Object-based Storage Devices

Tenth NASA Goddard Space Flight Center Conference on
Mass Storage Systems and Technologies

Nineteenth IEEE Symposium on Mass Storage Systems

Thomas M. Ruwart

tmruwart@ciprico.com

Ciprico, Inc.

April 15, 2002

Overview

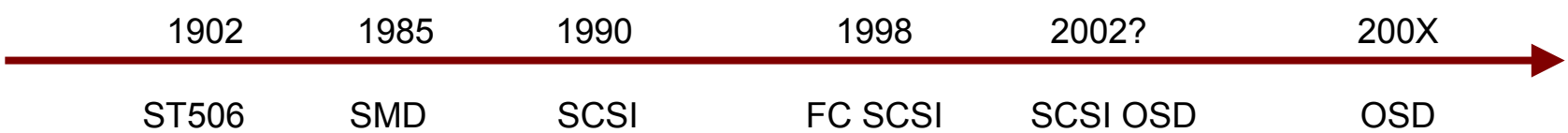
CIPRICO

- What is OSD?
- What problem(s) are we trying to solve?
- Why would you want to use OSD?
- Example OSD Scenarios
- Current OSD-like implementations
- Issues with OSD

What is OSD?

CIPRICO

- Object-based Storage Devices – An Enabling Technology
- Grew out of the Network Attached Secure Disks (NASD) project at CMU
- A flexible and powerful protocol used to communicate with storage devices
- Proposed as a protocol extension to the SCSI command set
- Actively being pursued by the OSD Technical Working Group in the Storage Networking Industry Association (SNIA)
- It is a natural step in the evolution of storage interface protocols
- For some however, it is very new and very different



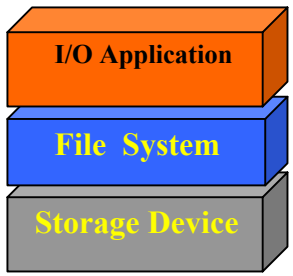
What OSD is NOT

CIPRICO

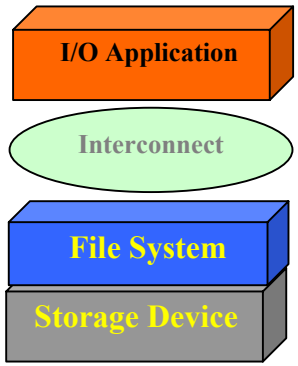
- It is not intended or expected that the object abstraction be a complete file system
- There is NO notion of
  - Naming
  - Hierarchical relationships
  - Streams
  - file system style ownership access control
- The omitted features are assumed still to be the responsibility of the OS file system

The General Application: Storage Architectures Today

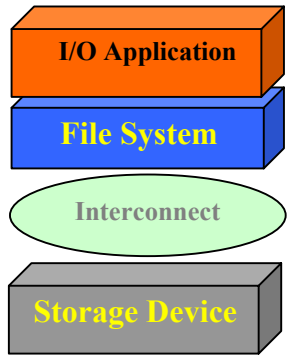
CIPRICO



Direct Attached Storage (blocks)



Network Attached Storage (files)

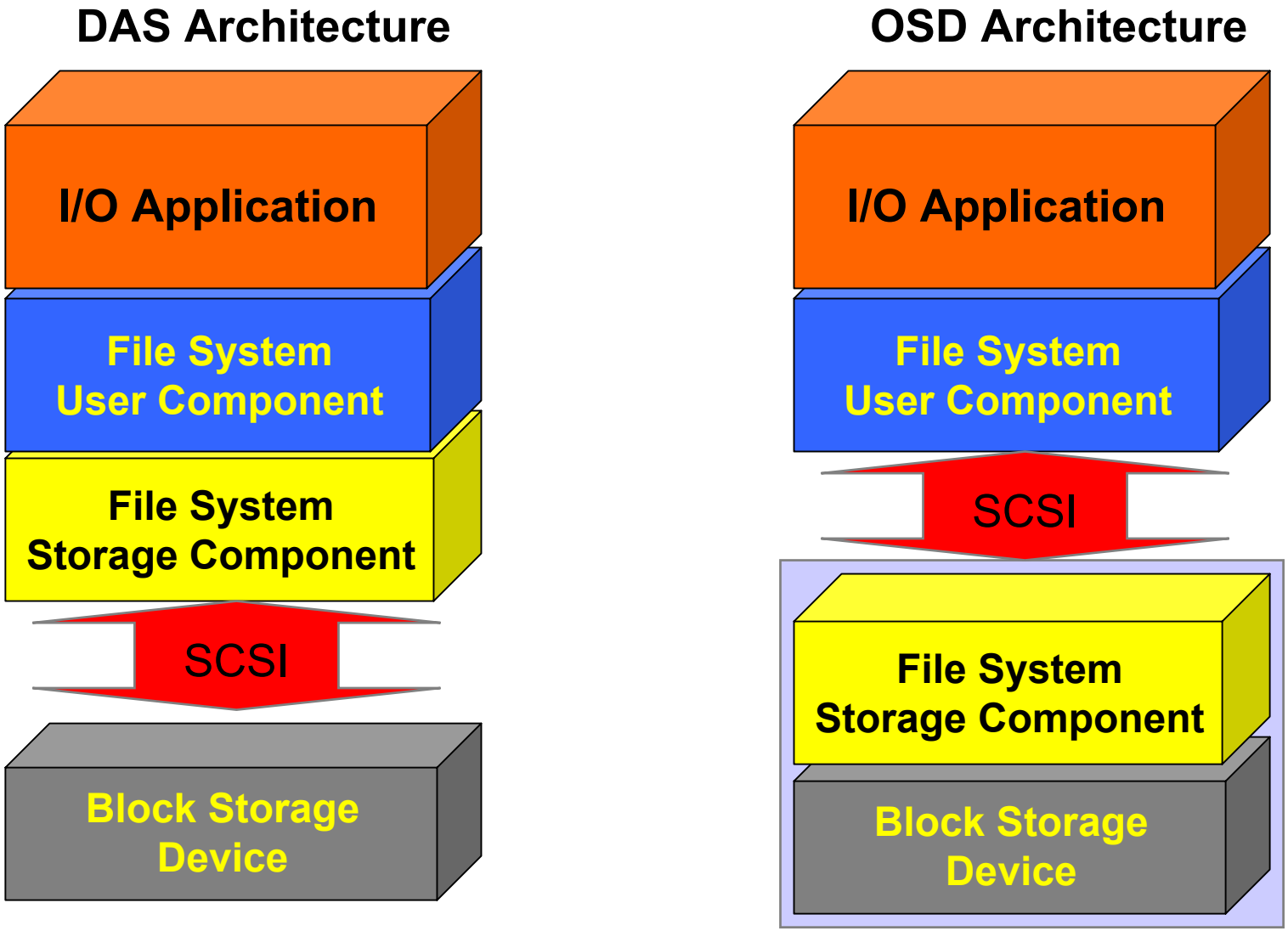


Storage Area Network (blocks)

Architecture defined by location of storage system & devices

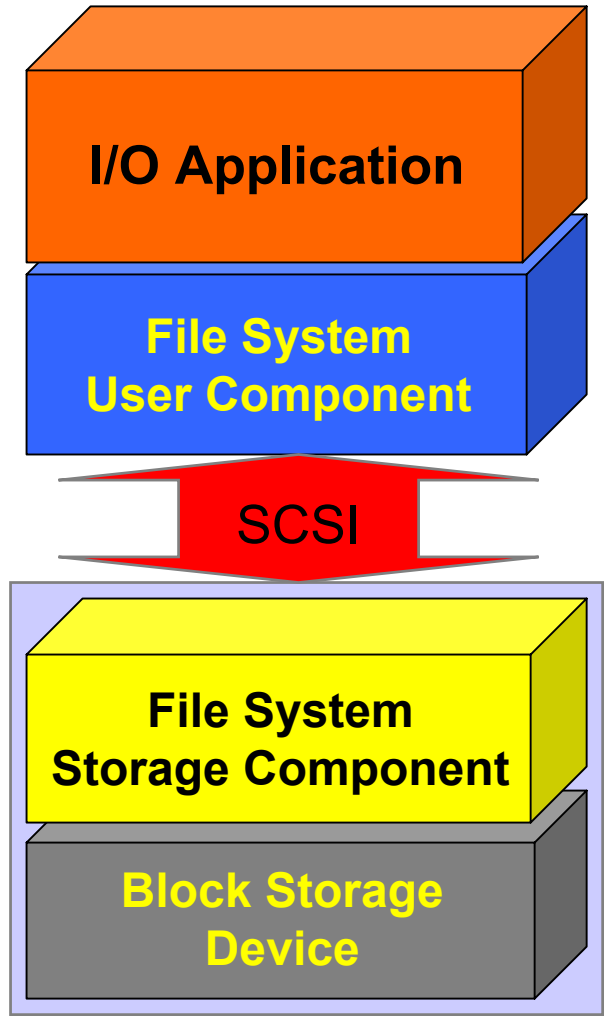
OSD System Architecture

CIPRICO

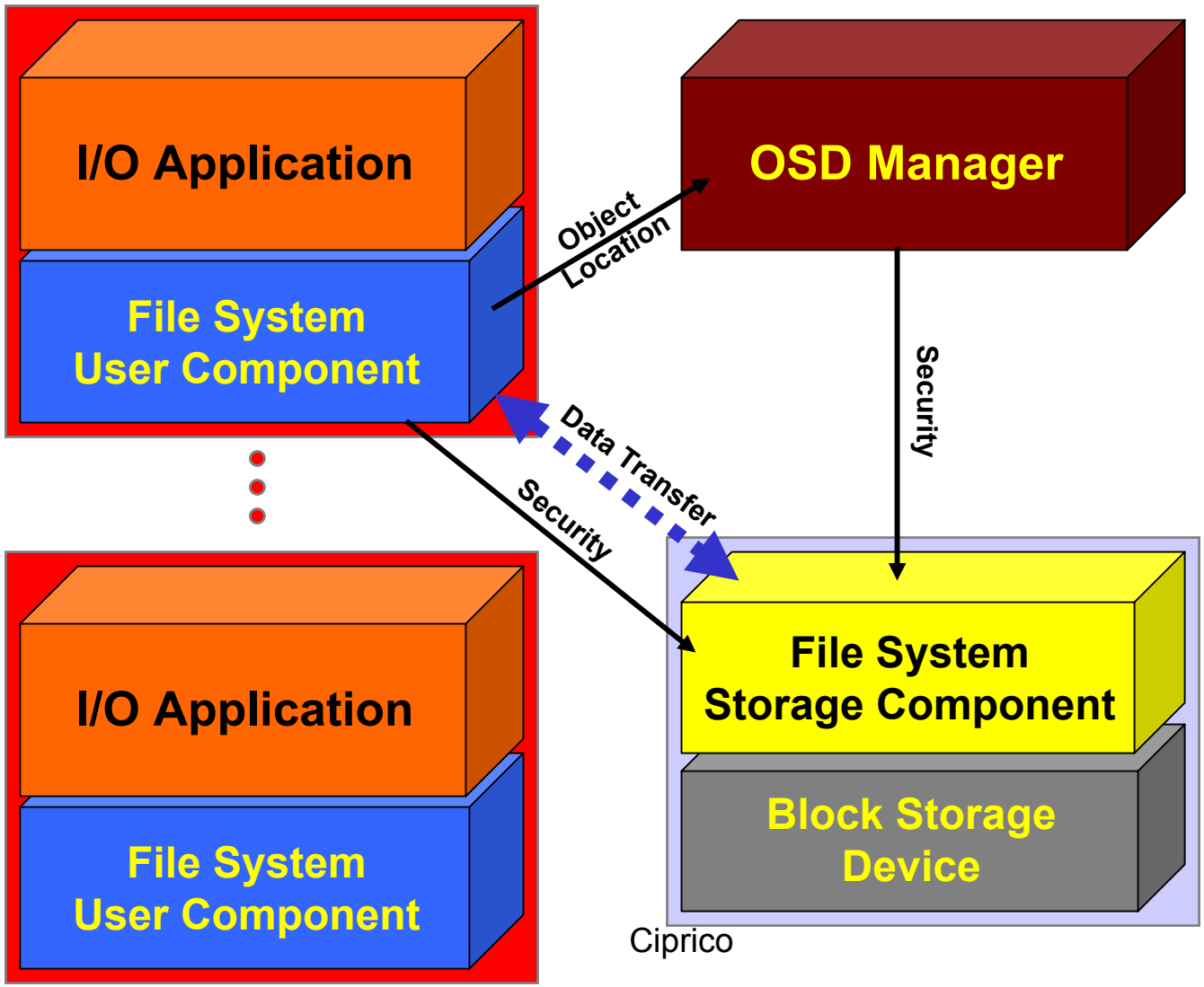


File System Components

- User File System Component
  - Hierarchy Management
  - Naming
  - User Access Control
  - Data Properties (Attributes)
- File System Storage Component
  - Free space management
  - Storage allocation for data entities
  - Attribute Interpretation



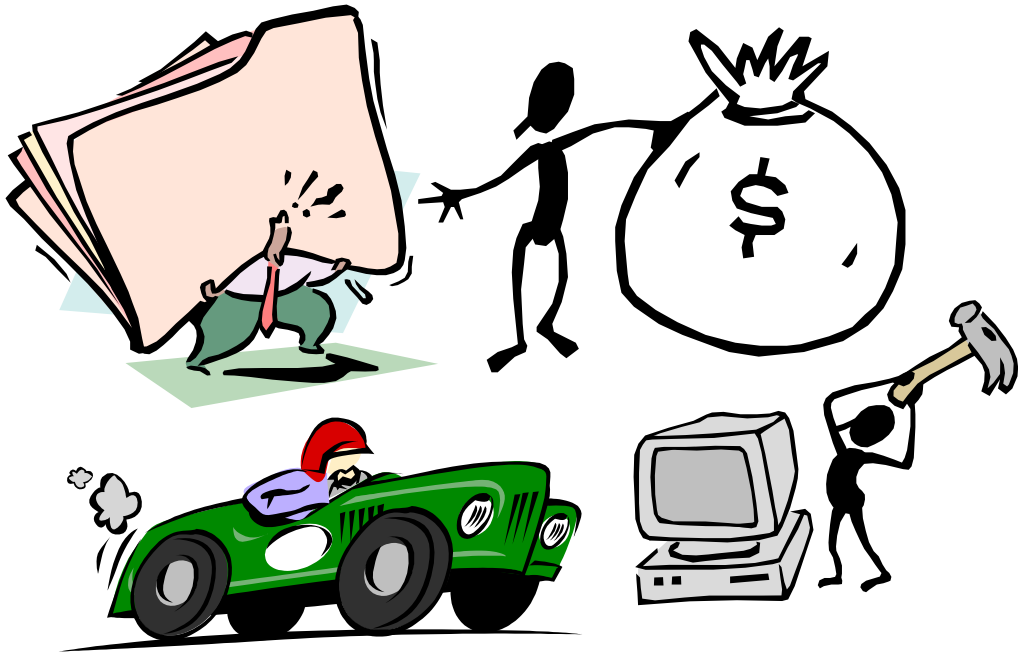
How OSD works



What problems are being solved?

CIPRICO

- Depends on the APPLICATION
- Different people are trying to solve different problems for different reasons
- Storage Device Utilization
- Data Management
- Cost
- Reliability
- Device Management
- Performance
- Security
- Availability
- Maintainability
- ***Extensibility***
- Restate the question: What problems CAN be solved with OSD?



What CAN be addressed by OSD

CIPRICO

- Improved storage management
  - Self-managed, policy-driven storage (e.g., backup, recovery)
- Improved device and data sharing
  - Shared devices and data across OS platforms
- Improved storage performance
  - Hints, QoS, Differentiated Services
- Improved scalability (and not just capacity)
  - Of performance and metadata (e.g, free block allocation)
- Current block-based access protocols and associated file systems are 30 years old (that's 210 in dog-years).
- OSD has the potential to make a significant impact on the ***Extensibility*** of a Storage System Architecture

Extensible Architectures

CIPRICO

- Density – the number of bytes/IOPS/bandwidth per unit volume
- Scalability – what does that word really mean?
    - Capacity: number of bytes, number of objects, number of files, number of actuators …etc.
    - Performance: Bandwidth, IOPs, Latency, …etc.
    - Connectivity: number of disks, hosts, arrays, …etc.
    - Geographic: LAN, SAN, WAN, …etc.
    - Processing Power
- Cost – address issues such as $/MB, $/sqft, $/IOP, $/MB/sec, TCO, …etc.
- Adaptability – to changing applications
- Capability – can add functionality for different applications
- Manageability – Can be managed as a system rather than just a box of storage devices
- Reliability – Connection integrity capabilities
- Availability – Fail-over capabilities
- Serviceability – Hot-plug capability
- Interoperability – Supported by many vendors – Heterogeneous by nature

Management

CIPRICO

- Self-managed devices
  - More autonomous because the devices are *aware* of the data objects they store
  - Space Management is simpler for the same reason
- Data management is easier because the task is offloaded onto the storage devices themselves
- Data Sharing
  - Heterogeneous OS support is implicit
  - Physical device sharing is implicit
- Policy-Driven backup, recovery, hierarchical storage management
- Managing objects is easier than managing blocks
- Managing Object-based Storage Devices is easier than managing block-based storage devices

Performance Virtualization

CIPRICO

- Three Performance Metrics
  - Bandwidth – number of sustained bytes per second
  - Latency – time to first byte of data
  - IOPS – number of sustained transactions per second
- Applications need only specify values for these three metrics as "attributes" of the object being created or accessed
- The Storage Device can then decide where/how best to store the object in order to meet the performance requirements (see Hybrid Storage Devices)
- Abstracts the physical storage device performance characteristics
- Attributes can also be used to make more informed decisions about cache usage
- Performance attributes can also be used to manage performance as a resource

Security

CIPRICO

- OSD has a Security Model built into it from the beginning rather than as an after thought
- The OSD Security Model enables a secure exchange and storage/execution of objects
- Using this security model Active Object Storage Devices can effectively implement encryption
- The inclusion of a Security Model gives OSDs more autonomy than plain disk drives

Technology Shifts

CIPRICO

- What happens when…
  - NEC Announces a 10 Terabit Memory Chip
  - MEMS devices bridge the gap between RAM and Disk
  - DVDR Replaces Tape
  - Disk densities hit 1 terabit/in^2
- Must Decouple the physical storage technology from the application(s) and the file systems
  - OSD is the ultimate virtualization technology but it is a ***standard***
- Underlying storage technologies can evolve independently of the data that they store and the protocols that access them

Active Disks

CIPRICO

- Normal disks or storage devices only Read and Write data
- An Active Disk is actually a Storage Device that understands the ***content*** and ***structure*** of the data it manages
- An object can be:
  - A simple block of data
  - A meta-object that is a dynamic collection of other objects
  - A method or executable procedure
  - Any or all of the above
- Active storage devices can be Hybrid devices made up of disks, tapes, DVDR, RAMDISK, Flash memory, …etc.
- Hybrid Active Storage devices can store data based on performance, security, or other attributes

Example OSD Scenarios

CIPRICO

- Block-device Emulation
  - An object is simply a sector or block on a disk
- NAS Filer Emulation
  - An object is simply a file with the normal file attributes of name, size, permissions, and ownership information
- Active Storage
  - A storage device that understands the content and structure of the data it manages
  - Allows for implicit or explicit caching algorithms
- Real-time environments and Quality of Service
  - OSD devices can more accurately schedule delivery of objects since the stored objects have explicit QoS attributes
  - Bandwidth/IOPS/Latency allocation is more explicitly defined

Common OSD-like Examples

CIPRICO

- Digital “Appliances”
  - Digital Cameras
  - MP3 Players
  - CD/DVD Players
- Systems
  - Napster, Morpheus, …etc.
  - Protocols and standards: Corba, UML, XML, …etc.

Issues

CIPRICO

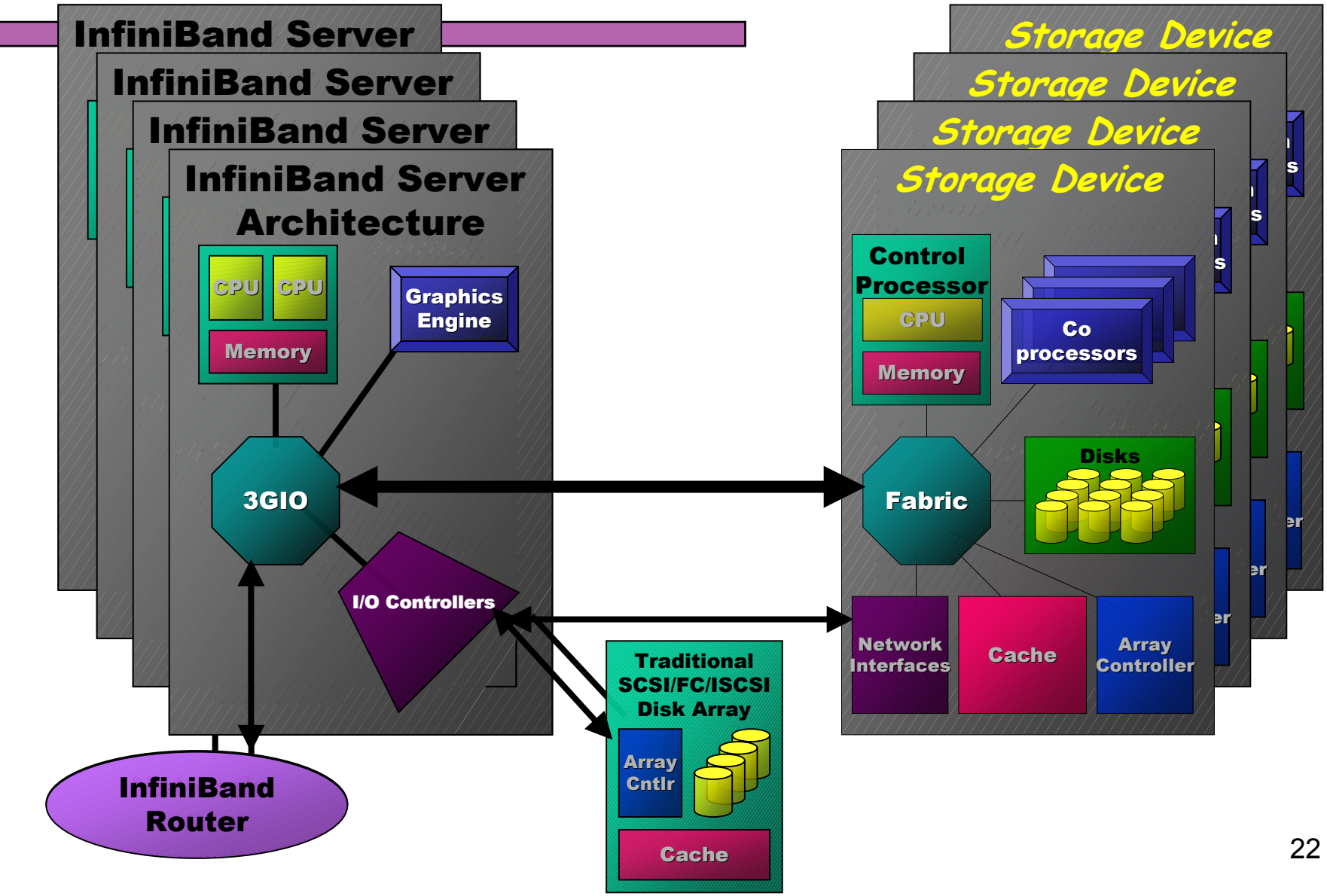
- Where is OSD implemented?
  - OSD on disk drives?
  - Disk arrays?
  - Removable media devices?
- Do you make a file system on top of OSDs or does OSD sit on top of a file system?
- Market Acceptance?
- How does OSD compete with ATA disks that are “good enough”?
- Support for legacy applications.
- Where does Microsoft fit into this picture?
- Where do the Software Application vendors fit into the picture?
- Where does Linux fit into this picture?
- Where do all the other OS vendors fit in?

Summary

CIPRICO

- Answer the original question: Is it crazy enough?
  - Yes it is – at least for the next 10 years.
- OSD is a building block focused on the storage devices. Higher level Object Oriented technologies must be applied in order to take full advantage of OSD's potential
- The benefits of OSD far outweigh the implementation issues
  - Can integrate with legacy system for a relatively smooth upgrade path
  - Allows for more "extensible" storage-centric systems
  - Provides enormous opportunities for product differentiation for storage vendors
  - Peer-to-Peer storage architectures are a natural by-product of OSD
  - It's cool.

CIPRICO


InfiniBand Server
InfiniBand Server
InfiniBand Server
InfiniBand Server Architecture
CPU
CPU
Memory
Graphics Engine
3GIO
I/O Controllers
InfiniBand Router
Traditional SCSI/FC/ISCSI Disk Array
Array Cntlr
Cache
Storage Device
Storage Device
Storage Device
Storage Device
Control Processor
CPU
Memory
Co processors
Fabric
Disks
Network Interfaces
Cache
Array Controller