Education

Capacity Optimization Technologies for the Cloud

Thomas Rivera, SEPATON

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.

About the SNIA DPCO

- 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
- 2011 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

Abstract

Capacity optimization technologies can significantly help IT organizations in the areas of physical storage acquisition, disaster recovery, and cloud storage.

This technical session will:

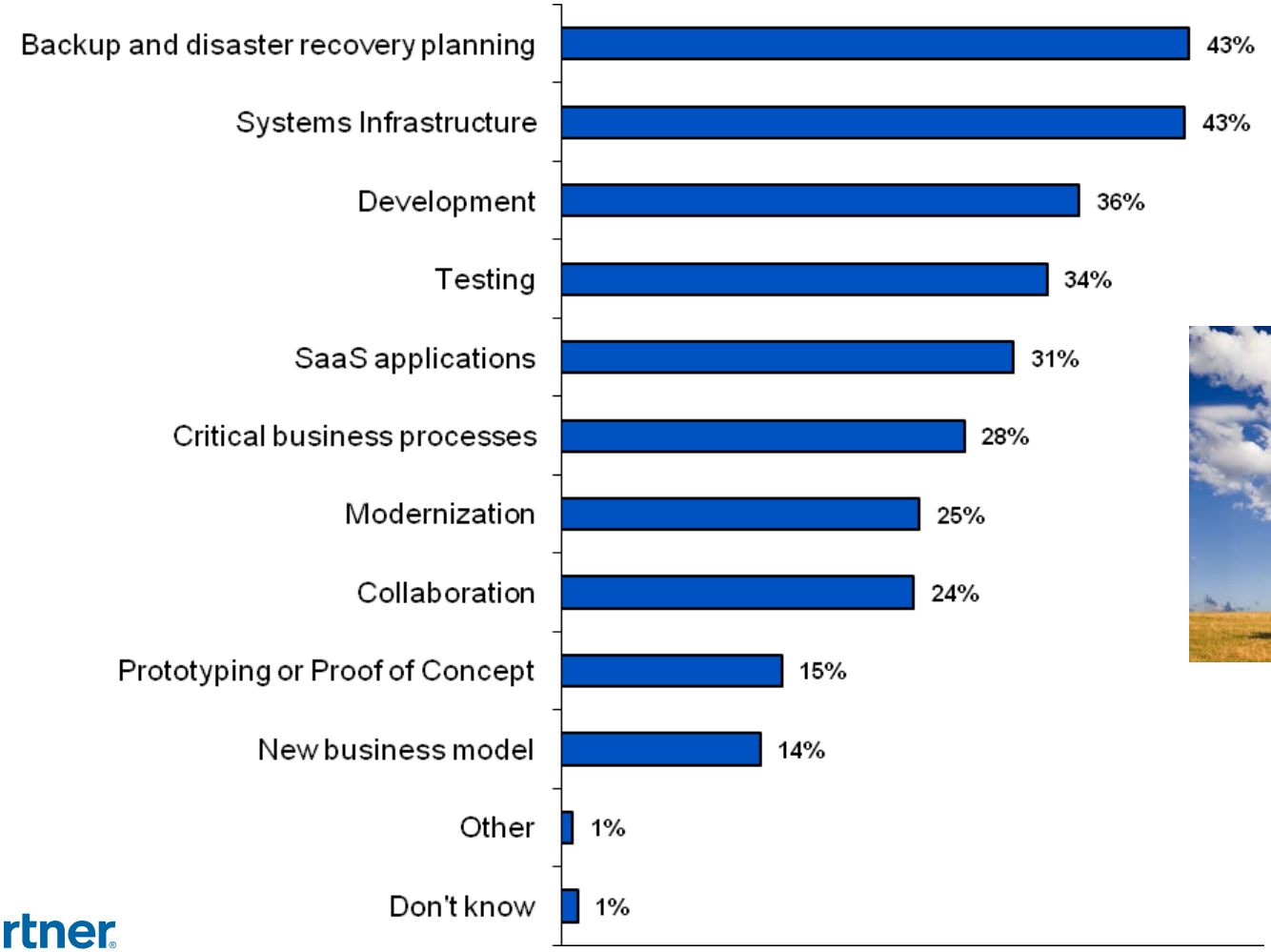
- Review what capacity optimization is, and the various technologies available
- Identify the various capacity optimization technologies' respective impact on DR/replication, and storage cost savings
- Address how the use of capacity optimization technologies affects Cloud Storage implementations

Why Capacity Optimization?

- The values of capacity optimization in the cloud:
  - Satisfy ROI/TCO requirements
  - Manage data growth
  - Increase efficiency of storage and backup
  - Reduce overall cost of storage
  - Reduce network bandwidth
  - Reduce operational costs including:
    - Infrastructure costs required space, power and cooling
    - Movement toward a greener data center
  - Reduce administrative costs

Cloud Direction

What types of cloud-computing projects is your organization planning/do you expect your organization to undertake within the next three years? *(Can choose multiple)*



Gartner

•From "Gartner Survey:
•Cloud Computing Strategy 3Q10"

Capacity Optimization Methods

- Compression
- Deduplication
- Delta snapshots
- Thin provisioning

Check out these SNIA Tutorials:

- **Understanding Data Deduplication**
- **Deduplication's Role in Disaster Recovery**
- **Advanced Deduplication Concepts**

Capacity Optimization Terminology

Capacity Optimization Methods [Storage System]

Methods which reduce the consumption of space required to store a data set, such as compression, data deduplication, thin provisioning, and delta snapshots

Data Deduplication [Storage System]

The replacement of multiple copies of data—at variable levels of granularity—with references to a shared copy in order to save storage space and/or bandwidth.

Subfile deduplication

A form of data deduplication that operates at a finer granularity than an entire file or data object

Single instance storage

A form of data deduplication that operates at a granularity of an entire file or data object

Capacity Optimization Terminology (Cont.)

Compression [General]

The process of encoding data to reduce its size. *Lossy* compression (i.e., compression using a technique in which a portion of the original information is lost) is acceptable for some forms of data (e.g., digital images) in some applications, but for most IT applications, *lossless* compression (i.e., compression using a technique that preserves the entire content of the original data, and from which the original data can be reconstructed exactly) is required.

Delta Snapshot [Data Recovery]

A type of point in time copy that preserves the state of data at an instant in time, by storing only those blocks that are different from an already existing full copy of the data.

Thin Provisioning [Storage System]

A technology that allocates the physical capacity of a volume or file system as applications write data, rather than pre-allocating all the physical capacity at the time of provisioning.

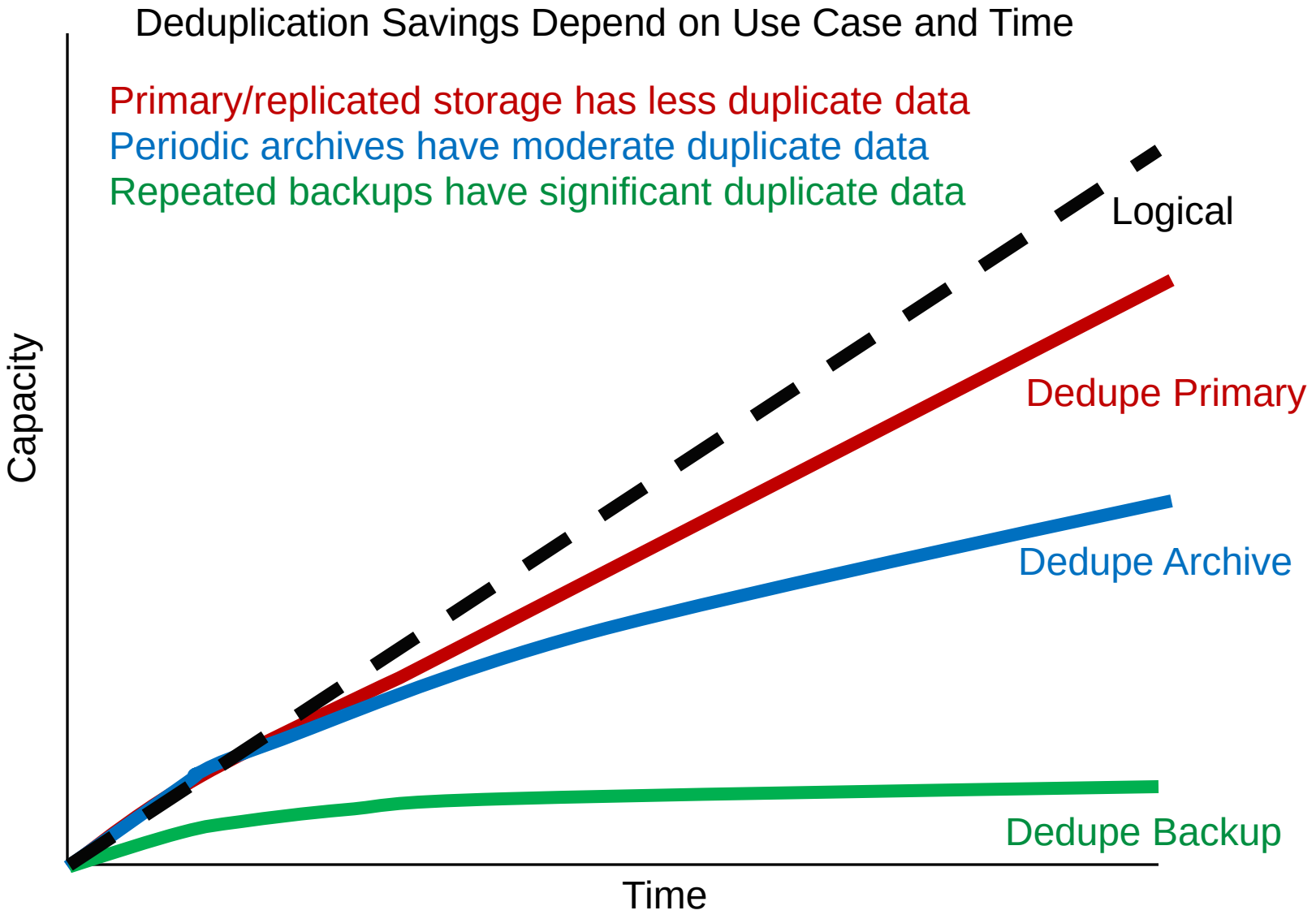
Compression with Deduplication

- Dedupe and compression are similar
  - Both are dependant on data patterns
  - Both consume system resources
  - Both can optimize required storage capacity
- Dedupe and compression are different
  - Some data can more effectively be optimized via dedupe
  - Some data can more effectively be optimized via compression
  - Some data can be optimized via dedupe **and** compression
  - Some data cannot be optimized at all
- Dedupe and compression are complementary
  - But some knowledge about the data pattern is required

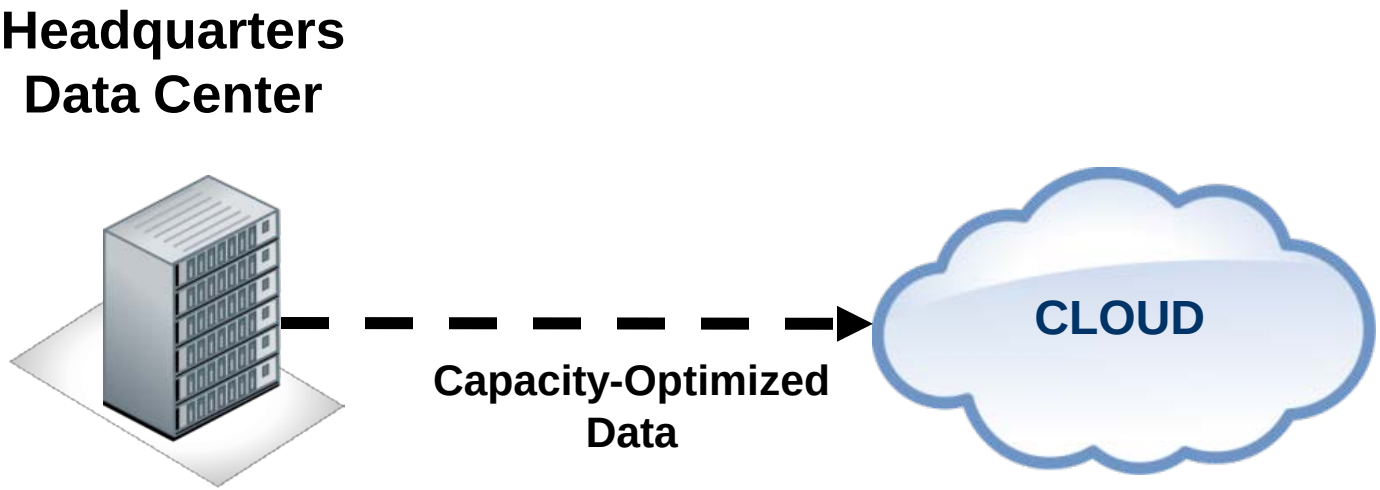
Capacity Optimization Use Cases

- Data Protection to the cloud
  - Backup to disk efficiently with longer retention periods
  - Long term retention and preservation
  - Disaster recovery
- Primary storage in the cloud
  - Store more logical data with less physical disk
- Data protection in the Cloud (also cross-cloud)
  - Backup to disk efficiently with longer retention periods
  - Long term retention and preservation
  - Disaster recovery

Deduplication Use Case Motivation

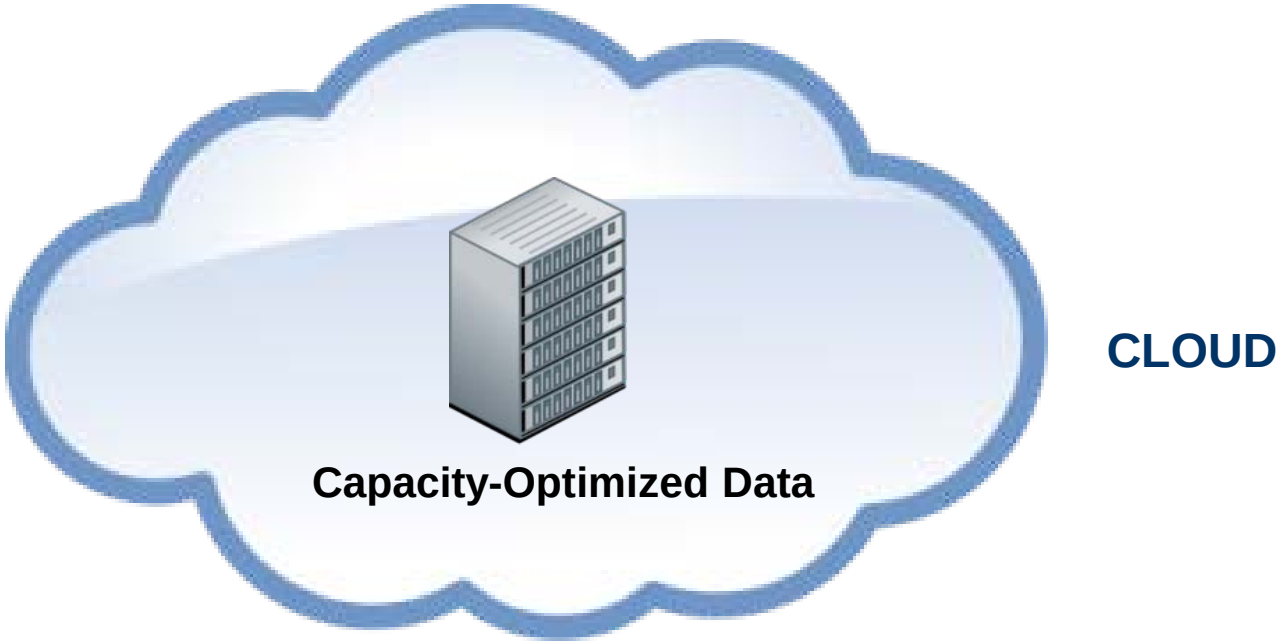


Use Case: Optimized Replication to the Cloud



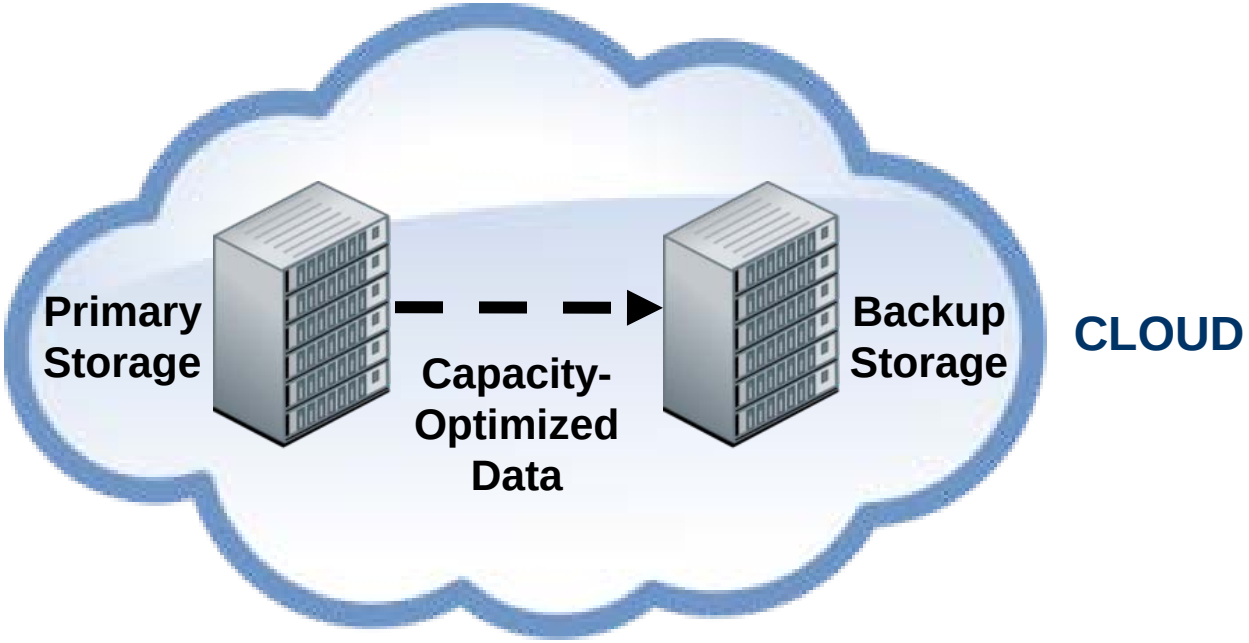
- Increased WAN Efficiency to the Cloud (Data Protection)
  - Reduced bandwidth
  - Reduced CPU utilization (if capacity optimization occurs "on-the-wire")
  - Shorten data transfer times by sending less physical data, and therefore also being able to send more logical data

Use Case: Optimized Primary Storage in the Cloud



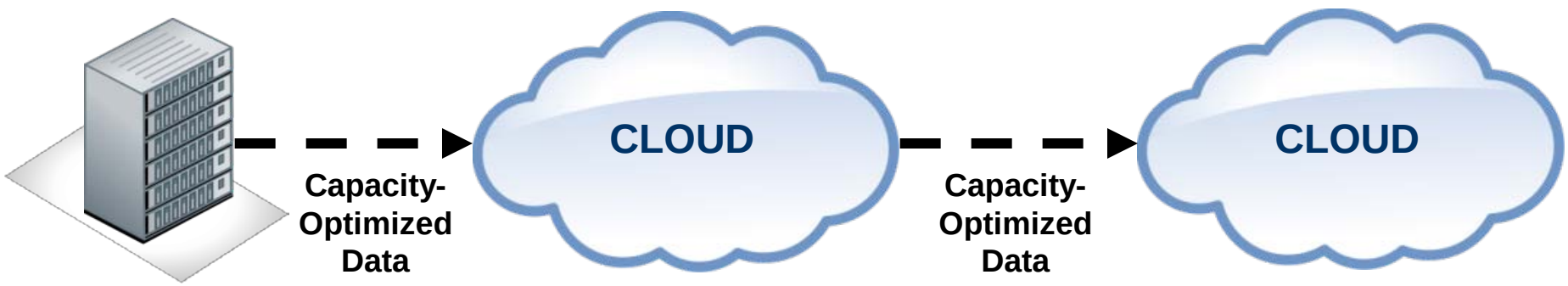
- Increased Storage Efficiency in the Cloud (Primary Storage)
  - Store more logical data with less physical disk
  - Less *usage* costs

Use Case: Optimized Data Protection in the Cloud



- Increased Storage Efficiency in the Cloud (Data Protection)
  - Store more logical data with less physical disk
  - Less *usage* costs (when paying for what is utilized)

Use Case: Optimized Replication Cloud-to-Cloud



- Increased WAN Efficiency for Cloud-to-Cloud replication
  - Reduced bandwidth
  - Reduced CPU utilization (if capacity optimization occurs "on-the-wire")
  - Shorten data transfer times by sending less physical data, and therefore also being able to send more logical data

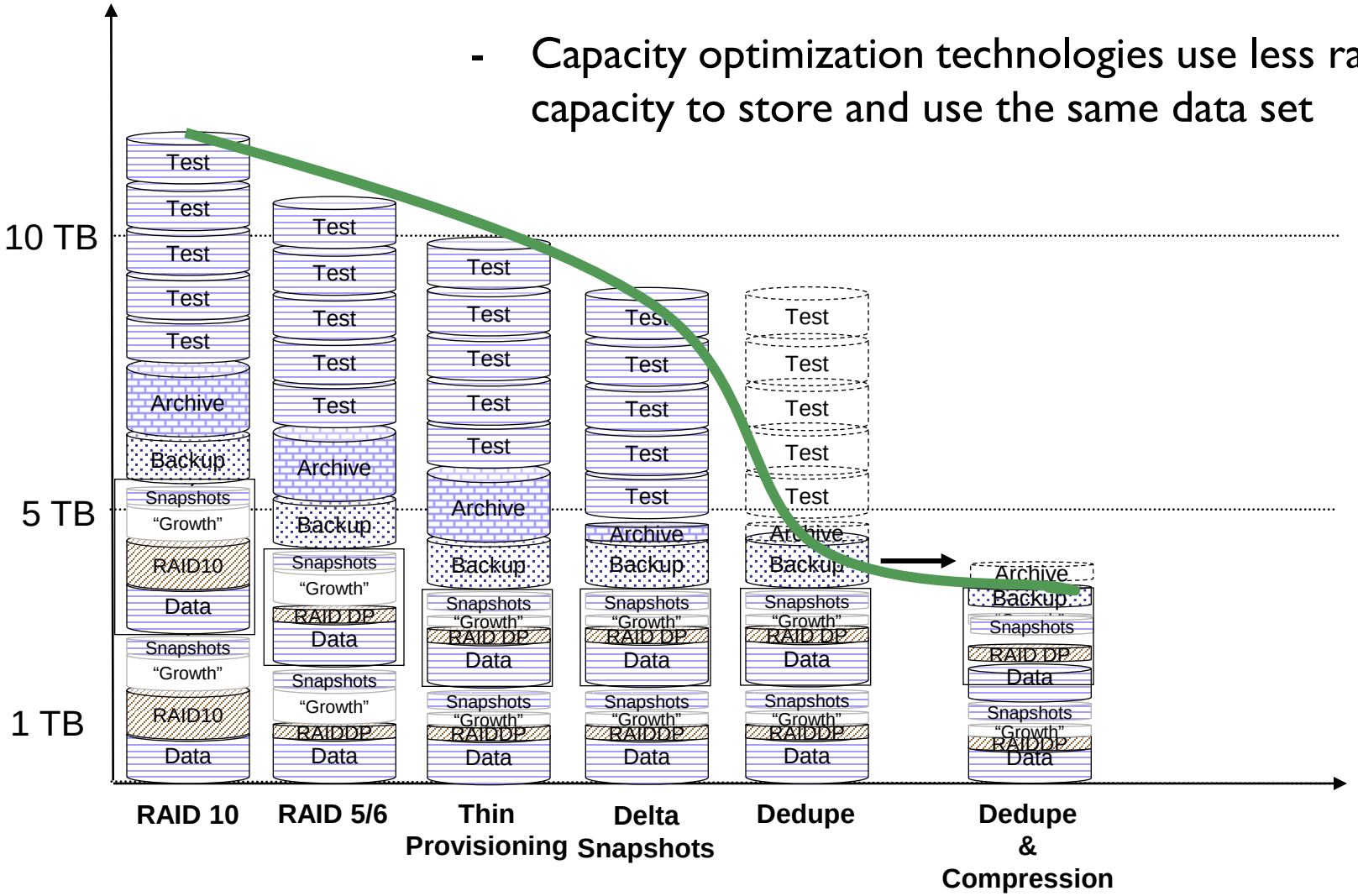
Capacity Optimization: Savings

- Every TB of disks you *don't* buy saves you
  - CAPEX for the equipment
  - CAPEX for the footprint
  - CAPEX and OPEX for power conditioning
  - OPEX for the power to spin the drives
  - OPEX for cooling
  - OPEX for storage management

Source: SNIA Green Storage Initiative

Capacity Optimization: Savings (Cont.)

- Capacity optimization technologies use less raw capacity to store and use the same data set



Typical Savings

- Compression
  - Highly data dependent
  - 0% – 75%
- De-duplication with compression
  - 40 – 95%, depending on dataset and time interval
  - ~ 40 – 50% average over time
- Thin provisioning (just-in-time storage provisioning)
  - 40 - 60%
  - Average 30% utilization ➔ over 80% utilization
- Re-sizeable volumes (expand or shrink as required)
  - 20 – 50%

Space Reduction Ratio & Percent

Space Reduction Ratio (In:Out)	Space Reduction Percent
2:1	1/2 = 50%
5:1	4/5 = 80%
10:1	9/10 = 90%
20:1	19/20 = 95%
100:1	99/100 = 99%
500:1	499/500 = 99.8%

- Ratios can meaningfully be compared only under the same set of assumptions
- Relatively low space reduction ratios still provide significant space savings

Summary

- Each capacity optimization technology makes tradeoffs to balance performance cost against reduction ratio:
  - May be able to reduce data significantly, but may be at a high performance cost
- May be tradeoffs for each capacity optimization technology
  - May not be able to mix and match most without "bottlenecks".
  - Since SW is often bound to HW, this can lead to a lot of deployment complexity
- Each capacity optimization technology tends to have a "sweet spot" with respect to data types it can reduce well
  - Structured versus non-structured data
  - Pre-compressed data
  - Encrypted data
  - Image data (jpeg, movies, x-ray's, etc.)

Capacity Optimization Options: How to Choose

- How do you know if a capacity optimization method will work for your types of data?
  - Ask cloud providers
  - Rough math
  - Data analysis tools
  - Sizing tools
  - Customer references
  - **Test using <u>your</u> data – with your cloud provider!**

Q&A / Feedback

- Please send any questions or comments on this presentation to SNIA: trackdatamgmt@snia.org

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

- SNIA Education Committee

Data Protection and Capacity Optimization (DPCO) Committee:

Mike Dutch
Michael Fishman
Larry Freeman
Bernd Henning
Gene Nagle
Ronald Pagani
Tom Pearce
Michael Peterson
Thomas Rivera
Tom Sas
Gideon Senderov
Sebastian Zangaro

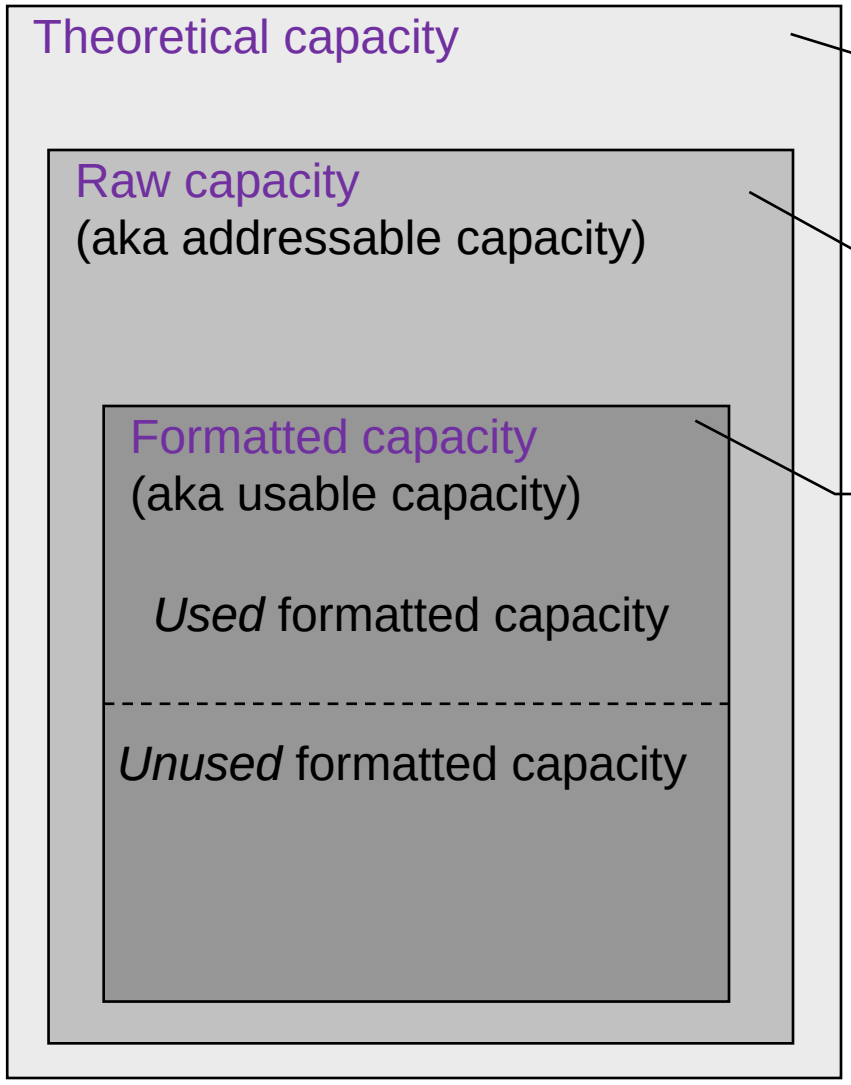
SNIA
Data Protection and Capacity Optimization Committee

It's easy to get involved with the DPCO !

- Find a passion
- Join a committee
- Gain knowledge & influence
- Make a difference

www.snia.org

Capacity Definitions



Includes unaddressable space, ECC data, remap areas, inter-sector gaps, etc.

Includes areas set aside for system use, spares, RAID parity, checksums, remapping, disk "right sizing", labels, etc.

Bytes available to be written after being formatted for use. May include areas that are normally reserved — such as snapshot setasides — if they can alternatively be configured for ordinary data storage by the storage administrator.

Effective capacity = amount of stored data + unused formatted capacity