



Education

# **MAID Technology Adoption into Data Centers**

Chris Santilli. COPAN Systems

- 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.
  - Neither the Author nor the Presenter is an attorney and nothing in this presentation is intended to be nor should be construed as legal advice or opinion. If you need legal advice or legal opinion please contact an attorney.
  - The information presented herein represents the Author's personal opinion and current understanding of the issues involved. The Author, the Presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.
- NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

## ➤ MAID Technology adoption into Data Centers

- ◆ Today the demand for more and more capacity is out-growing the ability to access the content. Or is it? The demand for consumer storage and home network is demonstrated by the 130+ million of disk drives shipped quarterly. The usage of the disk storage ranges from “transactional” to “archive/persistent” data. A new technology MAID (Massive Array of Idle Disks) provides for the power management of disks drives, thus creating a new service level for retaining and accessing “archive/persistent” data.
  - ◆ The Storage Administrator now has choices of storage platforms that match their data storage, retrieval and access requirements
- Refer to the FAQ for SNIA Tutorials “Information for Authors, Contributors, Presenters and Reviewers” at <http://www.snia.org/members/tutorials/faq>

## ➤ MAID Technology

- ◆ Technology for power-efficiency only...?
- ◆ What the benefits...
  - › Performance
  - › Reliability
  - › Capacity
  - › Power efficiency

## ➤ Use cases for MAID technology

- ◆ Transaction data versus Persistent data access models
  - › Persistence (as defined by wikipedia) - In [computer science](#), **persistence** refers to the characteristic of data that outlives the execution of the program that created it.
  - › Transactional (as defined by answers.com) - collection, storage, modification, and retrieval of [transactions](#) of an organization. A transaction is an event that generates or modifies [data](#) that is eventually stored in an information system.
- ◆ Applications that can utilize the MAID technology

# What is MAID?

- MAID (Massive Array of Idle Disks)
  - ◆ A storage system comprising an array of disk drives that are powered down individually or in groups when not required. MAID storage systems reduce the power consumed by a storage array.
- MAID References
  - ◆ SNIA Definition ([www.snia.org](http://www.snia.org))
  - ◆ Dennis Colarelli, Dirk Grunwald and Michael Neufeld, “The Case for Massive Arrays of Idle Disks (MAID),” Usenix Conference on File and Storage Technologies (FAST), January 2002, Monterey CA
- MAID Benefits
  - ◆ Power, Density, Address Space, Data Protection, Higher Reliability, Service Life
- MAID Design Parameters
  - ◆ Power, Bandwidth, Cost, Data Availability, Data Access Behavior, Serviceability

# Designing 1000-drive Disk Arrays

## ➤ 1000's of drives

- ◆ Connecting all drives
- ◆ Powering all drives
- ◆ Accessing all drives

## ➤ Access Patterns

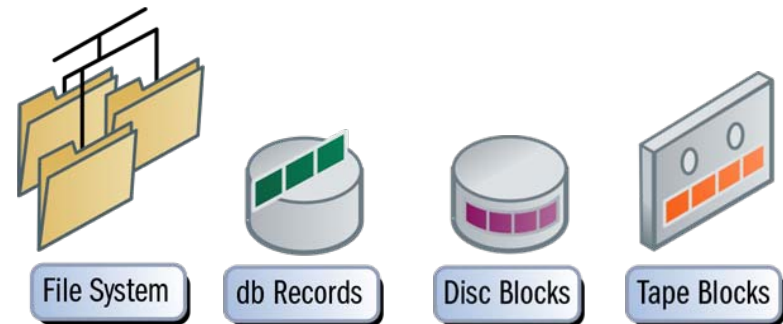
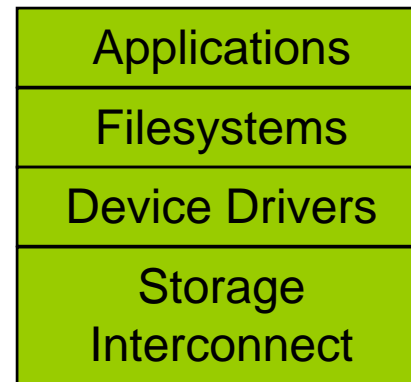
- ◆ Sequential Write/Read
- ◆ Random Write/Read
- ◆ Throughput
  - > MB/s/watt
  - > MB/s/U
  - > MB/s/\$

Attribute	Traditional	MAID
Controllers	1-2 per 30 drives	1-2 per 100 drives
Disk Bandwidth (seq = 170 MBps (rand = 200 IOPs)	5.4 GBps  6000 IOPs	18.0GBps  20,000 IOPs
Controller Bandwidth (dual port)	800MBps	800MBps
Capacities (1TB drives)	30TB	100TB

# Architecture Impacts with MAID

## ➤ Architecture Access Layers

- ◆ Applications (Backup/Restore, Archive, HSM)
- ◆ File Systems
- ◆ Drivers (Device, Failover, etc.)
- ◆ Storage Interconnect (FC, IP, etc.)
- ◆ Storage Media (Tape, CD, Fiche)



# Requirements for MAID Arrays

## ➤ Access Management

- ◆ Data objects require access models for long-term retention
- ◆ Access is not “micro” or “milli” seconds
  - > MTTFB and MTTLB
- ◆ Highest regard for storage and retrieval of objects over time (e.g. years)

## ➤ Data Management

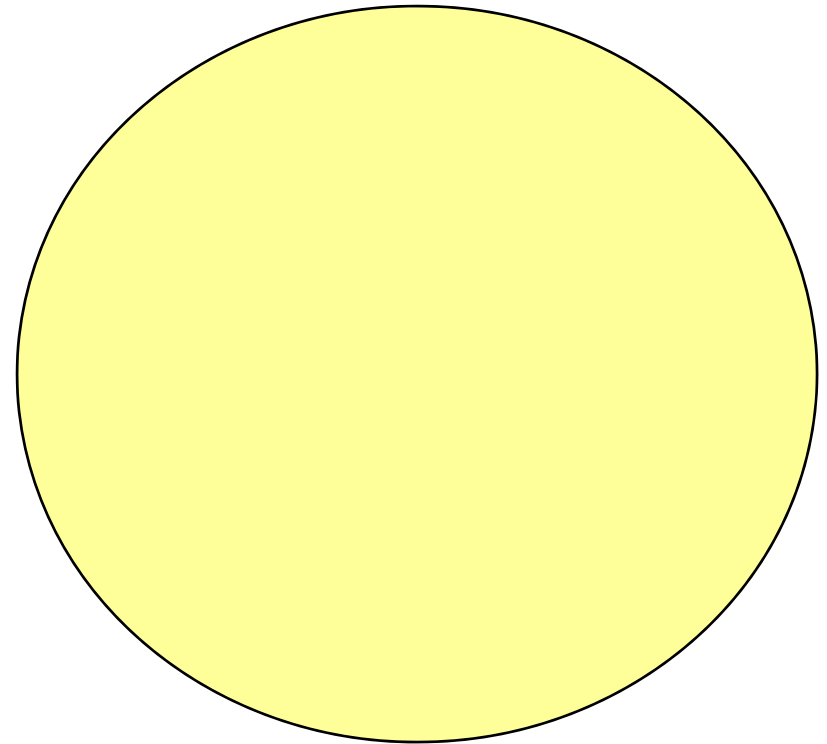
- ◆ Management of “massive” capacities
  - > Number of volumes, Number of Virtual Tapes, Number of Files

## ➤ Device Management

- ◆ Management of device utilization and allocation of appropriate resources during PEAK access
- ◆ Management of storage array resources (processors, cache, power, cooling, cables)

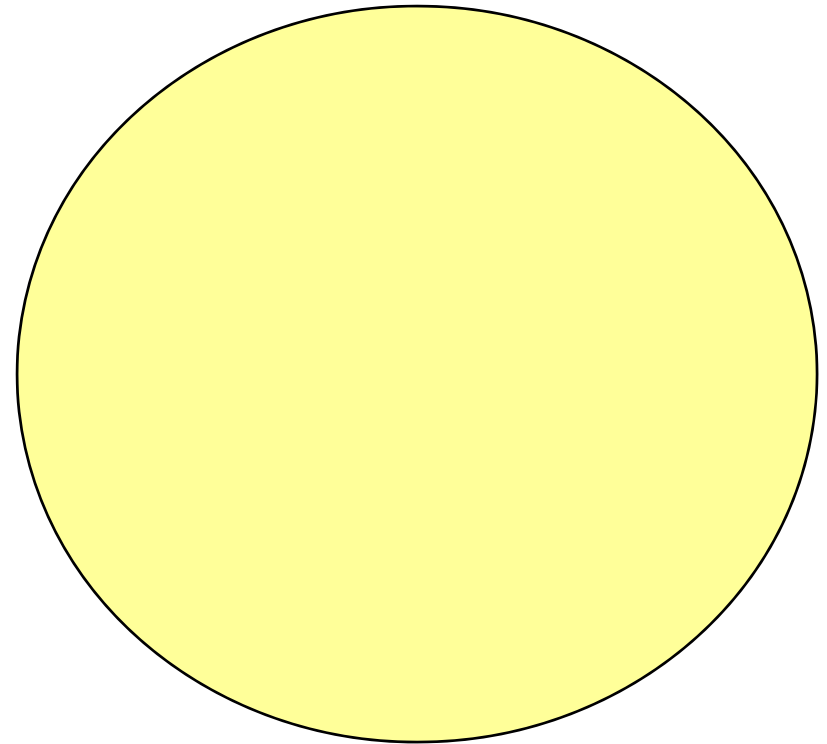
# MAID Array Approaches

- Access Management
  - ◆ Match Archiving and Backup/Restore Workloads (MTTFB, MTTLB)
- Data Management
  - ◆ Increase the reliability, availability and serviceability of traditional disk array subsystems
- Device Management
  - ◆ Catalog more efficient than tape (retrievals of data objects)
  - ◆ Monitor both active and inactive data (power on hours, power off hours)



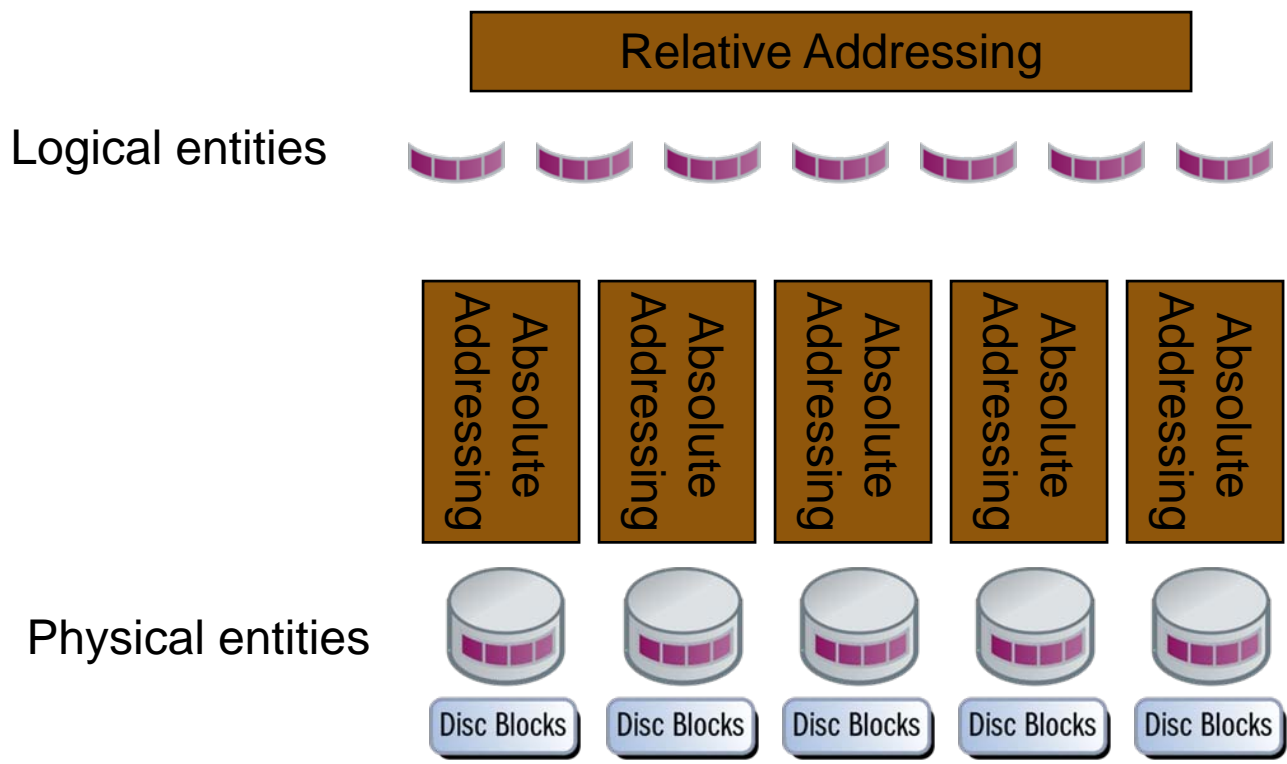
# Disk Array Approaches

- Access Management
  - ◆ Provides as a “one size” fits all.
  - ◆ Provided with a cache and tunable parameters to match applications needs (IOPS/MBps)
- Data Management
  - ◆ All resources (disk,power,cooling,cables, etc.) are dedicated to the read/writing (RAID) of data
- Device Management
  - ◆ Mapping of devices at discovery and at device failure



# Relative vs. Absolute Addressing

➤ When creating “massive” capacities... aggregation of “physical” devices



Why important?

The cost to  
Access data  
Power,  
Processor,  
Cables,  
Disks

# Power Management for Storage

## ➤ Power needed

- ◆ Create the data
- ◆ Update/Modify the data
- ◆ Access the data
- ◆ Maintain the data

## ➤ Types of storage (transactional/persistent)

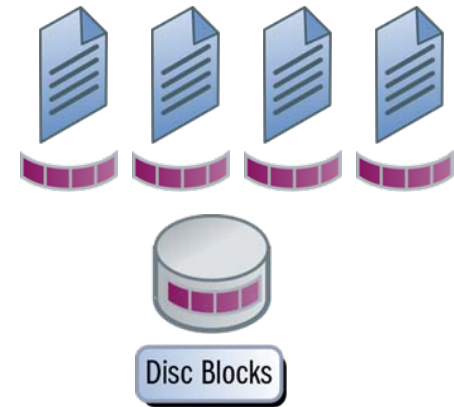
- ◆ DRAM, Flash Memory
- ◆ Disk, Tape
- ◆ CD, Optical, Fiche

# Data storage workloads

- Data Access
  - ◆ ALL of data or SOME of data
  - ◆ Modify SOME or ALL of the data (new copy)
  - ◆ Access it every SECOND or once a MONTH
- Traditional approach for Disk Arrays are for ALL data accessible ALL the time
- Data objects
  - ◆ Creation – new data
  - ◆ Modification – append, modify existing
  - ◆ Lock – allow no more modifications

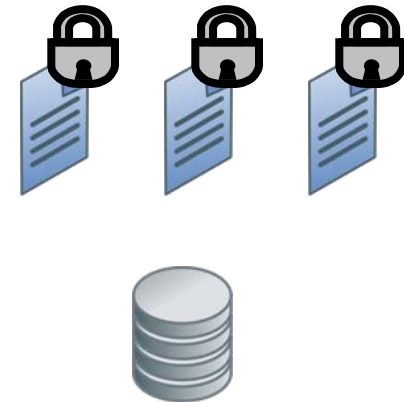
## ➤ Data is stored in objects

- ◆ Creation of data
  - > Byte access
- ◆ Retrieval of partial objects
  - > Byte access
- ◆ Modification of partial objects
  - > Byte access



## ➤ Objects retrieved

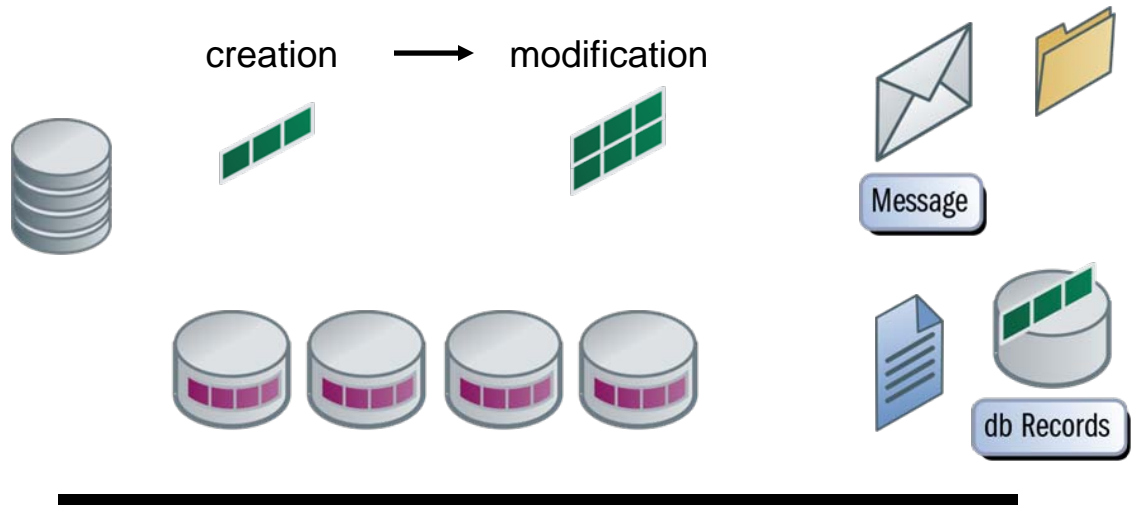
- ◆ Full objects retrieved
  - > Object access
- ◆ New versions updated
  - > Object access



# Access Models

## ➤ Transactional Access Model

- ◆ Creation of objects
- ◆ Modifications of objects



## ➤ Persistent Access Model

- ◆ Full objects stored
- ◆ Full/partial objects retrieved
- ◆ Updates are full object revisions



# Peak Access to Massive Capacities

## ➤ Access Models

- ◆ Transactional Models

- > Transactions workloads to disk that utilize power management show no power savings
    - Research references : (<http://csl.cse.psu.edu/publications/ispass03.pdf>)

- ◆ Persistent Model

- > MTTFB – Mean Time to First Byte
    - Time to from the request of the object/file to the transmission of the first byte
  - > MTTLB – Mean Time to Last Byte
    - The time from the first byte transmitted to the end of the last byte requested

Examples	Size of file/object	Time to complete request	
Example 1: 10s MTTFB @ 100MBps	1 GB	20s	(10s + 10s)
	10 GB	110s	(10s + 100s)
Example 2: 10s MTTFB @ 1000 MBps	1 GB	11s	(10s + 1s)
	10 GB	20s	(10s + 10s)
Example 3: 1s MTTFB @ 100MBps	1 GB	11s	(1s + 10s)
	10 GB	101s	(1s + 100s)

# System Reliability

- What contributes to the MTBF of a Disk Array Storage Enclosure?
  - ◆ Historical data has pointed to environmental parameters (e.g. heat, vibration, EMI, RFI, etc.)
  - ◆ What about workload factors?
- Workload factors of Disk Array Storage Enclosure
  - ◆ Are there “wear” factors that need to be considered?
  - ◆ Are there frequency of access, modification, updates that are considered?
  - ◆ Access of full/partial objects impact disk behavior
- System reliability and longevity aid in the archives for many years
  - ◆ SNIA Tutorial “ Retaining Information for 100 Years”

# User Applications

- Backup Restore
  - ◆ Instantaneous, Disaster Recovery
- Archive
  - ◆ Structured data
    - › Databases, Records
  - ◆ Unstructured data
    - › Filesystems, Tiers of Storage
- Ubiquitous to DAS, SAN, NAS data objects

# Complementary Technologies

- Searching
  - ◆ Indexing, Tables, Sequential Access
- Catalog
  - ◆ Metadata for locality, Metadata on content
- Capacity Optimizations
  - ◆ Deduplication (block, file)
  - ◆ Compression (block, file)
- Migration
  - ◆ Staging, Replication, Transformation of data
- Integrity
  - ◆ Data Scrubbing, Redundancy

# Future Power Efficient Technologies

- Power managed Disk Drives
  - ◆ Power improvements to the individual disk drives
- Use of Flash Memory for high density and lower power consumption
  - ◆ Solid State Disks
    - SNIA Tutorial “Solid State Storage for Enterprise”
  - ◆ Extended memory in servers
    - Use of DRAM and Flash memory in the same server
- Disk array enclosures to utilize DC power distribution
  - ◆ SNIA Tutorial “Building the Green Data Center: Towards Best Practices and Technical Considerations”

# Summary

- Data Centers are challenged with the growth of new data and the retention of data for longer periods of time
- Need an approach to data capacities, data access that do not “LINEARLY or EXPONENTIALLY” increase the cost and data center floor space
- **Different** types of data objects have ... **Different** storage platform needs

- Please send any questions or comments on this presentation to SNIA: [trackstorage@snia.org](mailto:trackstorage@snia.org)

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

**- SNIA Education Committee**

**Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here**

**Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here  
Name of contributor here**