

A decorative graphic consisting of multiple overlapping, wavy lines in shades of purple, blue, orange, and green, flowing from the left side of the page towards the right.

What's old is New Again – Storage Tiering

Gideon Senderov, NEC

Author:

SNIA - Data Protection & Capacity Optimization (DPCO) Committee

- ◆ The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.
- ◆ Member companies and individual members may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced in their entirety without modification
 - ◆ The SNIA must be acknowledged as the 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, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- ◆ The information presented herein represents the author's personal opinion and current understanding of the relevant 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.

About the SNIA DPCO Committee

- This tutorial has been developed, reviewed and approved by members of the Data Protection and Capacity Optimization (DPCO) Committee which any SNIA member can join for free
- The mission of the DPCO is to foster the growth and success of the market for data protection and capacity optimization technologies
 - ◆ Online DPCO Knowledge Base: www.snia.org/dpco/knowledge
 - ◆ Online Product Selection Guide: <http://sniadataprotectionguide.org>
- 2014 goals include educating the vendor and user communities, market outreach, and advocacy and support of any technical work associated with data protection and capacity optimization

Check out these **SNIA-DPCO Tutorials** at www.snia.org/education/tutorials



- **Data Protection in Transition to the Cloud**
- **Protecting Data in the Big Data World**
- **Understanding Data Deduplication**
- **Deduplication's Role in Disaster Recovery**

➤ Storage Tiering

- ◆ Physical tiering of storage has been possible for decades, but has seen limited adoption
- ◆ New interest in tiering has arisen due to new techniques and technologies that automate the movement of “hot” data to high performance storage tiers and “cool” data to low-cost tiers
- ◆ Other factors that influence tiering include: performance, capacity optimization, availability, economics and compliance
- ◆ This session will include:
 - › Fundamentals of storage tiering
 - › The use of the cloud as a tier(s)
 - › Recent innovations in tiering techniques

“It’s almost impossible to conduct a conversation with a storage vendor without raising the topic of “storage tiering” ... While there is tremendous user value to be gained from employing storage tiering, there is also at times a lack of clarity and of semantic accuracy in the market”

Mark Peters

Enterprise Strategy Group

What is Storage Tiering?



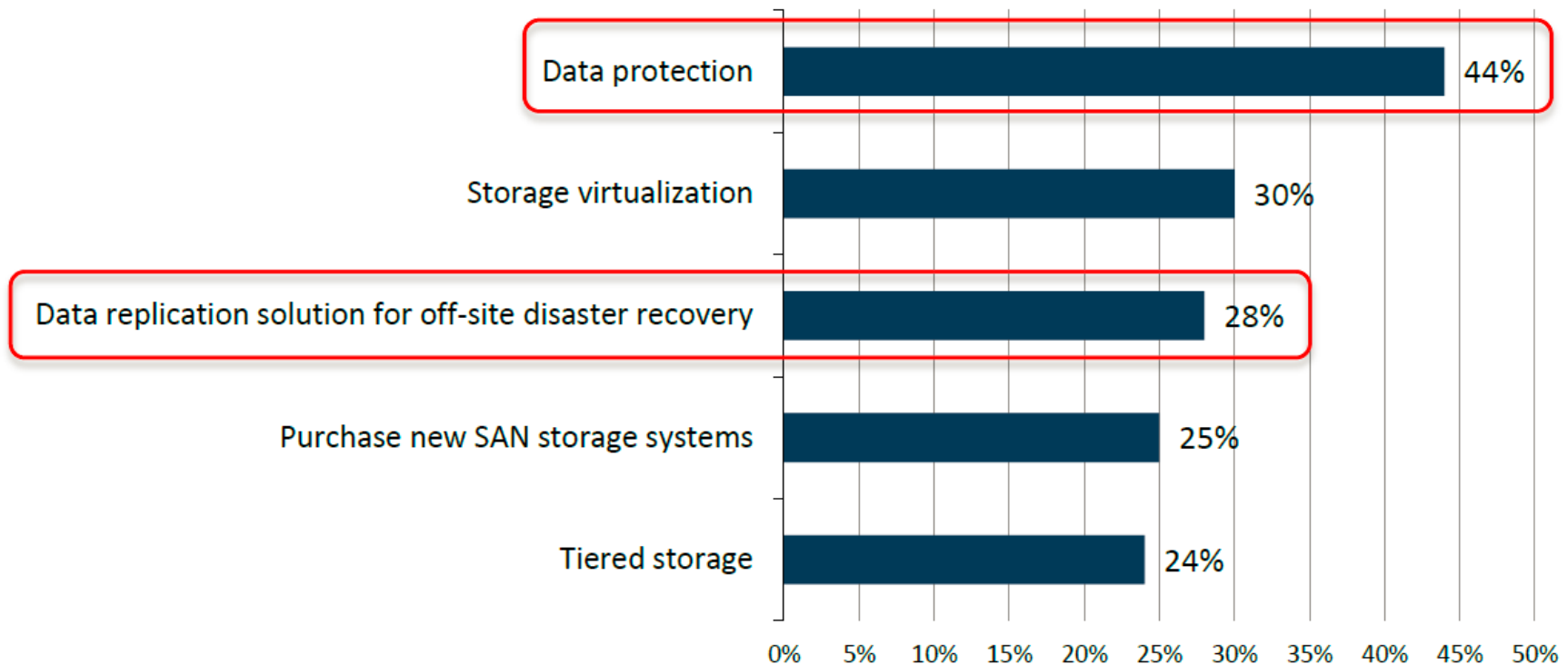
From the SNIA Dictionary:

- ▶ **Tiered storage [*storage System*]:** Storage that is physically partitioned into multiple distinct classes based on price, performance, or other attributes.

Data may be dynamically moved among classes within a tiered storage implementation based on access activity or other considerations.

Storage Infrastructure Spending

In which of the following data storage infrastructure areas will your organization make the most significant investments over the next 12 months?
(Percent of respondents, N=296)



Source: Enterprise Strategy Group, 2013.

➤ Class by Storage Technology

- ◆ Nand FLASH or DRAM for SSDs
- ◆ High Performance HDDs
- ◆ High Capacity HDDs
- ◆ High Capacity Magnetic Tape or Optical Media

➤ Class by Data Criticality

- ◆ Mission Critical
- ◆ “Hot” (warm) Data
- ◆ “Cool” Data

The Storage Hierarchy

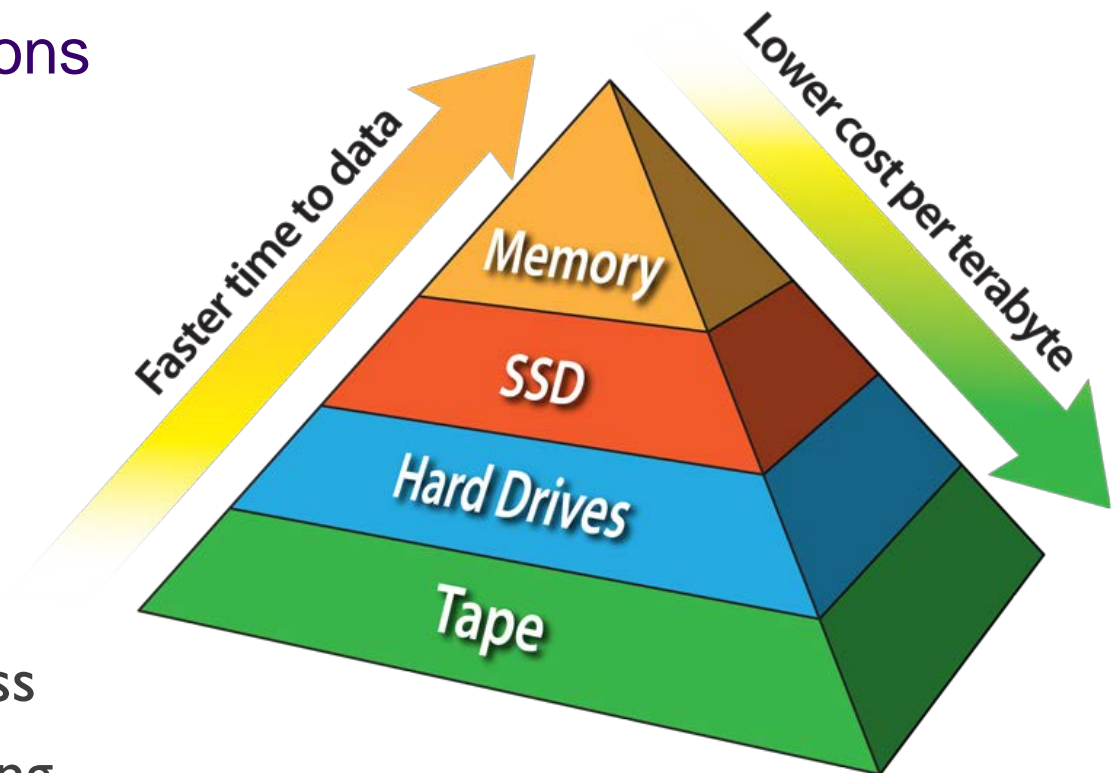
➤ Price / Performance tradeoffs

➤ Other Considerations

- ◆ Persistence
- ◆ Access Method
- ◆ Capacity
- ◆ Data Endurance

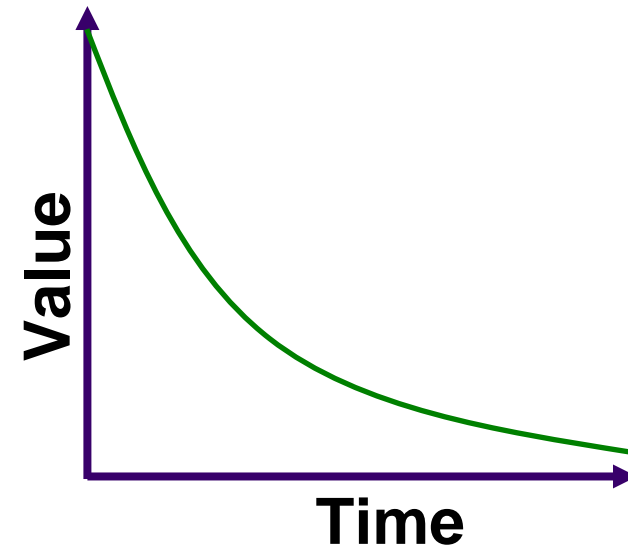
➤ Data Access

- ◆ In-memory access
- ◆ Caching / buffering
- ◆ Local / cloud tiering



Why is Tiering Important?

- All data is not created equal
 - ◆ Data structure and access profiles differ
- Data's value is dynamic over time
 - ◆ New data often loses value quickly
 - ◆ Older data may suddenly have high demand
- Data growth isn't showing signs of slowing down
 - ◆ One of the top IT pressures is meeting the increasing demands for storage
 - ◆ A "one-size-fits-all" data storage strategy will not scale long-term



Tiering Considerations

➤ Users

- ◆ Want everything on Tier 1 with the highest performance
- ◆ Don't typically pay for the storage they consume

➤ Business owners

- ◆ Want the data on the lowest tier for cost savings (if paying...)
- ◆ Often don't understand the performance difference between tiers

➤ SLAs – between business owners/IT and users

Tiering Considerations (Cont.)

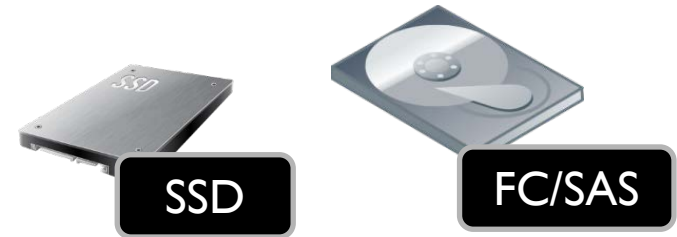
- ◆ Why not compromise and put everything on Tier 2 or 3?
 - ◆ Would be easier to administer, but...
 - ◆ The storage may not meet the needs of the data (applications)

- ◆ Data isn't always predictable
 - ◆ Some traditionally "Tier 1" data can reside on Tier 2 storage
 - ◆ Data generally drops tiers as it ages, but may become relevant again

Tiering Choices (Examples)

➤ Tier 1

- ◆ SSDs (sometimes called “Tier 0”)
- ◆ 10K & 15K RPM FC/SAS disk drives
- ◆ Mostly RAID 10 and some RAID 5



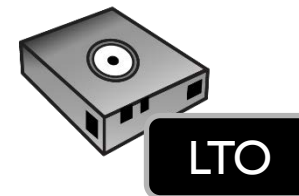
➤ Tier 2

- ◆ 7.2K RPM SATA/SAS disk drives
- ◆ RAID 5 and RAID 6



➤ Tier 3

- ◆ Automated tape storage
- ◆ Predominantly LTO today



➤ Public/Private cloud

- ◆ Outsourced off-site storage of “cool” data



Quick History of Storage Tiering

➤ 1980s

- ◆ Mainframe DFHSM (Data Facility Hierarchical Storage Manager)
 - Online disk was cost prohibitive
 - Automatic staging and retrieval from tape

➤ 1990s

- ◆ HSM (Hierarchical Storage Management)
 - Files migrated to dedicated HSM server
 - Policy engine determined what files were migrated



➤ 2000s

- ◆ ILM (Information Lifecycle Management)
 - Concept instead of product
 - Included compliance concerns

Tiering: Generations over Time

- ◆ Some common items across generations:
 - ◆ Freeing up more expensive storage
 - ◆ Returning data to original tier when used
 - ◆ Performance and availability considerations

➤ Automated tiering

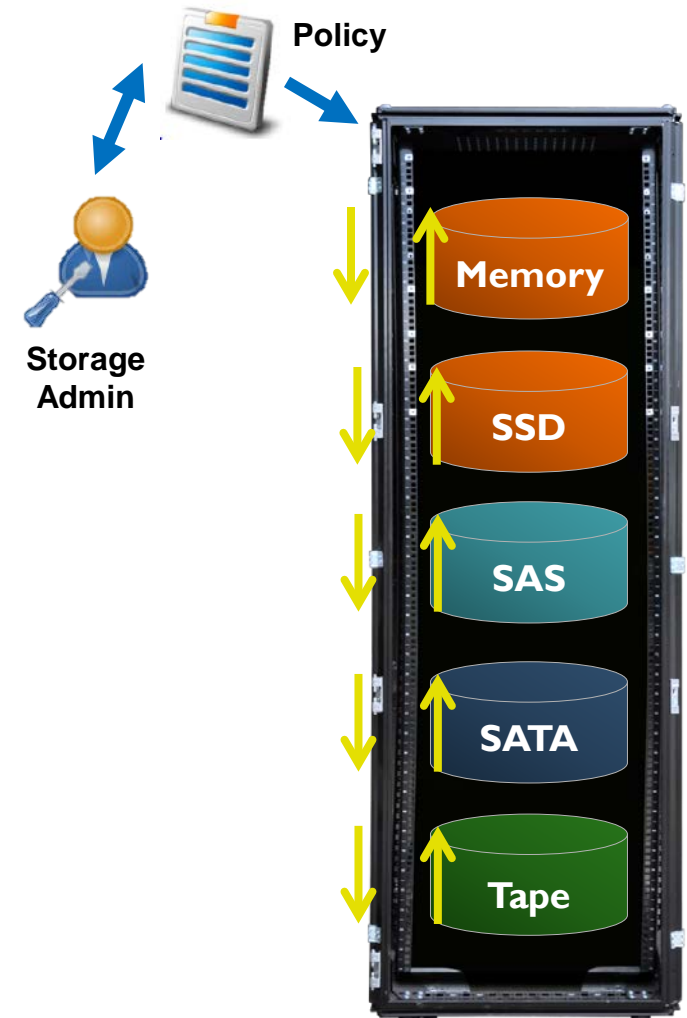
- ◆ File and sub-file migration performed transparently between tiers
- ◆ Policies and/or predictive analysis governed migration

➤ Cloud tiering

- ◆ Pushes “cool” data to the cloud (from a data protection perspective)
- ◆ Can use cloud-based storage as a backup tier

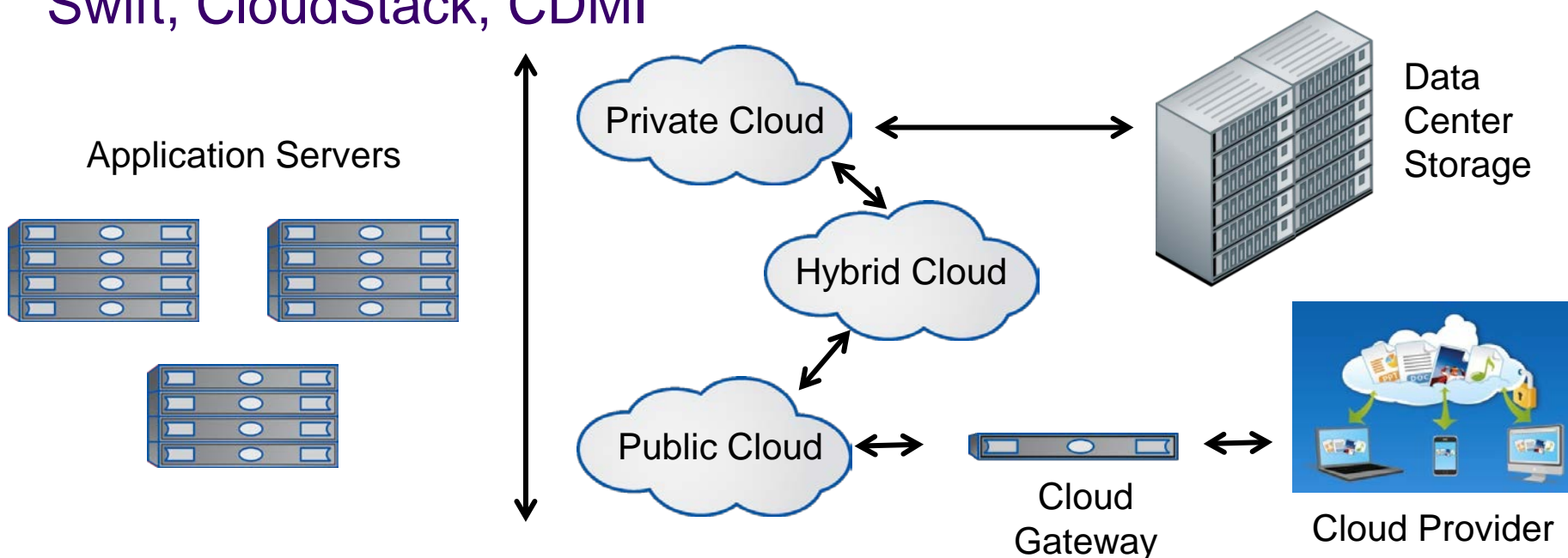
Automatic Tiering

- Data is migrated between tiers based on policy due to data usage patterns, e.g., time-based or file type
- Tiering can be on a single system or across multiple dissimilar systems
- Makes it more economical to deploy SSD storage



Cloud Tiering

- Cloud gateways are available for direct block and file access
- Cloud is being used for offsite backup and archive tiers
- Hybrid Clouds combine public & private clouds using a common set of data management protocols, e.g., OpenStack Swift, CloudStack, CDMI



Cloud Tiering: Pros & Cons

➤ Pros

- ◆ “Unlimited” scale
- ◆ Reduced data center foot print
- ◆ No physical storage management

➤ Cons

- ◆ Large data sets and bandwidth
- ◆ Security concerns
- ◆ Exit and migration
- ◆ Access and Control

Questions to Ask During Selection

❖ Tiering Scale and Granularity

- ◆ Is data tiered within a storage system – or across systems?
- ◆ Is data tiered at the file or sub-file/block level?

❖ Tiering Policies and Methods

- ◆ Reactive, Predictive or Realtime?
- ◆ How often is the data re-tiered?
- ◆ Can the Tiering policy be adjusted?

❖ Tiering Costs

- ◆ Total Software and Hardware costs?
 - › Purchase
 - › Implementation
 - › On going costs; operating costs, support
- ◆ Future Scaling costs?

- Knowing your objective and what questions to ask will make a tiered storage deployment more successful
- Today's tiering technologies are built on prior generations
 - ◆ Adopting concepts from HSM and ILM
 - ◆ But with significant differences
- New technologies and storage systems are bringing new capabilities

The SNIA Education Committee thanks the following individuals for their contributions to this Tutorial:

Authorship History

Data Protection and Capacity Optimization
(DPCO) Committee, March, 2012

Updates:

Data Protection and Capacity Optimization
(DPCO) Committee, September, 2012

Data Protection and Capacity Optimization
(DPCO) Committee, May, 2013

Data Protection and Capacity Optimization
(DPCO) Committee, August, 2013

Data Protection and Capacity Optimization
(DPCO) Committee, August, 2013

Additional Contributors

Ashar Baig
David A. Chapa
Kevin Dudak
Mike Dutch
Larry Freeman
David Hill
Gene Nagle
Ronald Pagani
Thomas Rivera
Tom Sas
Gideon Senderov
Alan Yoder

Please send any questions or comments regarding this SNIA Tutorial to tracktutorials@snia.org