Rightsizing Tiered Storage Systems

Octavian Paul ROTARU, PhD
ACMS

SNIA Legal Notice

- 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.

Abstract

A multi-tiered storage system with automated data movement provides the best solution for managing the data explosion IT is experiencing. While tiered storage strategies can cut enterprise data storage costs and address storage capacity issues, rightsizing the storage tiers is a difficult exercise in many environments.

The purpose of this lecture is to go over the commonly used tiering estimating algorithms and methods (usually based on IO skew calculation) and explain their shortcomings in different workload contexts (cyclical data workloads specific to telecom industry, high performance workloads, etc.), as well as propose a new storage tiering estimation methods which attempts to solve these issues and provide more accurate estimates.

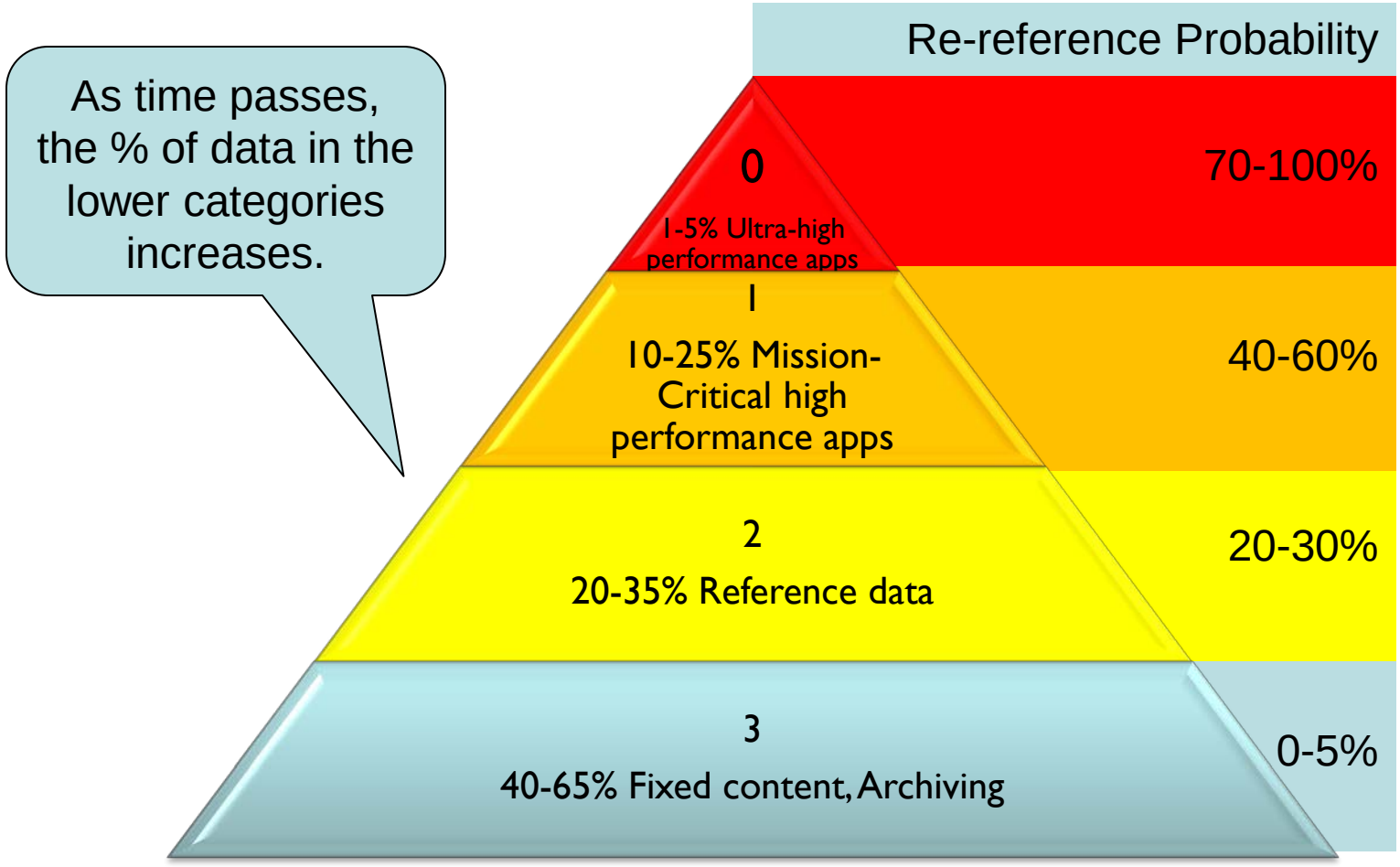
Unpredictable rate of storage growth and fluctuations in data rates often lead to performance issues. Automated storage tiering software can solve this problem and optimize storage allocation for performance if the sizing of the tiers is estimated right.

The impact of the following factors that are usually overlooked on the tiering mix will be discussed in this lecture: data movement speed and tiering overheads, storage based replication, snapshots and clones, IO size, Sequential vs. Random IO, etc.

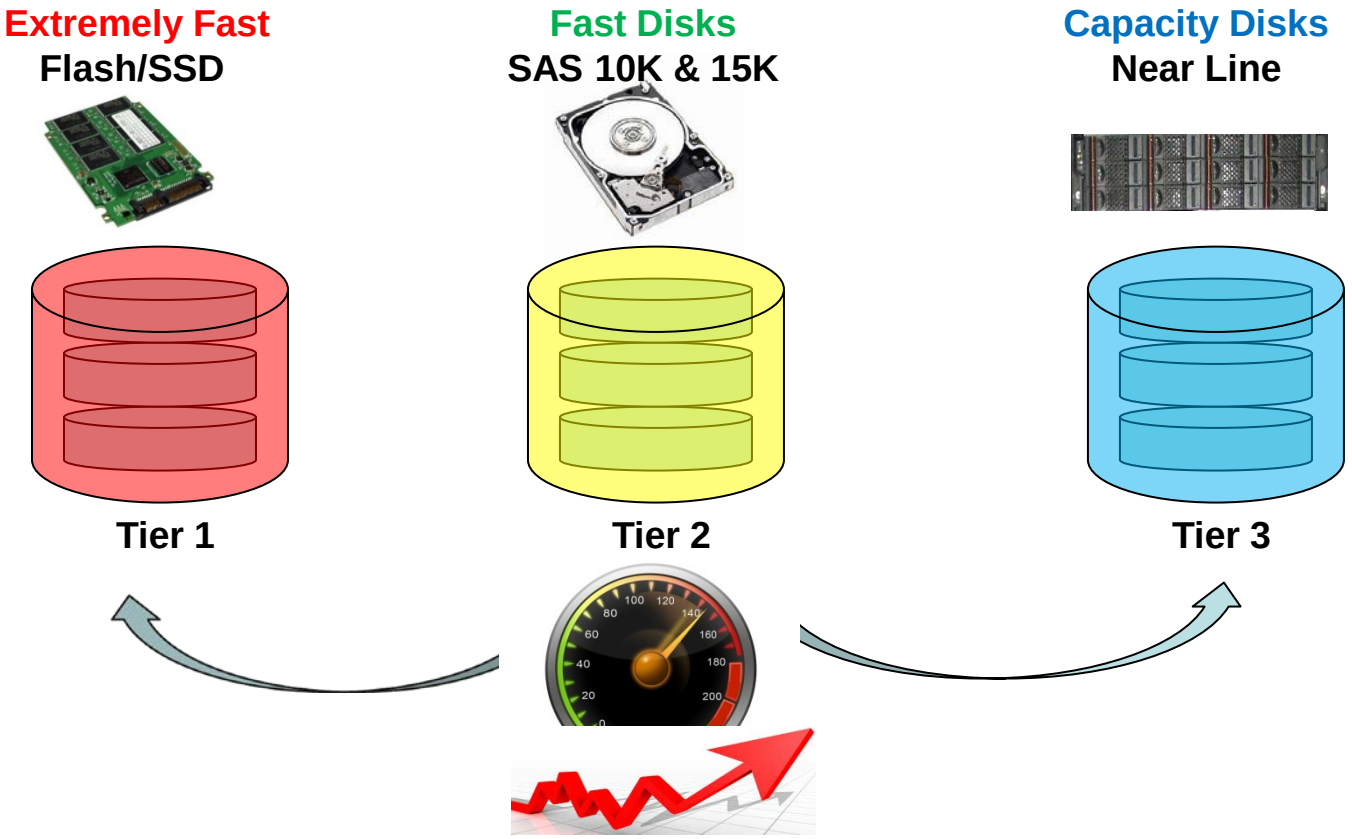
What is a Storage Tier?

- A tier is a type of storage media with certain performance characteristics
- Tiered storage is physically partitioned into multiple distinct classes based on price, performance, or other attributes
- Automated storage tiering moves data dynamically among classes/tiers within a tiered storage configuration based on access activity or other considerations

Data Classification



Automated Storage Tiering



Why use tiering?

- Automated Tiering Software (ATS) dynamically moves data between different disk types and RAID levels to meet space, performance and cost requirements
- For most of the IT departments tiering is seen as a cost saving method – used to reduce overall cost requirements for storage
- In the same time, tiering is perceived as an effective way to meet performance needs by automatically placing each segment of data on the type of storage that meets its needs

Storage sizing is a balancing game

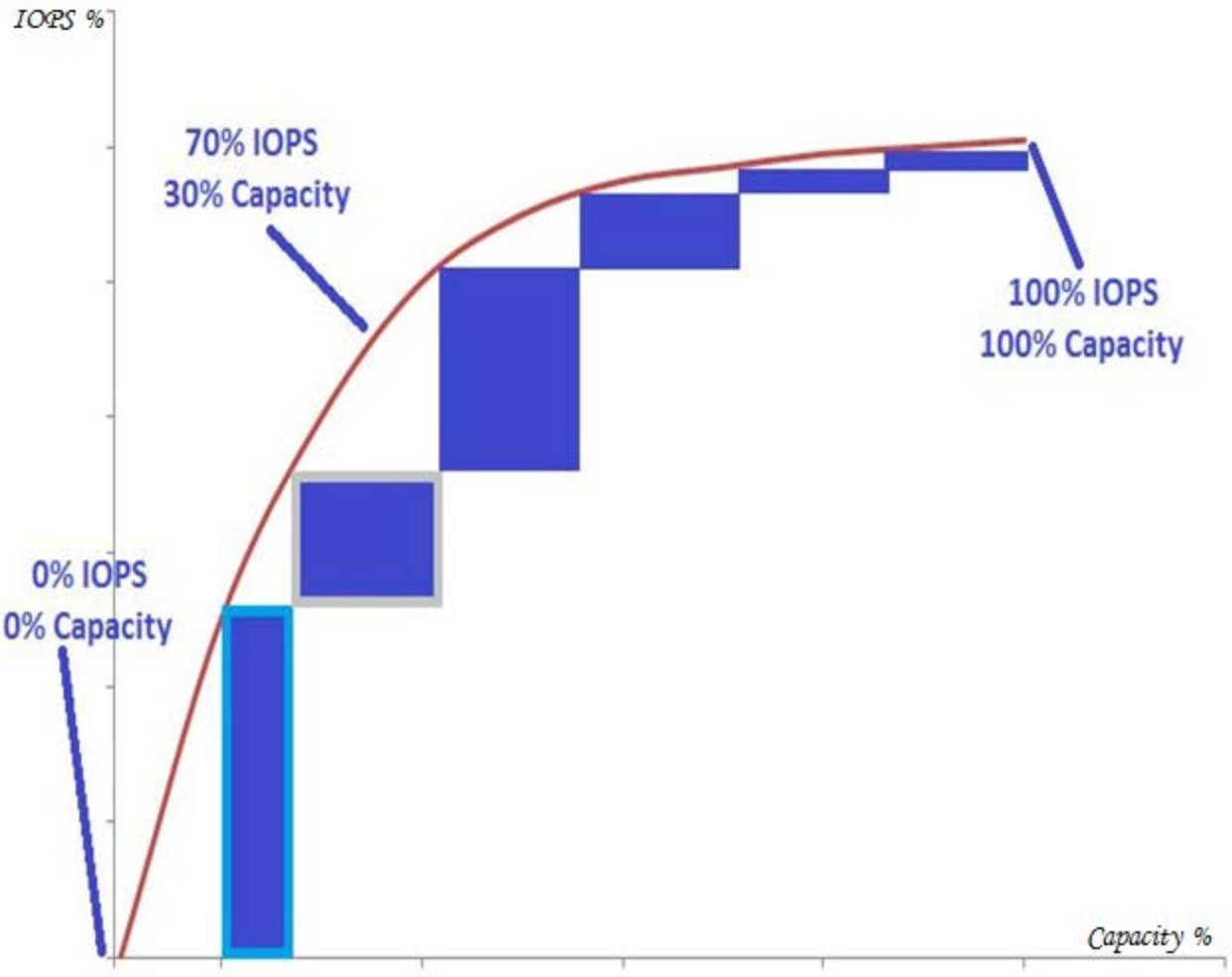
- Store too much data on too few spindles → Performance Issues
  - No matter what its capacity is, a disk will be able to handle a fixed number of IO requests per second
  - IO capabilities of a spindle disk are given by its type (15K rpm, 10K rpm, 7.2K rpm), seek time and latency
- Store too little data on many disks → Cost Impact
  - Space must be kept unused in order to match the IO requirements
- How do we do it right?

Data Skew

- There is a point on the data distribution curve where:

  IOPS % + Capacity % = 1

- This is what we call the skew point
- The skew point is the point on the curve where the tangent has the value 1 (the tangent intersects the x-axis at a 45 degrees angle)
- In this example the skew point is at 70%, meaning that 70% of the IOPS is generated by 30% of the data



Data Skew Continued

- If the skew point is at 50%, this means that all data is equally active – and tiering makes no sense in this context
- The higher the skew point is, the more effective tiering will be
- If the data distribution and the associated skew point are known, then the right mix of storage tiers can be easily chosen
- The problem is that in most of the cases the IO distribution for the data is unknown

IO Density

- IO Density = IO Capabilities / Capacity
- IO Density = IOPS / GB
- 300 GB 15K SFF disk → IO Density = 0.55
  - A 15k rpm SFF disk is capable of up to 190 IOPS
  - When disk utilization is too high the response time escalates rapidly and is no longer predictable – so we will use 150 IOPS maximum per 15K SFF drive (maximum 80% utilization)
  - 300 GB = 274 GiB
  - IO Density for 1 such disk: 150/274 = 0.55
- 146 GB 15K SFF disk → IO Density = 1.12
- 600 GB 15K SFF disk → IO Density = 0.28

How do vendors size tiered configurations?

- Most vendors rely on calculating the data skew to estimate the tiering mix
- Data skew calculation is an empirical method of estimation that works in many cases, but not in all
  - Calculating data skew cannot be done without having a good understanding of the data – and this requires input from the end-users
- In most of the cases the IO distribution of data is not known, and at best is estimated – which usually results in unrealistic numbers for data skew (usually too high)
- For most vendors, a high data skew number is an indication that more cheap storage can be used – and as a result making the price more attractive.

When does automated tiering fails?

- Insufficient knowledge of the data and its usage patterns
- Cyclic Data
- Overutilization
- Over-optimistic estimation

Insufficient knowledge about data

- Happens more often in green field installations when a tiered configuration is being purchased without knowing how the workload profile will evolve
  - The tiering configuration is chosen based on a hypothetical workload profile and skew point – which proves wrong
  - In green field projects it is usually recommended to start without tiering – only with Tier 2 (middle tier) storage, and add tiering at a later stage when data is growing and the workload profile can be analyzed
- Unfortunately this can also happen in storage replacement projects – due to insufficient data analysis

Cyclic data

- Cyclic data is typical for telecommunication industry (other sectors too, but telco is the textbook example of cyclic data)
  - Telco providers split users into billing cycles.
  - Data for each cycle accumulates throughout the month
  - Each cycle is processed starting on a certain date of the month
- Tiering estimation
  - If data is collected for a couple of days, the skew point will appear to be very high and perfect match for tiering (over 90%)
  - Most of the IO will happen on a limited segment of data (1-2 billing cycles out of let's say 28 – so less than 7.5% of the data)

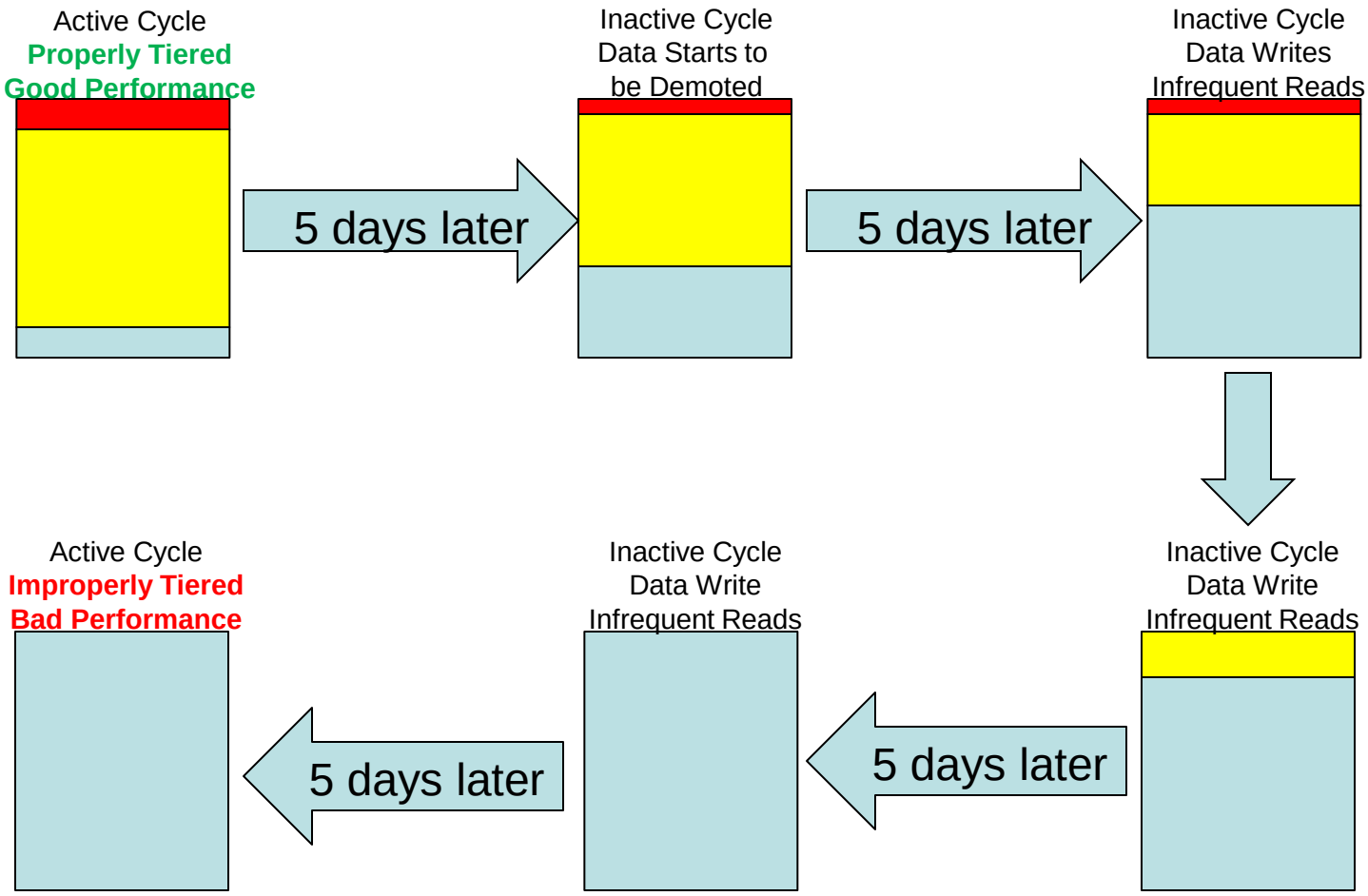
Cyclic data continued

- Automated Tiering Behavior with Cyclic data
  - Inactive cycles will have limited IO – only data being accumulated → Will be demoted over time to the lowest tier
  - Active cycles will have high IO – will be promoted to the highest tier
  - Once a cycle switches from inactive to active, the response time escalates
  - Active Data is located now on the lowest tier (high capacity and low IO) and unable to meet the performance requirements
  - Automatic Tiering software cannot move large amounts of data fast enough in order to prevent this situation

Cyclic data continued

- Cyclic data is better suited for 2-tier configurations (usually the 3rd tier is excluded)
- Even if the data is cold/warm for 26 days a month, during the time the billing cycle is processed, a large amount of data becomes hot all of a sudden.
- Automatic tiering software increases the back-end utilization of the array – and cannot move data fast enough to cope with such situations.

Cyclic Data Wrong Tiering Example



Tier 3 – Performance Considerations

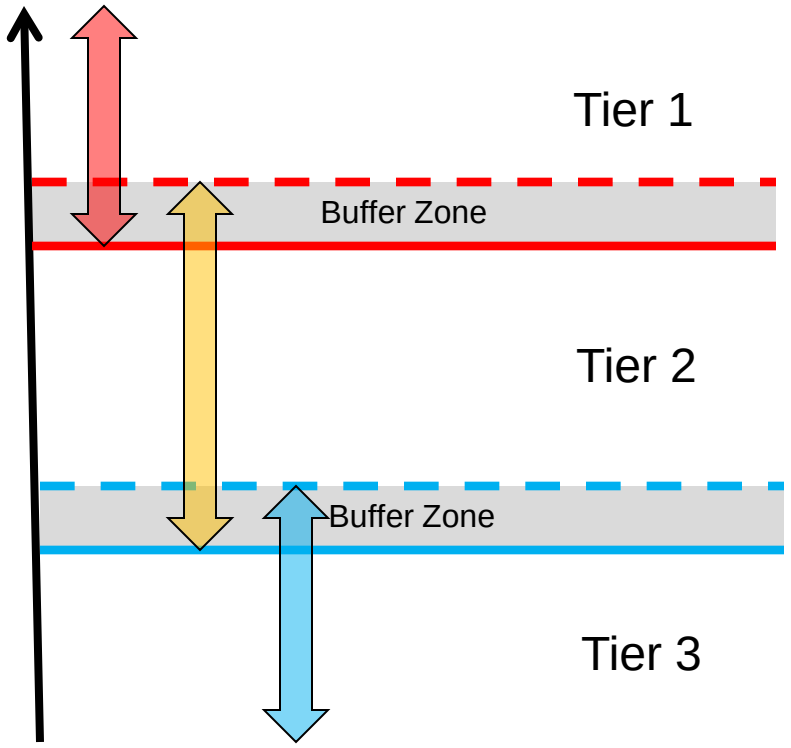
- Relocating dormant data to the third tier can:
  - **Help** (release space in the upper tiers that can be used for hot data)
  - **Hinder** (if data becomes active tier 3 must perform well enough to be able to handle both the host IO and the relocation IO – otherwise performance will be impacted)
- The performance of 3rd tier is important and cannot be ignored
- In most ATS implementations host IO takes priority over relocation IO
  - Data relocation in and out of Tier 3 will freeze if the utilization is very high
  - If dormant data becomes suddenly hot and utilization of Tier 3 become very high – ATS will not be able to handle the situation

Overutilization

- Automated Storage Tiering utilizes system resources
  - Moving data around means reading from one tier and writing it into another → IO overheads
  - When the storage array is under heavy load automated tiering is usually suspended
  - In such circumstances, response time will increase – since data placement is not updated based on access frequency
- Automated Storage Tiering requires available space
  - Available space is required in every storage tier for Automated Tiering to be able to move data

Data Thrashing

- Irrespective of what parameter or combination of parameters is used to decide when to promote or demote data, the border between tiers should not be a fixed value – but a buffer zone.
- A fixed value will lead to thrashing – and as the value will slightly change up and down of the threshold, data will keep on bouncing between tiers
- ATS implementations should take into account buffer zones and allow data movements only when the buffer zone is exceeded



How to Improve tiering estimates

- Data skew is not enough to identify the right tiering mix
- IO density has to be taken into account in order to find the right tiering configuration
- If enough data is available, IO density can be calculated in 5% or 10% intervals (associated to the data skew graph)
- Knowing the IO density for each interval will help identify the right drive mix and RAID configuration
- Knowing the workload profile and data characteristics is crucial in identifying the right tiering mix

How do we do it right?

- Choose the type of storage based on the IO density
- Tier if the data skew allows it
- Take into account the overheads of tiering
- Don’t just buy capacity – buy the kind of capacity that matches your performance needs

What is missing?

- ATS should be able to decide where to place data depending on the % of writes
- ATS should take into consideration IO density in conjunction with the data skew
- ATS should avoid thrashing
- ATS should be able to prioritize relocation IO over host IO in certain conditions (see slide 19)

Attribution & Feedback

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

Authorship History

Name/Date of Original Author here:
Octavian Paul Rotaru , March 2014

Updates:

Additional Contributors

Alex MacDonald, NetApp
Jean-Christophe Perrelet, Iron Ore Company of Canada

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