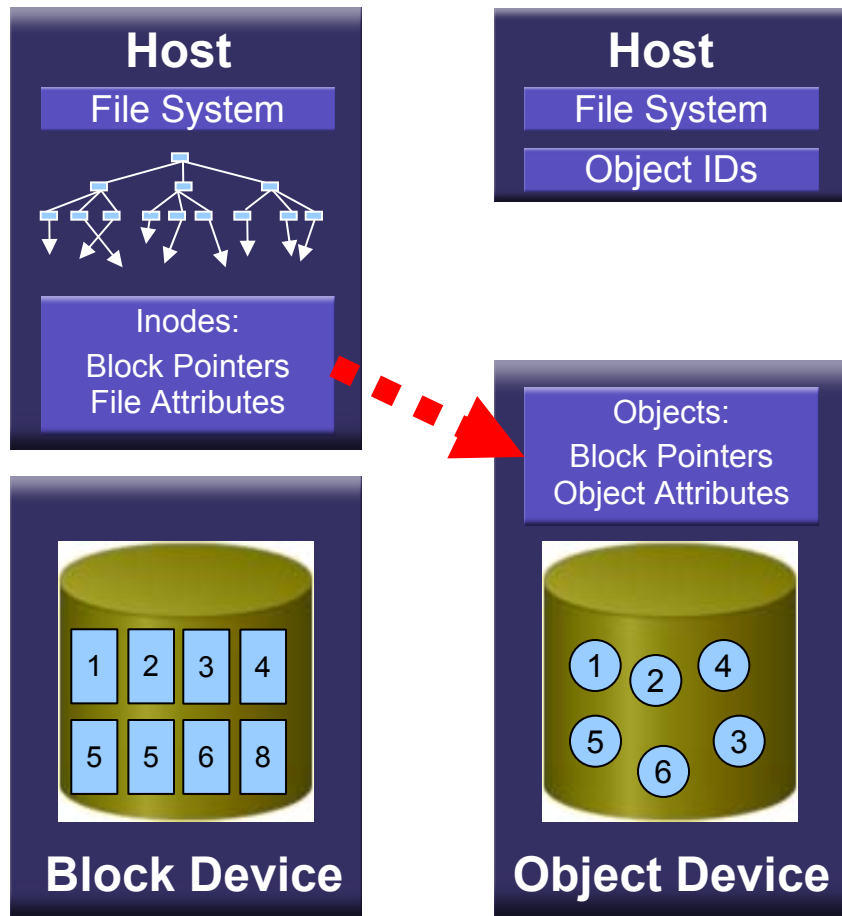
Attributes – inferred:

access patterns, content, indexes...

Attributes – user supplied:

retention, QoS...

Object Autonomy

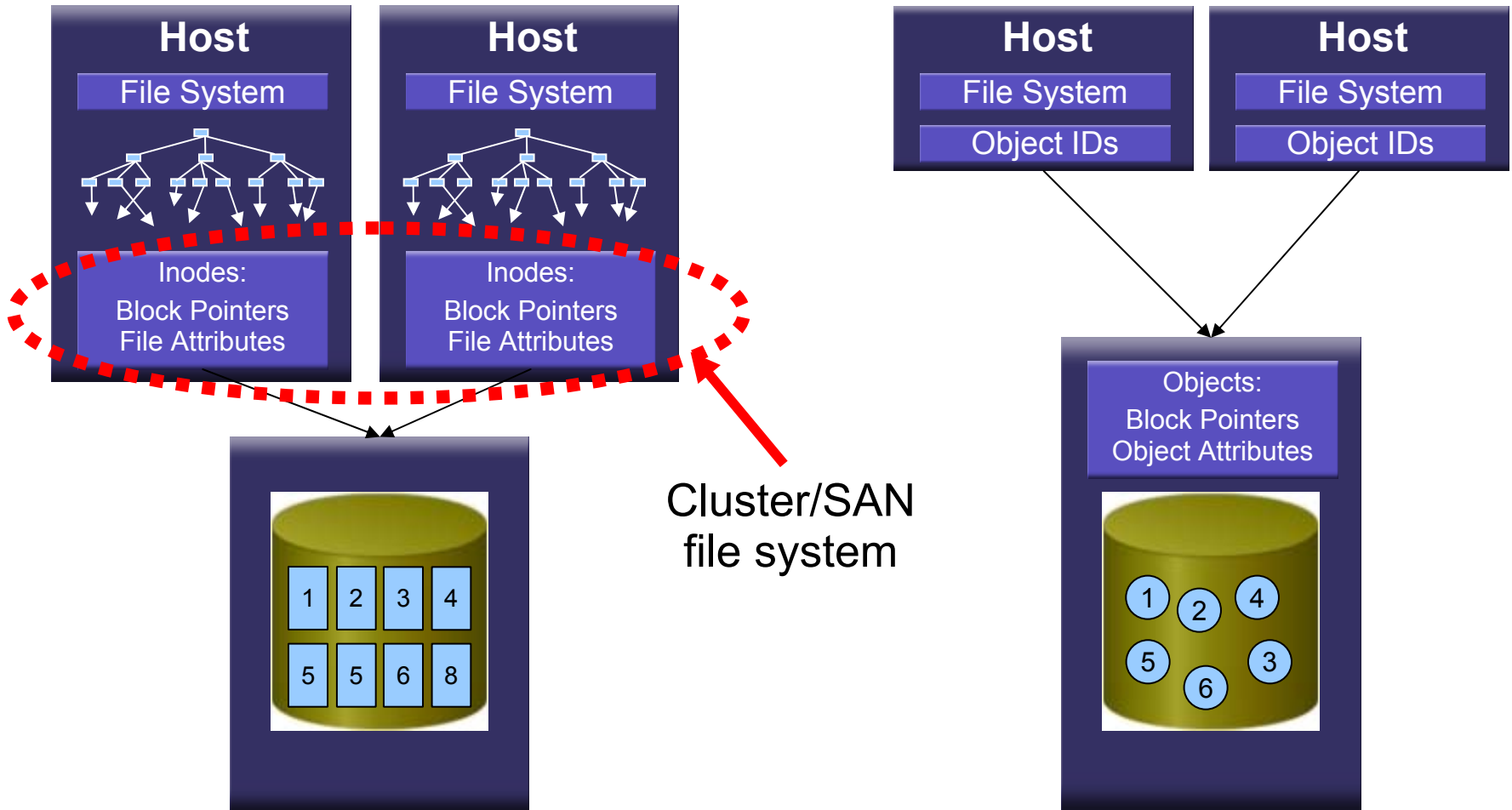


Storage becomes autonomous:

- ◆ capacity planning
- ◆ load balancing
- ◆ backup
- ◆ QoS, SLAs
- ◆ understand data/object grouping
- ◆ aggressive pre-fetching
- ◆ thin provisioning
- ◆ search
- ◆ compression/de-duplication, commonality factoring
- ◆ strong security, encryption
- ◆ compliance/retention/secure delete
- ◆ availability/replication
- ◆ Audit
- ◆ self-healing
-
-
-

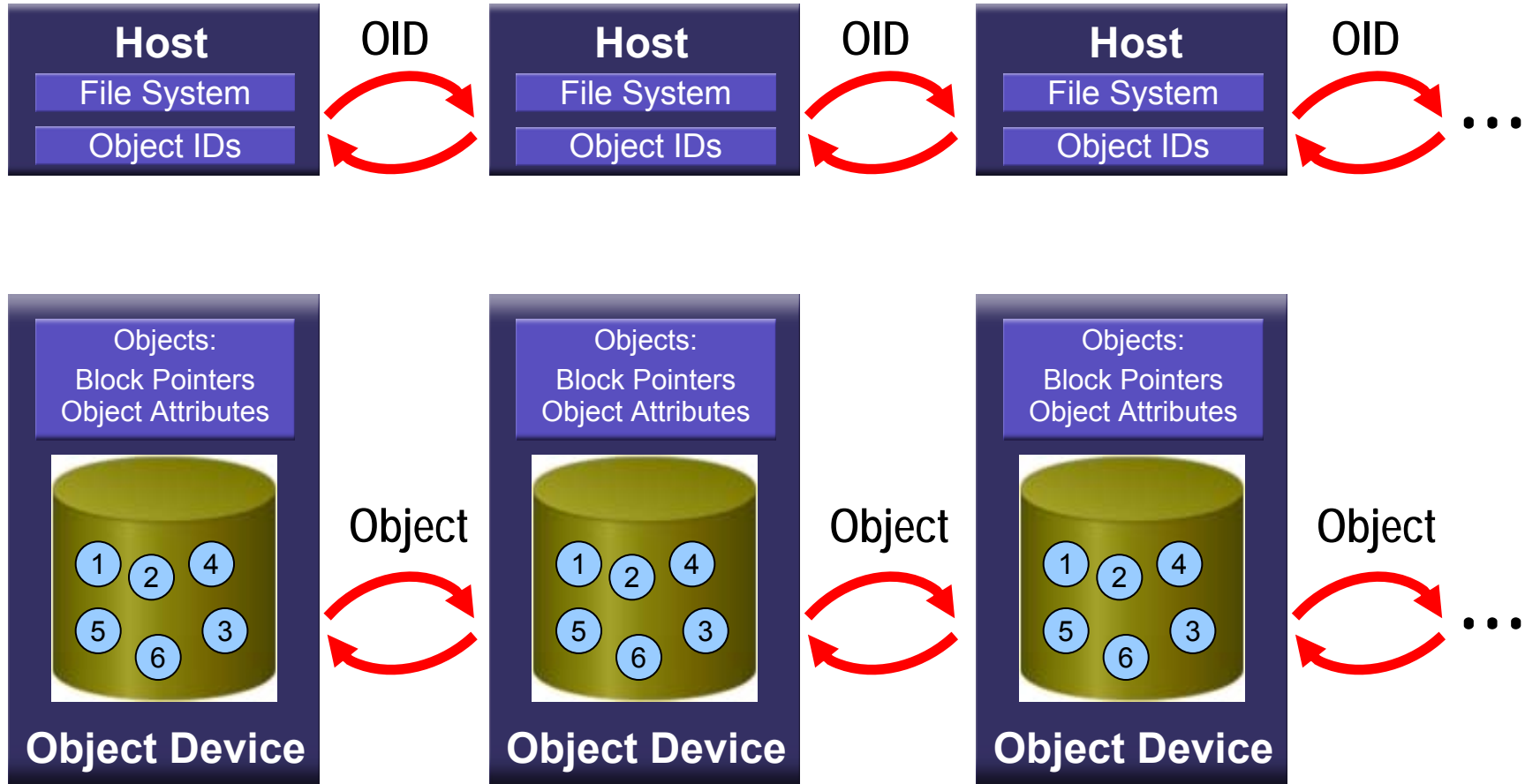
Data Sharing

Homogeneous/Heterogeneous

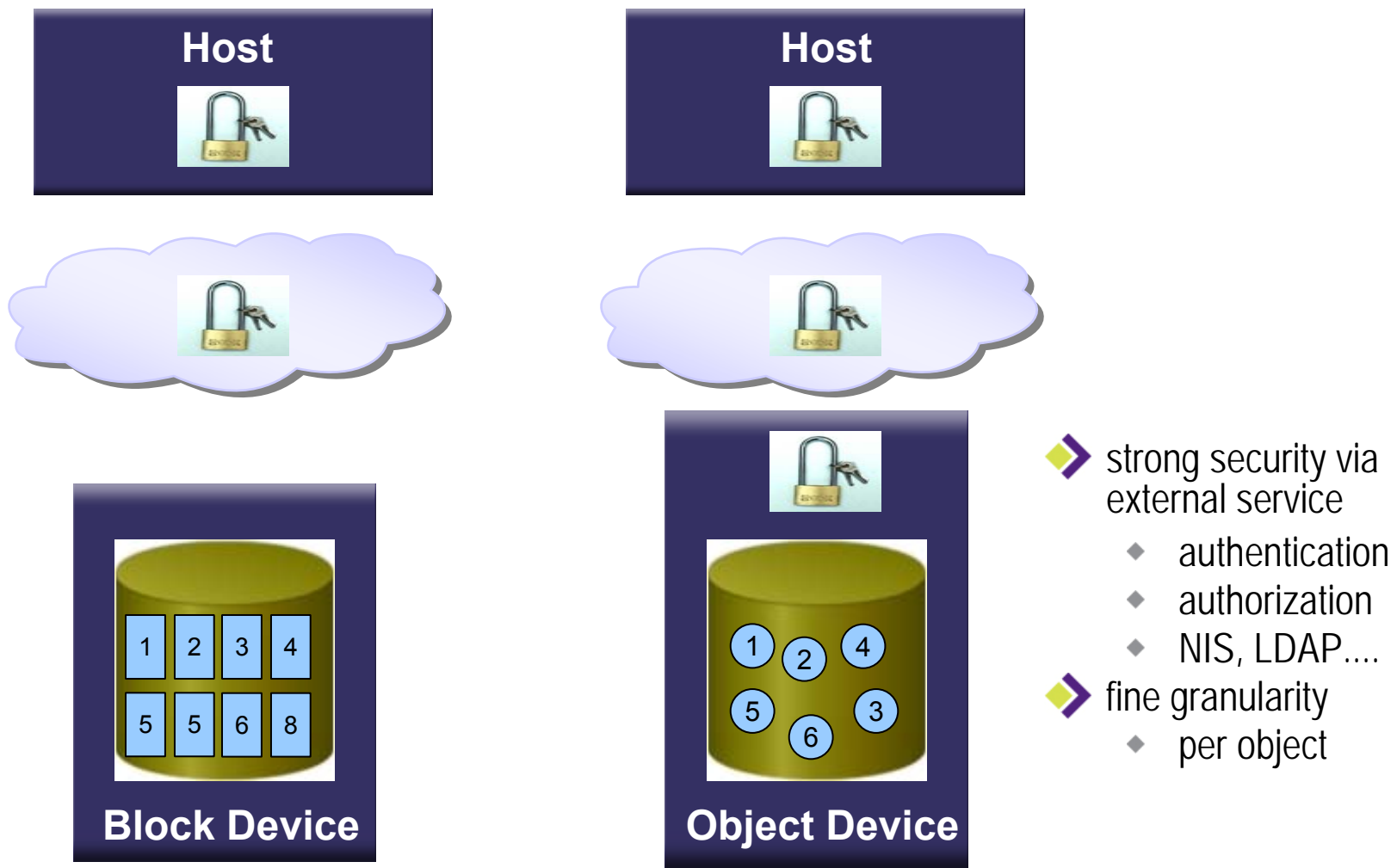


Data Migration - ILM

Homogeneous/Heterogeneous

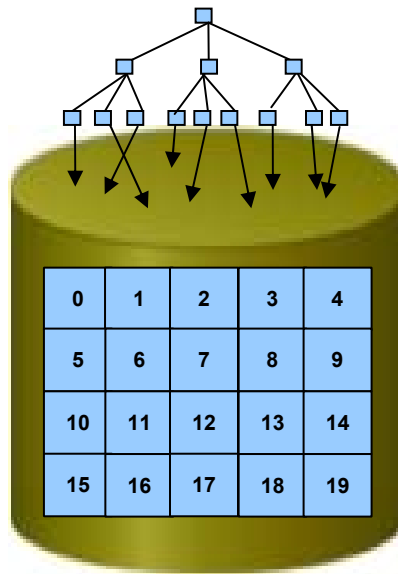


Additional Layer of Security



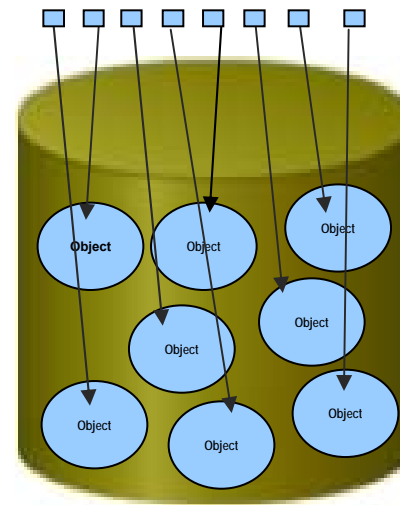
Living in a Flat Namespace

File names / inodes



Traditional
Hierarchical

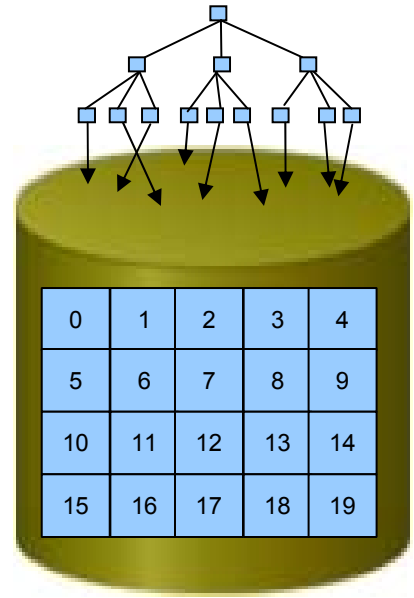
Objects / OIDs



Flat

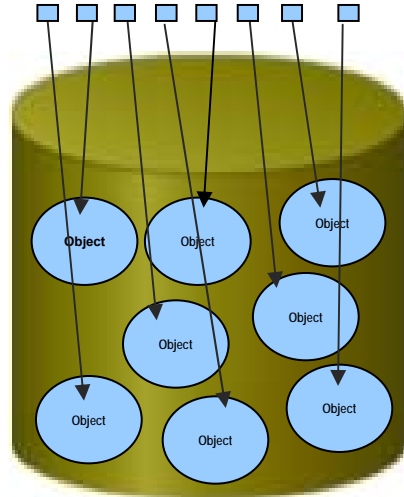
Virtual View / Virtual File Systems

File names / inodes



Traditional

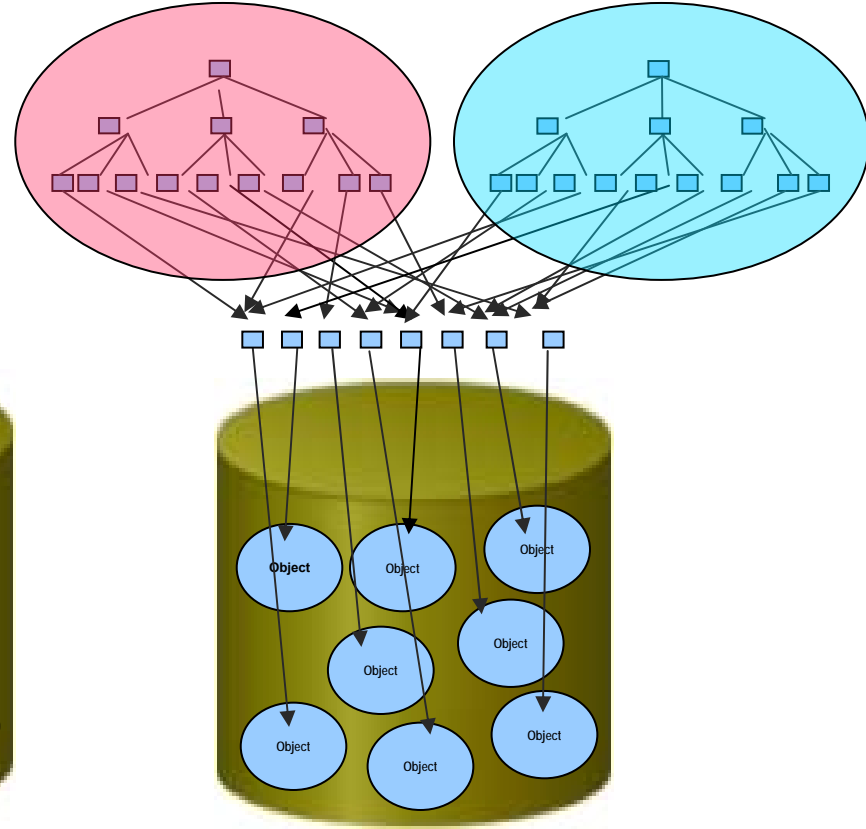
Objects / OIDs



Flat

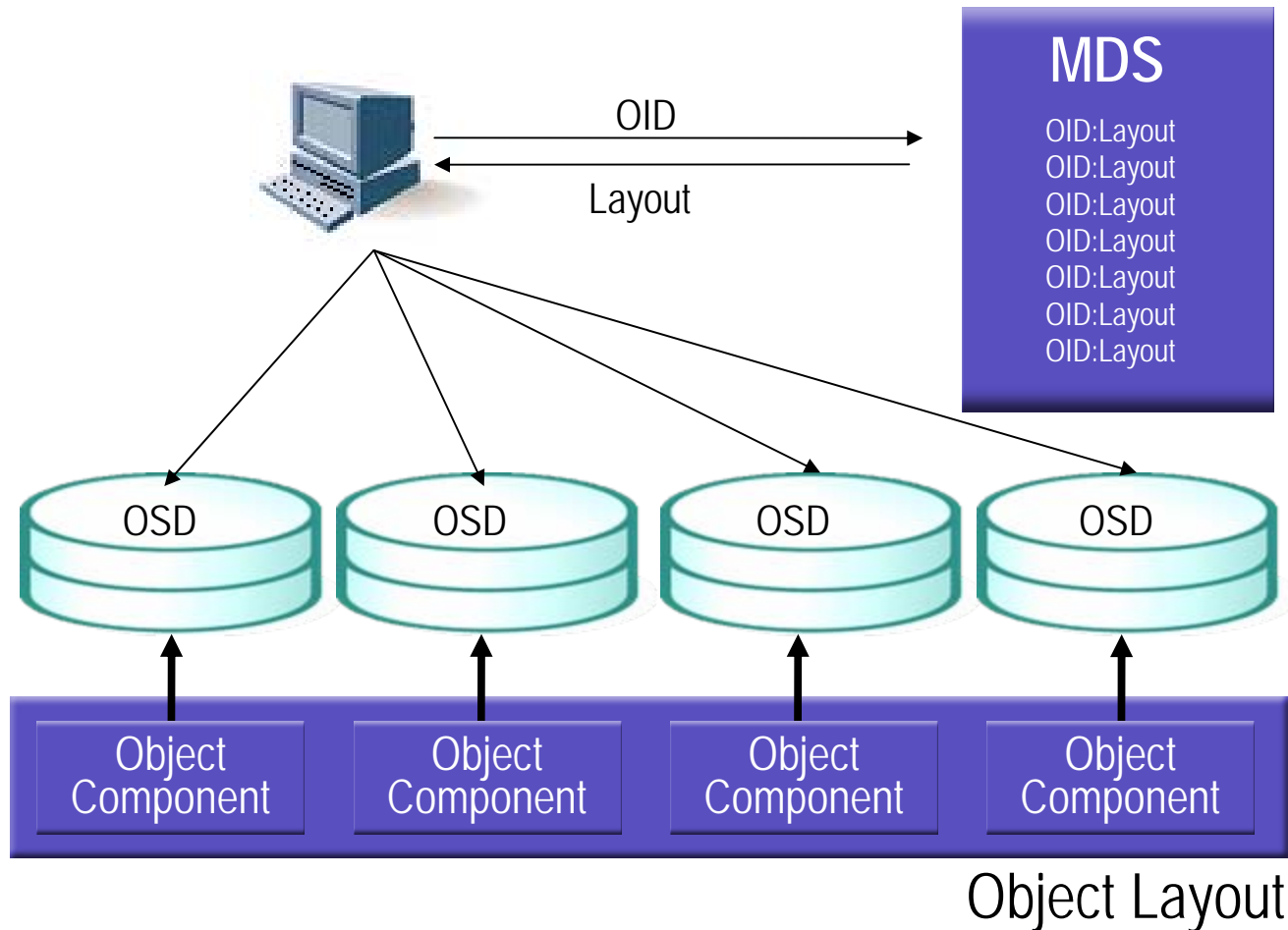
Virtual View A

Virtual View B



Virtual

Object Decomposition

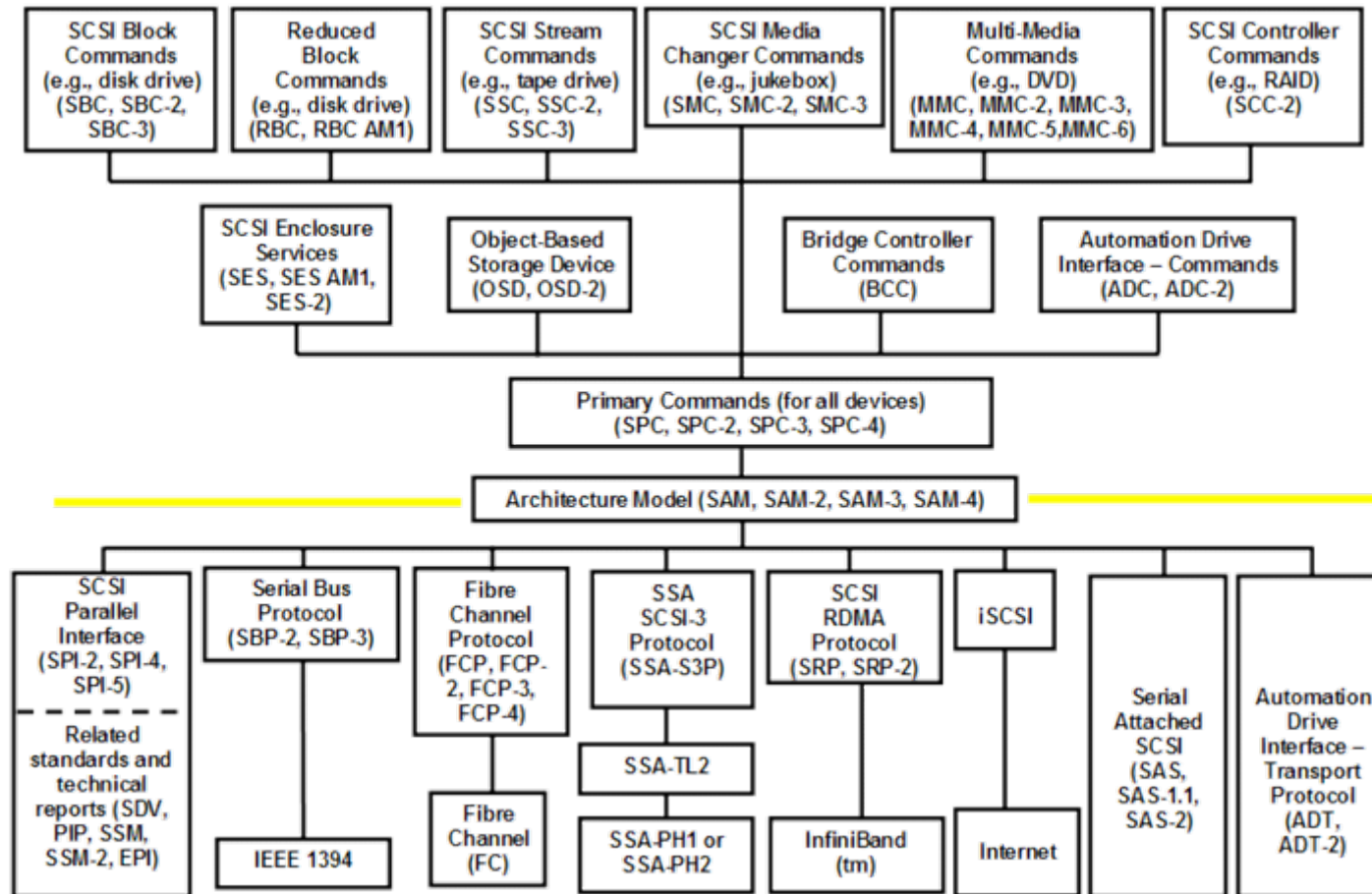


- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)



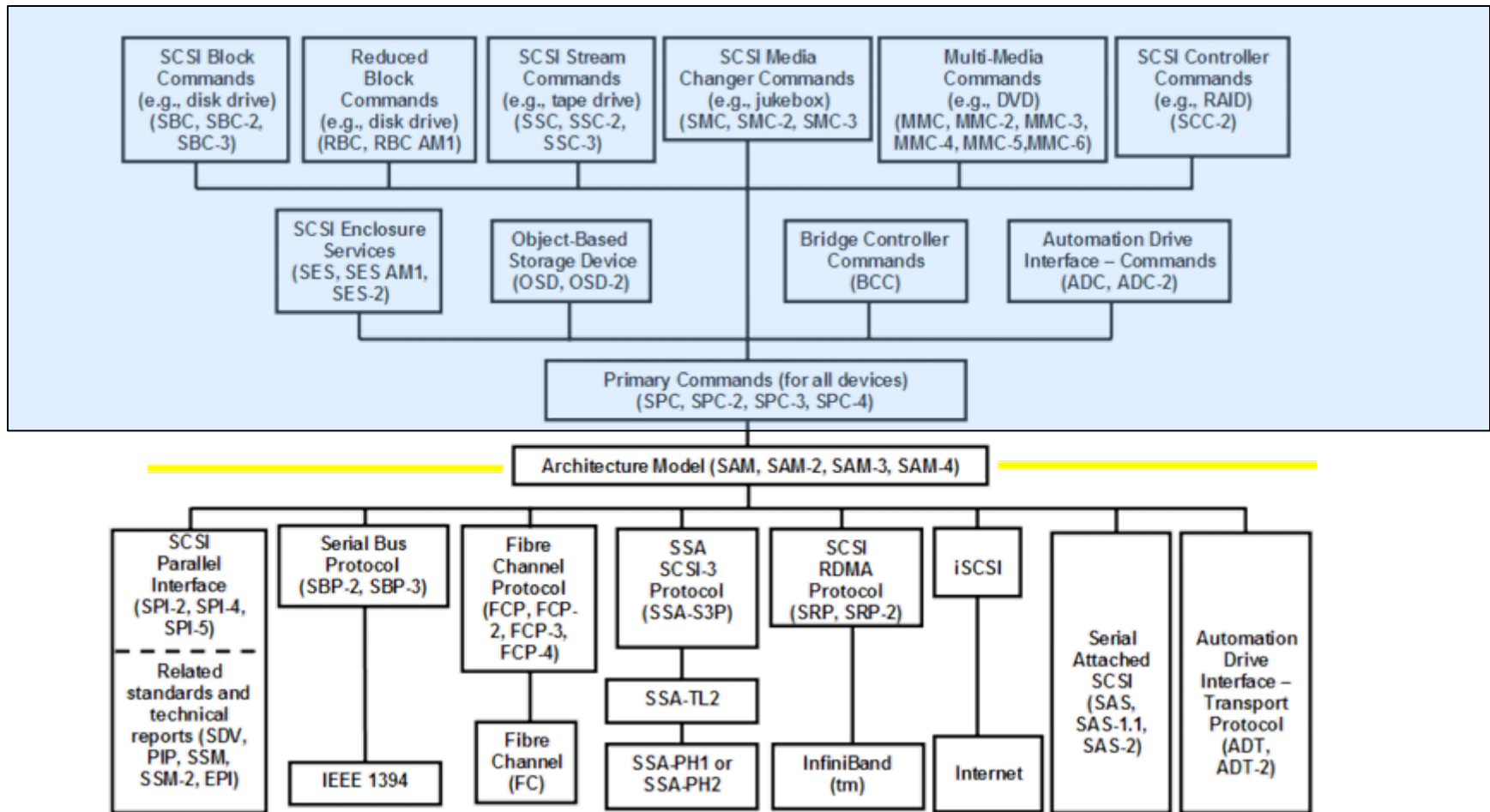
Check out **SNIA Tutorial:**
Object-Based Storage Devices

SCSI Standards Architecture



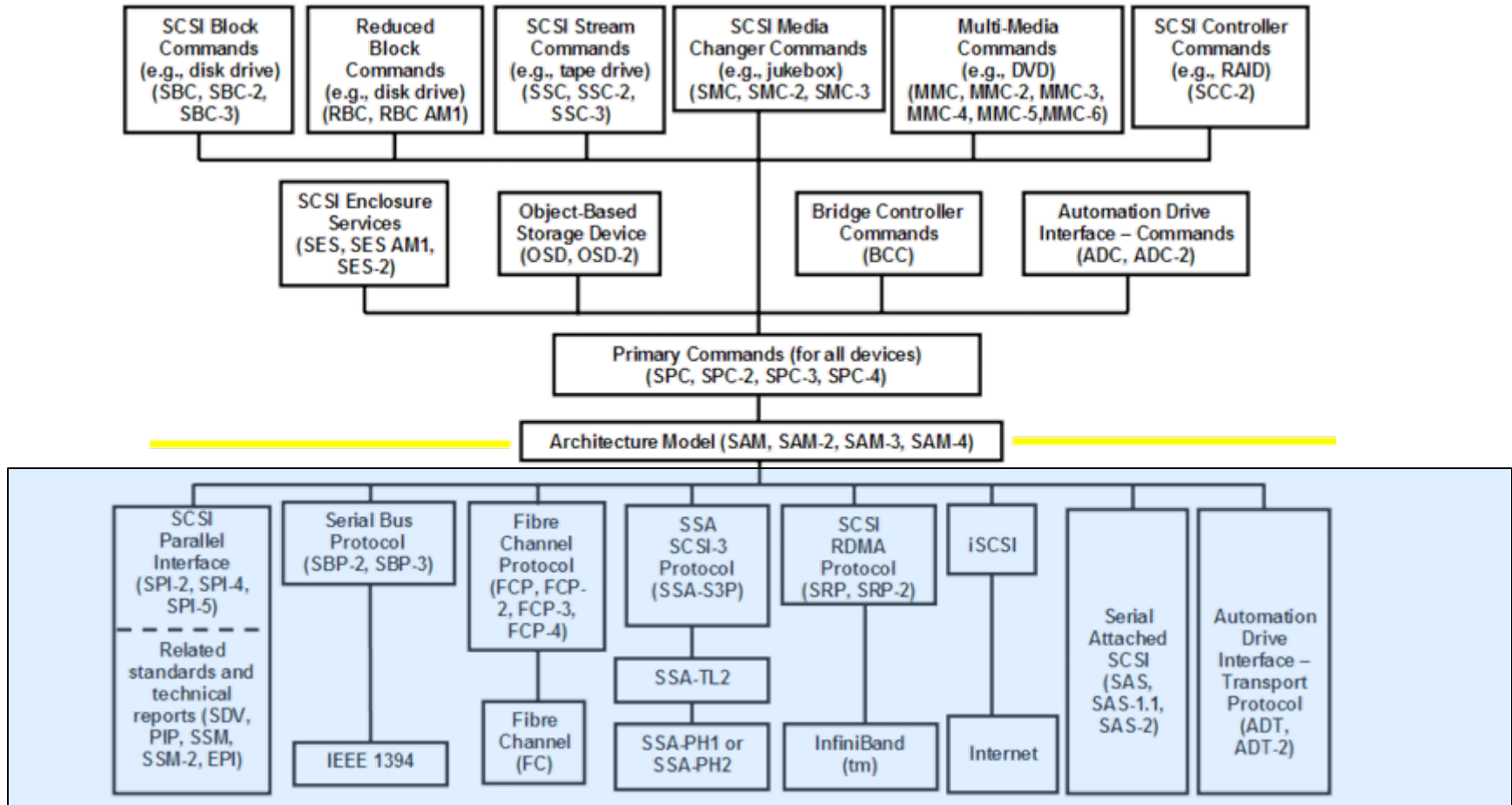
Source: www.t10.org

SCSI Commands



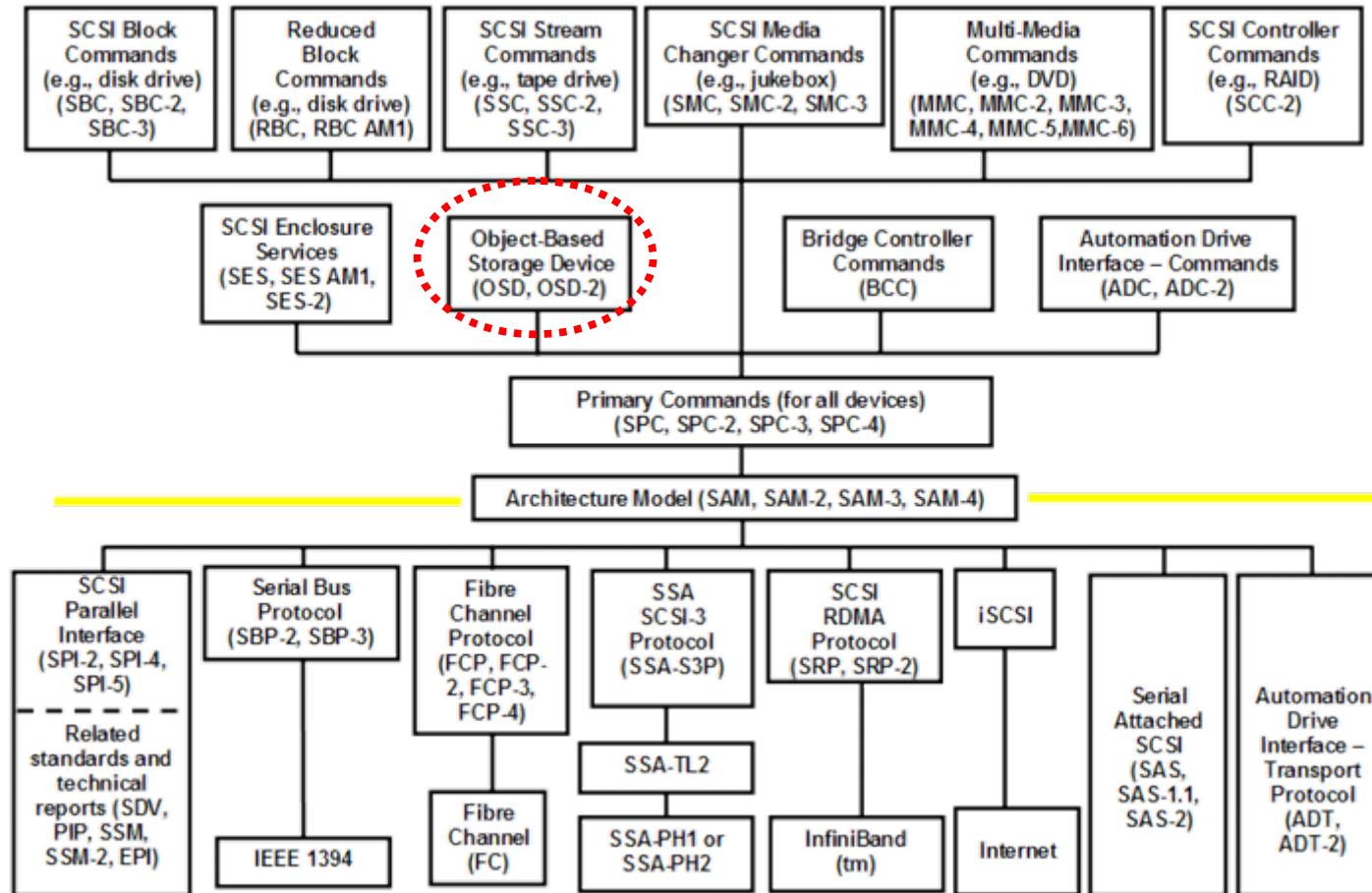
Source: www.t10.org

SCSI Transports



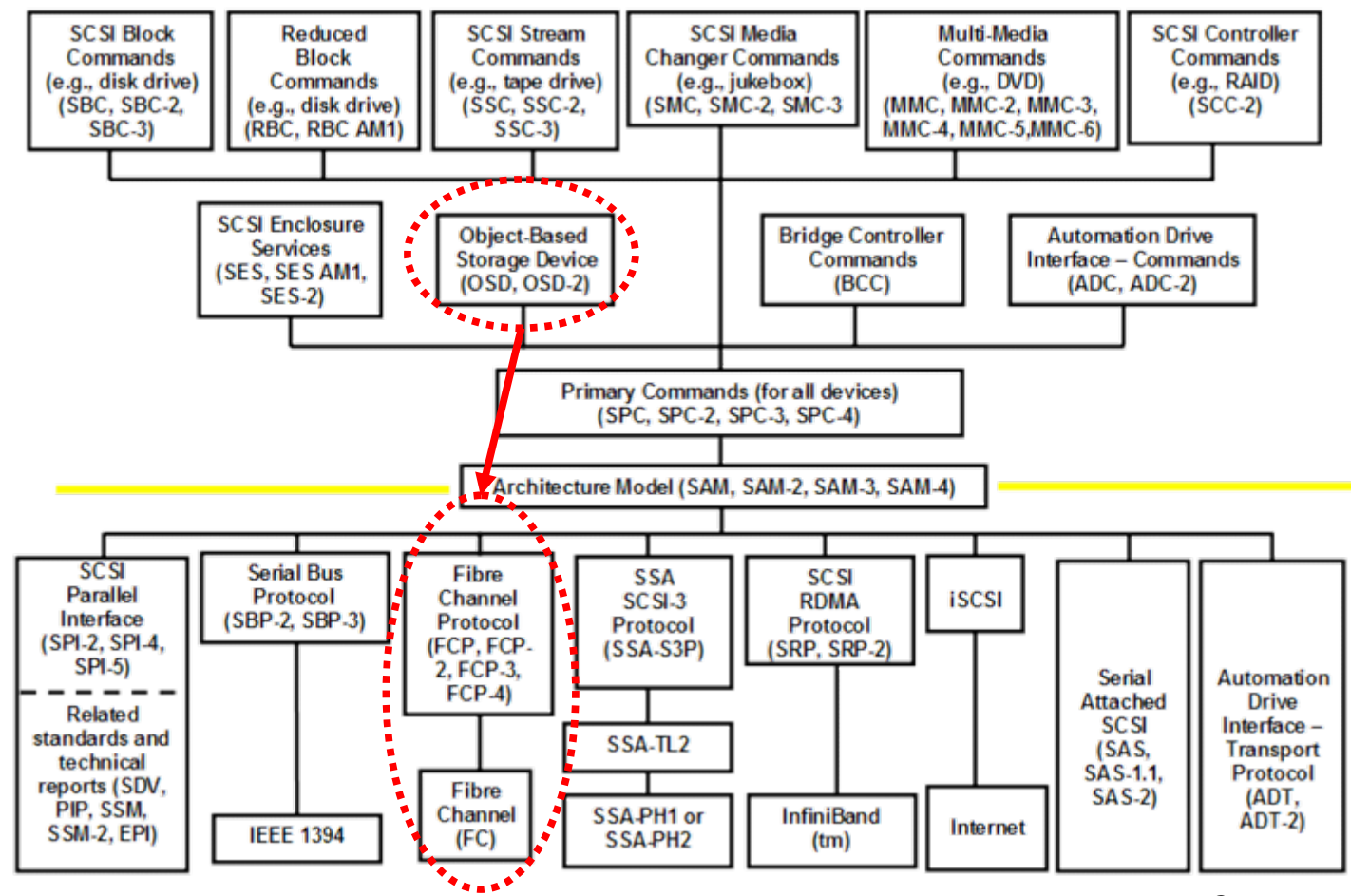
Source: www.t10.org

SCSI OSD/OSD-2



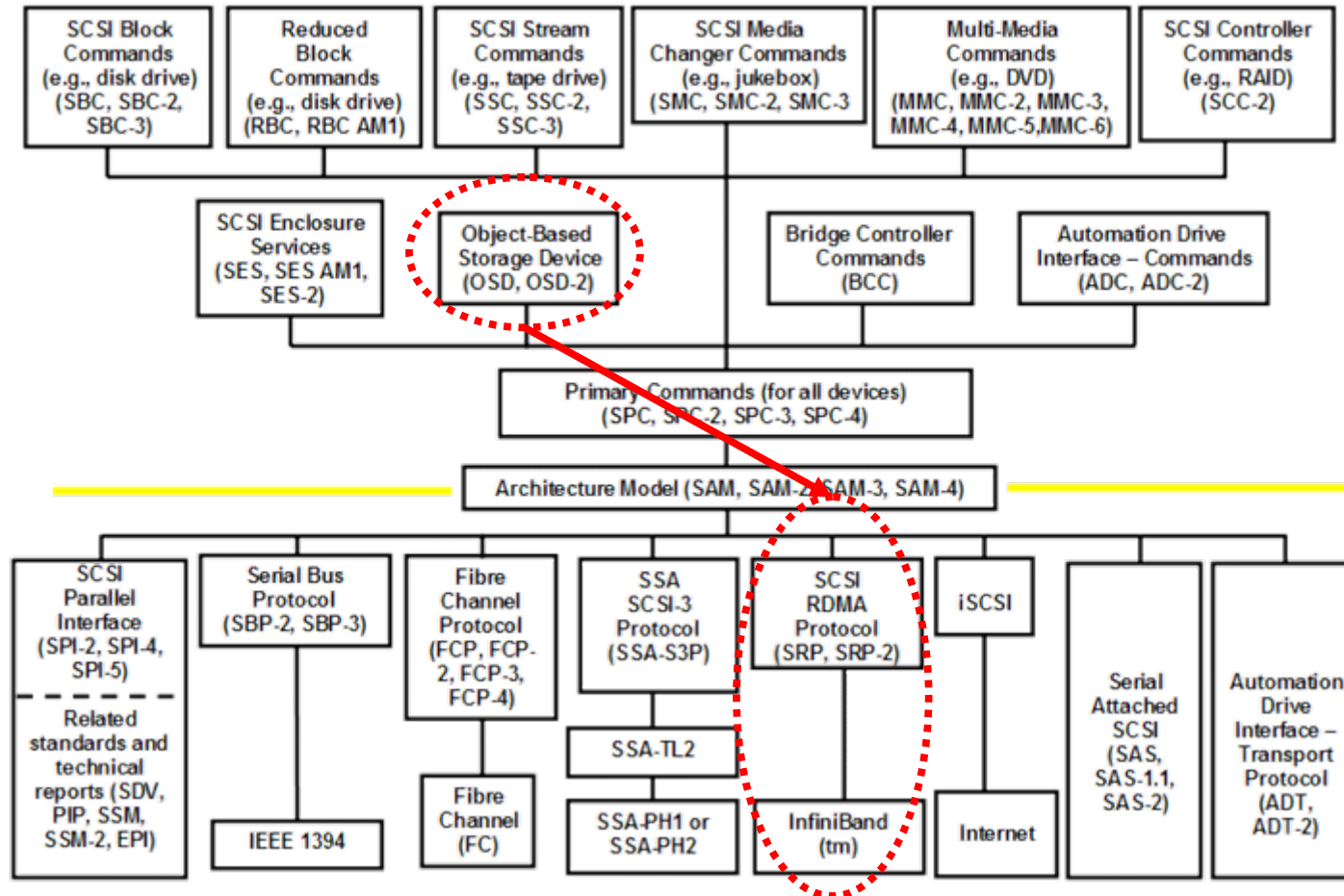
Source: www.t10.org

SCSI OSD/OSD-2 over Fibre Channel SNIA



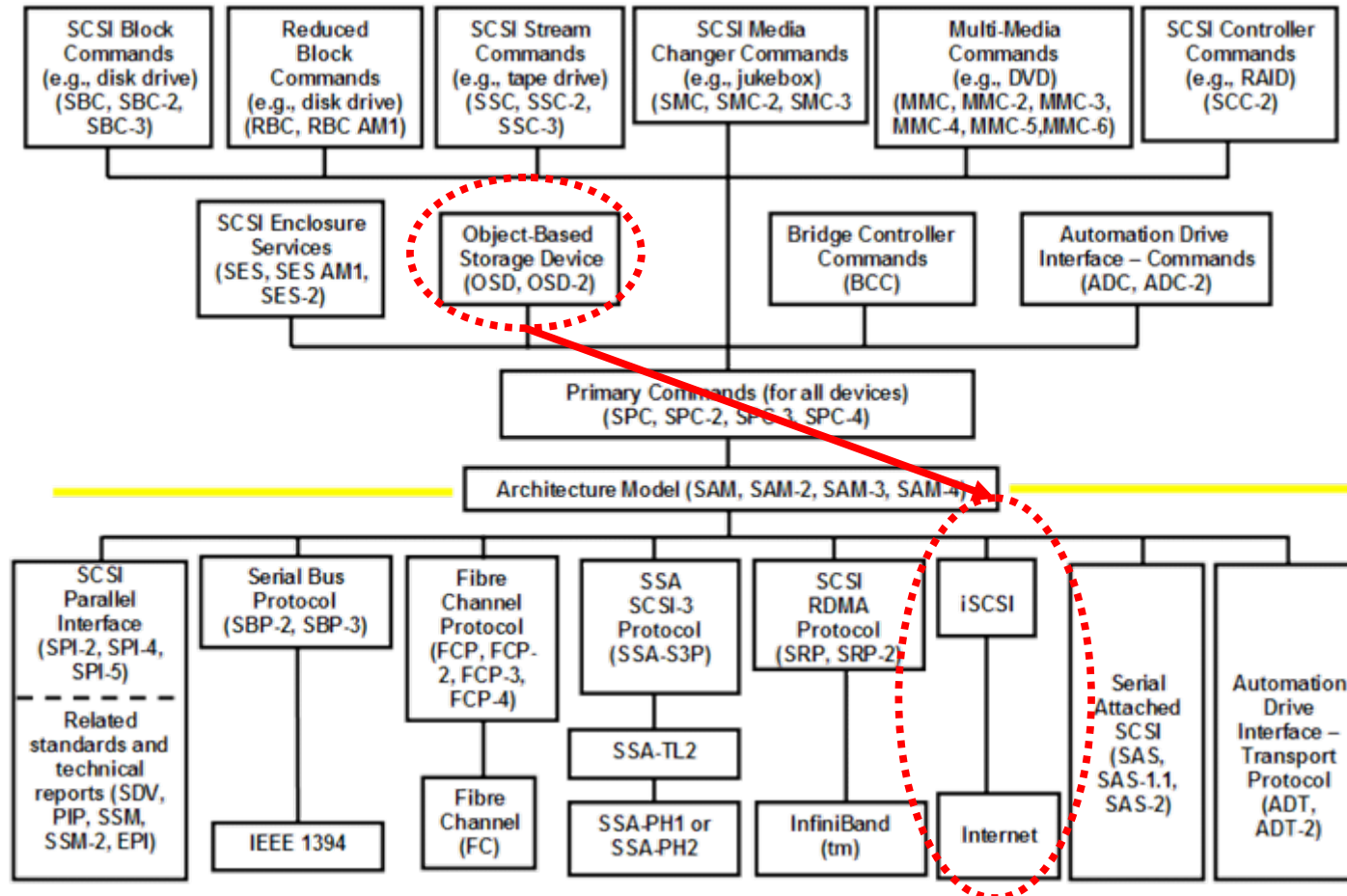
Source: www.t10.org

SCSI OSD/OSD-2 over InfiniBand



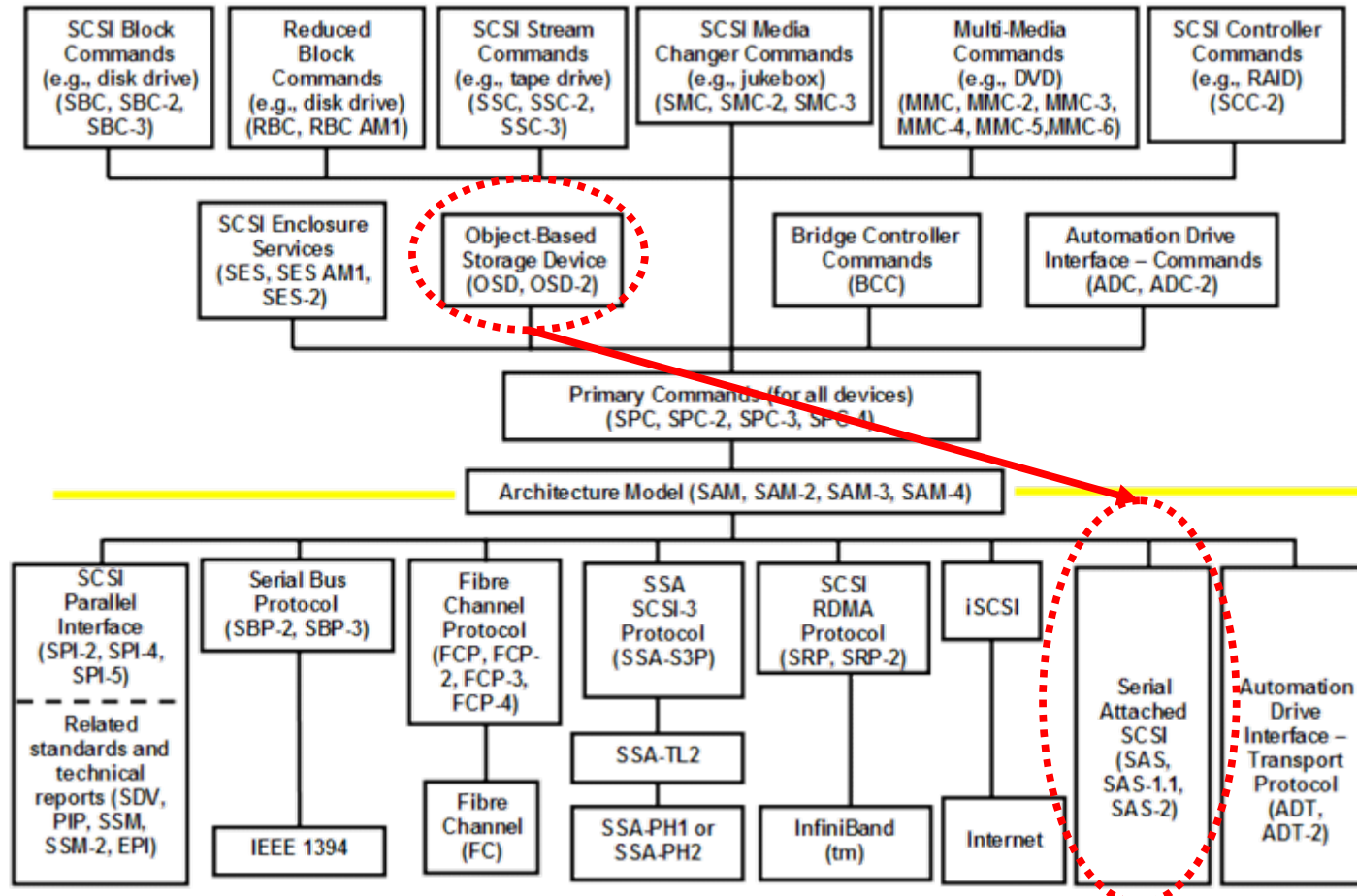
Source: www.t10.org

SCSI OSD/OSD-2 over Ethernet



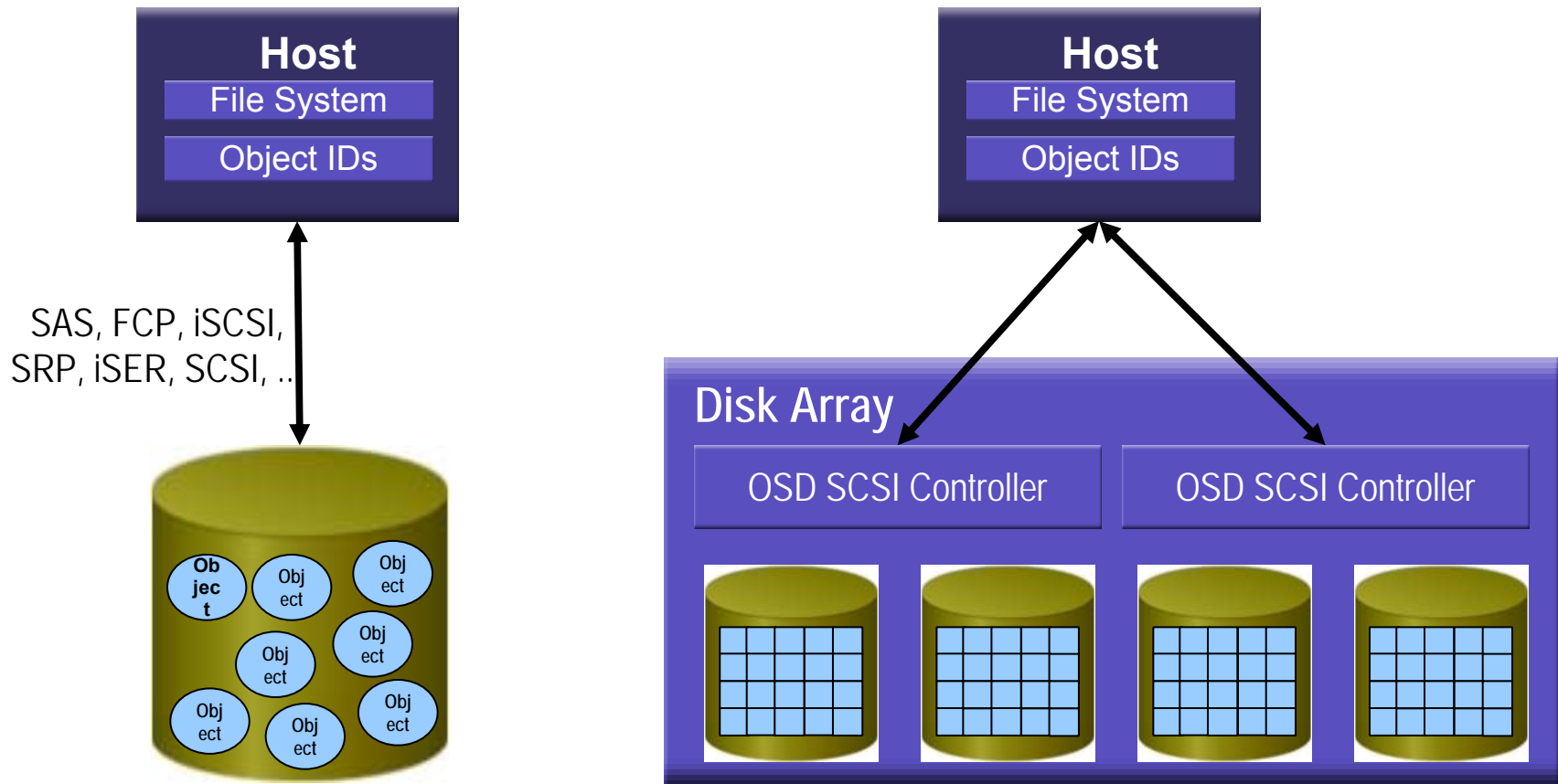
Source: www.t10.org

SCSI OSD/OSD-2 over SAS



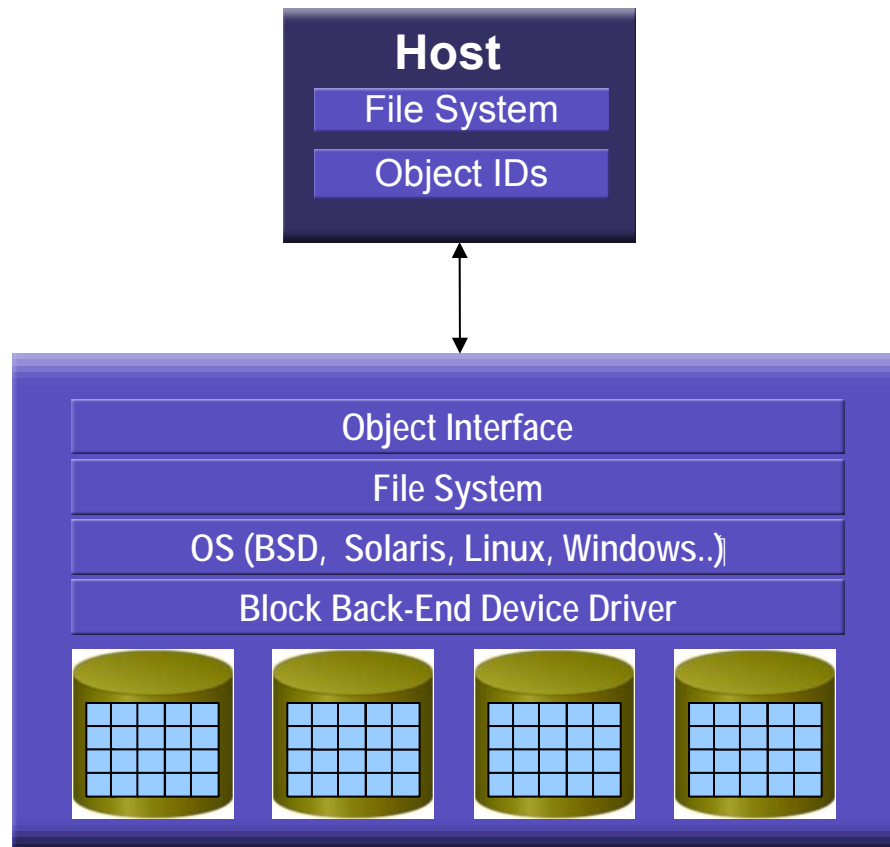
Source: www.t10.org

ANSI T10 OSD SCSI Targets



- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - > Object Storage Server (OSS)
 - > Content Addressable Storage (CAS)
 - > Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

Object Storage Server - OSS

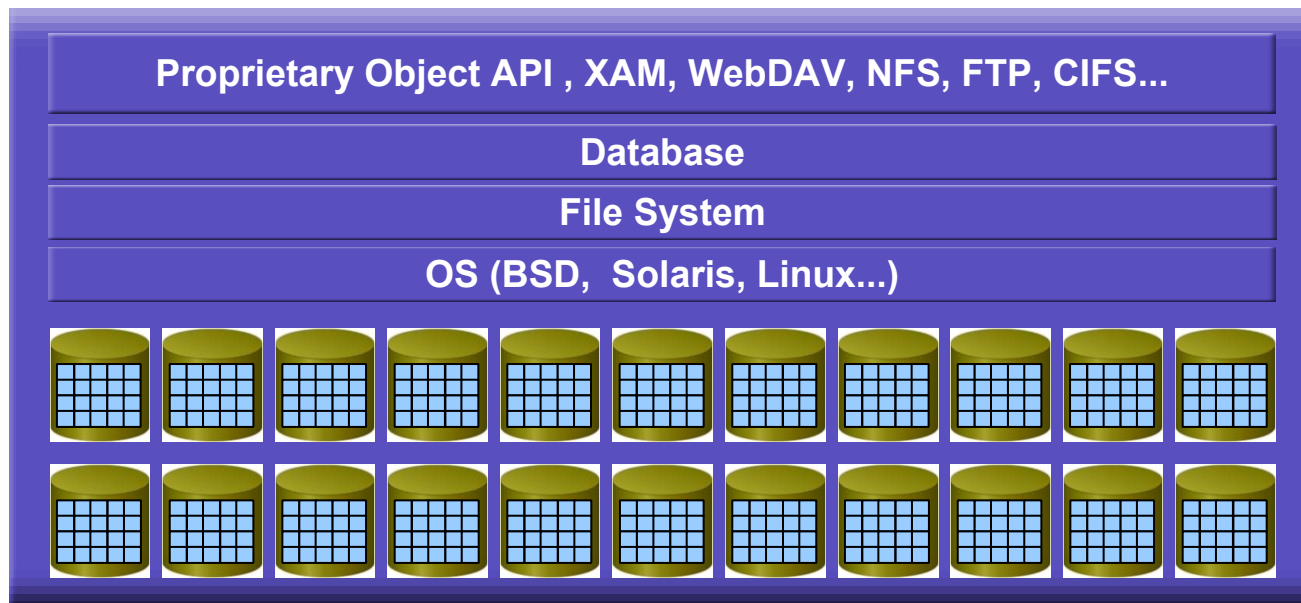


OSS could be a migration path to provide object technologies to legacy block devices

- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › **Content Addressable Storage (CAS)**
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

Content Addressable Storage - CAS

- OIDs are hash values derived from the objects' content
- Used as digital archive systems for long term fixed content data
- ECM applications used as data injection machines



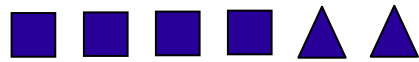


Single Data Image

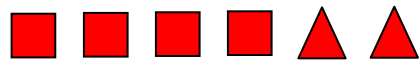
Data Placement

Improved Resilience

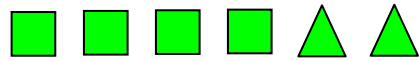
Object 1



Object 2

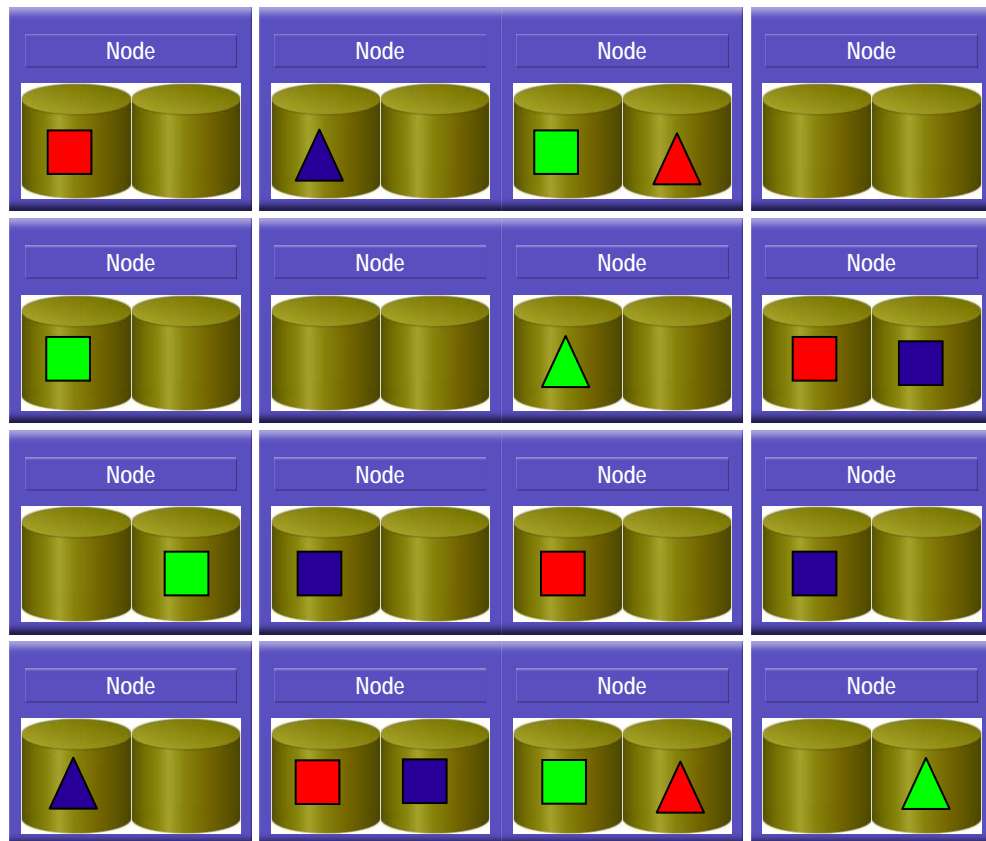


Object 3



□ = Data

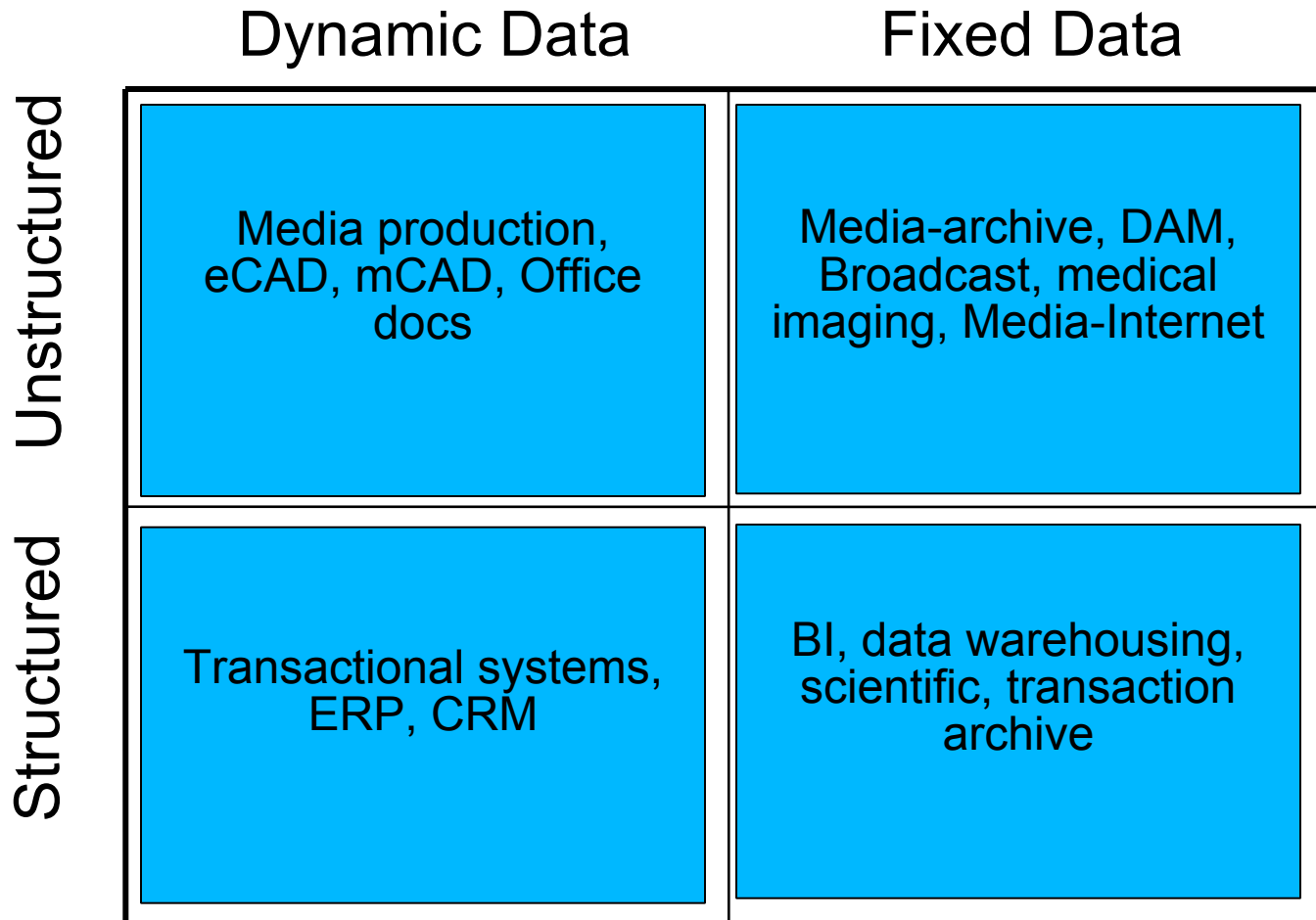
△ = Parity



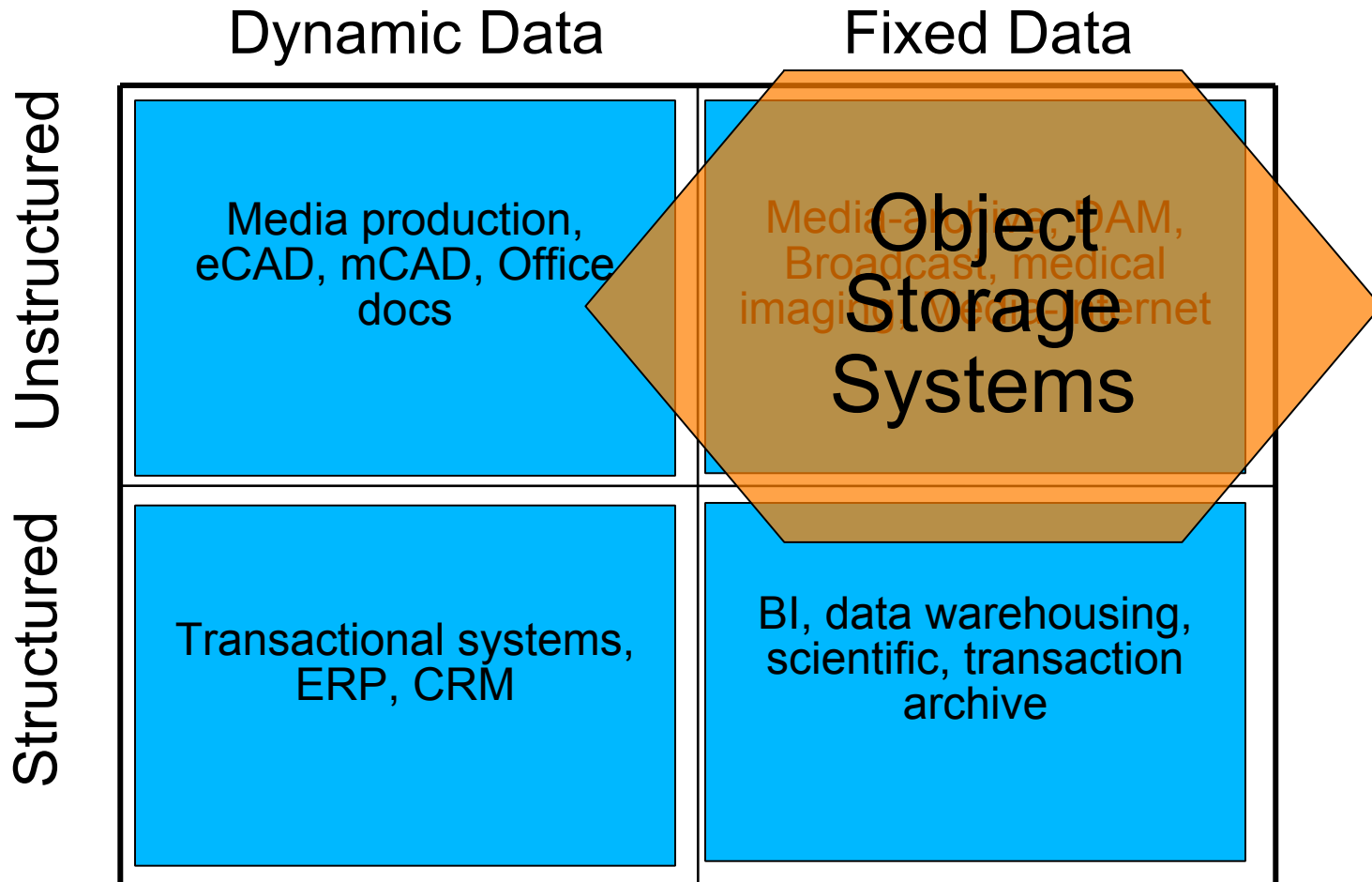
Storage Market Segmentation

	Dynamic Data	Fixed Data
Unstructured		
Structured		

Storage Market Segmentation

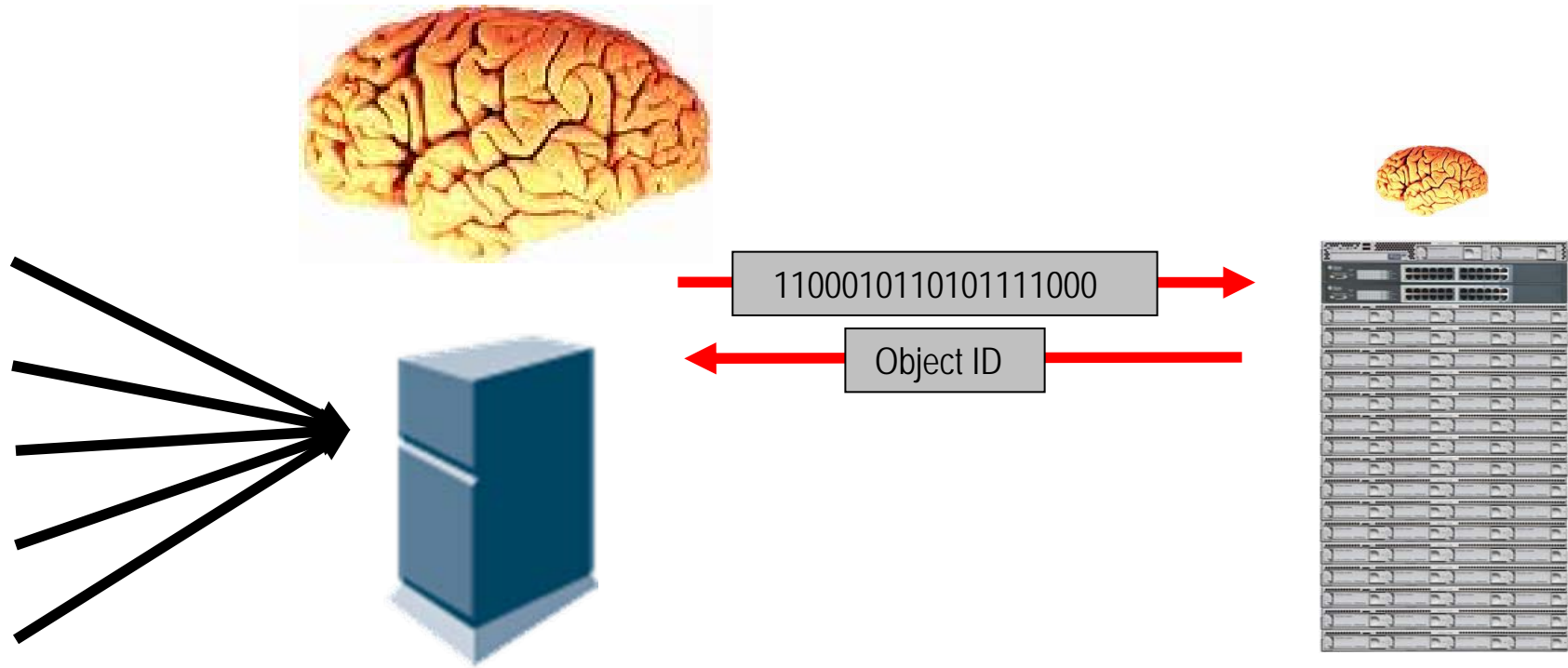


Storage Market Segmentation



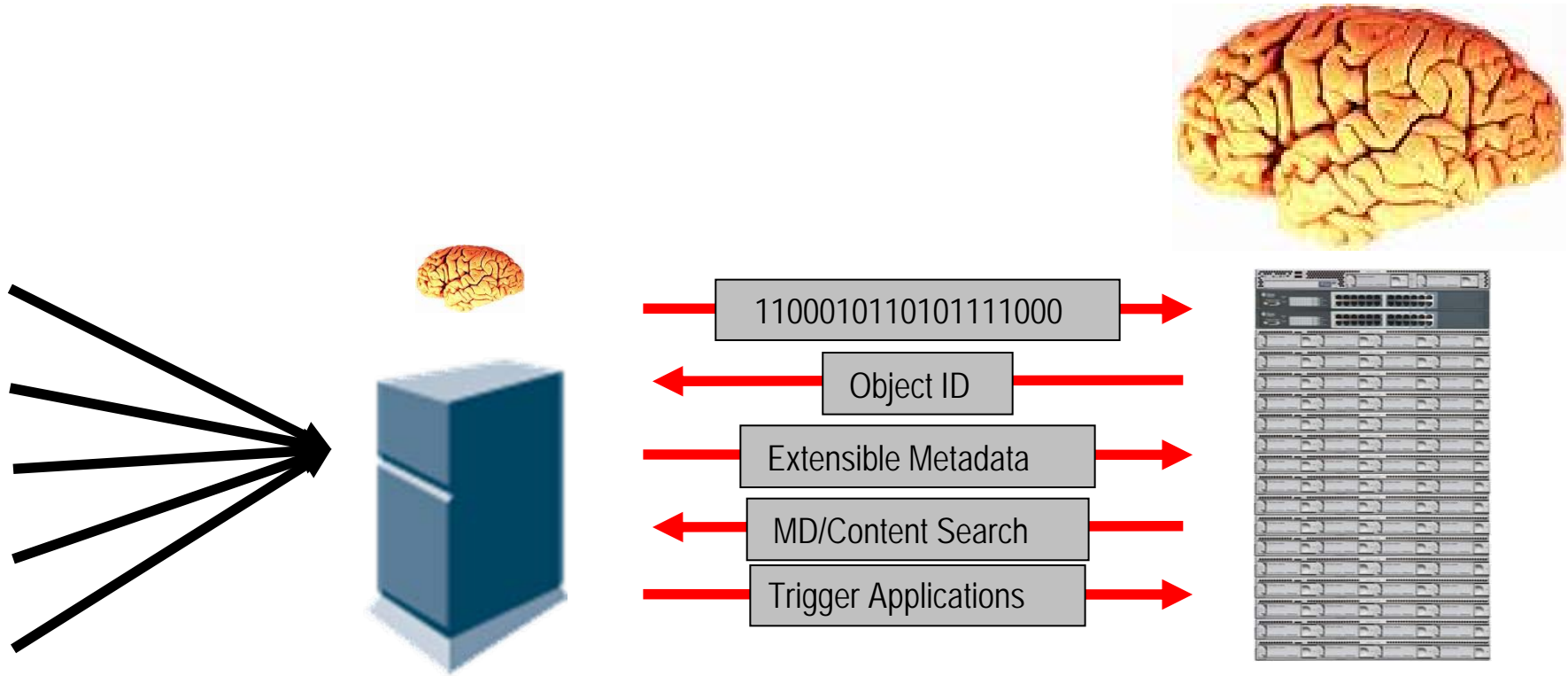
- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › **Content Aware Storage (CAS)**
- Intelligent Storage Nodes (ISN)

CAS: “Content Addressable Storage” SNIA



Enterprise Content Management
Injection Engine

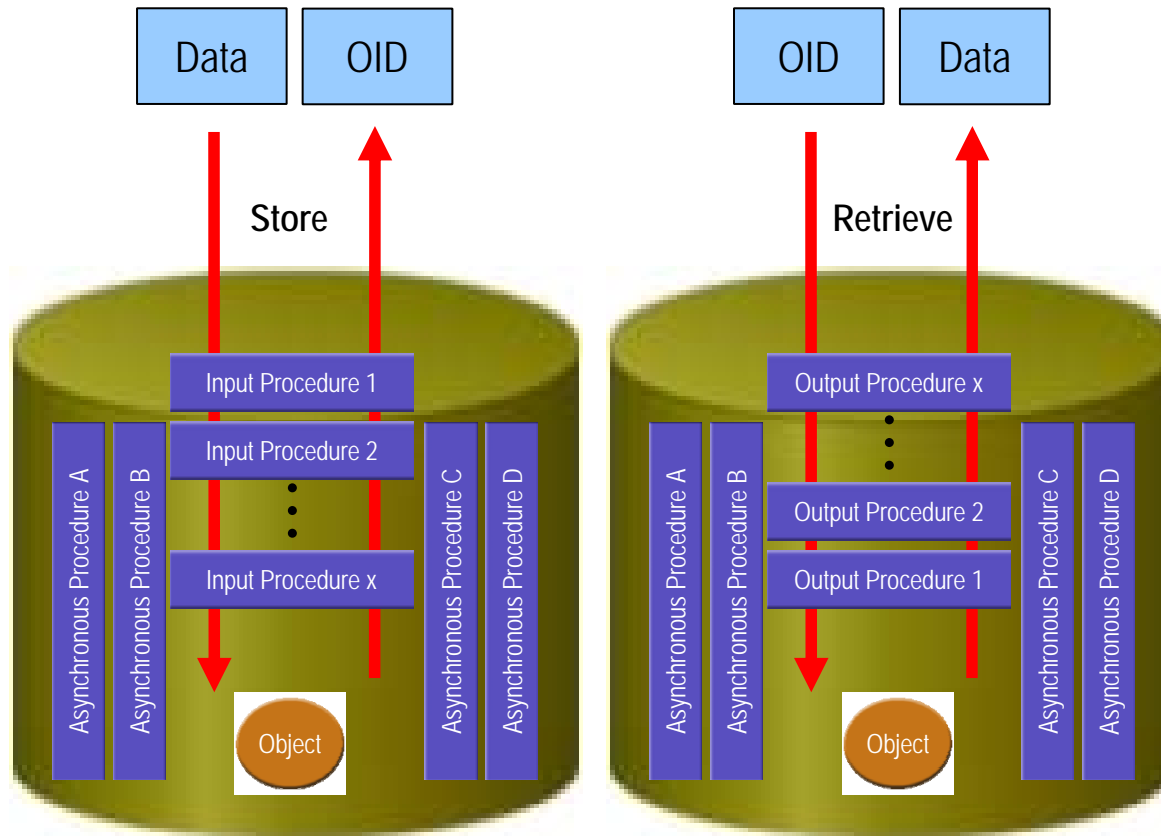
CAS: “Content Aware Storage”



Enterprise Content Management
Injection Engine

Content Aware Storage








Flexibility



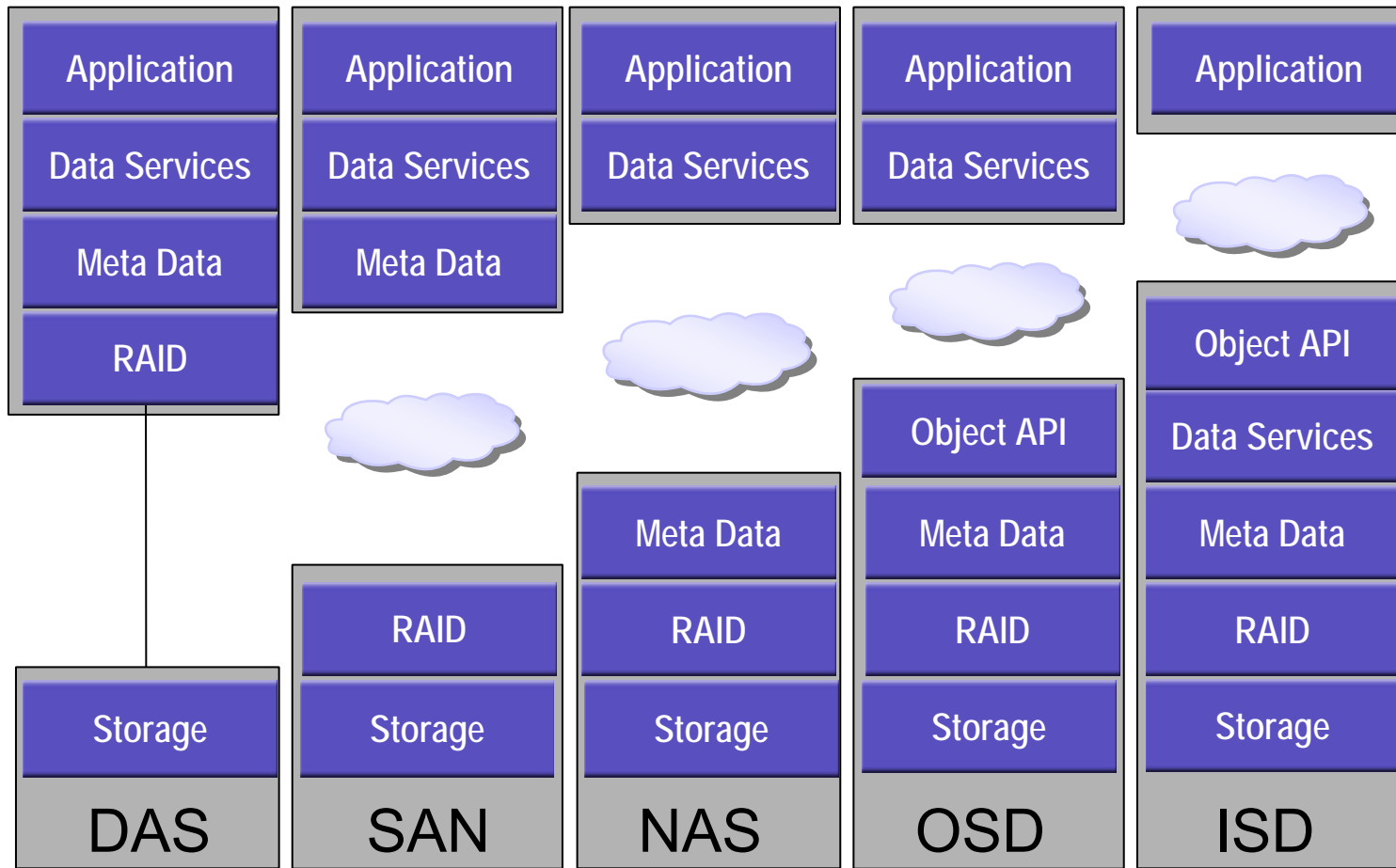
- User-developed trigger apps
- Synchronous:
 - ◆ modify the behavior of store/retrieve/query/delete
 - ◆ e.g. transcode, downsample, filter, watermark, extract metadata from file, headers, encrypt, audit log...
- Asynchronous:
 - ◆ process data at rest
 - ◆ e.g. capacity optimization, scrubbing, migration, sanity check...

- Block-Based Data Access
- File-Based Data Access
- Object-Based Data Access
 - ◆ Object-Based Storage Devices (OSD)
 - ◆ Object Storage Systems
 - › Object Storage Server (OSS)
 - › Content Addressable Storage (CAS)
 - › Content Aware Storage (CAS)
- Intelligent Storage Nodes (ISN)

The New Challenges

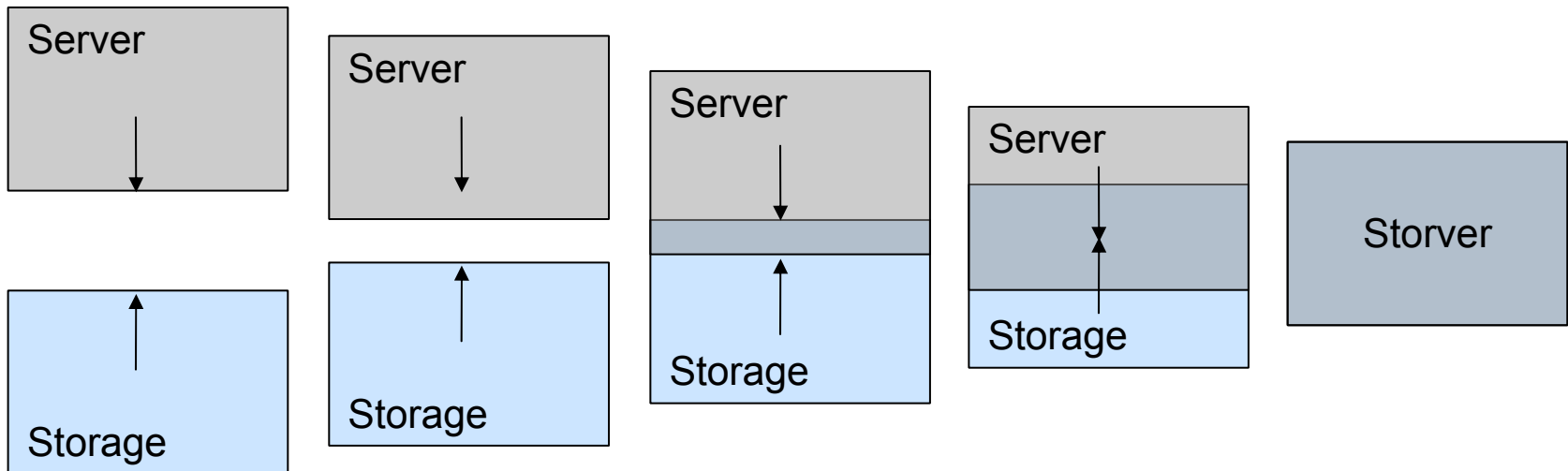
- I/O Performance
- Availability
- Reliability
- Easiness to store, organize, retrieve and dispose 
- Complex data operations 
 - ◆ aggregate, join, view, sort, convert, encrypt... 
- Enhanced search operations 
- Flexibility to present data 
- Customized storage behavior 
- Reduced administration costs 

The Evolution of Data Storage:



The Vertical Consolidation

- Storage and server
- Migration of data processing applications
- No I/O is best I/O



Further Reference

- http://www.snia.org/tech_activities/workgroups
- <http://www.snia.org/apps/org/workgroup/osd/>
- <http://www.snia.org/apps/org/workgroup/fcastwg/>
- <http://www.snia-dmf.org/>
- <http://www.t10.org/ftp/t10/drafts/osd>
- <http://www.t10.org/ftp/t10/drafts/osd2>
- <http://ietf.org/html.charters/webdav-charter.html>
- <http://ietf.org/html.charters/nfsv4-charter.html>
- <http://www.snia.org/education/tutorials/>

- Please send any questions or comments on this presentation to SNIA:
trackstorage@snia.org

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

SNIA Education Committee

**Christian Bandulet, Sun Microsystems
Rob Peglar**