Storage Tutorial For Content Creation and Distribution

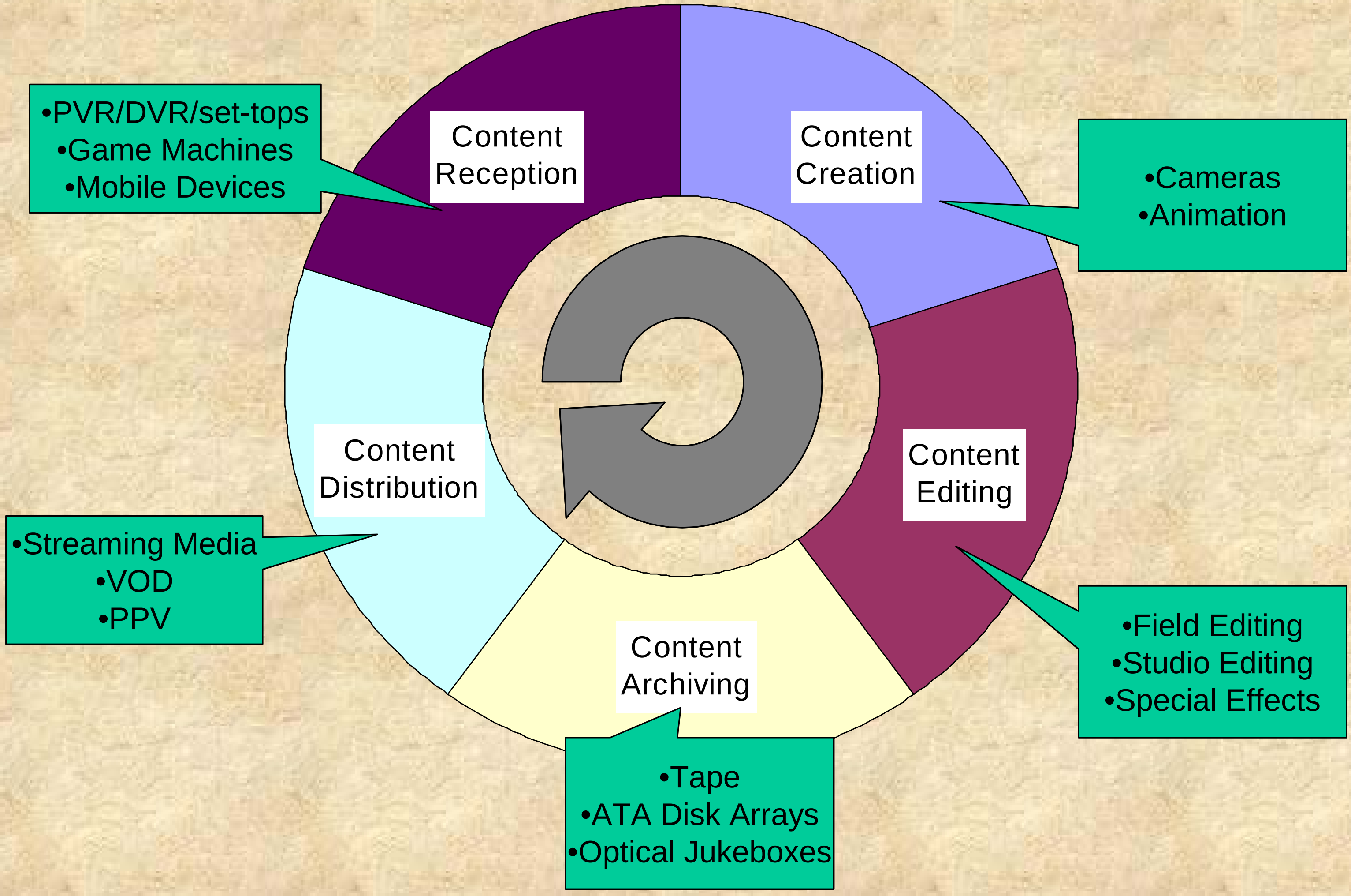
Tom Coughlin
Coughlin Associates
www.tomcoughlin.com

Outline

- Content Value Chain
- Storage Demand for Entertainment Applications
- Storage Devices
- Storage Systems
- Digital Storage Applications for Entertainment Media
- Conclusions

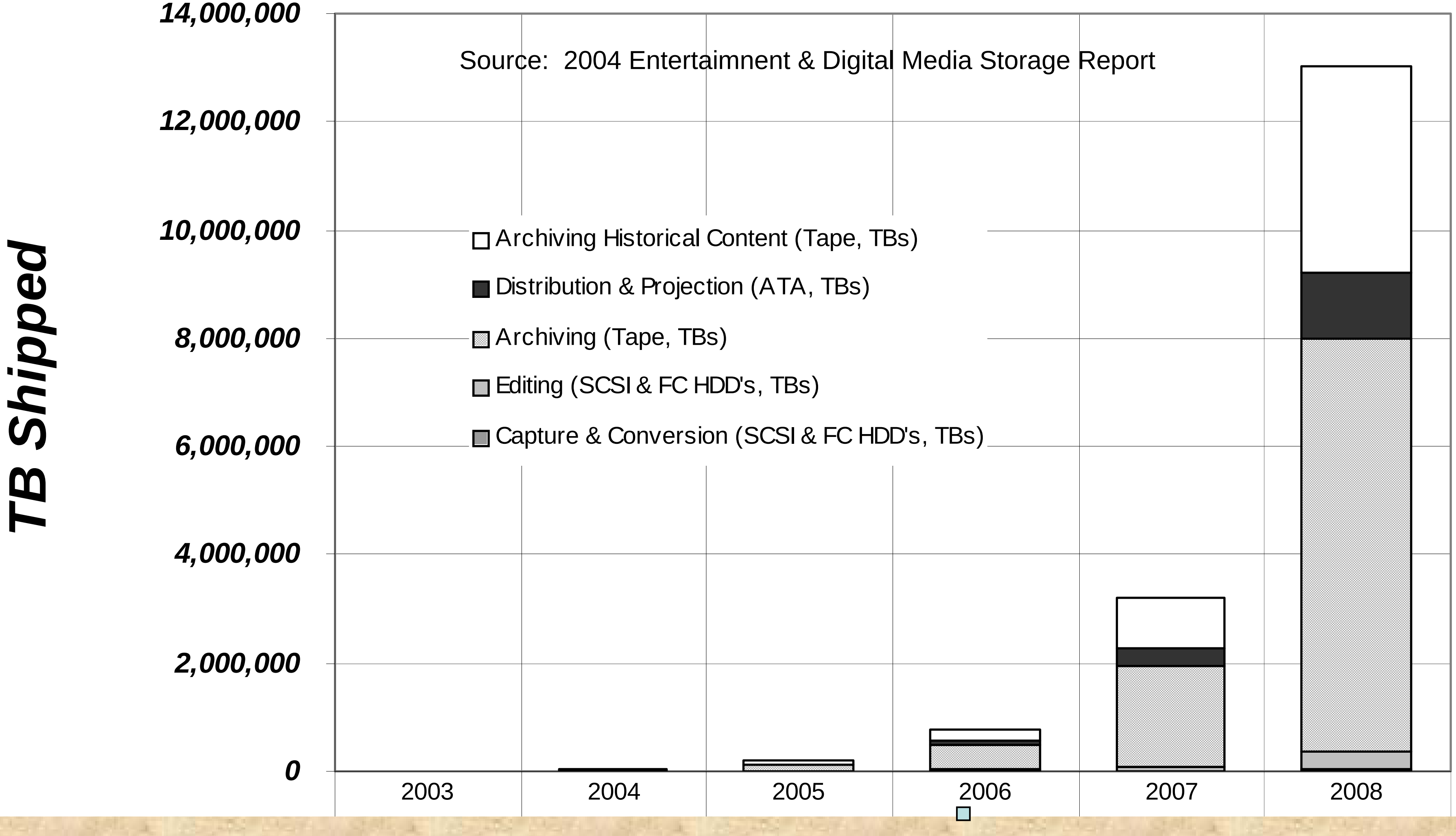
STORAGE MAKES ME HAPPPY!

Digital Content Value Chain



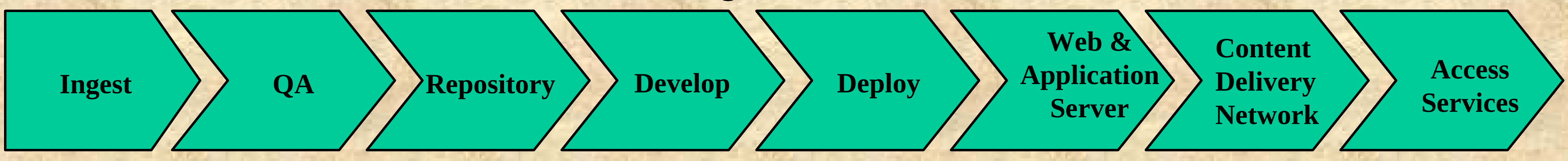
Content Creation, Distribution, & Archive Market

(SCSI/FC HDD, ATA HDD, & Tape)



Activities in Content Creation, Editing, Archiving and Distribution

Content Creation　Content Management　Content Distribution



Content Creation

- Acquire Content
- Field Editing
- Ingest
- Digitization
- Capturing Meta-data
- Auto-Logging
- Proxy Creation

Content Management

- Library and Archive Storage
- Search Tools
- Edit and Re-Purpose material
- Metadata Management
- Rights Management
- Distribution Scheduling
- Content Import/Export
- Web access

Content Distribution

- Various Delivery Networks
- Point to Point & Point to Multipoint
- Push and Pull modes
- Real-time and Scheduled
- Dedicated Infrastructure
- Format/Destination Selection
- Caching and Edge Storage

Uncompressed Video Production Storage Needs (Raw DPX 10 bit log files).

Resolution	Frames/sec	MB/second	Capacity/minute (GB)
SD	1.7	38.4	2.3
1K	3.2	76.8	4.6
HD	8.2	197	11.8
2K	12.5	300	18.0
4K	50	1.2	72.0

Digital Production and Distribution Rules!

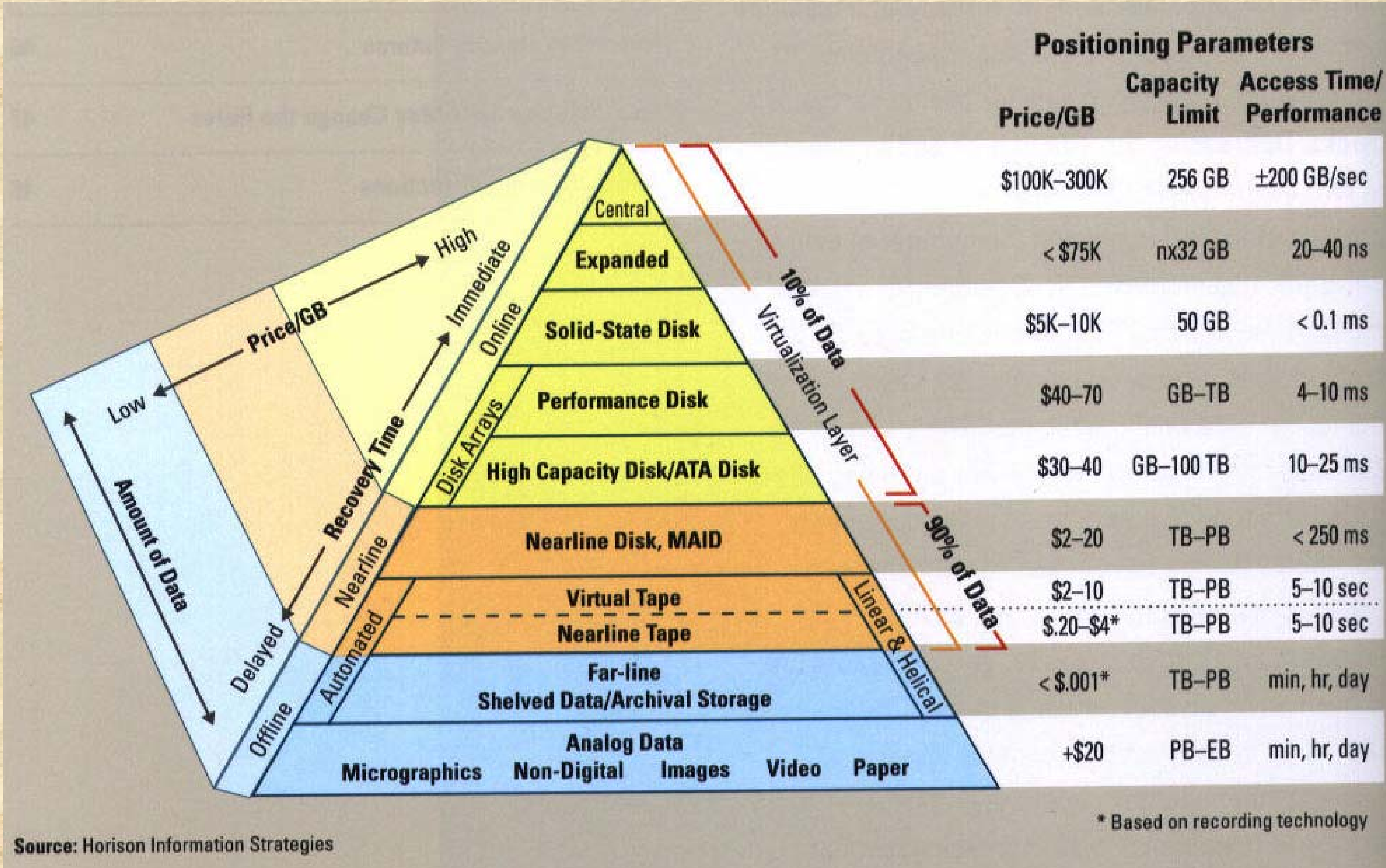
- Save more than a factor of 50 in video capture and editing costs vs. traditional film
- Special effects and editing possible with digital production can’t be matched with older analog techniques
- Save 80% on digital theatre distribution vs. film distribution

Adventures in Archiving

- Demand huge, and growing
- Long term storage formats
- Format obsolescence
- Need for format transfer planning—archiving will not be merely static
- Enormous need for good metadata tagging and data search and access improvements

Storage Devices

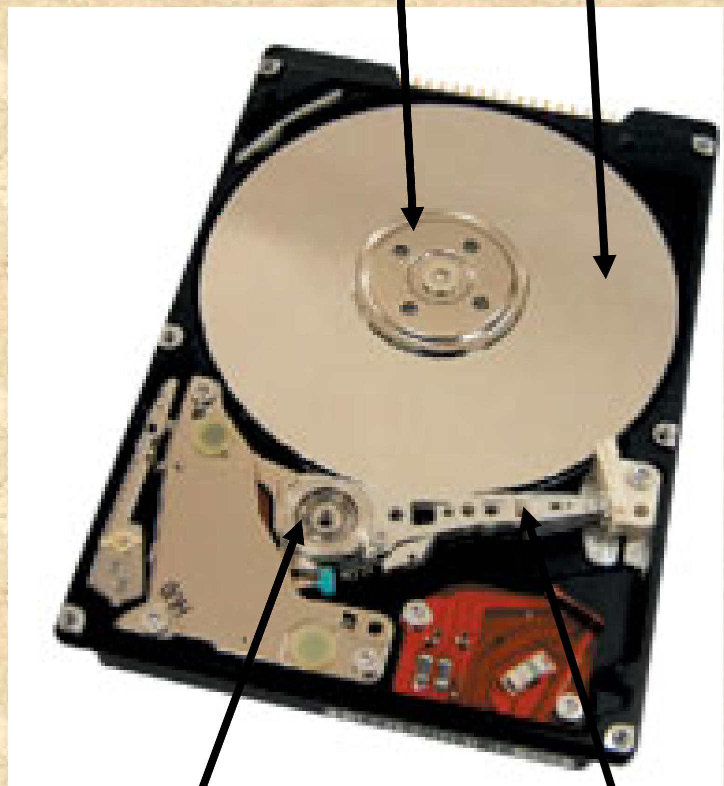
Storage Hierarchy



Magnetic Rigid Disk Drives (HDD)

Spindle Motor

Disk



Head Actuator

Head Suspension

15k RPM FC Drive

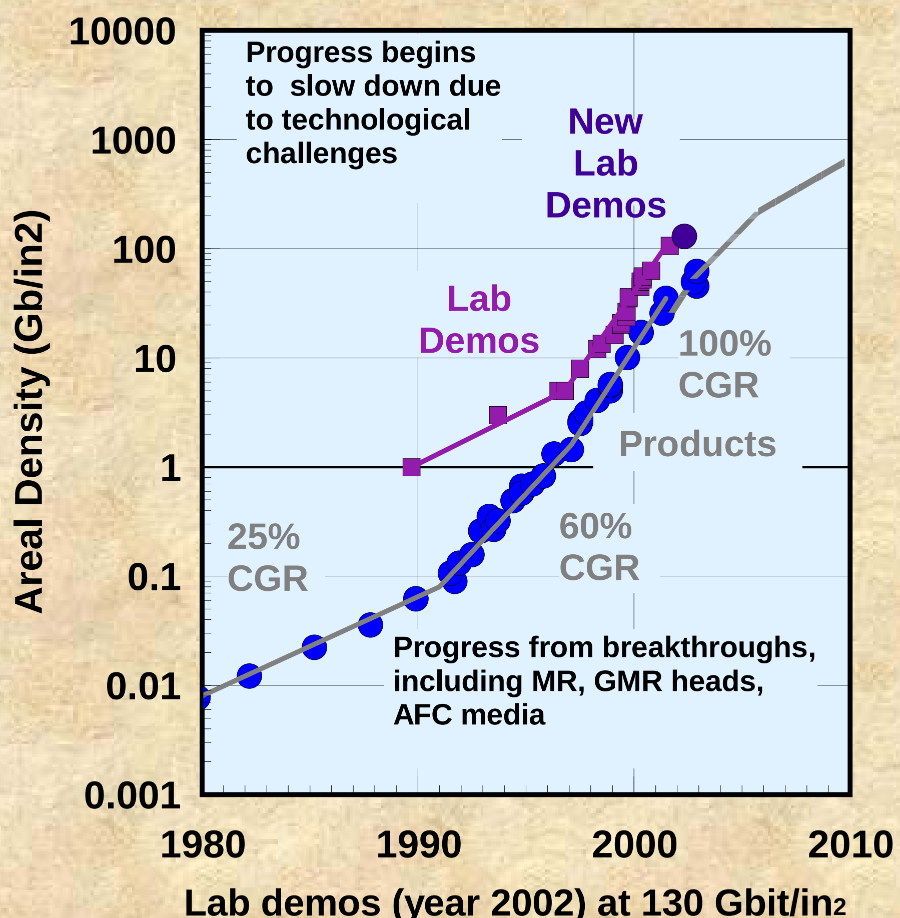
2.5 Inch Mobile Drive

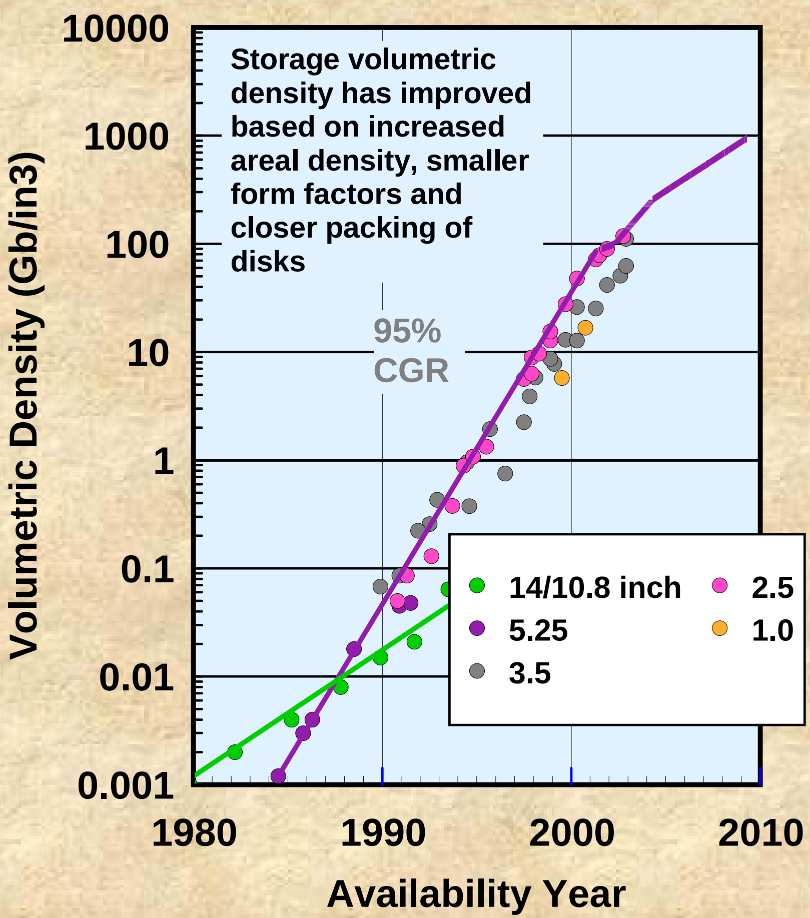
High Capacity ATA Drives
(now up to 400 GB)

Toshiba 0.85 inch HDD

HDD Areal and Volumetric Density Growth

Storage areal density CGR starts to slow from 100% per year near 100 Gbit/in^2. Volumetric density follows at similar rate.





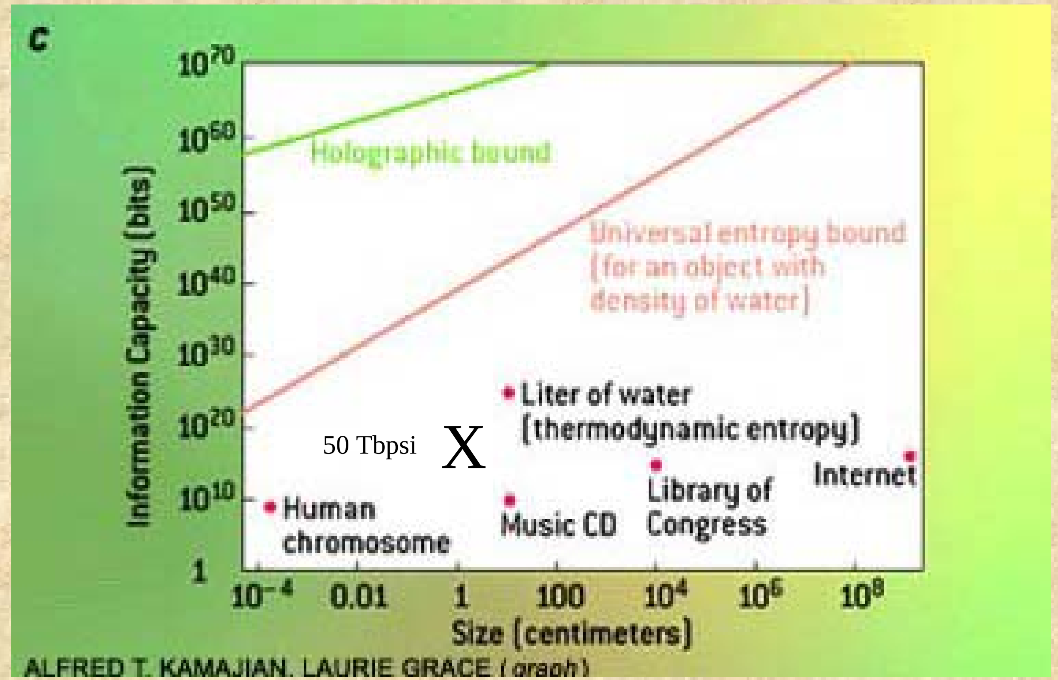
From Clod Barrara, IBM
March 2004

SHIPPING PRODUCT DISK CAPACITY PROJECTIONS

Year	95mm Mainstream Capacity Per Platter
2002	40
2003	80
2004	120
2005	180
2006	270

By 2006 we could have four disk 3.5-inch disk SATA drives with storage capacities of over 1 TB.

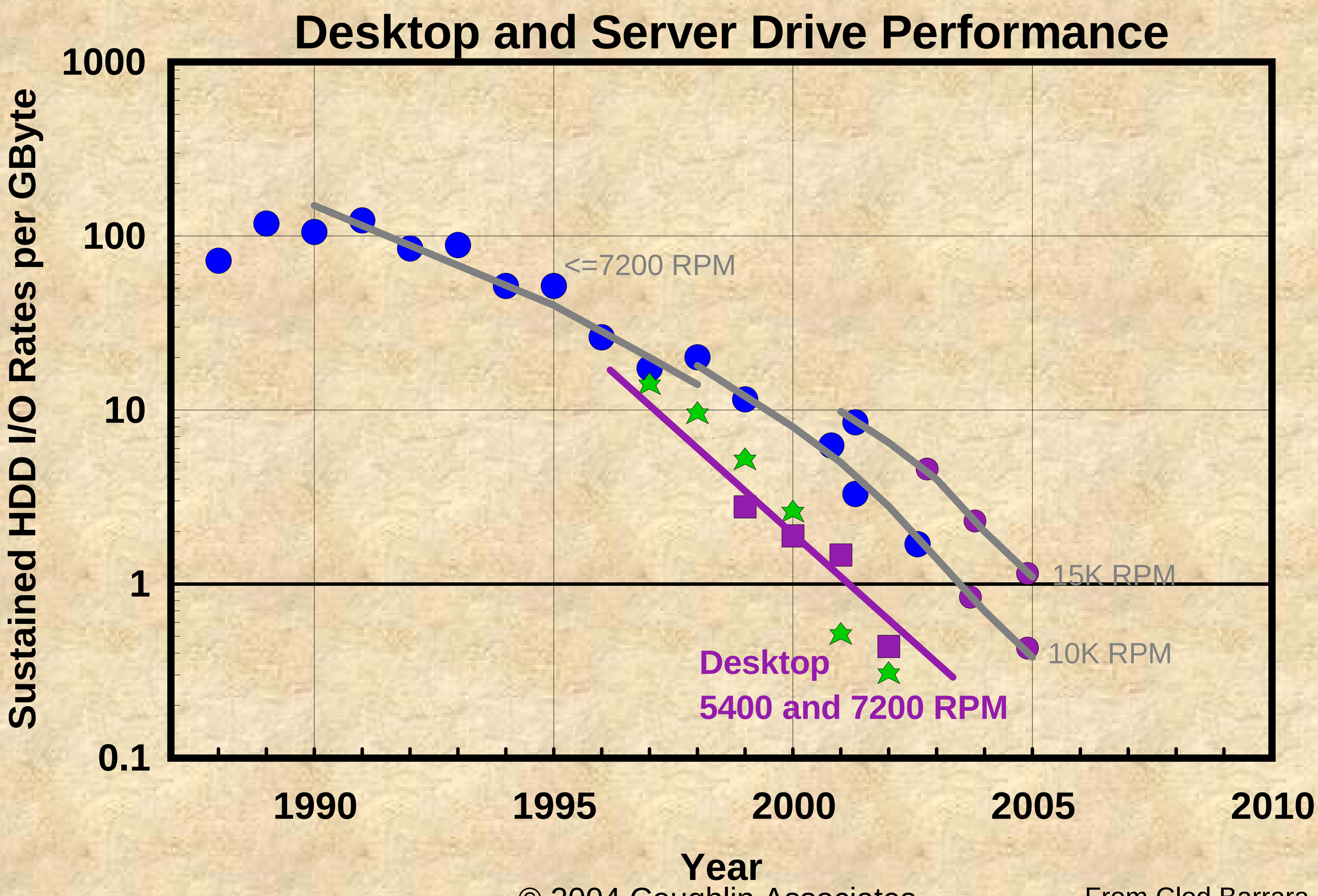
The Universe Still Beats Us by Far in Information Capacity



Source: Information in the Holographic Universe, August 2003 Scientific American

- The holographic and universal information bounds are far beyond the data storage capacities of any current technology!
- Magnetic recording technology may allow up to 50 Tbpsi (50 X 10^{18} bpsi)

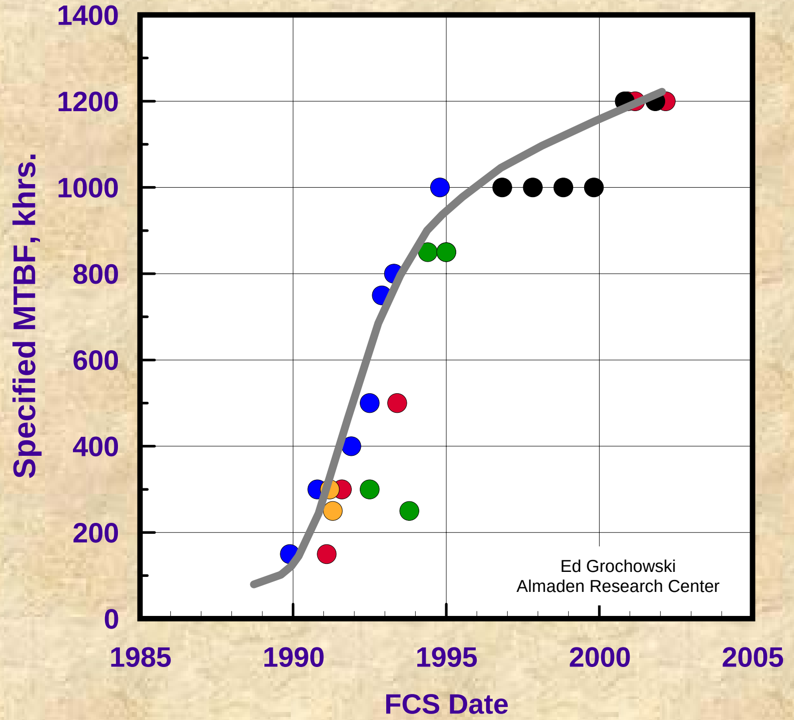
HDD Access Density



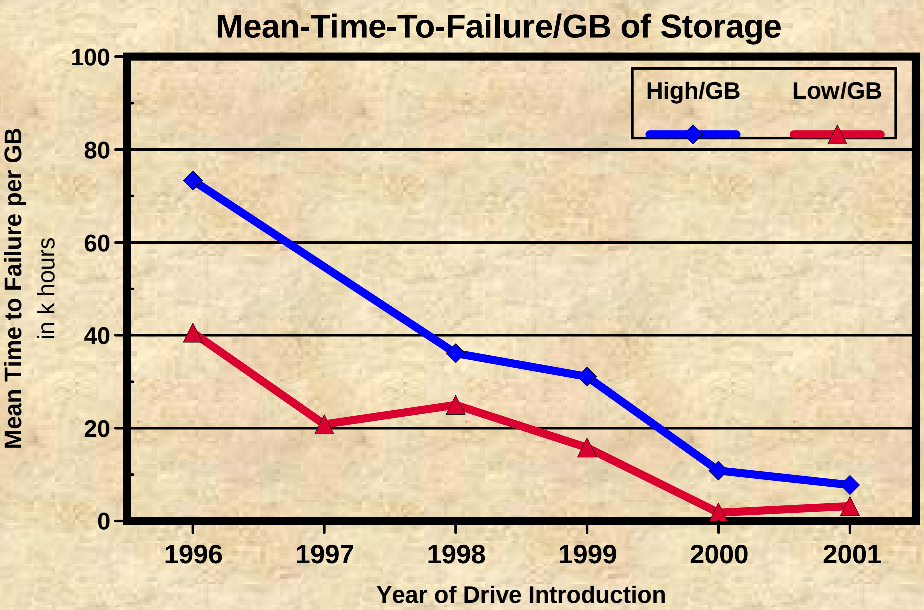
From Clod Barrara, IBM
March 2004

HDD Reliability Trends

- MTBF/GB falling – drive rebuild times growing
- Multi-parity RAID a required aggregation technology



HDD MTBF Manufacturer Specifications



From Clod Barrara, IBM
March 2004

Popular Digital Tape Formats
(All ½ inch tape cartridge technologies)

SAIT

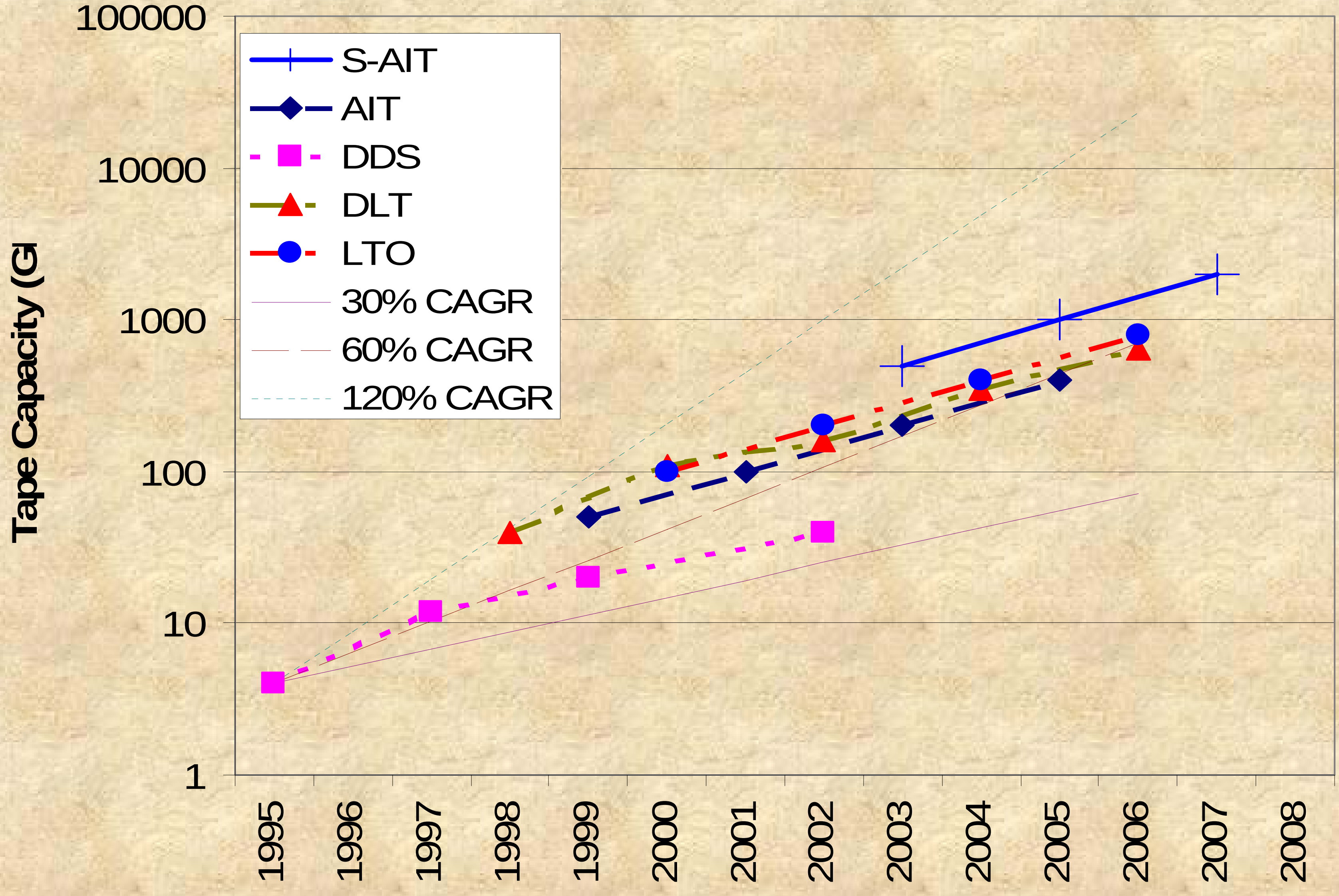
LTO

S-DLT

Tape with Tape Drive

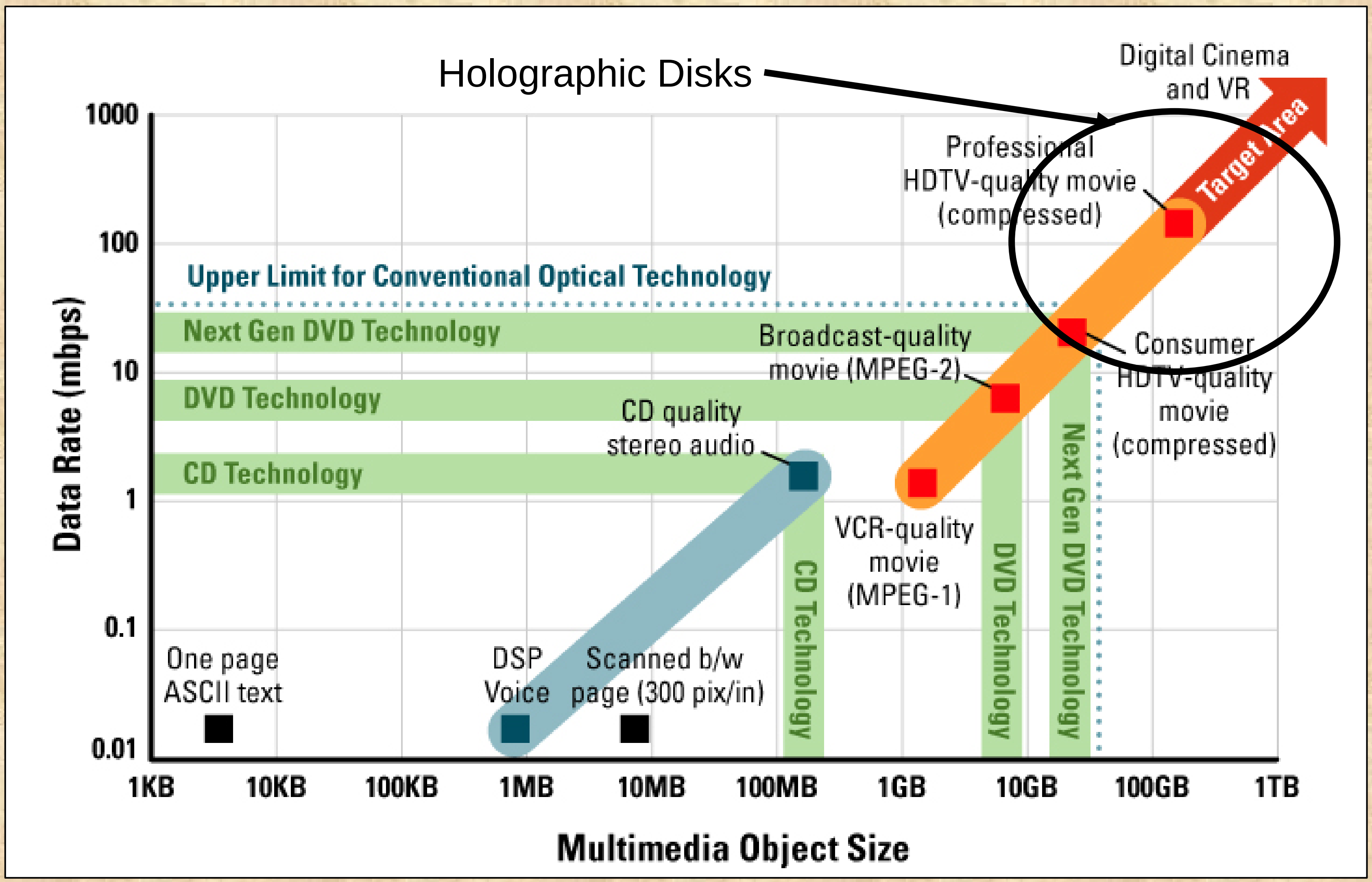
- Tape is still digital archive media of choice
- Tape data access is on the order of minutes vs. milliseconds or seconds for disk
- Tape media costs have been somewhat underwritten by VCR tape production, implications for future of tape costs
- ½ inch tape capacities of up to 10 TB projected

Active tape format CAGRs are about 40%. Disk Drive CAGRs are expected to be ~60%



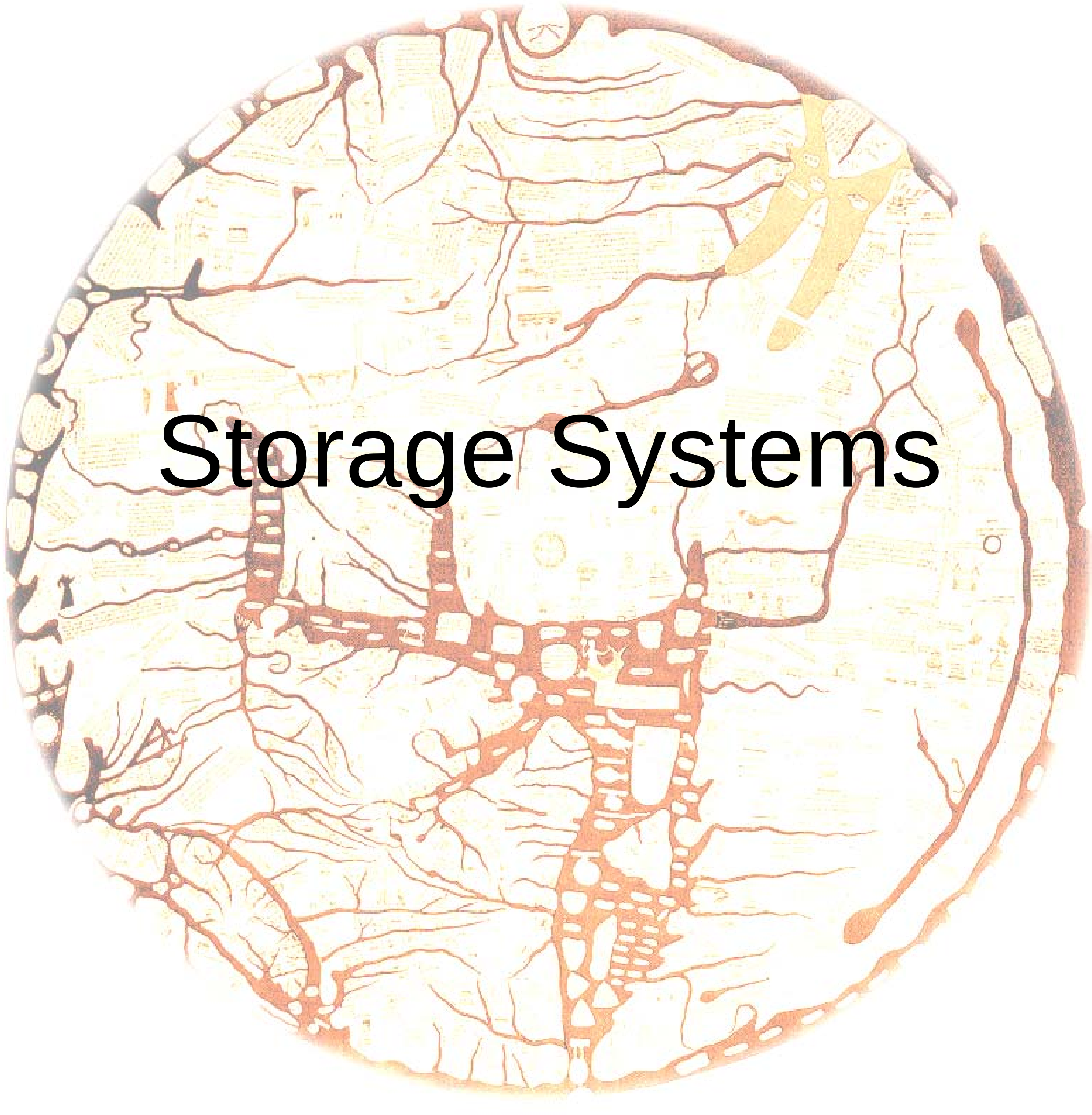
Blue Ray Optical Disks and Drive

MultiMedia Object Size/Bandwidth

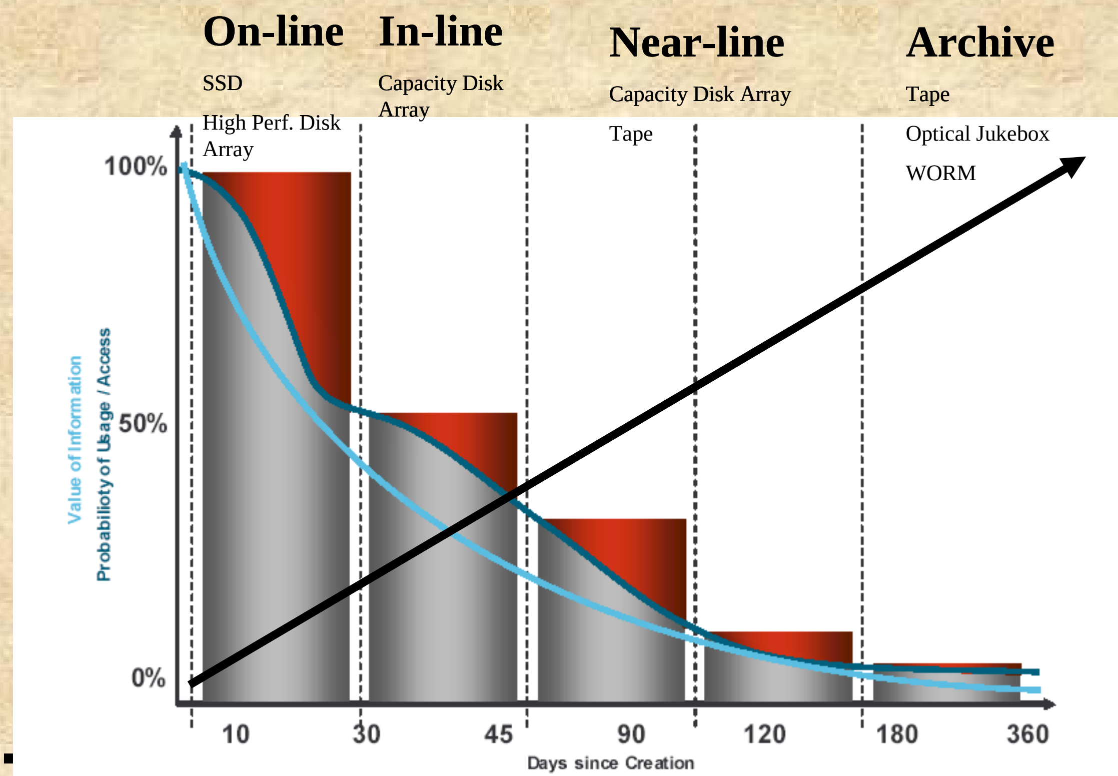


Source: Telcordia 3/03

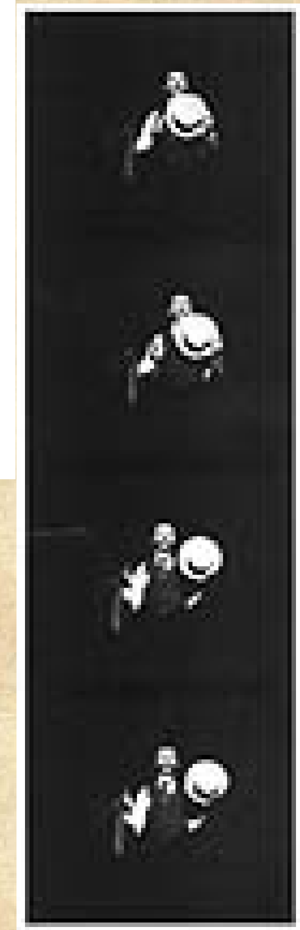
Storage Systems



Digital Content Lifecycle in Production and Distribution.

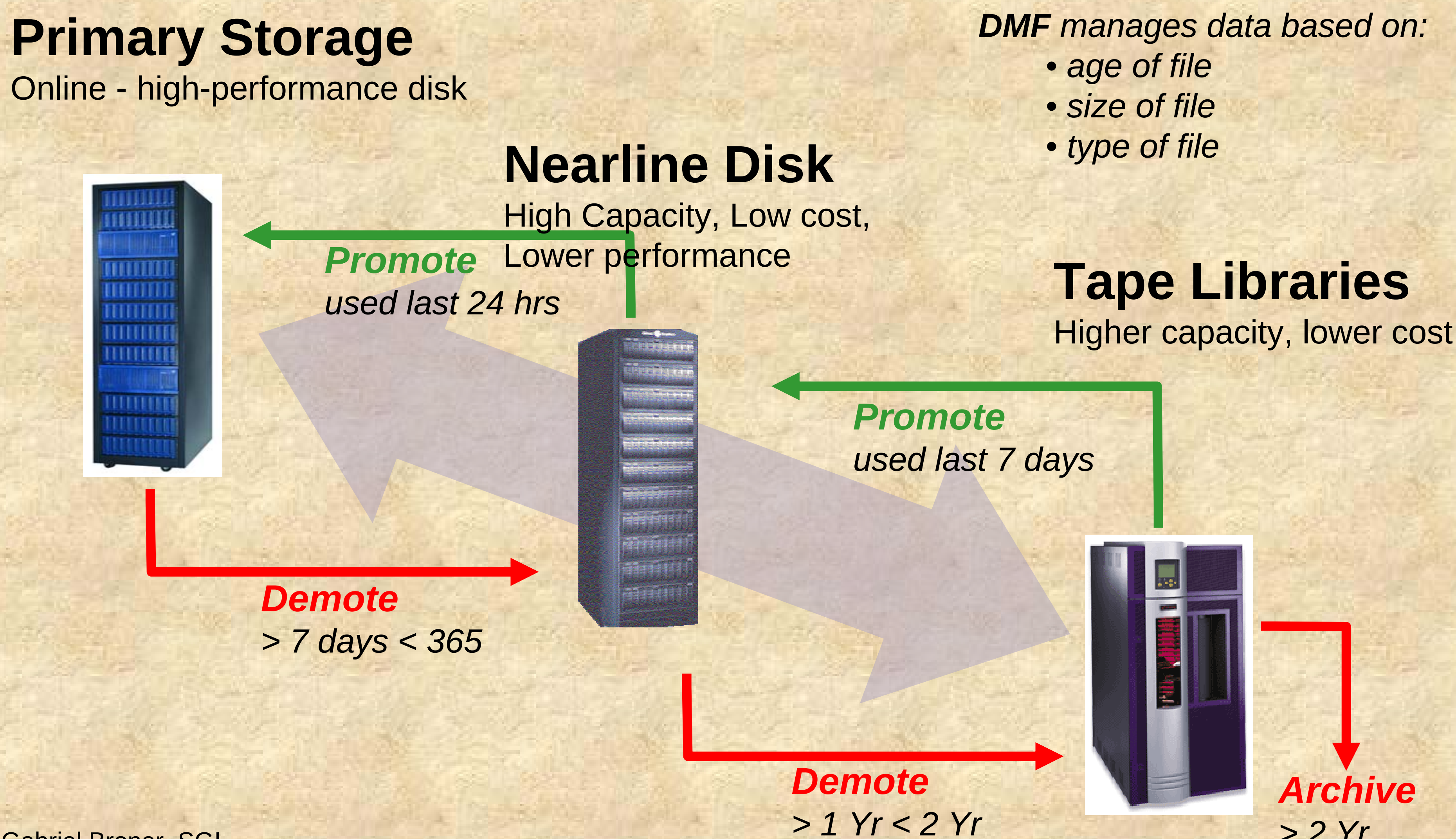


Storage Requirements



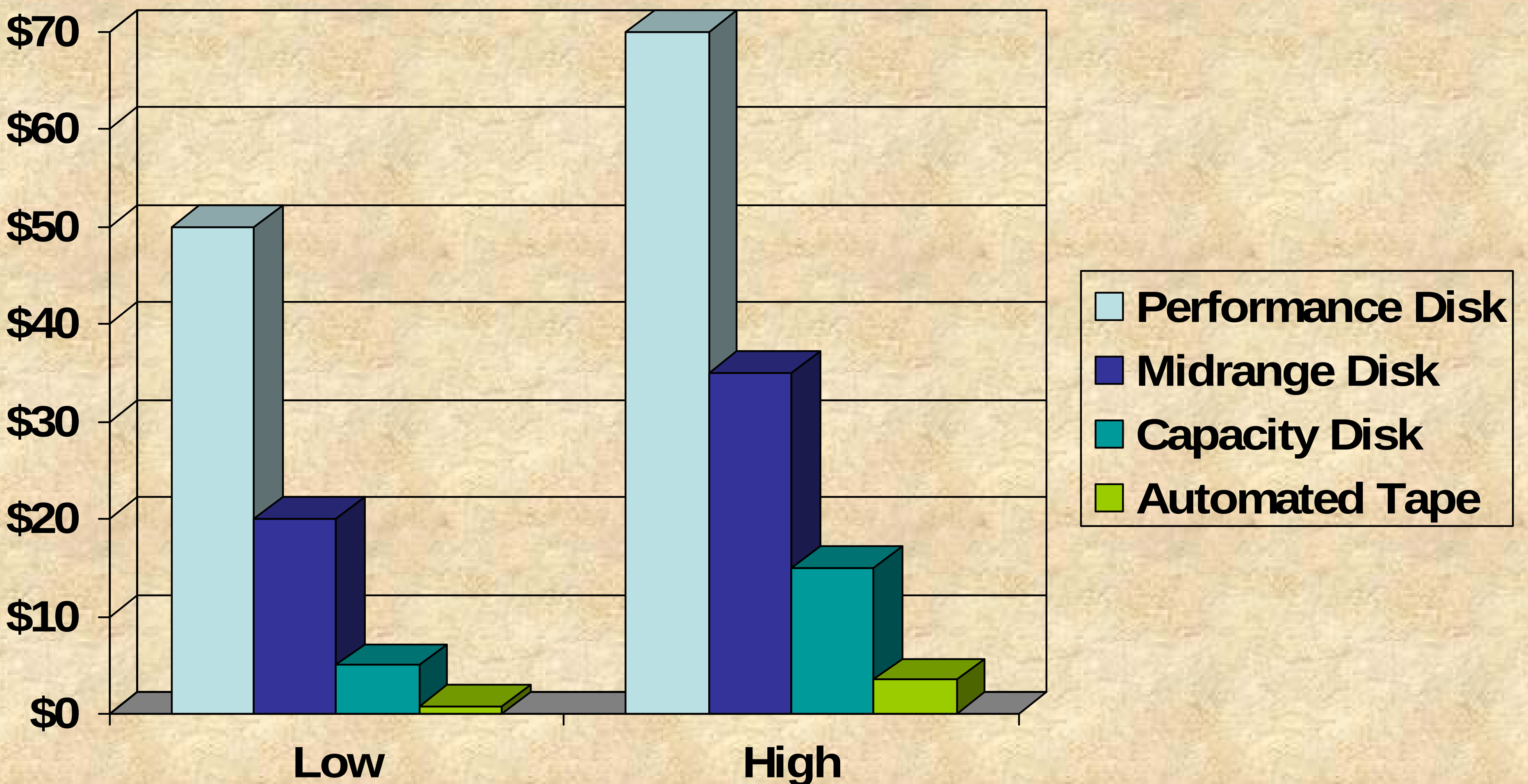
Content Stage	Frequent Changes	Frequent Accesses to Fixed Content	In-Frequent Accesses to Fixed Content
Production	Non-linear editing	Production Viewing	Archiving
Distribution	N/A	Distribution Viewing, PVR/DVR	DVDs, VHS, Content Downloads

SGI InfiniteStorage DMF
Data Life Cycle Management

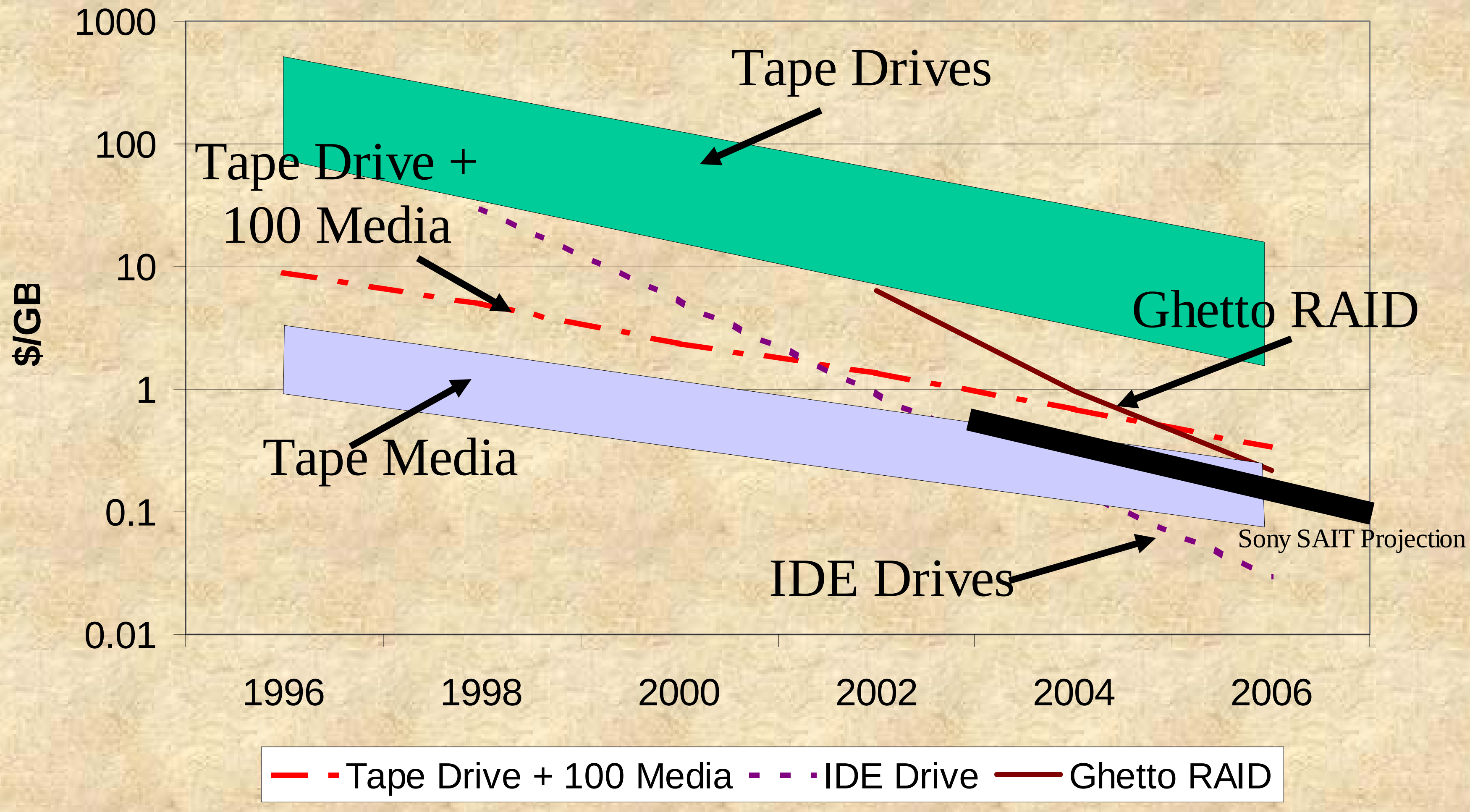


Gabriel Broner, SGI,
Jan. 2004

Comparative Prices of Storage Systems

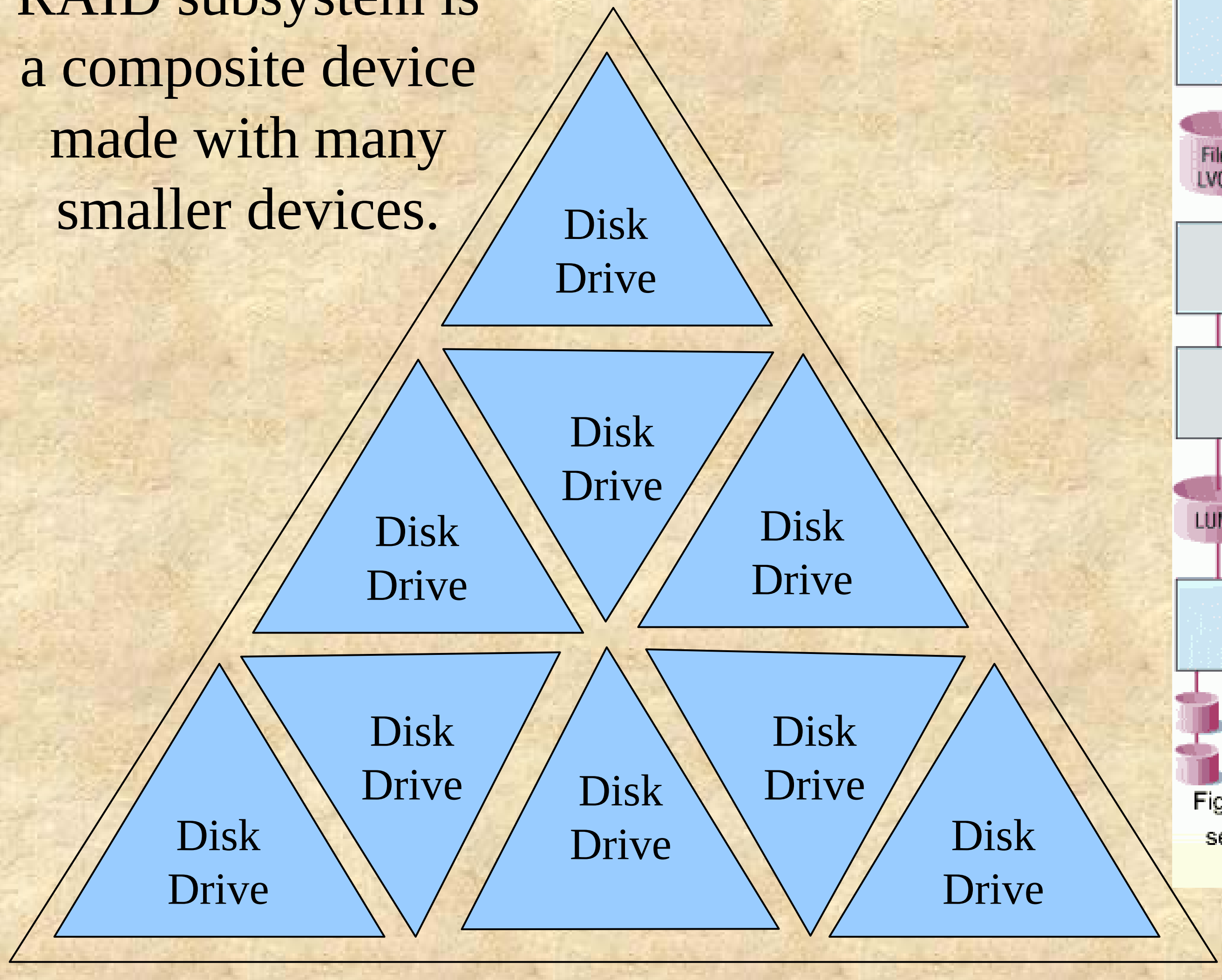


Comparison of Tape and ATA Disk Storage Economics



Device Virtualization

RAID subsystem is a composite device made with many smaller devices.



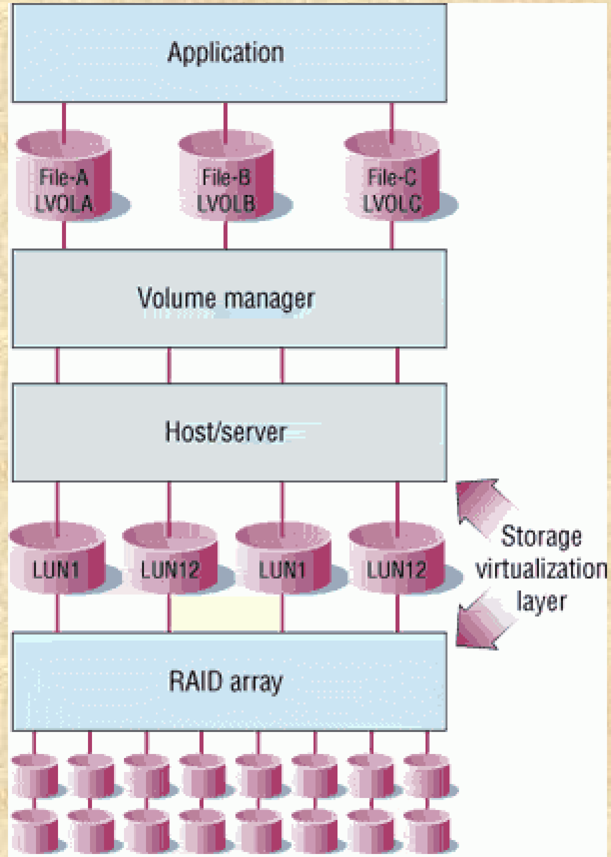
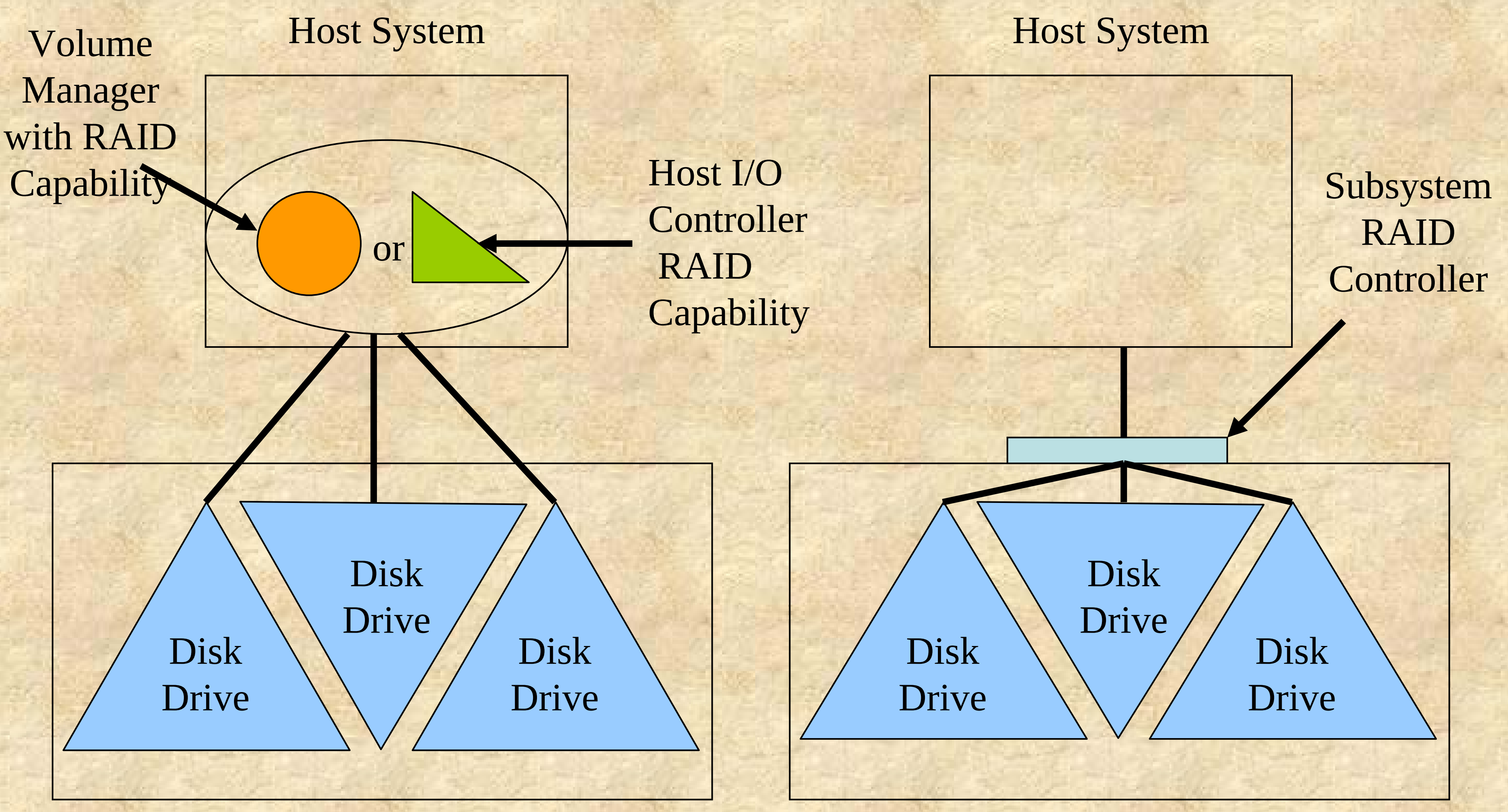


Figure 3: The host system and volume manager see four LUNs presented by an external RAID array.

RAID Systems



JBOD with each device addressed individually by host-resident RAID

RAID subsystem acting as a single virtual device

RAID Advantages

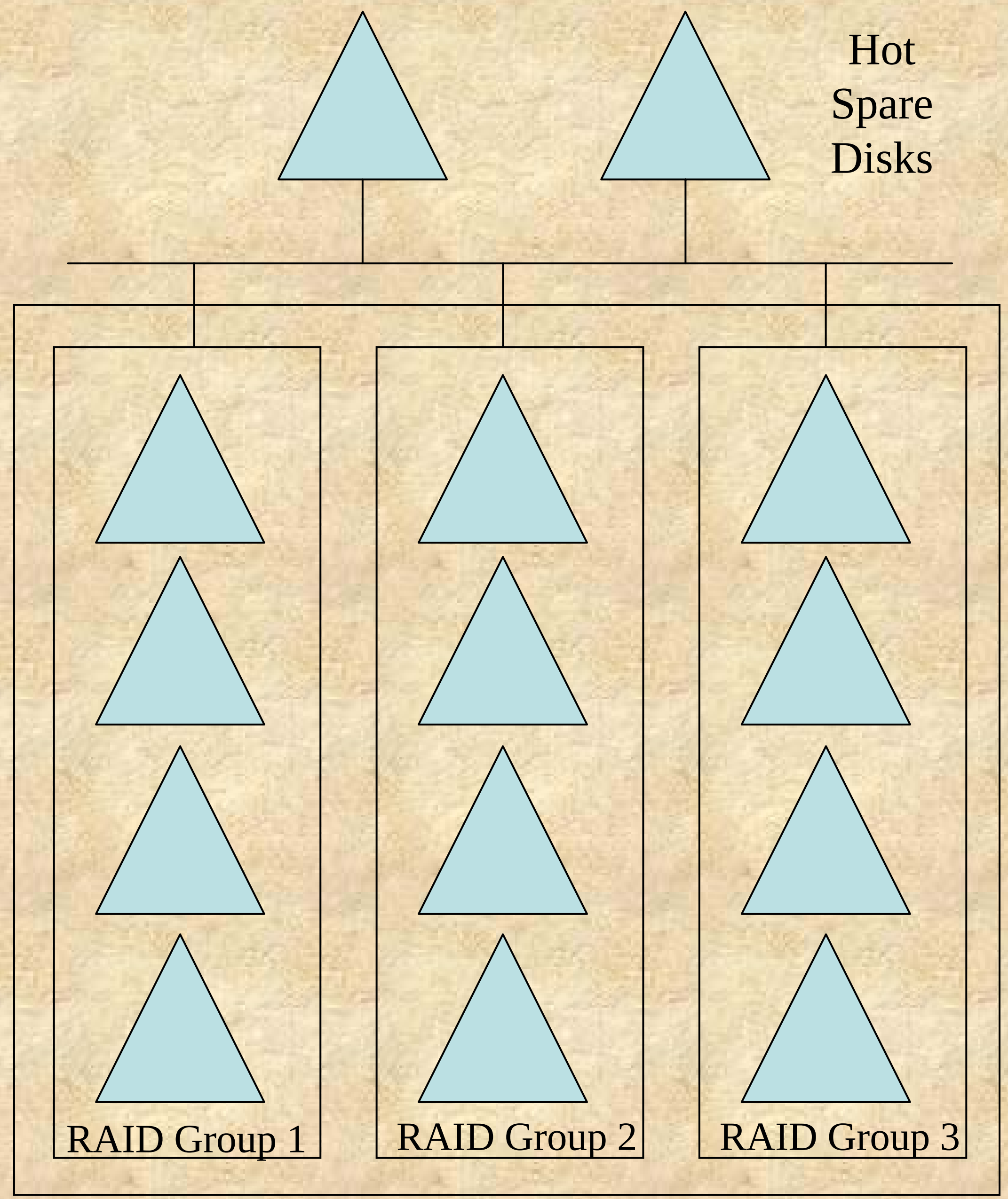
- Can allow for more reliable data and/or improved system performance
- A RAID requires fewer host I/O controller slots. Also a RAID can use a single network (e.g. SCSI) address rather than individual addresses for each drive
- By creating a virtual drive with 1 file system there is no need of the host to manage separate file systems on the individual drives

Characteristics of RAID Levels

	RAID 0	RAID 1	RAID 5
Usable disk space	100%	50%	67-93%
Parity and Redundancy	None	Duplicate data	Parity distributed over each drive
Minimum number of disks	2	2	3
I/Os per Read	1 Read	1 Read	1 Read
I/Os per Write	1 Write	2 Write	2 Reads + 2 Writes
Performance	Best	Good	Worst for Writes
Fault Tolerance	Worst	Best	Good
Cost	Best	Worst	Good
Characteristics	Best over all performance, but data is lost if any drive in the logical drive fails. Uses no storage space forfault tolerance	Tolerant of multiple, simultaneous drive failures. Higher write performance than RAID 5. Uses the most storage capacity for fault tolerance.	Tolerant of single drive failures. Uses the least amount of storage capacity for fault tolerance

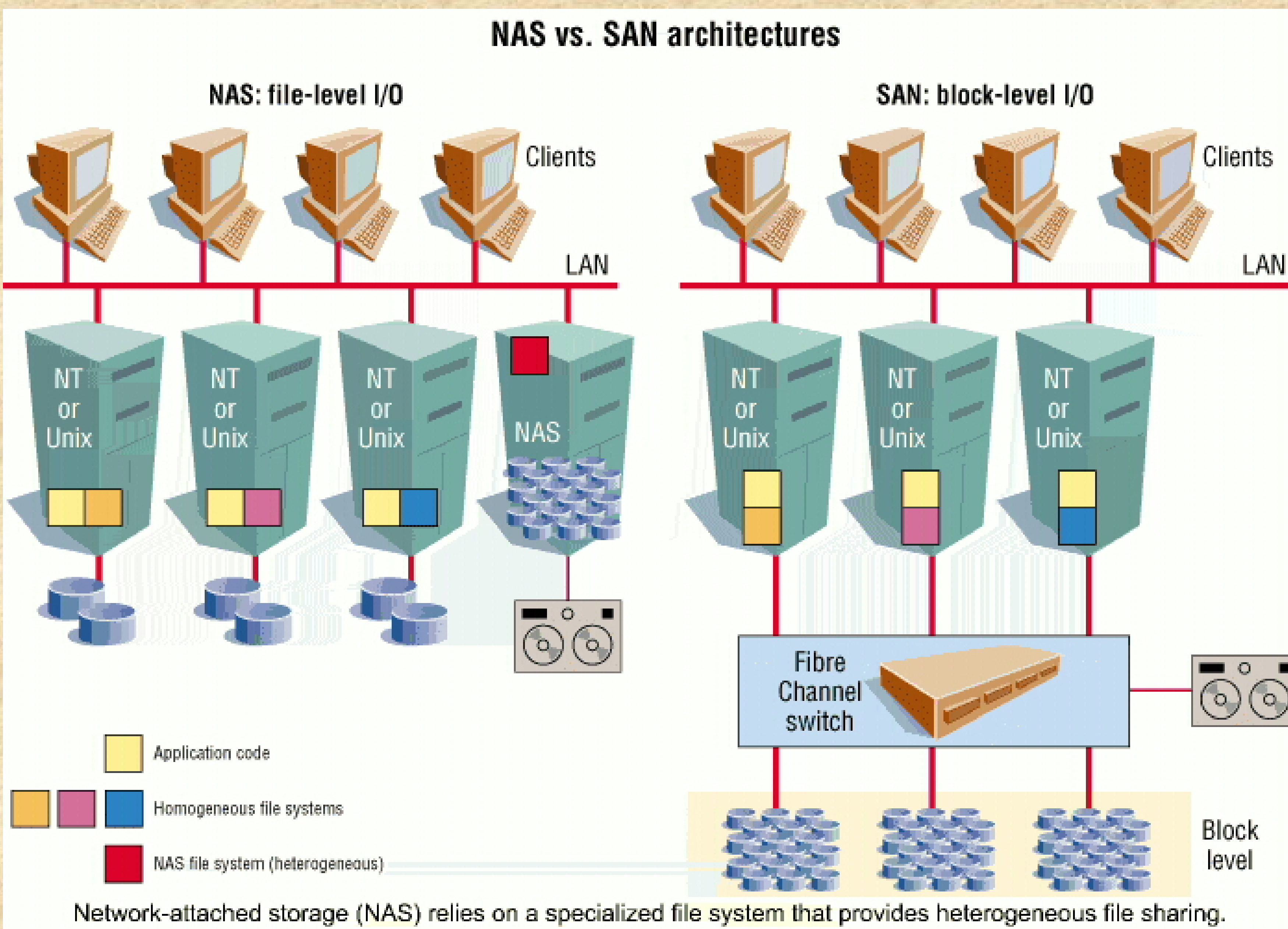
RAID Reliability

- Redundancy
  - drives (hot spares)
  - power supplies
  - fans
  - controllers
- Automatic fail-over to spares



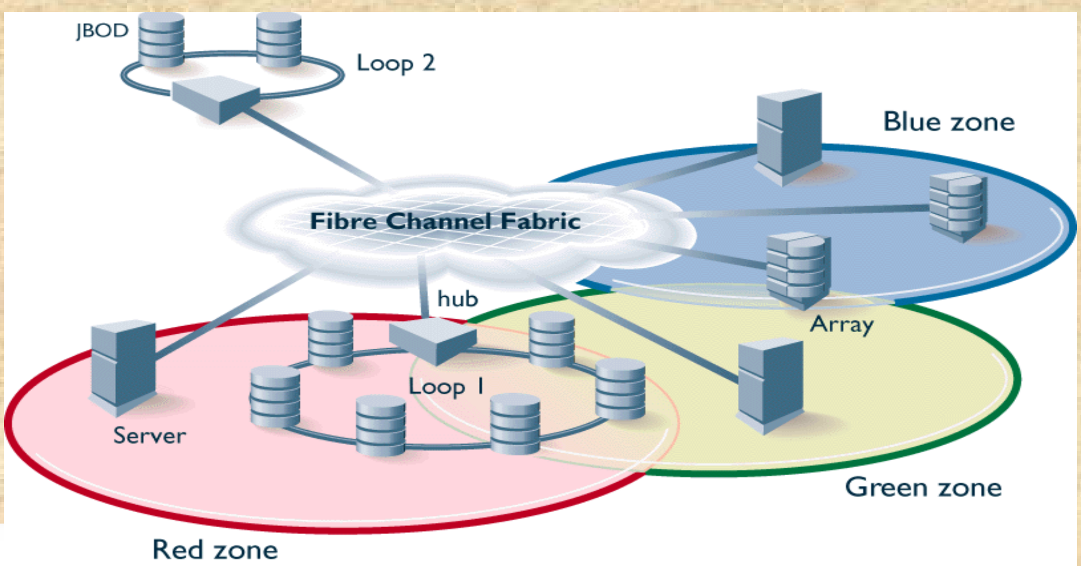
Direct Attached vs. Networked Storage

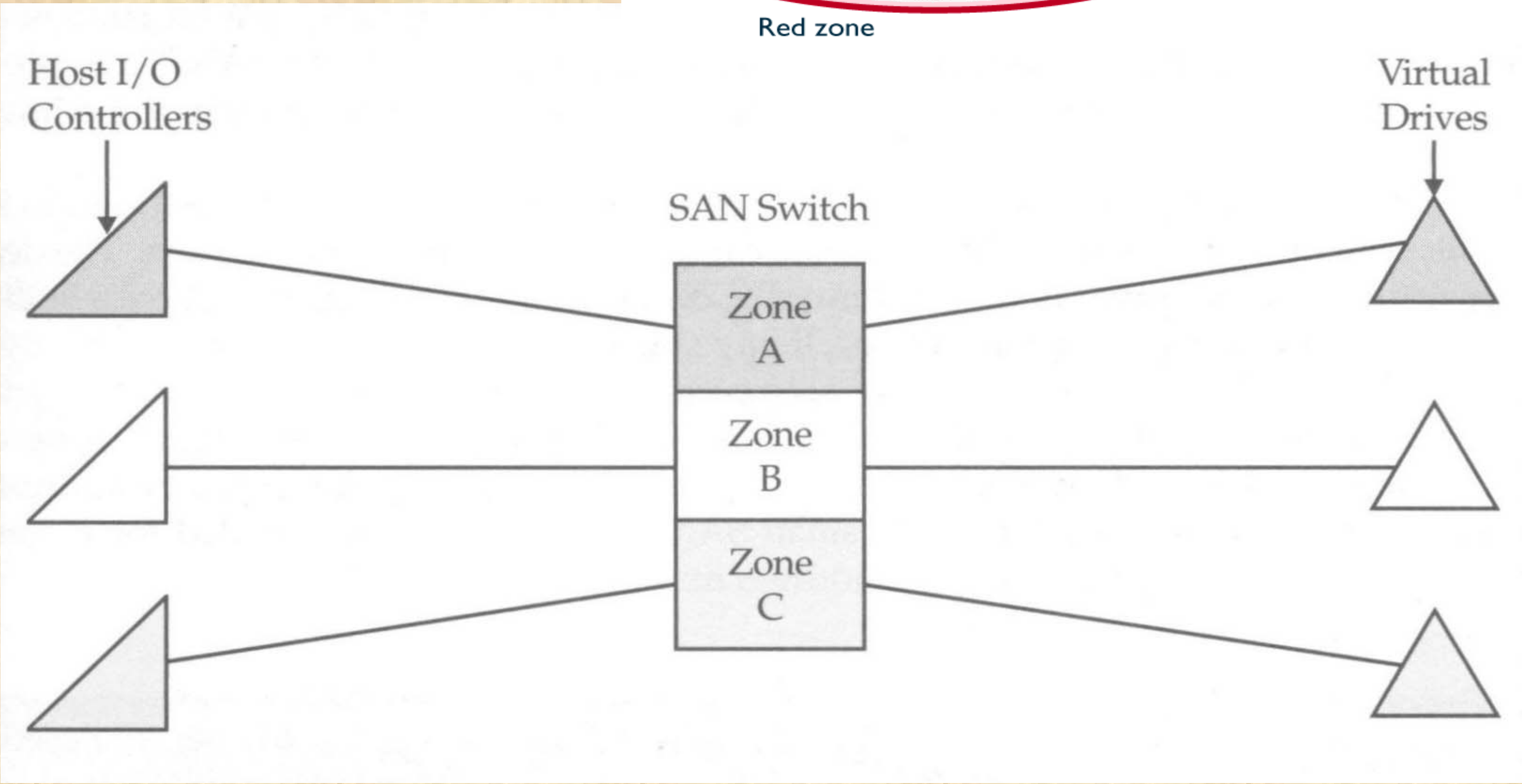
- In **DAS** (Direct Attached Storage) data storage can be incrementally added to a computer system and is subservient to the computer host.
- A **SAN** (Storage Area Network) is a “network storage” system in which storage is accessed through a separate storage network.
- A **NAS** (Network Attached Storage) is an independent aggregated system that can be attached to an existing LAN network in order to increase network available storage.



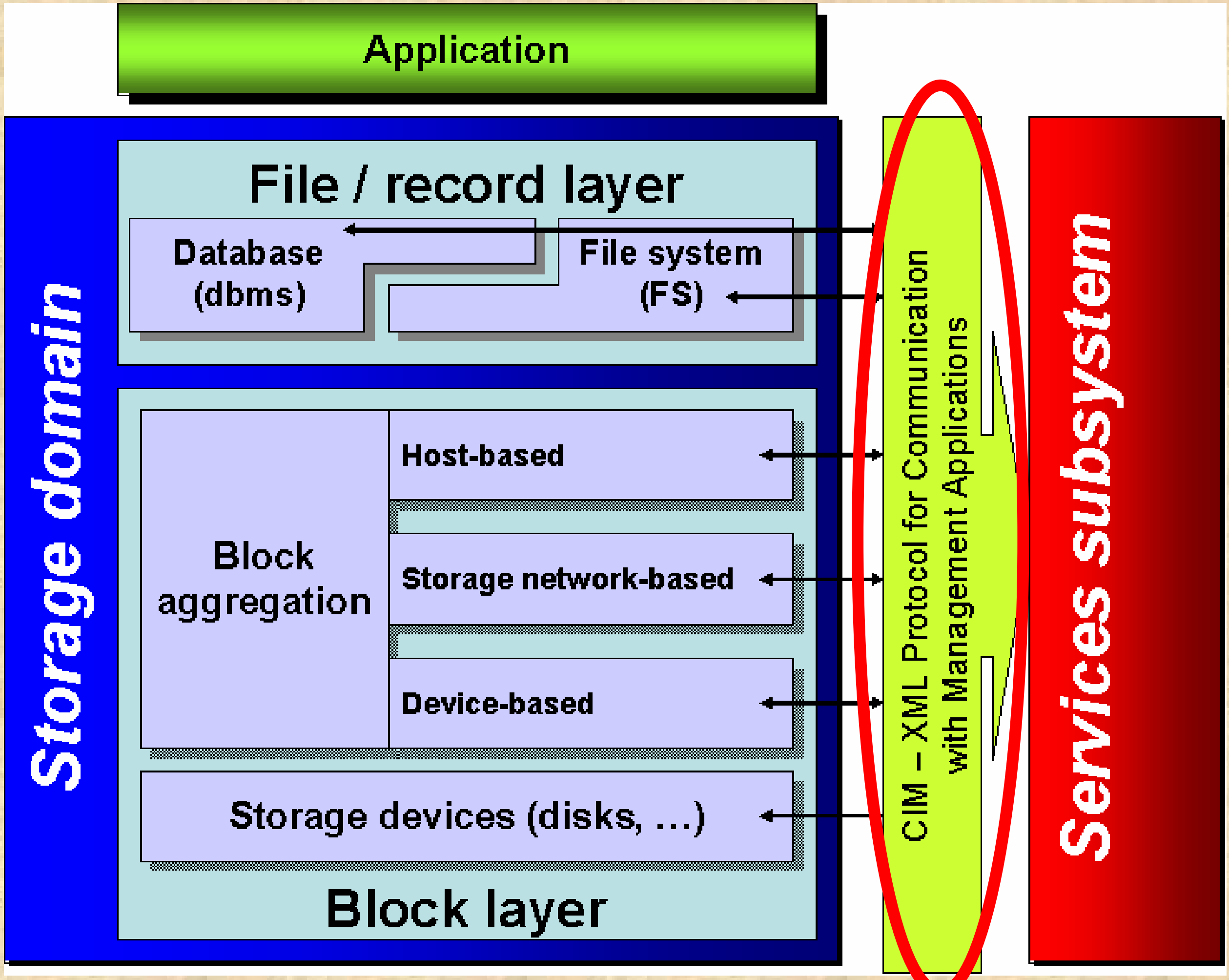
NAS and SAN Architectures, InfoStor, December 2000

Switch Zoning





Standard Device Management Interfaces – SMIS (SNIA Std.)



Examples of ATA-based Storage Systems (Popular for Static Content Storage Systems)

Isilon IQ 3-Node 4.3 TB

Quantum DX30
The DX30 separates backup functions from archive functions to optimize the data protection process.

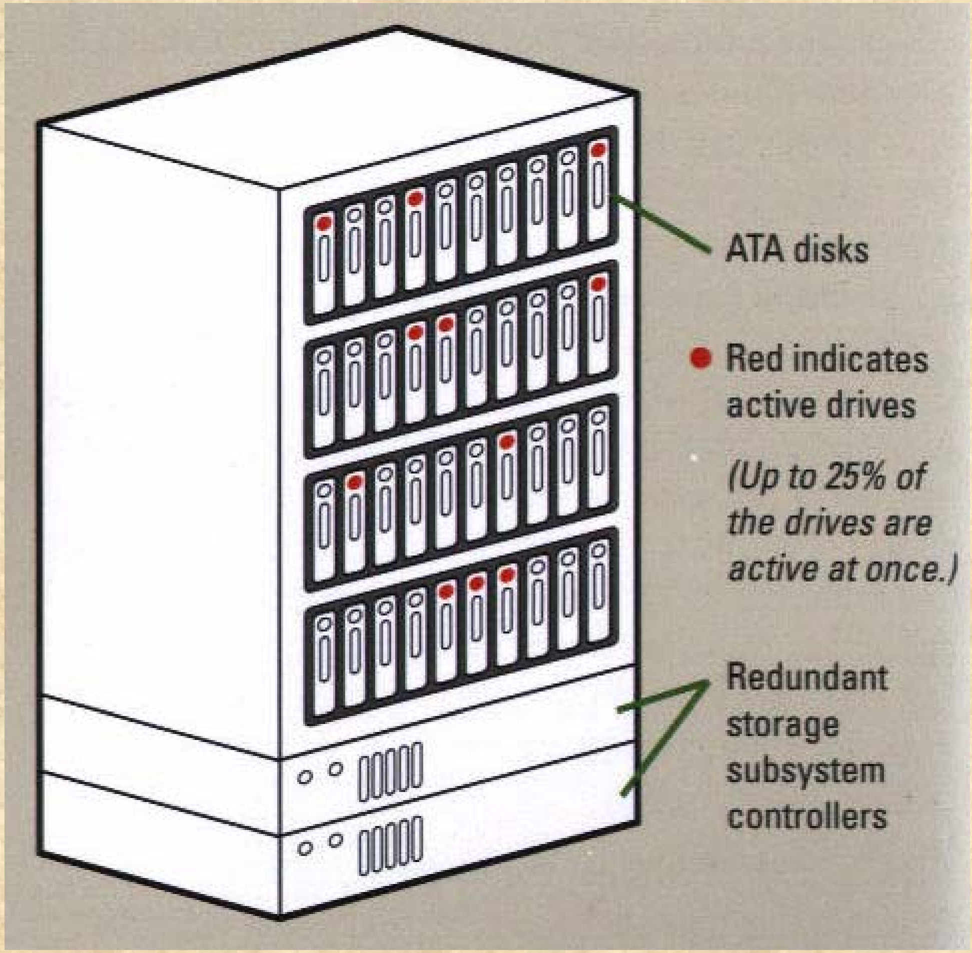
Nexsan ATABeast Nexsan's 14 TB for 7 cents a MB.

(See new introductions at 2004 NAB)

STK Bladestore product uses 3.5 inch drives on blade acting as one drive to a fibre channel output

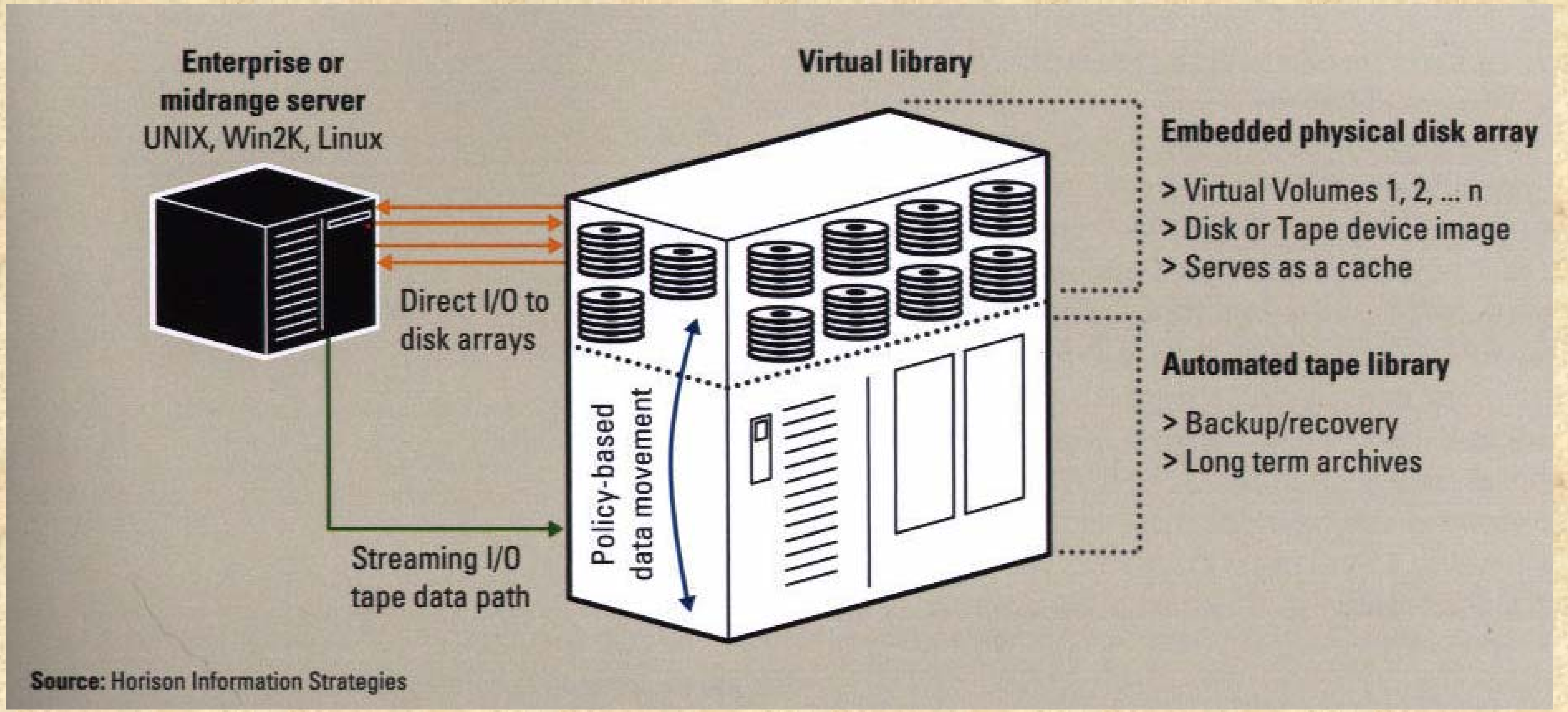
R200 now offers up to 96 TB

MAID (Massive Array of Inactive Disks)



- Disks inactive most of the time (only about 25% active at any time)
- Can be RAID or JBOD
- Workload is mostly writes, seldom read
- Reduced costs since components shared
- Low power
- Field replaceable drives
- Start-ups ?? offering MAID systems

Virtual Tape Cache for Backup



Tape-based Digital Content Storage System

Sony Petasite Tape Library

Content Software

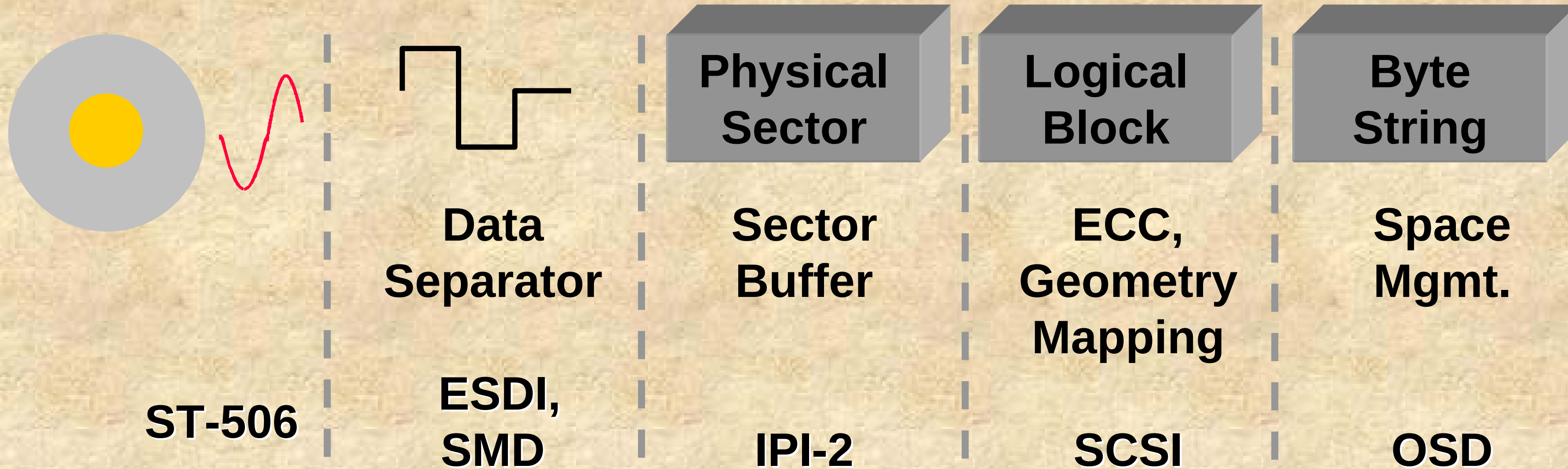
- SGI
- Veritas
- Exanet
- SANbolic
- Kasenna
- Context Media
- Many Others

Connection Interfaces and Protocols

- SCSI
- Serial Attached SCSI
- Fibre Channel
- FATA (Seagate and HP)
- ATA
- Serial ATA

- TCP/IP and variations
- iSCSI
- FC over IP
- Infiniband

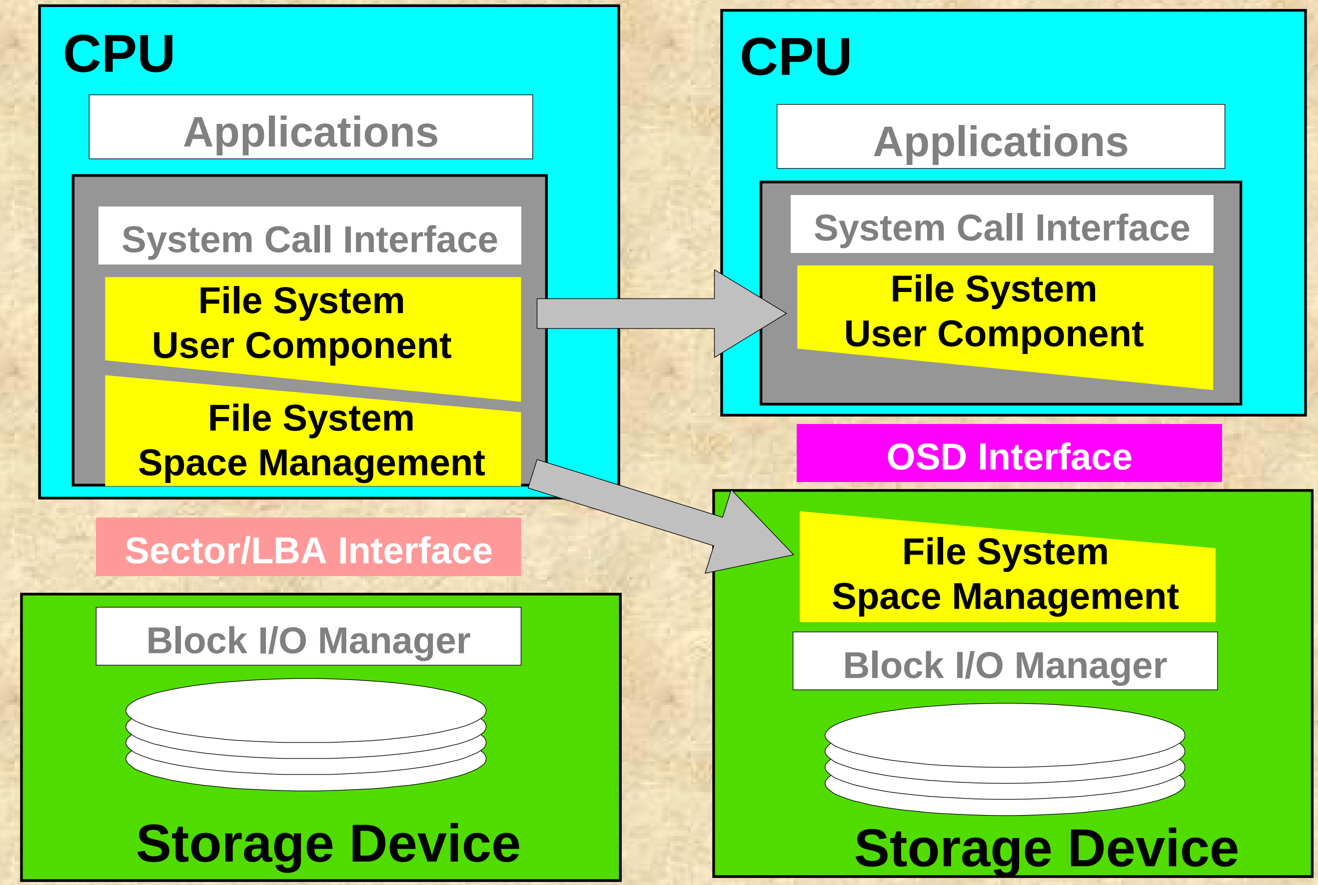
Storage Interface Progression



- Each change represents intelligence moving from host to drive
- Each advancement was met with resistance
- Eventually advantages of new intelligence were compelling

From Seagate Technology, 2003

OSD: A New Standard Interface



Completes Device Abstraction

From Seagate Technology, 2003

Object Storage Systems

Expect wide variety of Object Storage Devices

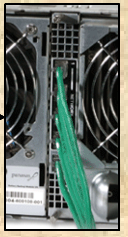
- Disk array subsystem
- Ie. LLNL with Lustre

- “Smart” disk for objects
- 2 SATA disks – 240/500 GB

- Prototype Seagate OSD
- Highly integrated, single disk

- Orchestrates system activity
- Balances objects across OSDs

- **Stores up to 5 TBs per shelf**
- Battery-backed redundant power



16-Port GE Switch Blade

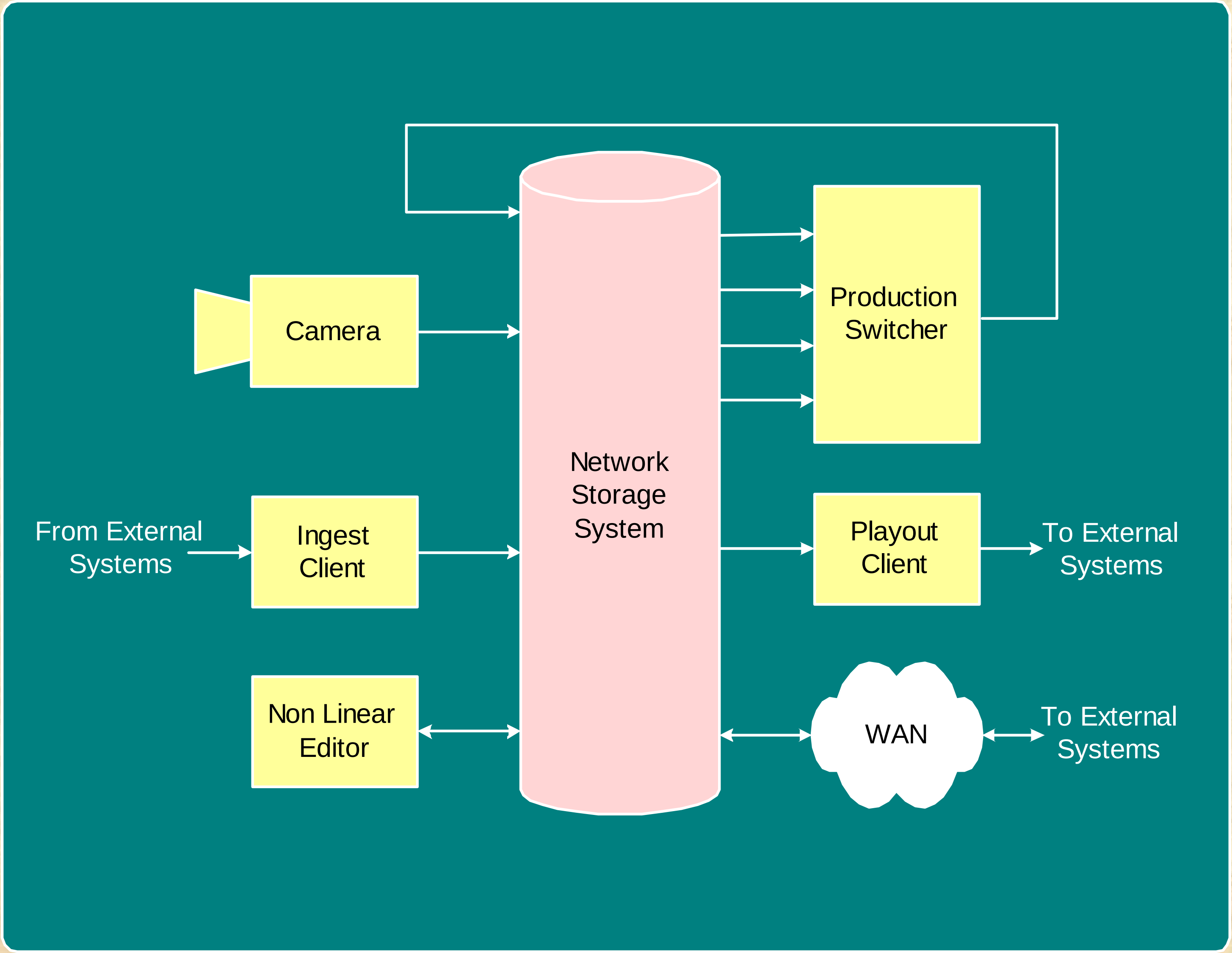
- 4 Gb/sec per shelf to cluster

From Seagate Technology, 2003

Applications for Entertainment Content Storage

Professional Digital Camera
(Storage System for Content Capture)

Asynchronous packet switched architecture



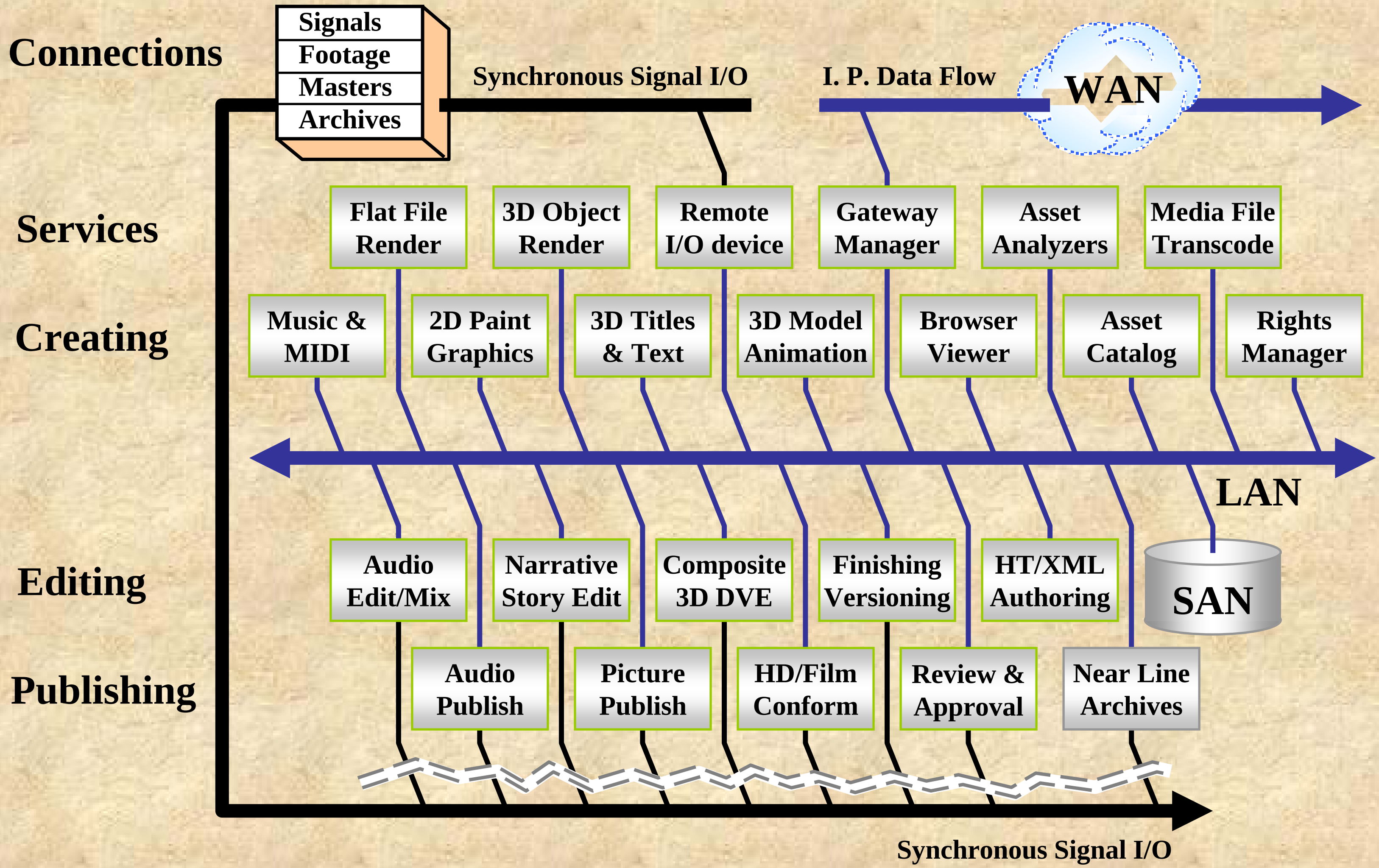
Material Exchange Format (MXF)

- International standard
- Designed to enable distribution of A/V files over IT infrastructures
- License free open source wrapper for video, audio and metadata
- Real time streams or non real time file transfers
- Wrapper can contain various metadata such as DRM

Material Exchange Format (MXF)

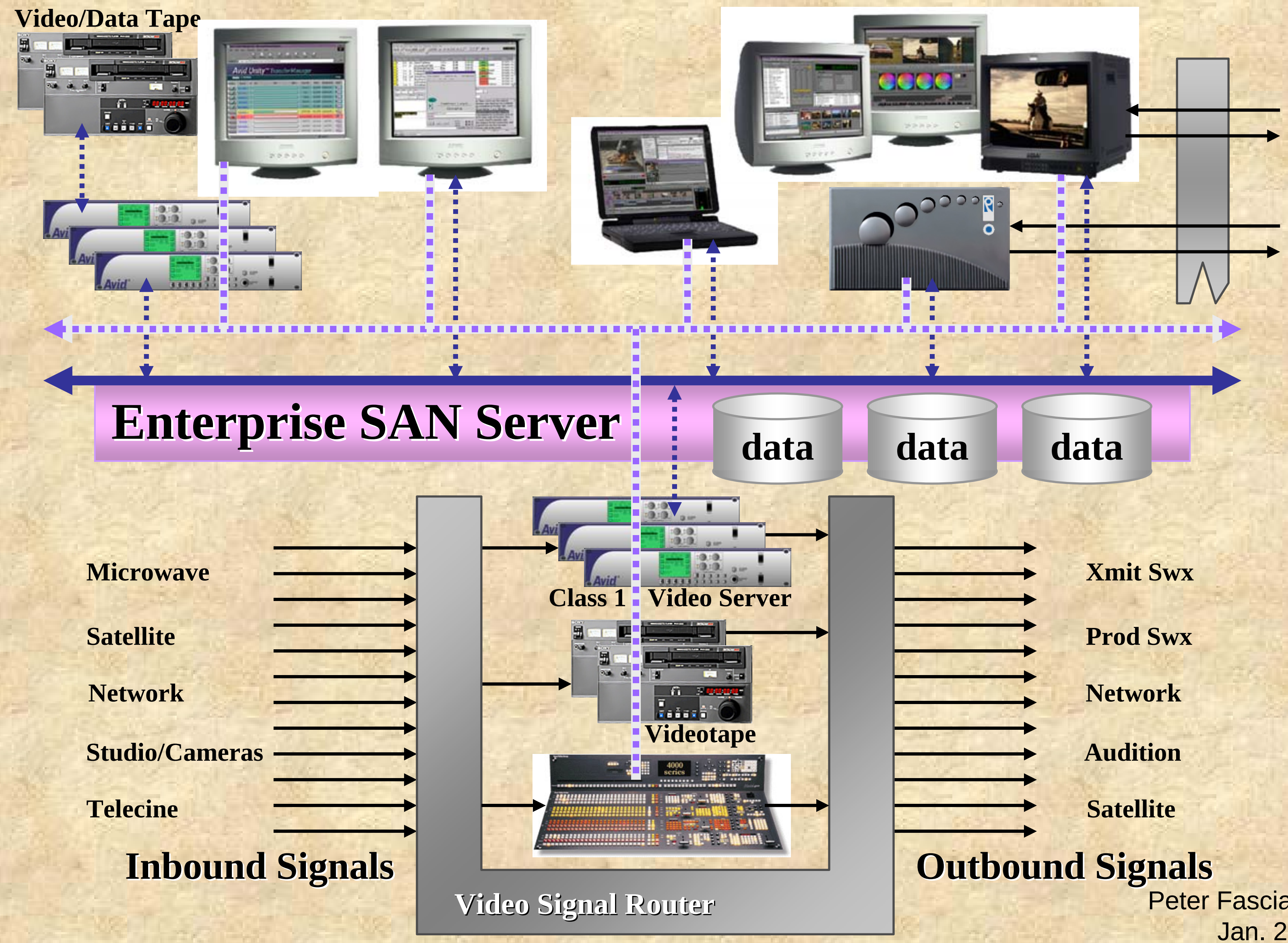
- Partitions enable files to be read while being written
- Files can also be tuned for file system
  - KLV Alignment Grid (KAG)
  - KAG specifies file system logical block size
- Standardized index tables
  - Enable fast access to edit units and partitions

Nonlinear Editing System



Peter Fasciano, Avid, Jan. 2004

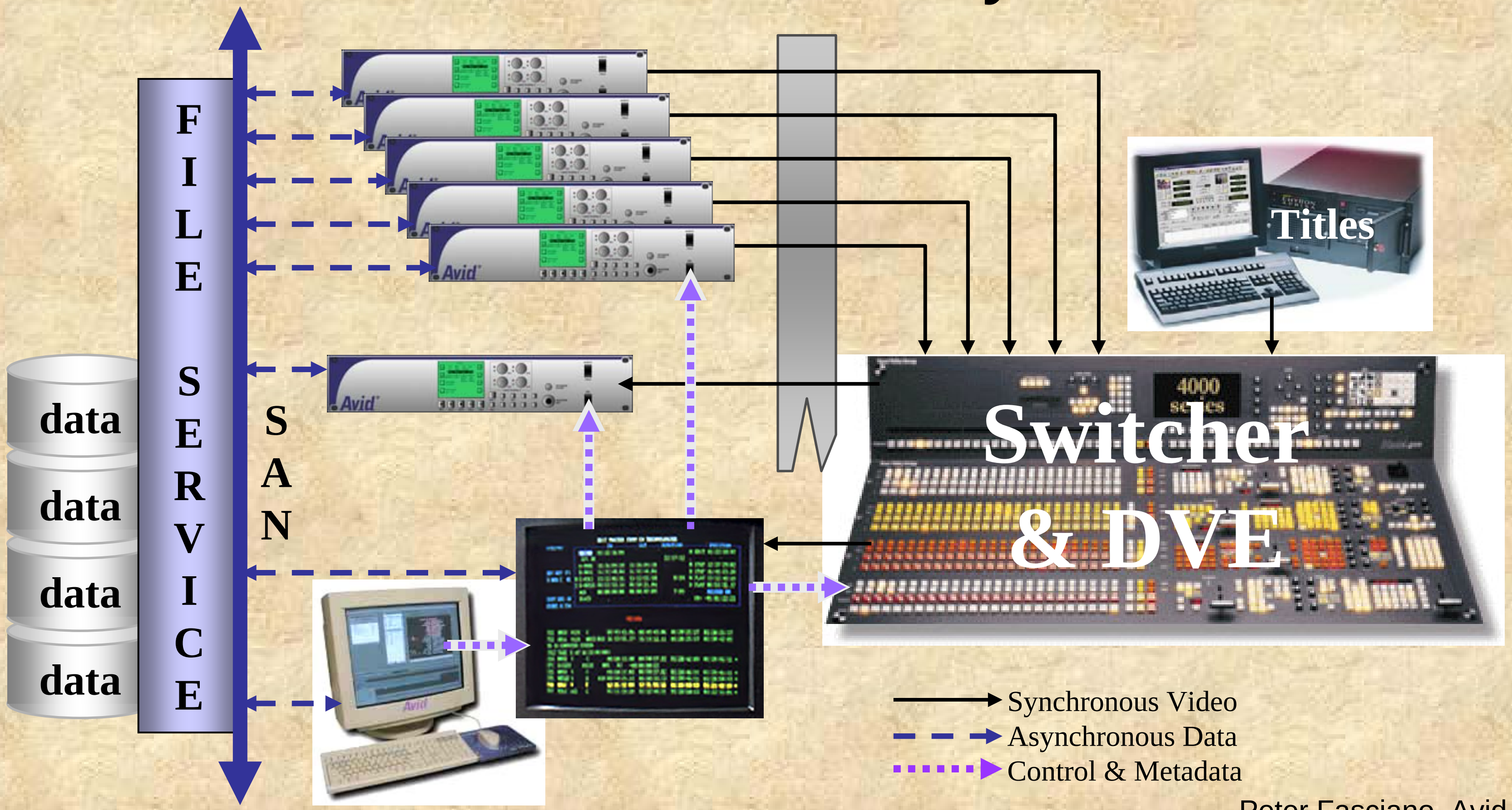
Nonlinear System Design (Avid)



Peter Fasciano, Avid, Jan. 2004

Nonlinear System Design (Avid)

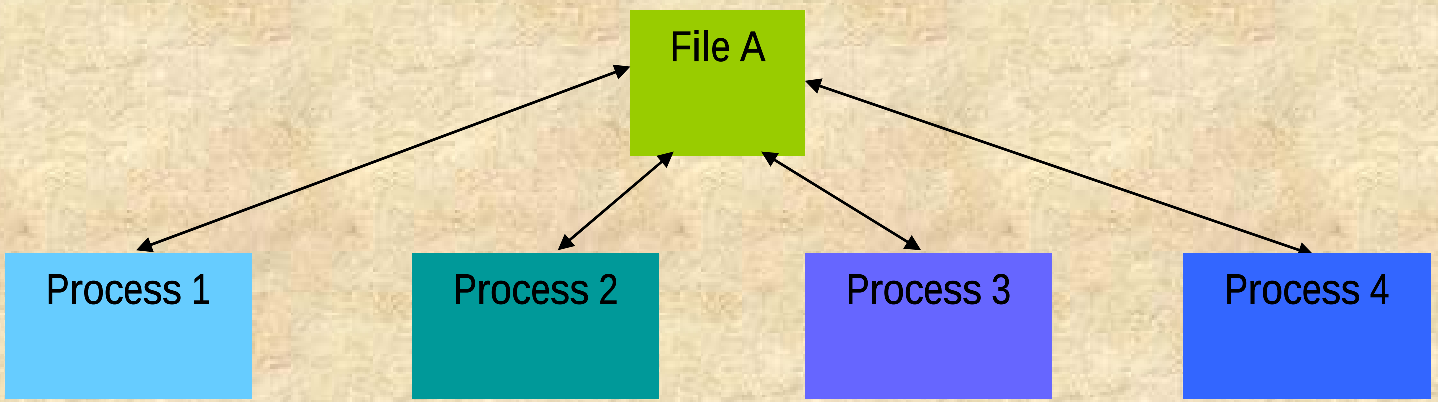
- The online real-time effects system



Peter Fasciano, Avid, Jan. 2004

Workflow with File Sharing (SGI)

Near-instantaneous access for data-intensive workflows

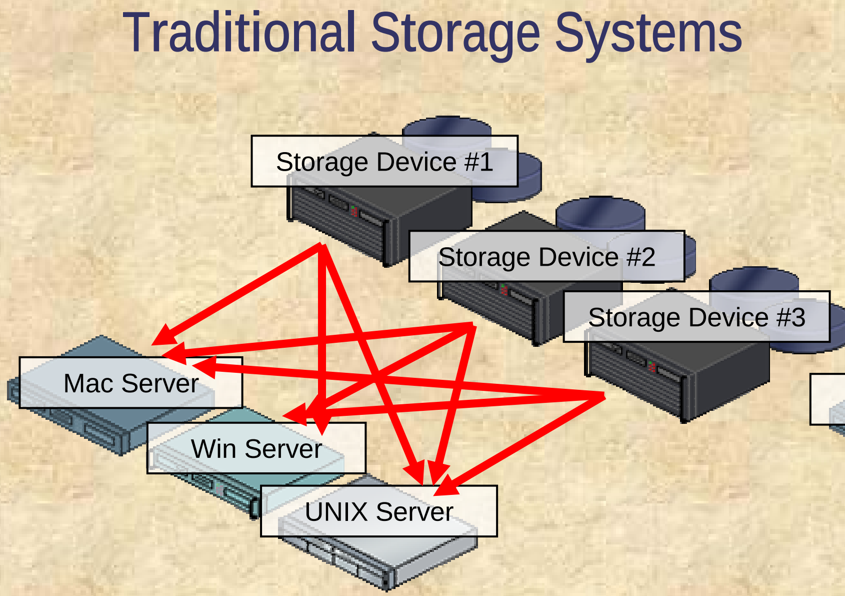


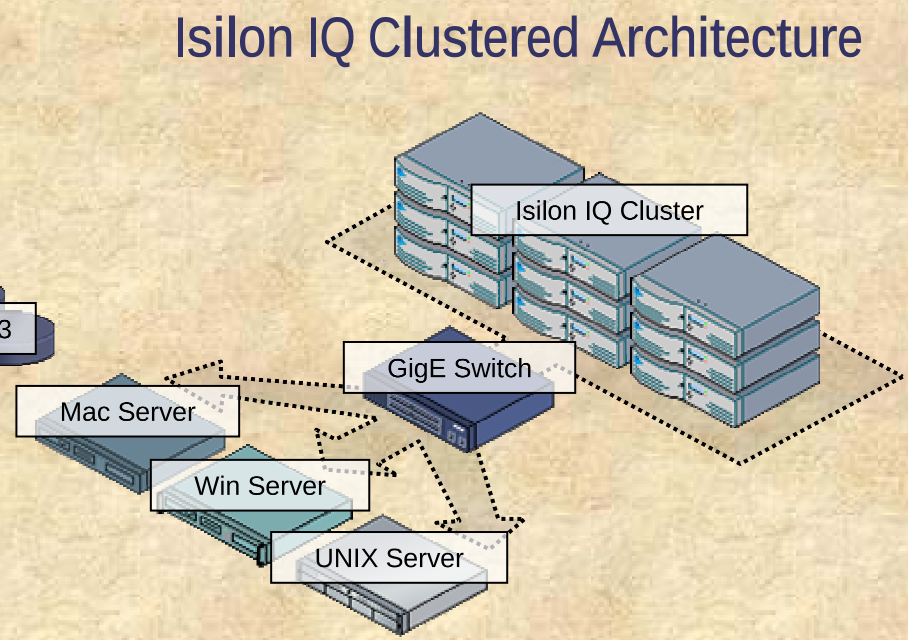
Media	Digitization	Color Correcting	Effects	Compositing
Mfg.	Design	Visualization	Structural Analysis	Crash Analysis

File sharing means large files don't have to be moved over the network—saving time, speeding workflow.

Gabriel Broner, SGI, Jan. 2004

Clustered Digital Content Storage





Acute Pain with Digital Content

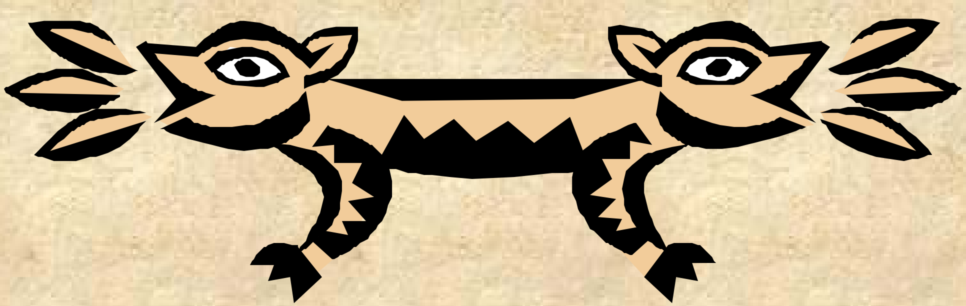
- Separate islands of storage
- Complex & hard to grow
- Server performance bottlenecks
- Inherent single points of failure

Isilon IQ Eliminates Customer Pain

- One single pool of storage
- Simple, easy, & modular to grow
- Cluster eliminates server bottlenecks
- No single points of failure

Brett Goodwin, Isilon,
Jan. 2004

Conclusions



- Digital content creation and distribution will require large volumes of storage
- Storage devices and requirements vary throughout the content value chain.
- Storage device and architecture development enables ever lower and more capable digital content creation and distribution!

Acknowledgement: Much of the material from this presentation was created while researching the **2004 Entertainment and Digital Media Storage Report**, Authors: Tom Coughlin, Pat Hanlon, and Dennis Waid. For more information see www.tomcoughlin.com.

Digital Storage Will Entertain a New Generation!