



Education

Technologies for Green Storage

Jiří Král, riaditeľ pre technický rozvoj
STORYFLEX a.s.

- 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.

➤ Technologies for Green Storage

- ◆ Hardware efficiencies are essential to reducing the amount of power used by storage. Equally real savings are obtained by reducing the number of copies of your data that must be made, kept and managed. This talk presents a number of technologies, ranging from thin provisioning and virtualization to flywheel UPSs, that each address part of the problem, and illustrates the impact that each technology can have on your data center footprint.

- Objectives
- Need for redundancy (extra copies of data)
- Software technologies for reducing excessive redundancy
- Other technologies for energy saving
- Data center technologies
- Typical savings

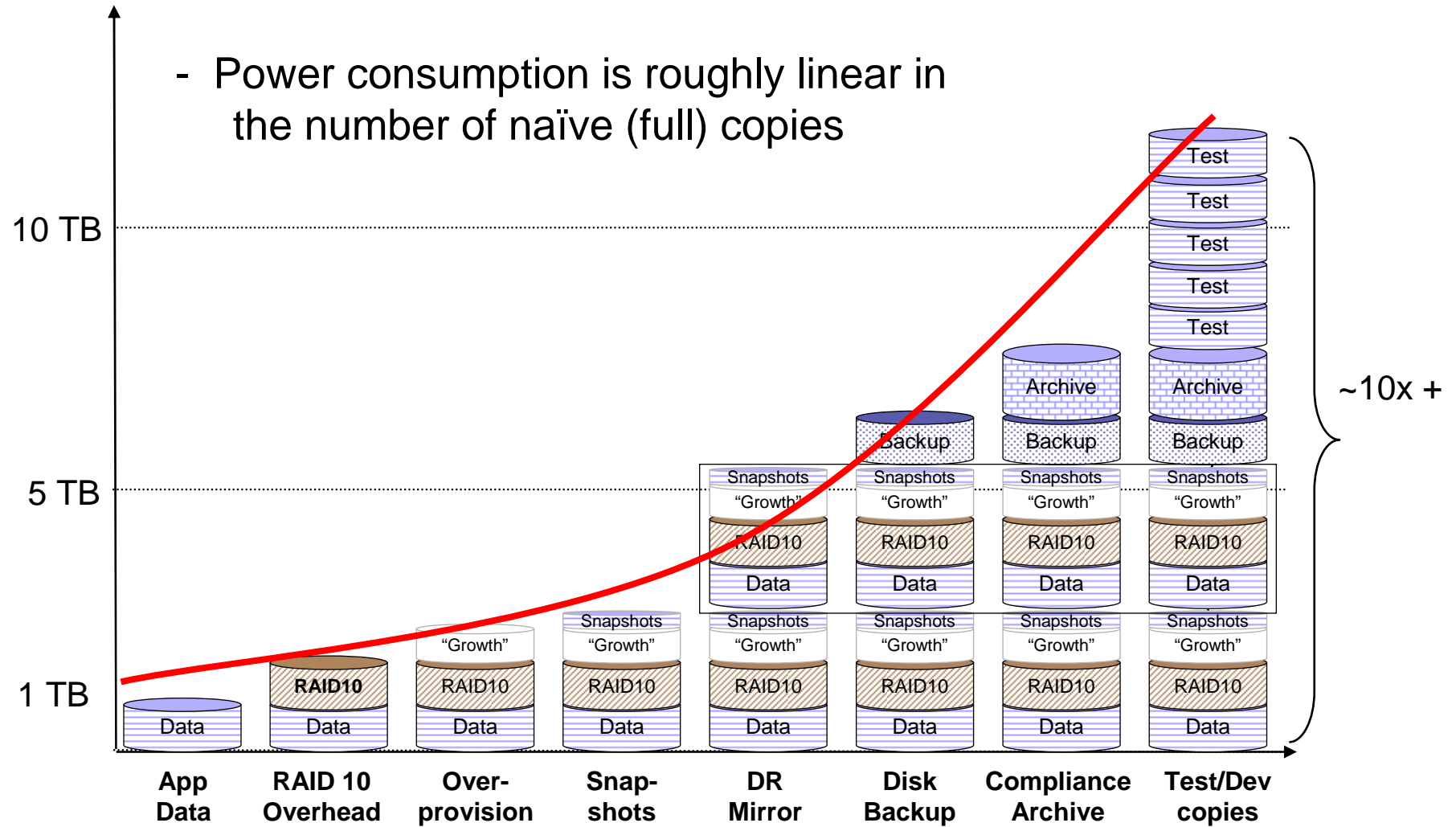
- ◆ Get more work done
 - ◆ for less money
- ◆ Translation: reduce data center footprint
 - ◆ in space
 - > less storage equipment to buy
 - ◆ in energy
 - > more energy-efficient equipment
 - > less equipment to cool
 - > better cooling methodologies
 - > better power management
 - ◆ in administrative costs
 - > less storage equipment to manage

Need for redundancy

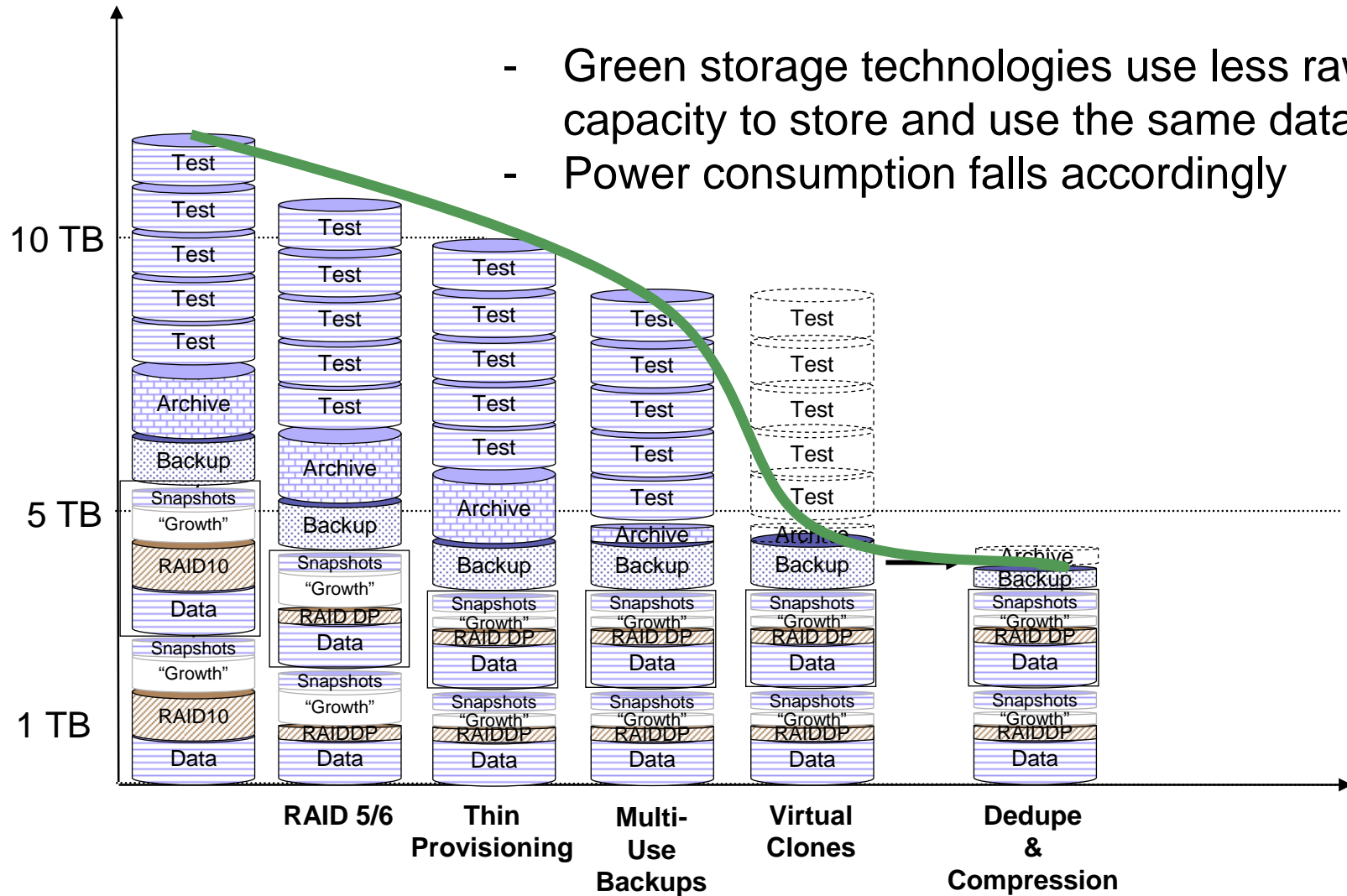
- RAID 10 – protect against multiple disk failures
- DR Mirror – protect against whole-site disasters
- Backups – protect against failures and unintentional deletions/changes
- Compliance archive – protect against heavy fines
- Test/dev copies – protect live data from mutilation by unbaked code
- Overprovisioning – protect against volume out of space application crashes
- Snapshots – quicker and more efficient backups

Result of redundancy

- Power consumption is roughly linear in the number of naïve (full) copies



Effect of green technologies



- Green storage technologies use less raw capacity to store and use the same data set
- Power consumption falls accordingly

Green Storage Technologies

- Enabling technologies
 - ◆ Storage virtualization
 - ◆ Storage capacity planning
- Green software
 - ◆ Compression
 - ◆ Snapshots
 - ◆ Virtual (writeable) clones
 - ◆ Thin provisioning
 - ◆ Non-mirrored RAID
 - ◆ Deduplication and SIS
 - ◆ Resizeable volumes



Green storage technologies (cont.)

- Other storage technologies and power saving techniques
 - ◆ Capacity vs. high performance drives
 - ◆ ILM / HSM
 - ◆ MAID
 - ◆ SSDs
 - ◆ Power supply and fan efficiencies
- Facilities-side technologies
 - ◆ Hot aisle/cold aisle
 - ◆ Water & natural cooling
 - ◆ Flywheel UPSs



Enabling technologies

- Storage virtualization
- Storage capacity planning



Green software technologies

- Compression
- Snapshots
- Virtual (writeable) clones
- Thin provisioning
- Non-mirrored RAID
- Deduplication and SIS
- Resizeable volumes



Compression

► Compression

- ◆ Old and venerable
- ◆ Origins in signaling, number and coding theory
 - > Motivated by limited bandwidth and lossiness of satellite communications
- ◆ Scattered throughout the data stack
- ◆ Many formats already compressed
 - > JPEG, MPEG, MP3, etc.
 - > Lossless compression (LZW) necessary for unknown data types
- ◆ Configuration matters
 - > Compress before encrypting, decrypt before decompressing
- ◆ Difficult in block-based environments
 - > Therefore usually done at the stream or object level

Snapshots

➤ NOT just wholesale copies of the data

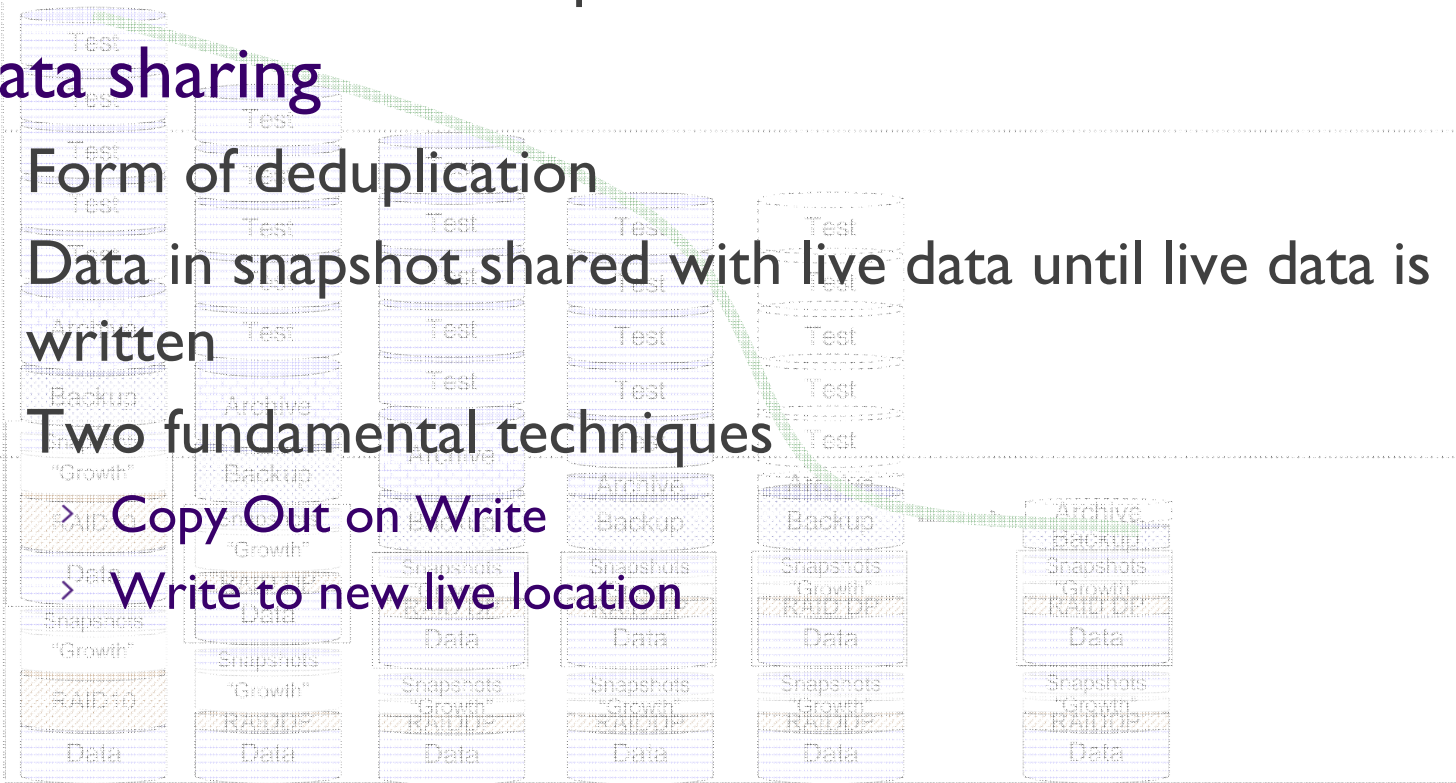
- ◆ We call those PIT copies

➤ Data sharing

- ◆ Form of deduplication
- ◆ Data in snapshot shared with live data until live data is written
- ◆ Two fundamental techniques

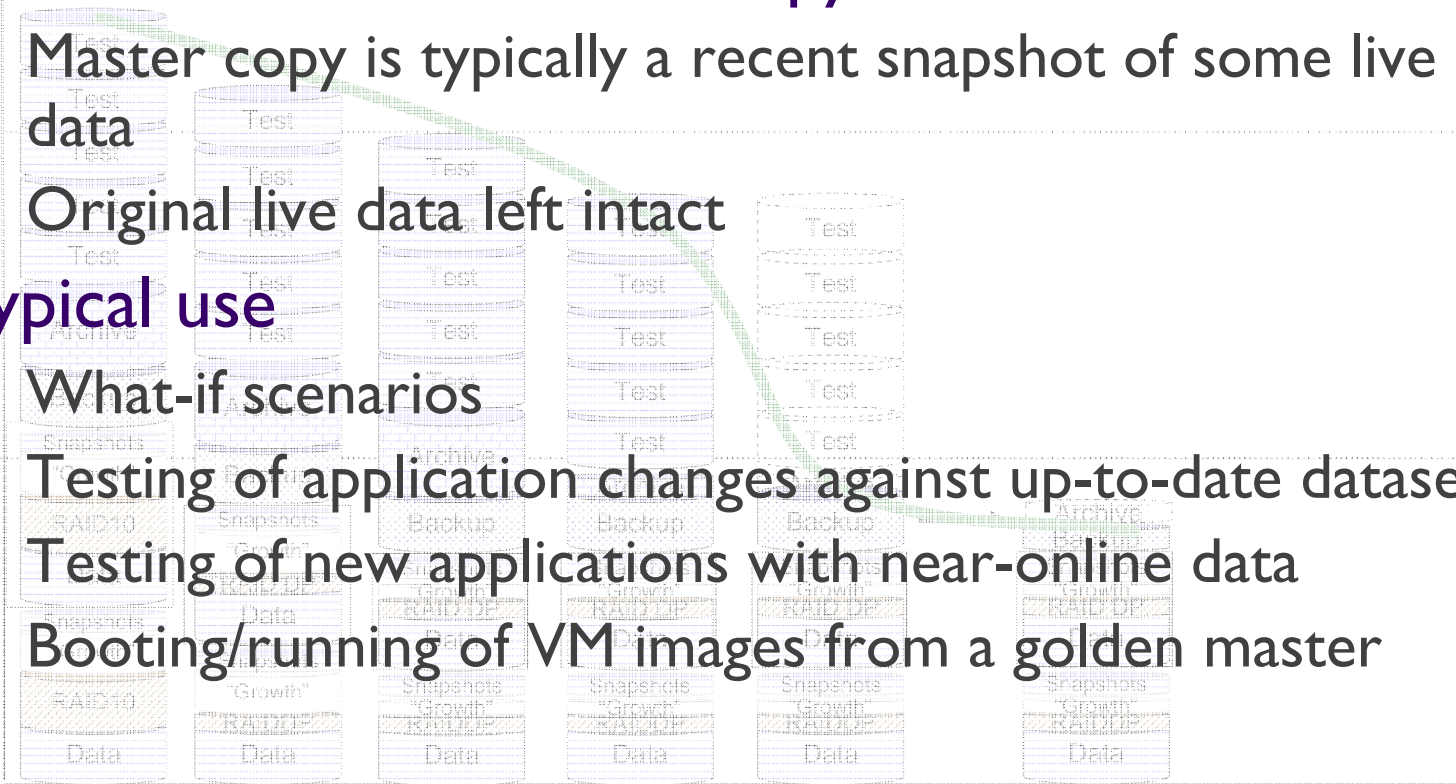
> Copy Out on Write

> Write to new live location



Virtual clones

- Writeable snapshots, essentially
- Data shared with master copy until written
 - ◆ Master copy is typically a recent snapshot of some live data
 - ◆ Original live data left intact
- Typical use
 - ◆ What-if scenarios
 - ◆ Testing of application changes against up-to-date datasets
 - ◆ Testing of new applications with near-online data
 - ◆ Booting/running of VM images from a golden master

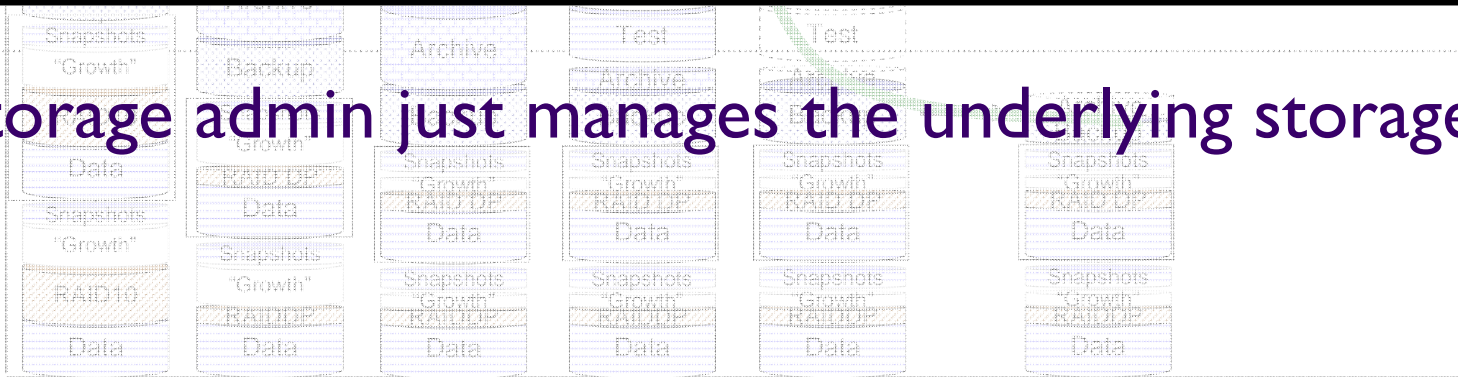


Thin provisioning

- Similar in concept to filesystem quotas

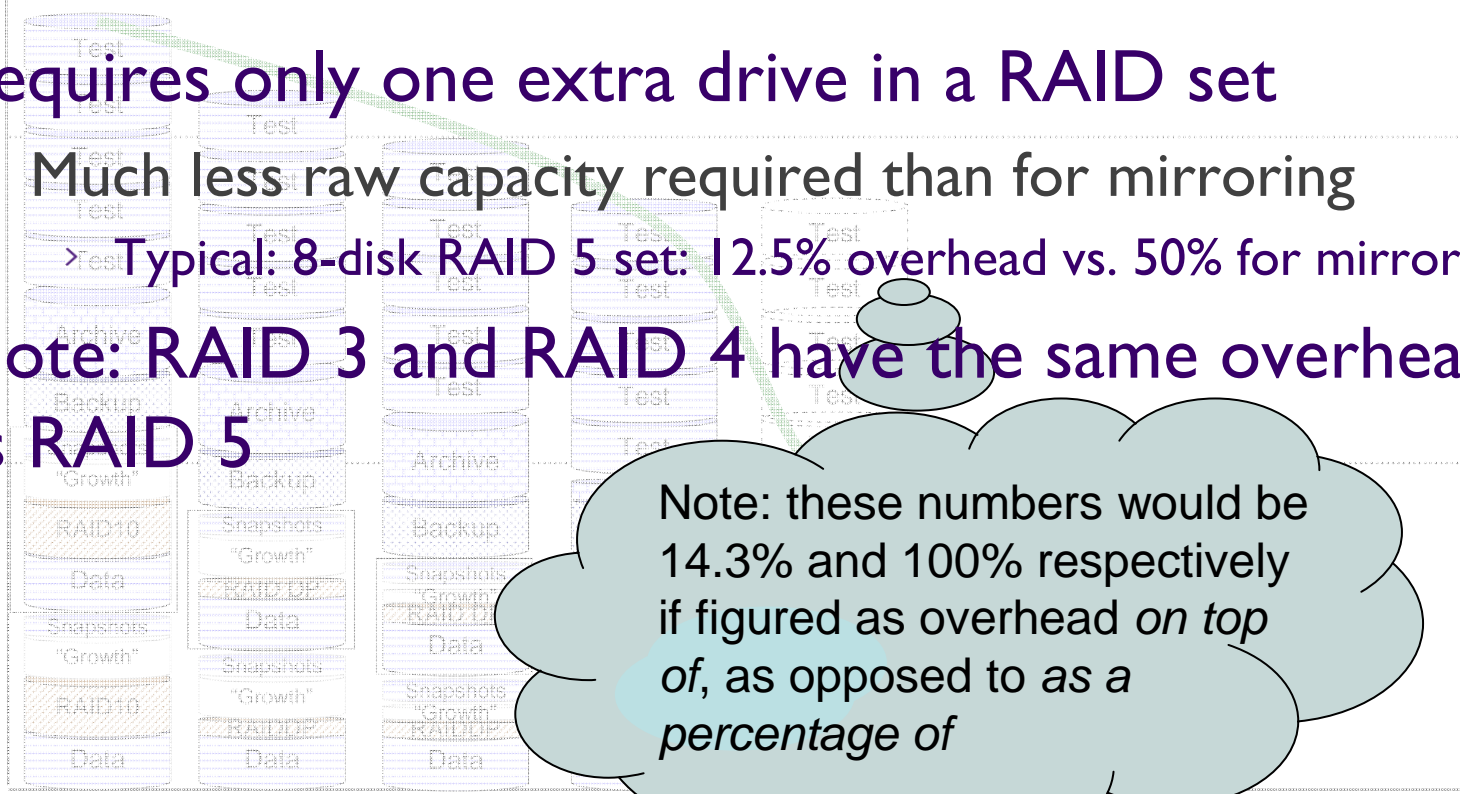
Quotas	Thin provisioned volume
Quota limit	Nominal volume size
Hit quota limit: writes fail	Fill up volume: writes fail
Increase quota	Increase volume size
Need more physical storage	Need more physical storage

- Storage admin just manages the underlying storage



RAID 5

- ▶ Allows any (one) drive in a RAID set to fail without data loss
- ▶ Requires only one extra drive in a RAID set
 - ◆ Much less raw capacity required than for mirroring
 - ▶ Typical: 8-disk RAID 5 set: 12.5% overhead vs. 50% for mirroring
- ▶ Note: RAID 3 and RAID 4 have the same overhead as RAID 5



Note: these numbers would be 14.3% and 100% respectively if figured as overhead *on top of*, as opposed to *as a percentage of*

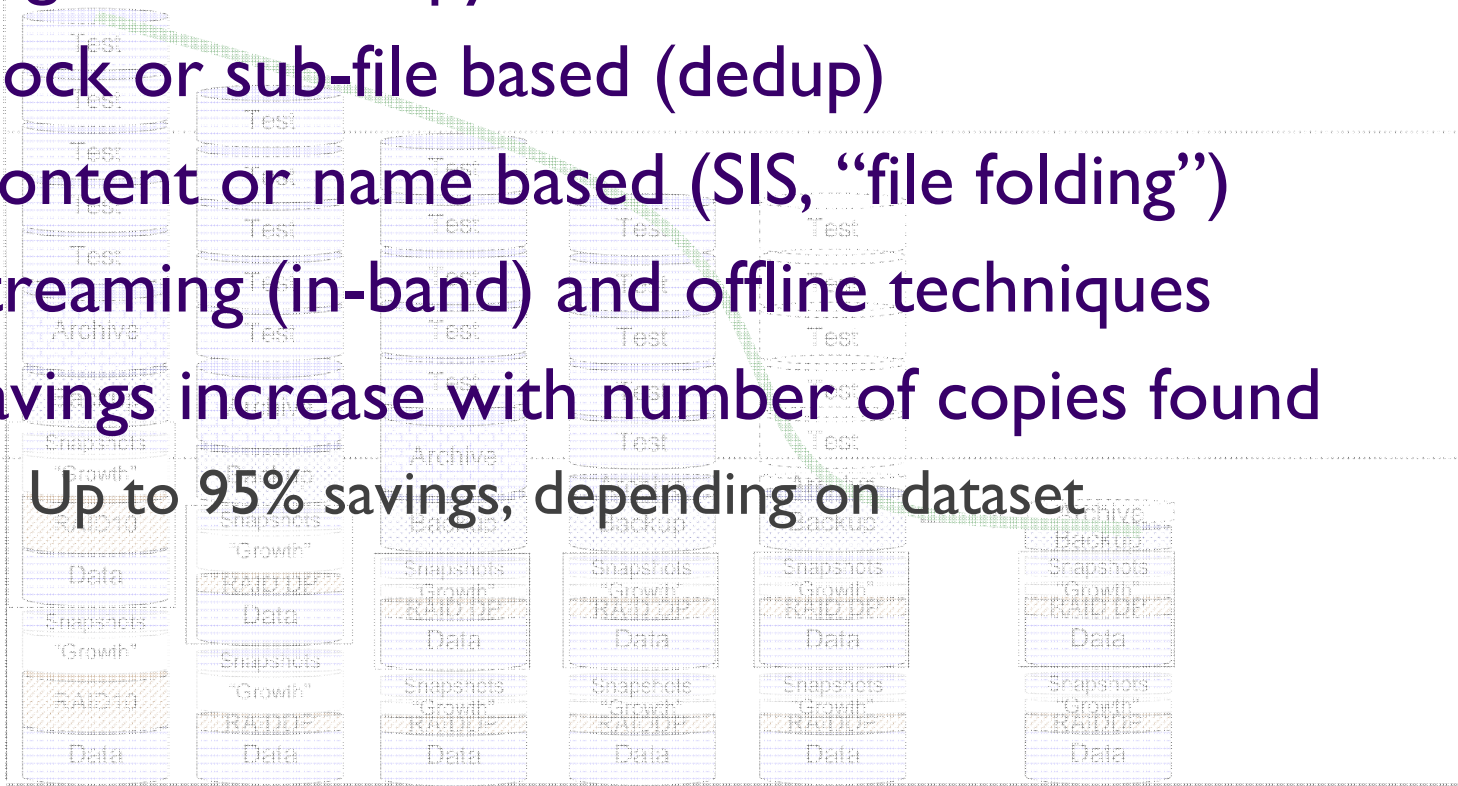
RAID 6

- More dependable than mirroring
 - ◆ Mirroring: can survive two failures in a disk group if they're not in the same mirrored pair
 - ◆ RAID 6: can survive failure of any two drives in the group
- Requires two extra drives per RAID set
 - ◆ However, typically somewhat larger RAID sets
- Necessary as drive sizes increase
 - ◆ Probability of a disk failure during RAID 5 parity reconstruct is getting too high
- More green than mirroring
 - ◆ 50% overhead in RAID 1 mirroring
 - ◆ 14.3% overhead in a 14-disk RAID 6 raidset

16.7% and 100% if figured on top of

Deduplication and SIS

- Find duplicates at some level, substitute pointers to a single shared copy
- Block or sub-file based (dedup)
- Content or name based (SIS, “file folding”)
- Streaming (in-band) and offline techniques
- Savings increase with number of copies found
 - ◆ Up to 95% savings, depending on dataset



Resizable volumes

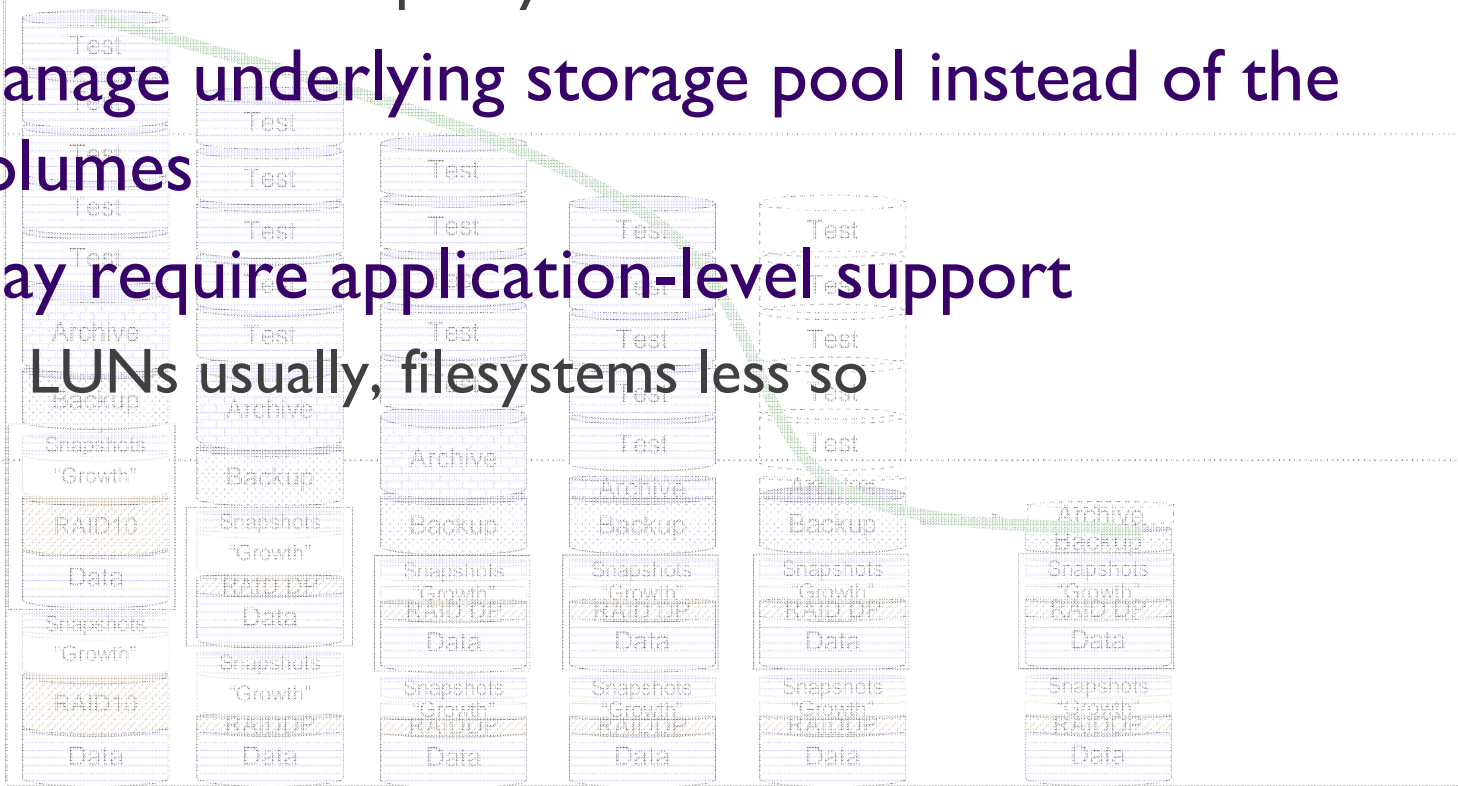
➤ Does away with need for overprovisioning

- ◆ Run at 70% capacity instead of 20-30%

➤ Manage underlying storage pool instead of the volumes

➤ May require application-level support

- ◆ LUNs usually, filesystems less so



Other technologies

- Capacity vs. high performance drives
- ILM / HSM
- MAID
- SSDs
- Power supply and fan efficiencies



Capacity vs. High Performance

➤ Capacity

- ◆ focused on GB/watt at rest
 - > 1 TB SATA: 15W
 - > 4 x 250 GB FC: 64W
- ◆ also tend to have better \$/GB
- ◆ NOTE: power use is quadratic with respect to rotational speed
 - > Use the slowest drives that will fit your needs

➤ Performance

- ◆ focused on seek time
 - > 1 TB SATA: 12 – 15 ms
 - > 300 GB FC: 3 – 4 ms
- ◆ also designed for higher RAS environments

➤ Exploit cost differences between storage tiers

- ◆ Idea: automatically move data to an appropriate storage platform at each period in its lifetime
- ◆ Tier change must have substantial value to make the overhead worth it
 - > Cost of ILM/HSM system
 - > Cost of administration
 - > Cost of data movement



**Check out SNIA Tutorial:
The Secret Sauce of ILM
Assessment**

➤ Practice

- ◆ Storage declines in value as it ages
- ◆ Manual movement of data sets

➤ New wrinkle

- ◆ Footprint considerations preclude keeping older gear

MAID (Massive Array of Idle Disks)

➤ Idea: spin down disks when not in use

◆ Pros

- › Disks use no power when spun down
- › > 50% power savings at idle

◆ Cons

- › Most data near-online (access times of several seconds)
- › Background disk housekeeping difficult
- › Often the same data center requirements (power supplies, CRAC units, PDUs etc.), but these are used at lower efficiencies
- › Impending competition from SSDs
- › Competition from....

► Our old “buddy”

◆ Pros

- › Tapes use no power when inactive
- › > 90% power savings at idle

◆ Cons

- › Data is at best near-online (access times of several seconds)
- › Not a random access format
- › Lack of true resilience to format failure
 - RAIT?



Check out SNIA Tutorial:

**Introduction to Data Protection:
Backup to Tape, Disk and Beyond**

SSDs (Solid State Disks)

➤ Usually refers to FLASH-based disks

- ◆ Pros
 - › Great READ performance
 - › At rest power consumption = 0
 - › No access time penalty when idle (cf. MAID)
 - › No need to keep some disks spinning (cf. MAID)
- ◆ Cons
 - › WRITE performance < mechanical disks
 - › Cost >> mechanical disks except at very high perf points
 - › Wear leveling requires a high space overhead
- ◆ Note: these dynamics changing rapidly with time
 - › SSSI – SNIA Solid State Storage Initiative

Power supply and fan efficiencies

➤ Efficiency of power supply an up front waste

- ◆ Formerly 60-70%
- ◆ Nowadays 80-95%
 - Climate Savers
 - 80+ group
 - Note: Efficient PSs *are* more expensive

➤ Variable speed fans

- ◆ Common nowadays
- ◆ Software (OS) control
- ◆ Energy consumption is quadratic in the rotational speed

Facilities-side technologies

- Monitoring
- Hot aisle – cold aisle technologies
- Water and natural cooling
- Flywheel UPSs



Savings calculations

- Space savings
- Equipment power savings
- Facilities power savings



Savings calculations (storage)

- Calculations herein are for space savings
- Relationship of space to \$\$ is loose
- But every TB of disks you don't buy saves you
 - ◆ CapEx (Capital Expenditure) for the equipment
 - ◆ CapEx for the footprint
 - ◆ CapEx for power conditioning and cooling
 - ◆ OpEx (Operational Expenditure) for equipment power
 - ◆ OpEx for power conditioning and cooling
 - ◆ OpEx for storage management
 - ◆ OpEx for service contract fees

Typical savings

➤ Compression

- ◆ 20 – 40%
- ◆ Remember, no savings from already compressed file formats

➤ Snapshots

- ◆ 90 – 95% per snapshot, compared to raw PIT copies
- ◆ Only data written since snapshot needs to be copied

➤ Writeable clones

- ◆ 80 – 95%, compared to raw PIT copies
- ◆ Only data that is written needs to be duplicated

Typical savings

- ▶ Thin provisioning
 - ◆ 40 - 60%
 - ◆ Average 30% utilization → over 80% utilization
- ▶ RAID 6
 - ◆ 35%
 - ◆ For 14-disk RAID 6 set, compared to RAID 1/10
- ▶ Deduplication
 - ◆ 40 – 95%, depending on dataset and time interval
 - ◆ ~ 40 – 50% average over time
- ▶ Resizeable volumes
 - ◆ 20 – 50%

Are these savings multiplicative?

➤ Sometimes yes

- ◆ RAID 6 + writeable clone
 - › Assume 1000GB writeable clone = **2000GB** needed for a raw writeable copy on RAID 1 storage
 - › 90% writeable clone savings takes us to 200GB
 - › 35% RAID 6 savings takes us from there to 130GB →
 - › **130GB** needed for a writeable clone on RAID 6

➤ Sometimes no

- ◆ Thin provisioning + resizeable volumes
 - › Similar effects, but you only get the savings once

➤ Sometimes maybe

- ◆ Snapshot + deduplication
 - › Can't dedup readonly snapshots
 - › Snapshots are a form of deduplication, so there's less to dedup
 - › OTOH, already deduped data can be snapshotted efficiently

Savings matrix

Savings multiply in combinations with checkboxes

	C	SS	VC	TP	R	DD	RV
Compression (C)		✓			✓	✓	
Snapshots (SS)	✓		✓	✓	✓	✓	✓
Virtual Clones (VC)		✓					
Thin Provisioning (TP)		✓			✓	✓	
RAID (R)	✓	✓		✓		✓	✓
Deduplication (DD)	✓	✓		✓	✓		✓
Resizable Vols (RV)		✓			✓	✓	

Equipment power savings

➤ Server virtualization

- ◆ up to 80% savings
- ◆ much depends on load

➤ Power supply efficiencies

- ◆ 10 – 20%
- ◆ difficult to do anything about this in a vacuum
- ◆ probably okay to just ride industry trends

➤ Variable speed fans

- ◆ up to 80% savings
- ◆ power consumption quadratic in rotational speed
- ◆ increased data center cold aisle temperatures are key

- Please send any questions or comments on this presentation to trackgreenstorage@snia.org

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

- SNIA Education Committee

**SNIA GSI Marketing Committee
SNIA reviewers**

Larry Freeman – NetApp