



Education

What's Old is New Again – Storage Tiering

Kevin Dudak, Spectra Logic

- 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.

➤ 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.
- ◆ The session will include:
 - › Fundamentals of Storage Tiering
 - › Recent innovations in tiering techniques
 - › The cloud as a tier

- The SNIA Data Protection and Capacity Optimization (DPCO) Committee was created to foster the growth and success of the market for data protection and capacity optimization technologies.



<http://www.snia.org/forums/dpco>



Check out **SNIA DPCO Tutorials:**

- **Trends in Data Protection and Restoration Technologies**
- **Advanced Data Reduction Concepts**

“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

July 12, 2011

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.

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.

➤ Class by Storage Type

- ◆ SSD
- ◆ High Performance HDD
- ◆ High Capacity HDD
- ◆ High Capacity Tape

➤ Class by Data Type

- ◆ Mission Critical
- ◆ Hot Data
- ◆ Cool Data

The Storage Hierarchy

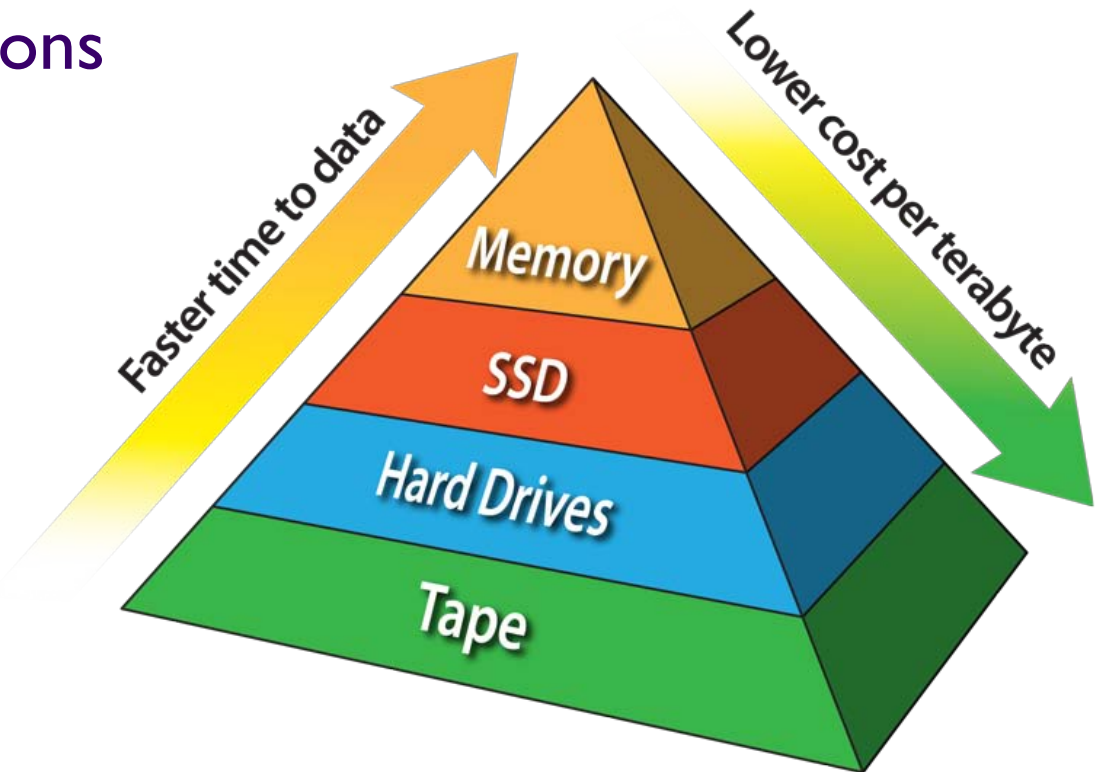
➤ Price/Performance tradeoffs

➤ Other Considerations

- ◆ Persistence
- ◆ Access Method
- ◆ Capacity

➤ Data Access

- ◆ In memory access
- ◆ Caching/buffering
- ◆ Tiering



Why is Tiering Important?

- All data isn't 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 any 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 does will not scale forever

➤ Users

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

➤ Business Owners

- ◆ Want data on the lowest tier for cost savings
- ◆ Often don't understand the difference in performance between tiers

- Why not compromise and put everything on Tier 2 or Tier 3?
 - ◆ It would be easier to administer
 - ◆ The storage would not meet the performance match the performance needs of the data

- 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

Traditional Tiers

➤ Tier 1

- ◆ 10K and 15K FC/SAS Disk Drives
- ◆ Mostly RAID 10 and some RAID 5



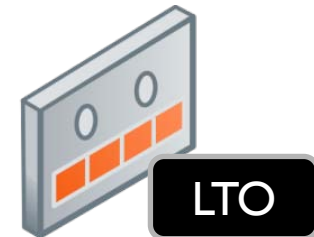
➤ Tier 2

- ◆ 7.2K SATA/SAS Disk Drives
- ◆ RAID 5 and RAID 6



➤ Tier 3

- ◆ Automated Tape Storage
- ◆ Predominantly LTO today



➤ Tier 0 (?)

- ◆ SSD Drives



➤ Public Cloud

- ◆ Outsourced off site storage of cool data



➤ 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



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

Each generation built on the ones before.

Some common things across generations include

- Freeing up more expensive storage
- Returning data to original tier when used

➤ Automated Tiering

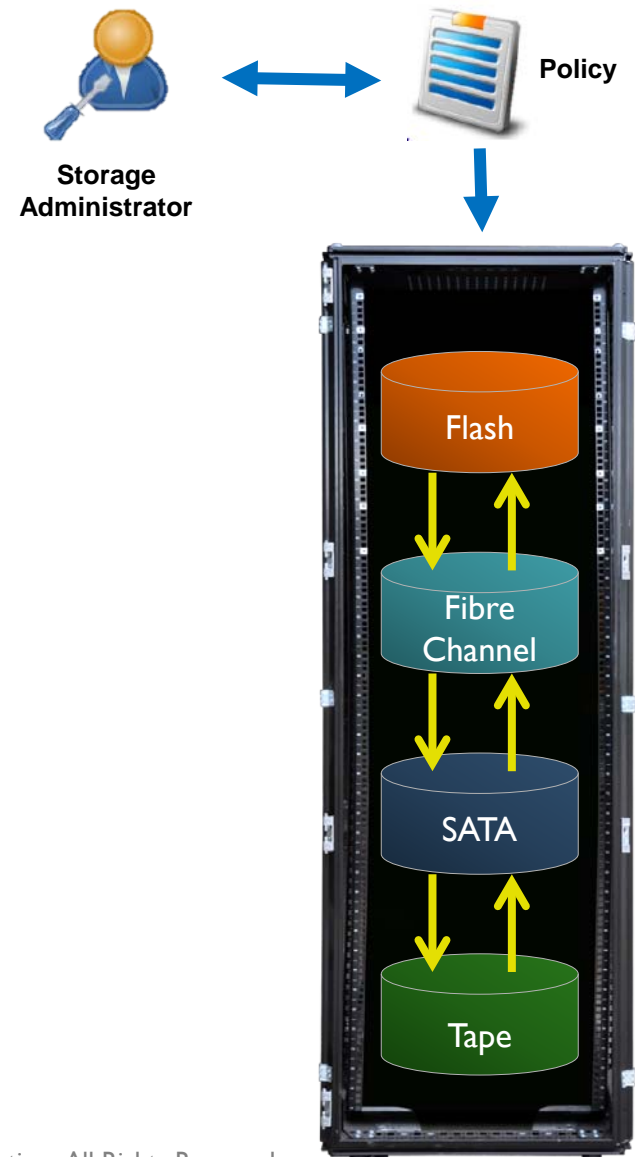
- ◆ File and sub-file migration performed transparently between tiers
- ◆ Used policies and/or predictive analysis for migration

➤ Cloud Tiering

- ◆ Pushes cool data to the cloud
- ◆ Can use public cloud as backup tier

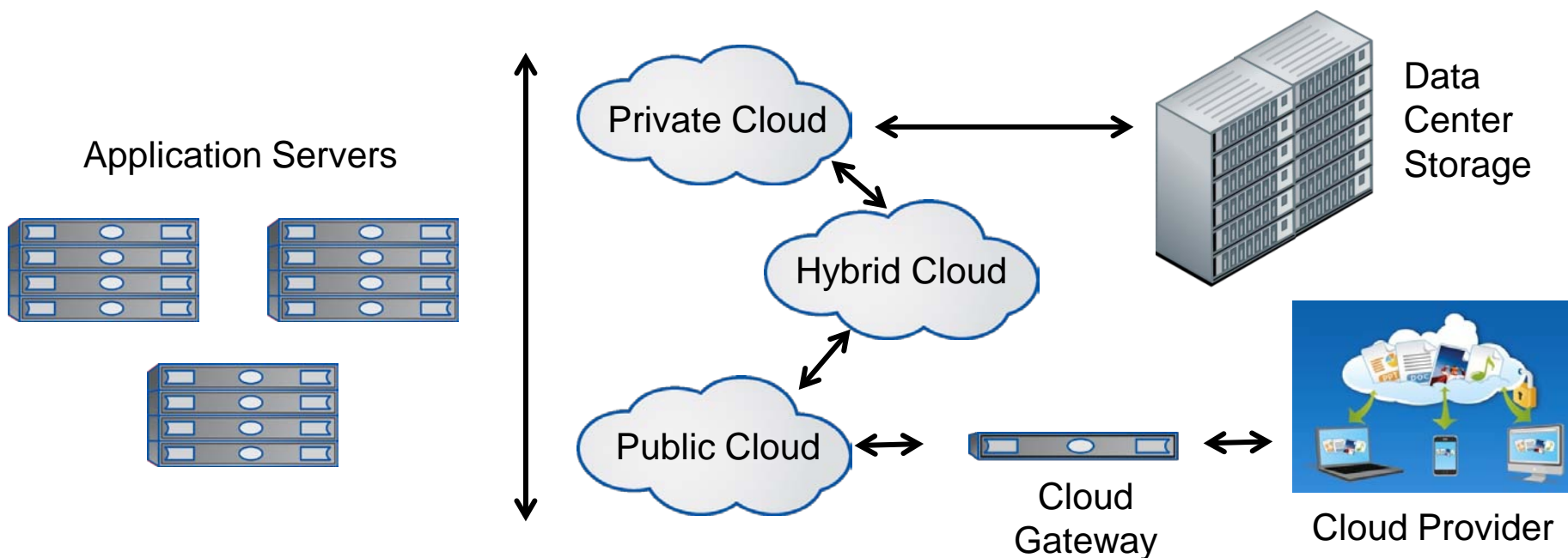
Automatic Tiering

- Data is migrated between tiers based on policy and data usage patterns
- Tiering can be on a single system or across multiple dissimilar systems
- Makes it more economical to deploy SSD storage
- Will high performance HDDs still be needed?



Cloud Tiering

- Cloud Gateways are developing for direct block and files access
- Cloud is being used for offsite backup and archive tiers
- Hybrid Clouds will combine public and private clouds using a common data management protocol (CDMI)



Cloud Tiering

➤ Pros

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

➤ Cons

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

➤ Tiering Scale and Granularity

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

- 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

- Knowing your objective and what questions to ask will make a tiered storage deployment more successful

The SNIA Education Committee would like to thank the following individuals for their contributions to this Tutorial.

Authorship History

DPCO, Spring 2012

Updates:

DPCO, Fall 2012

Additional Contributors

Kevin Dudak

Mike Dutch

Larry Freeman

Gene Nagle

Thomas Rivera

Tom Sas

Alan Yoder

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