SNIA

Education

GREEN STORAGE II: Metrics and Measurement

Erik Riedel, Ph.D.
Seagate Technology

SNIA Legal Notice

- 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.

Abstract

Education SNIA

- **GREEN STORAGE II: Metrics and Measurement**
  - This tutorial will cover storage-specific topics related to energy-efficiency and outline the current state of the industry. We will discuss a range of technologies that are currently considered “green storage” and the components of modern storage systems that impact energy consumption. We will discuss potential metrics for measuring, managing and designing for power in storage systems. We will also briefly outline ongoing efforts in the SNIA Green Storage TWG to standardize metrics for measuring storage systems, including our recently released specification for measuring idle storage.

Tutorial Overview

Education SNIA

- Storage-specific topics related to energy consumption and efficiency
  - current state of the industry
- Comparing the components of “green storage”
  - metrics for measuring, managing & designing for power
- Ongoing efforts in the SNIA
  - SNIA Green Storage Power Measurement Specification
  - Green Storage Technical Work Group (TWG)
  - collaboration with EPA on ENERGY STAR for Storage
  - partnerships with The Green Grid, the DMTF and others

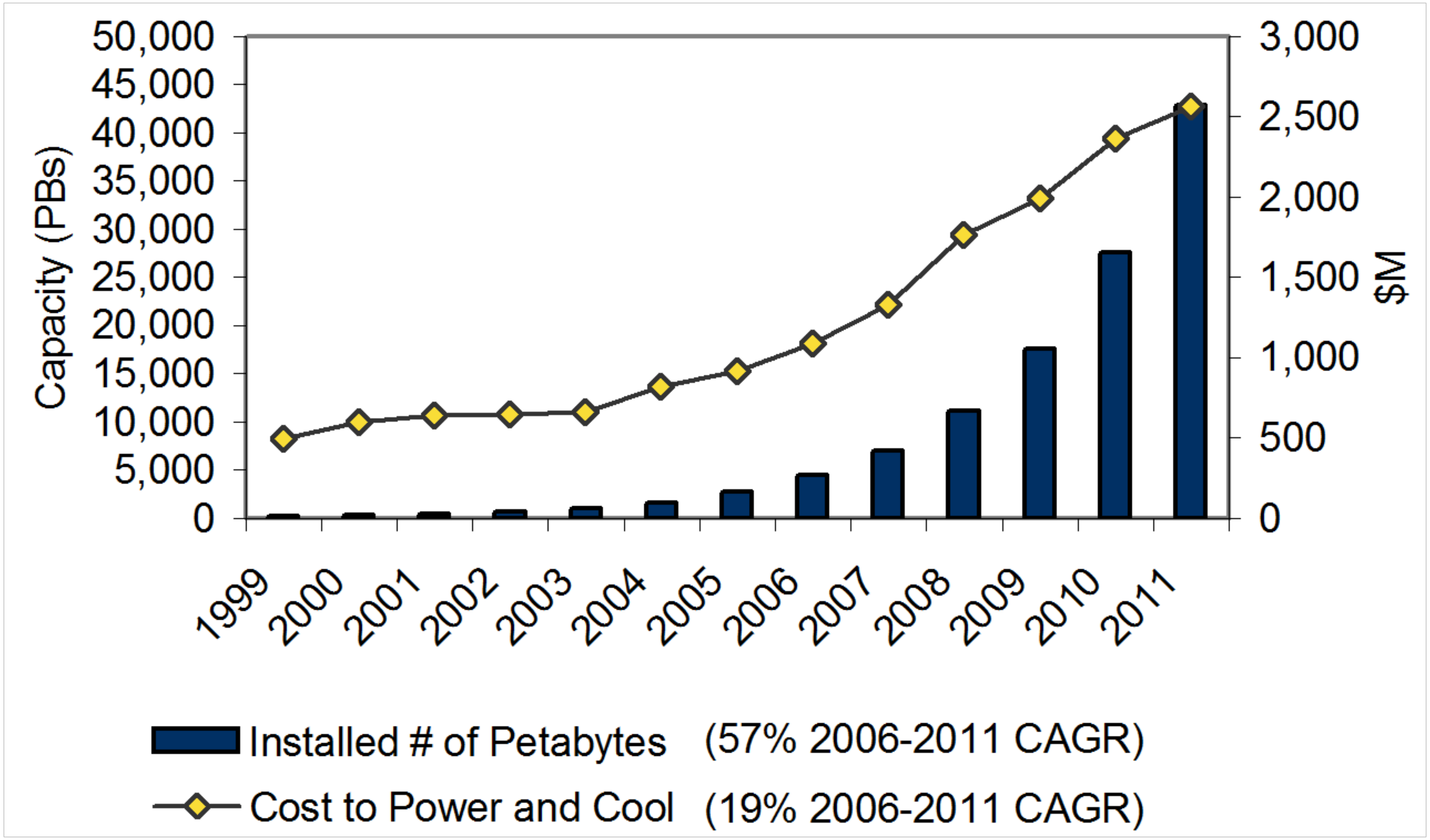
Background and Objectives

- Background
  - basic storage concepts
  - energy & engineering topics covered in "Green Storage I"
- Objectives
  - understand the storage-specific aspects of "green" (especially energy and power) for storage components and aggregated products (e.g. arrays, controllers, appliances)
  - outline considerations in how to measure and model storage energy and power
  - identify technologies that might reduce the ongoing energy costs of reliably maintaining data
  - overview ongoing industry efforts to establish metrics and standards for more energy efficient storage

Outline

- Current state of affairs and industry trends
- Power measurement
  - storage subsystems
  - idle and active modes
  - power supply loading / efficiencies
  - power measurement & monitoring equipment
- Green metrics and taxonomy
  - measuring green-ness
  - storage product categories
- Technologies for efficiency & power reduction
  - hard drives, solid state, tape
  - de-duplication, thin provisioning, MAID
- SNIA green storage efforts
  - unplugged fests, green standards, white papers / workshops, alliances

Cost of Data Storage



IDC #212714, “The Real Costs to Power and Cool All the World's External Storage” – June 2008 Dave Reinsel
Chart used by permission of IDC

What impacts power consumption

Education SNIA

- Storage capacity / usage efficiency
  - increasing data → larger capacity → more disks
  - redundant copies → magnify capacity needs
  - variability in usage and utilization → inefficient allocation of space
  - What is valuable data? What is the retention policy?
- Data transfer rate / access speed
  - high I/O bandwidth → higher rotational speed; striping across many drives
  - low access times → faster actuators; higher rotational speeds; caches
  - How fast and immediate must data be available? (time-to-data)
- Data integrity
  - 25% of "digital universe" is unique, but 75% are replicas / duplicates
  - partly to ensure data integrity and survivability; partly wasteful
- Data availability / system reliability
  - RAID uses extra drives, plus redundant power supplies, fans, controllers,
  - How valuable is data? How likely are failures? How fast must data be available?

Potential paths to "green"

must be driven by metrics / standards / guidelines

- Improve usage efficiency
  - De-duplication
  - Thin provisioning
- Minimize energy consumption
  - Improved component designs – high-efficiency power supplies, advanced & flexible drives
  - Variants of MAID – idle and spin-down
- New technologies
  - Solid state storage
  - Alternative + hybrid system designs (opportunity to rethink)

Anatomy of a Storage System

Education SNIA

- System design, complexity and redundancy vary depending on applications & usage
- Component designs, software features, and workload affect power consumption and efficiency

Apps

Software

Switches

Appliances

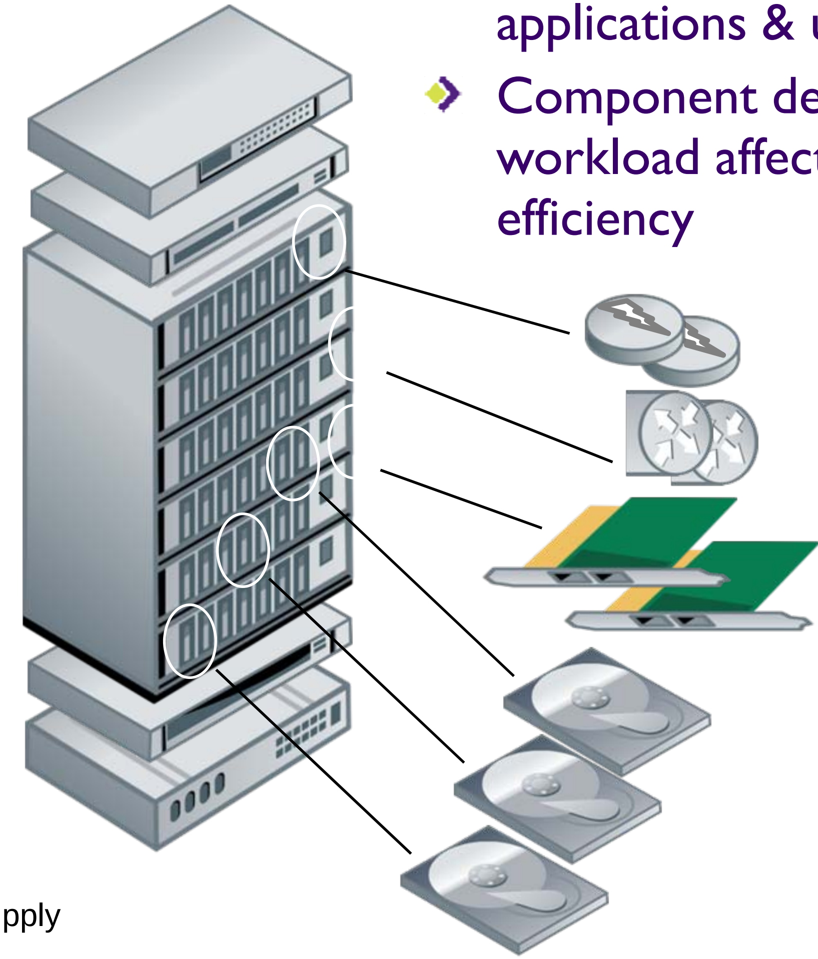
Disk Arrays

PDUs

Power Distribution Unit

UPSs

Uninterruptible Power Supply



Power Supplies

Fans

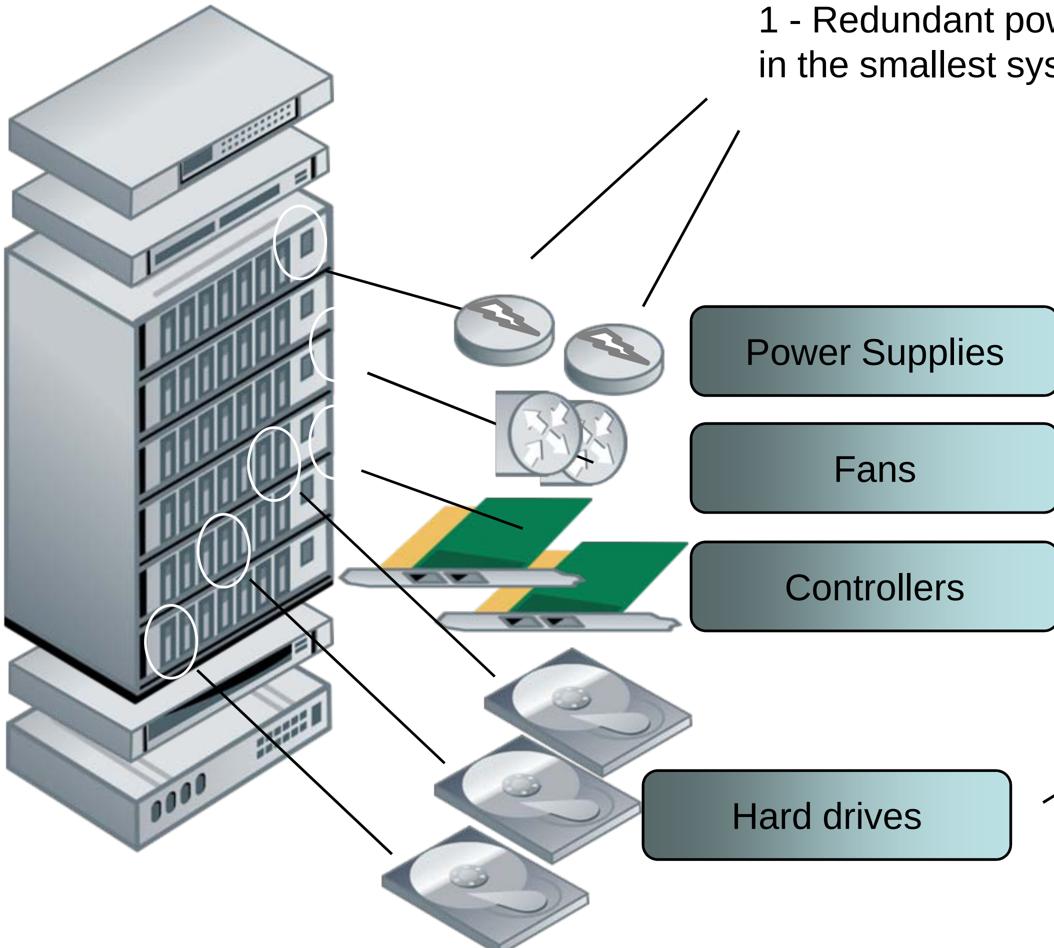
Controllers

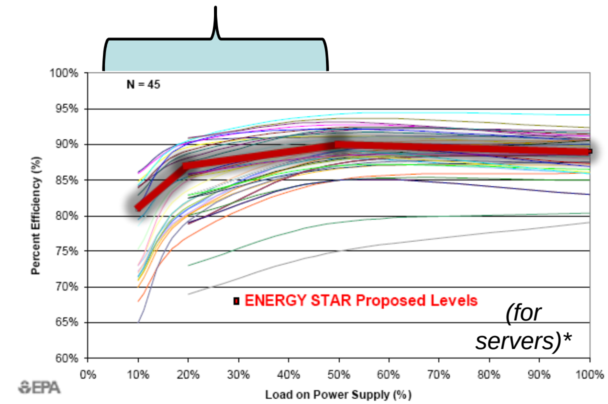
Hard drives

Storage – Power Supply Efficiency

Education SNIA

1 - Redundant power supplies are standard, except in the smallest systems





*presented by EPA at ENERGY STAR Computer Server Stakeholder Meetings; July 2008

2 - Significant mechanical components, require dual-output power supplies (12V, 5V)

3 - Power supplies often custom-designed for reliability

Idle Power vs. Active Power

Education SNIA

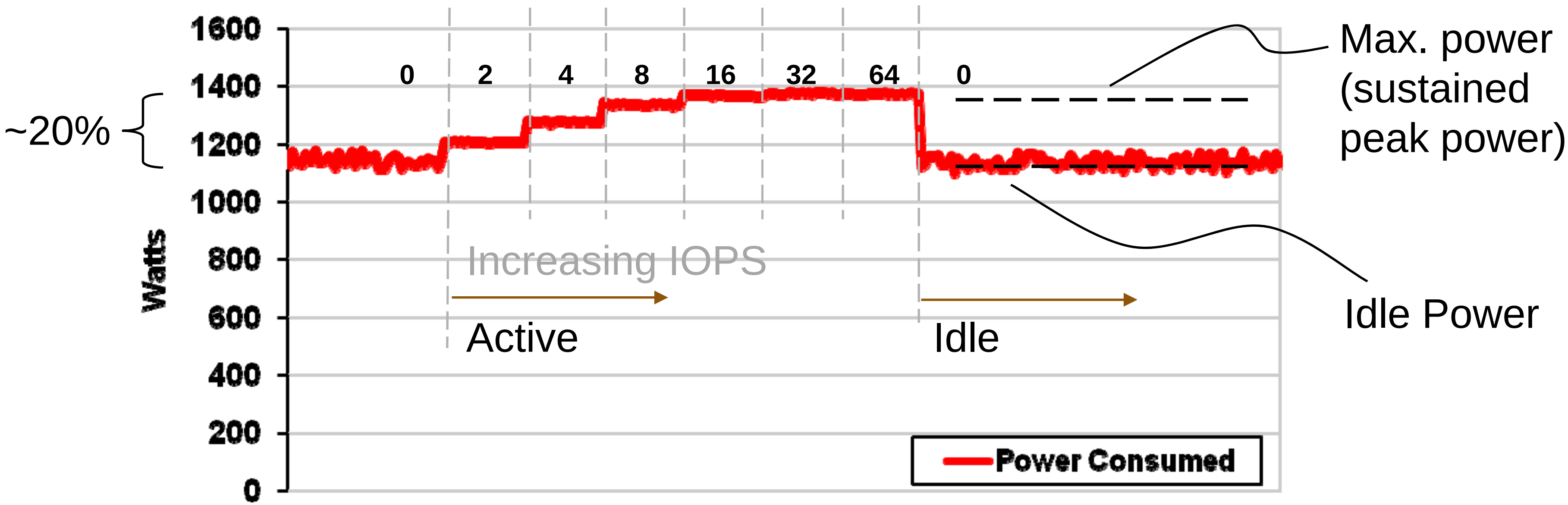
Idle Mode

- storage system is protecting data, ready to process IOs
- background maintenance & optimization tasks on-going
- factors: time-to-data, overhead electronics, fan, maintenance
- systems are idle large fractions of the time

Active Mode

- storage system is carrying out IOs
- background tasks continue in parallel
- factors: workload (seq/random), response time, throughput
- evaluate a variety of workloads, plus sustained peak power

Example of Power Measurement



- Ideally, systems consume minimum power in all modes
  - Example system consumes **significant power in idle (80% of max)**
- % of time in Idle versus Active depends on storage type, application and workloads; available optimizations will vary
- Power consumed is not linearly proportional to workload (indicates potential room for improvement)

Measurement Tools

- Variety of power monitoring & measurement tools available
  - rack-mounted, networked PDUs for continuous operational monitoring
  - more accurate power meters w/ data logging capabilities are preferred for system characterization and benchmarking
  - select a tool based on accuracy, features, Amp/V/Watt levels
  - refer to SPEC website for recommended measurement devices and settings www.spec.org/power_ssj2008/device-list.html
- Measure operating conditions (temp, humidity, altitude) w/ power to establish baselines and understand system behaviors
- Both total and sub-system power consumed are valuable info.

Low-current / voltage power meter

Networked, instrumented rack-mounted PDU

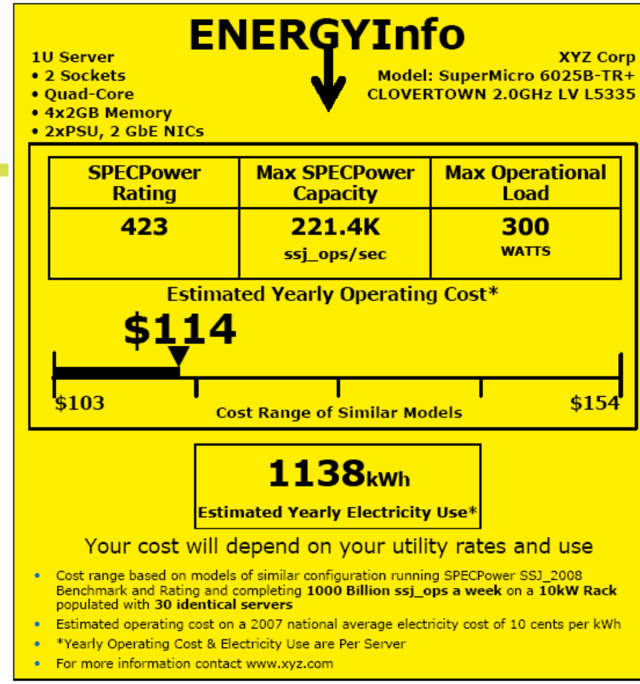
Power meter with data logging and 0.1% accuracy

Green Metrics – purpose / challenge

- Need scientific measures and common vocabulary to assess “green” performance
  - assist in data center design, operational monitoring/tuning and regulatory compliance
- Storage systems have many modes and outputs
  - a single metric (such as GB/Watt) may not reflect the characteristics or capability of the whole system
  - some system “outputs” are not all easily quantifiable
- Multiple metrics may be weighed and combined to form a single metric (e.g. annual energy bill)
  - will vary with usage and system; your mileage will vary

Metrics motivation

Department of Energy labeling program



Computer Server example

Workload considerations

- **Data at rest – Idle power (GB/Watt)**
- Data on the move – Throughput (MB/s)
- Data at work – Performance (IOPS)

Potentially useful metrics

- **GB per Watt;** MB/s per Watt; IOPS per Watt
- Power supply efficiency; CO_2 footprint
- Total annual energy bill (ultimately determined by usage)

Reliability / availability / serviceability considerations

- Latency (time-to-date)
- Redundancy level (RAID efficiency, resilience to failures)

EPA ENERGY STAR certification program

Storage taxonomy

- Need a taxonomy (product classification) to enable fair comparisons among similar storage products
  - e.g. for motor vehicles – motorcycles, cars, trucks
- Similar green metrics may apply to all product categories, but different values establish best-in-class
- Unique considerations apply to special categories
  - e.g. amphibious cars, skid steer loaders, tanks
- Clear taxonomy will simplify comparisons and aid regulatory efforts

Measurement Standard

- Storage taxonomy
- Measurement conditions
- Idle metric
- Reporting results

SNIA
Advancing storage & information technology

SNIA Green Storage Power Measurement Technical Specification

WORKING DRAFT

Version 0.0.18

20 January 2009

Publication of this Working Draft for review and comment has been approved by the Green TWG. This draft represents a "best effort" attempt by the Green TWG to reach preliminary consensus, and it may be updated, replaced, or made obsolete at any time. This document should not be used as reference material or cited as other than a "work in progress."

Taxonomy – Categories

Attribute	Category					
	Online	Near online	Removable Media	Virtual Media Library	Appliance	Interconnect
Access Pattern	Random	Random	Sequential write	Sequential write		
MaxTTD (t)1	t < 80 ms	t > 80 ms	t > 80 ms t < 5 min	t < 80 ms	t < 80 ms	t < 80 ms
User accessible data	Required	Required	Required	Required	Prohibited	Prohibited

- Six categories, covering most storage industry products

Taxonomy – Categories

Attribute	Category					
	Online	Near online	Removable Media	Virtual Media Library	Appliance	Interconnect
Access Pattern	Online 1	Near Online 1	Removable 1	Virtual 1	Appliance 1	Interconnect 1
MaxTTD (t)1	Online 2	Near Online 2	Removable 2	Virtual 2	Appliance 2	Interconnect 2
User accessible data	Online 3	Near Online 3	Removable 3	Virtual 3	Appliance 3	Interconnect 3
	Online 4		Removable 4			
	Online 5		Removable 5			

- 22 total "buckets" covering the breadth of the industry

Taxonomy – Online

- Most common storage systems

Attribute	Classification				
	Online 1	Online 2	Online 3	Online 4	Online 5
Access Pattern	Random	Random	Random	Random	Random
Connectivity	Not specified	Connected to single or multiple hosts, but not shared	Network-connected	Network-connected	Network-connected
Storage Protection	Optional	Optional	Required	Required	Required
FBA/CKD Support	Optional	Optional	Optional	Optional	Required
Maximum Configuration 1	4	> 4	> 20	> 100	> 1000
MaxTTD (t)	t < 80 ms	t < 80 ms	t < 80 ms	t < 80 ms	t < 80 ms
No SPOF	Optional	Optional	Optional	Required	Required
Integrated PDU and UPS	Optional	Optional	Optional	Optional	Required
Rackmount	No	Yes	Yes	Yes	Yes
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Required
User-Accessible Data	Required	Required	Required	Required	Required

Taxonomy – Near Online

Education SNIA

Attribute	Classification		
	Near Online 1	Near Online 2	Near Online 3
Access Pattern	Random	Random	Random
Connectivity	Network connected	Network connected	Network connected
Maximum Configuration 2	4	> 4	> 100
MaxTTD (t)	t > 80 ms	t > 80 ms	t > 80 ms
No SPOF	Optional	Optional	Required
Non-Disruptive Serviceability	Optional	Optional	Required
User-accessible Data	Required	Required	Required

Taxonomy – Removable

Education SNIA

Attribute	Classification				
	Removable 1	Removable 2	Removable 3	Removable 4	Removable 5
Access Pattern	Sequential write	Sequential write	Sequential write	Sequential write	Sequential write
MaxTTD (t)	80 ms < t < 5 m	80 ms < t < 5 m	80 ms < t < 5 m	80 ms < t < 5 m	80 ms < t < 5 m
No SPOF	Optional	Optional	Optional	Optional	Required
Robotics	Prohibited	Required	Required	Required	Required
Maximum Drive Count	Not specified	≤ 4	≥ 5	≥ 25	≥ 12
User-accessible Data	Required	Required	Required	Required	Required

Taxonomy – Virtual

Attribute	Classification		
	Virtual 1	Virtual 2	Virtual 3
Access Pattern	Sequential write	Sequential write	Sequential write
Connectivity	Network connected	Network connected	Network connected
FICON Support	Optional	Optional	Required
Maximum Configuration 1	100	> 100	> 100
MaxTTD (t)	t < 80 ms	t < 80 ms	t < 80 ms
No SPOF	Optional	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Required
User-accessible Data	Required	Required	Required

Taxonomy – Appliance & Interconnect

Attribute	Classification		
	Appliance 1	**Appliance 2**	**Appliance 3**
Connectivity	Direct or Network connected	Direct or Network connected	Network connected
Maximum Configuration 1	20	> 20	≥ 100
MaxTTD (t)	t < 80 ms	t < 80 ms	t < 80 ms
No SPOF	Optional	Optional	Required
Non-Disruptive Serviceability	Optional	Optional	Required
User-accessible Data	Prohibited		

Attribute	Classification		
	Interconnect 1	**Interconnect 2**	**Interconnect 3**
Maximum Port Count (n)	Switch: n ≤ 32	Router: n ≤ 4 Switch: 32 ≤ n ≤ 128 Extender: ≤ 4	Router: > 4 Switch: > 128
No SPOF	Optional	Optional	Required
User-accessible Data	Prohibited	Prohibited	Prohibited

Storage Power – Idle

Equation 6-1: Average Idle Power

$$P_i = \frac{\sum W_i}{n}$$

Where:

- P_i *is average idle power*
- W_i *is power in watts measured in each sampling interval i*
- *n is the number of samples gathered by the power meter during the measurement interval.*

- Idle Metric

Equation 7-1 SNIA Idle Power Metric

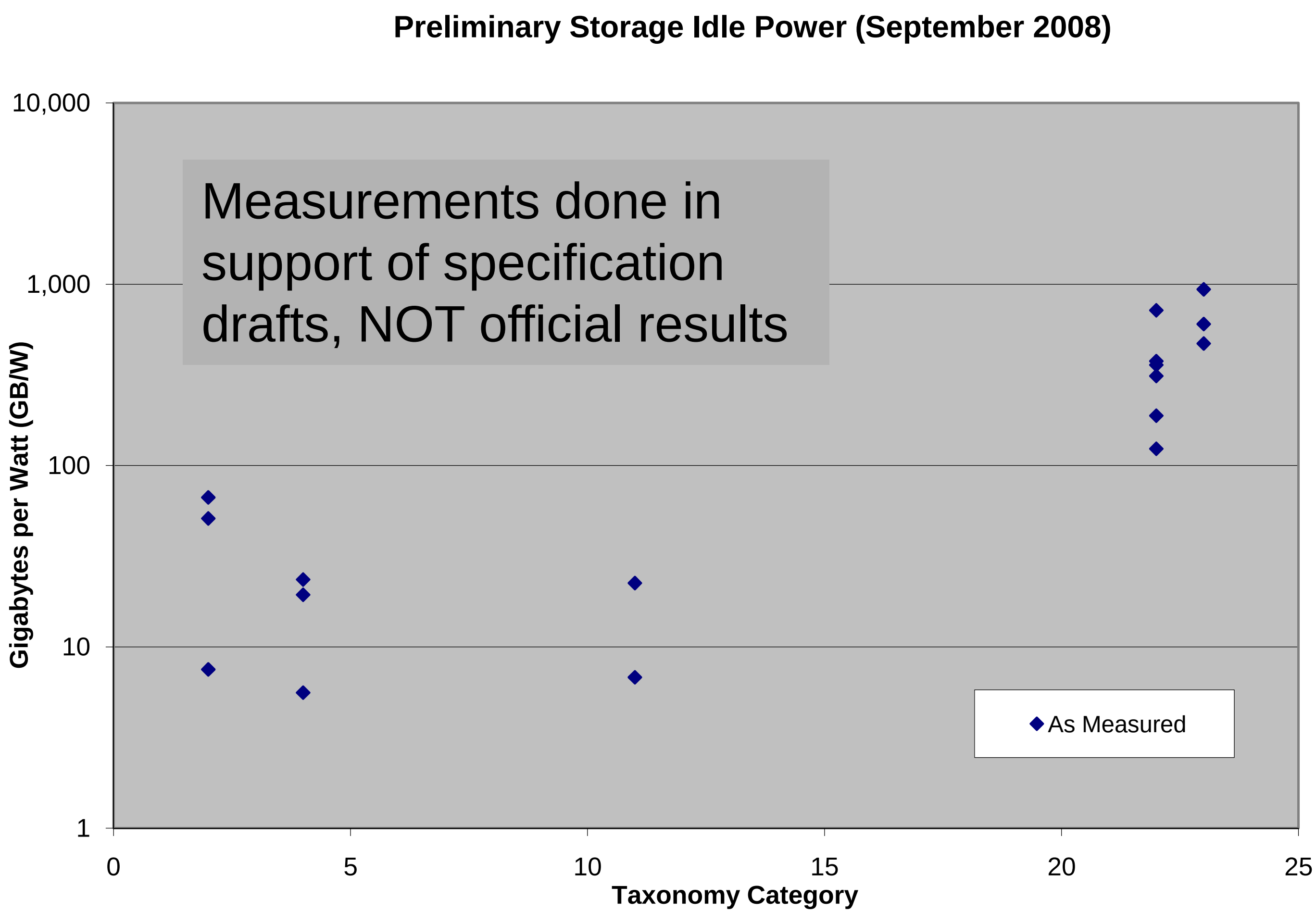
$$P = \frac{C}{P_i}$$

Where:

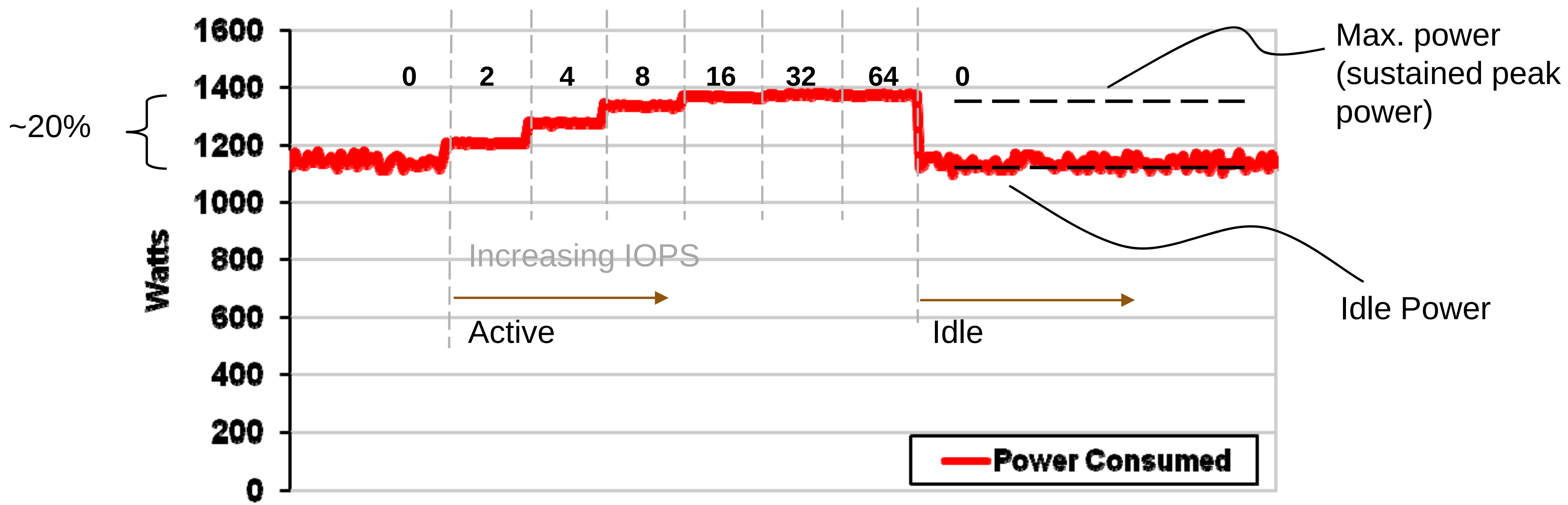
- *P is the SNIA Idle Power Metric*
- *C is the total capacity of the SUT*
- P_i *is the average idle power*

Storage Power – Idle (2)

- Wide range of GB/W values across taxonomy categories



What's Next – Active Power



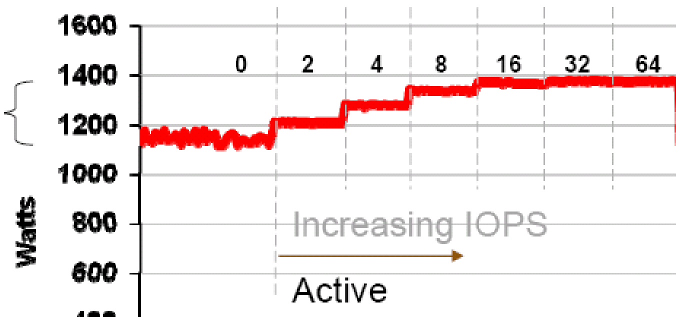
- Many variations in workloads
- Many variations in system configuration
- Takes us into the realm of benchmarking

Desired Metric – "Productivity"

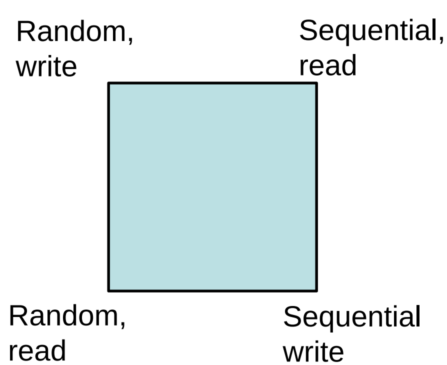
Education SNIA

Many possible definitions – must balance simplicity against applicability

- "typical workload", with levels



- "four corners", maximum performance, maximum power



- detailed performance benchmark – results/W

Storage Performance Council

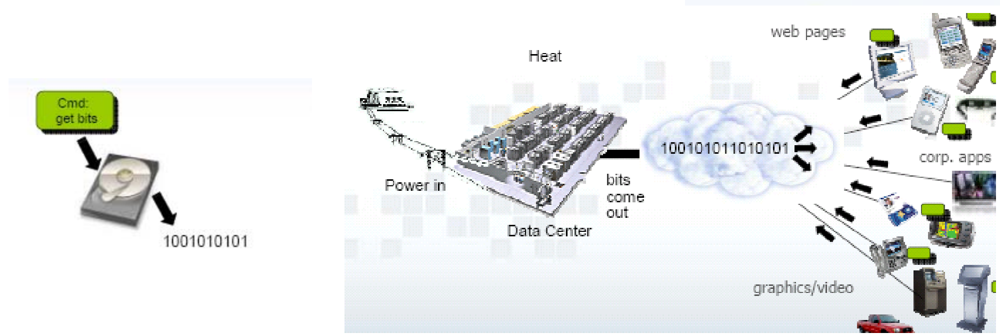
Defining, administering, and promoting industry-standard, vendor-neutral benchmarks to characterize the performance of storage products

spec

Standard Performance Evaluation Corporation

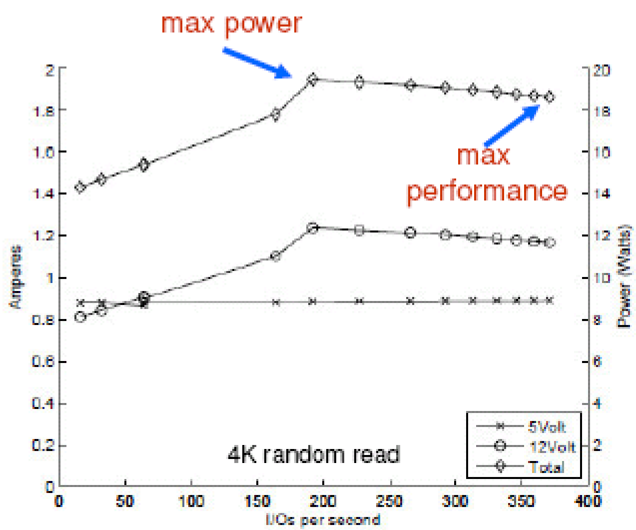
- The Green Grid Productivity Proxy Proposals example – Proxy #4 – bits/kilowatt-hour

the green grid SM
get connected to efficient IT



Complications

- Max power =/= Max performance



Single disk drive power profile

Storage Modeling for Power Estimation

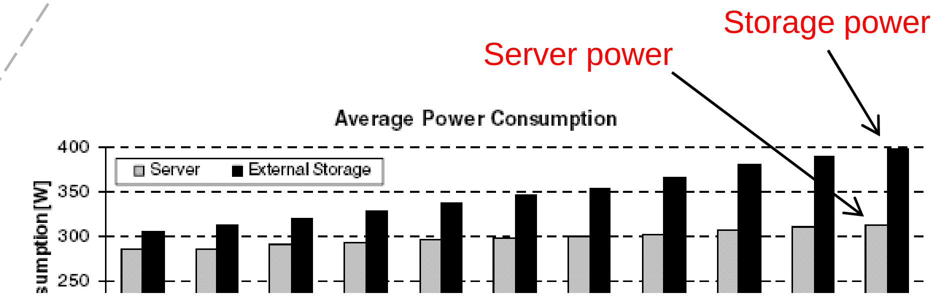
Miriam Allalouf • Yuriy Arbitman • Michael Factor •
Ronen I. Kat • Kalman Meth • Dalit Naor •

IBM Haifa Research Labs

ABSTRACT

Power consumption is a major issue in today's datacenters. Storage typically comprises a significant percentage of datacenter power. Thus, understanding, managing, and reducing storage power consumption is an essential aspect of any efforts that address the total power consumption of datacen- ... decade. We observe that the power consumption of disks is composed of fixed and dynamic portions. The fixed portion is consumed in the idle state and includes items such as the power consumed by the spindle motor. The dynamic factors are affected by the I/O workload and include items such as the power for data transfers

"Storage Modeling for Power Estimation", Miriam Allalouf , Yuriy Arbitman, Michael Factor, Ronen I. Kat, Kalman Meth, and Dalit Naor; IBM Haifa Research Labs; manuscript; March 2009



SPECweb 2005 (banking) + storage

- Significant whole-system considerations

The Next Frontier for Power/Performance Benchmarking: Energy Efficiency of Storage Subsystems

Klaus-Dieter Lange

Hewlett-Packard Company, 11445 Compaq Center Dr. W, Houston, TX-77070, USA
Klaus.Lange@hp.com

Abstract. The increasing concern of energy usage in datacenters has drastically changed how the IT industry evaluates servers. The energy conscious selection of storage subsystems is the next logical step. This paper first quantifies the possible energy savings of utilizing modern storage subsystems by identifying inherent energy characteristics of next generation disk IO subsystems. Additionally, the power consumptions of a variety of workload patterns is demonstrated.

Keywords: SPEC, Benchmark, Power, Energy, Performance, Server, Storage, Datacenter.

1 Introduction

Today's challenge for datacenters is their high energy consumption [1]. The demand for efficient real estate in datacenters has moved to more power efficient datacenters. This increasing concern of energy usage in datacenters has drastically changed how the IT industry evaluates servers. In response, the Standard Performance Evaluation

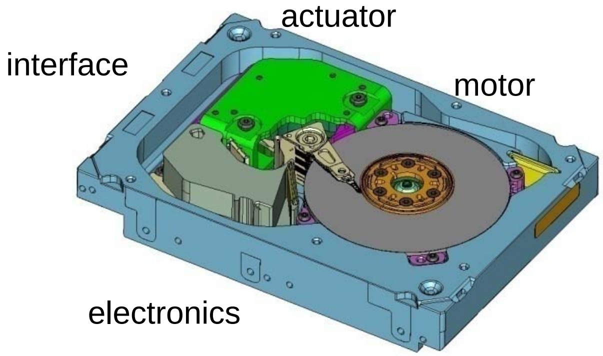
"The Next Frontier for Power/Performance Benchmarking: Energy Efficiency of Storage Subsystems" Klaus-Dieter Lange; SPEC Benchmark Workshop 2009; January 2009

Outline

- Current state of affairs and industry trends
- Power measurement
  - storage subsystems
  - idle and active modes
  - power supply loading / efficiencies
  - power measurement & monitoring equipment
- Green metrics and taxonomy
  - measuring green-ness
  - storage product categories
- Technologies for efficiency & power reduction
  - hard drives, solid state, tape
  - de-duplication, thin provisioning, MAID
- SNIA green storage efforts
  - unplugged fests, green standards, white papers / workshops, alliances

Hard Disk Drives

Education SNIA



7200 RPM

high capacity

15K RPM

high performance

3.5 inch HDD

5400 RPM

laptop

2.5 inch HDD

- Besides component improvements, power reduction achievable if you
  - Select large capacity drives
    - high GB/spindle or GB/drive → high GB/Watt
  - Employ small form-factor (SFF, 2.5”) drives
    - low-energy spindle & fast seek → high IOPS/Watt
  - Utilize drive power management features (idle and stand-by modes)
    - power saving possible without major impact on performance or time-to-data
- Component improvements
  - Efficient motors, workload-optimized actuators, flexible electronics

Solid State Drives

Education SNIA

- Definition
  - SSD is "a semiconductor-based block storage device that behaves as a virtual HDD and appears to the host device as a disk drive"

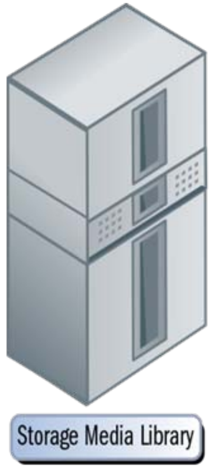
    IDC, Worldwide Solid State Drive 2007-2011 Forecast and Analysis: Finding Space in the Expanding Digital Universe, Doc # 207739, July 2007.
  - based on a variety of possible technologies – NAND flash (SLC or MLC); DDR RAM + battery; future Fe / M / R / PC RAM
- Characteristics (flash SSD compared to HDD)
  - fast read/write (~15 ns), low latency for random IO
  - good $$/IOPS but high $$/GB
  - low power consumption, unproven reliability
- Opportunities for energy efficiency
  - target high IOPS access for mixed systems, optimizing workload
  - provide an additional power-efficient layer of caching

Check out SNIA tutorial:

Solid State Storage for the Enterprise by W. Hutsell, B. McKean

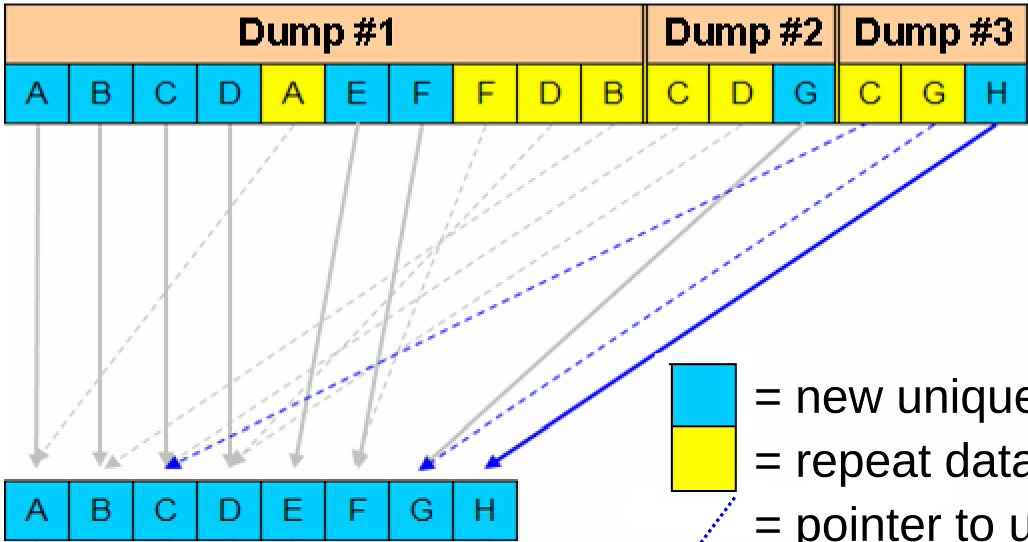
Tapes

- Proper utilization of tiered storage can reduce power consumption (OpEx) as well as purchases (CapEx)
- Clear benefit of tape is low to zero power consumed while data is stored in high volumetric density
- Sequential technology – not suitable for fast, random access, low time-to-data
- Great for secondary backup; tradeoffs for primary backup between HDD (archive storage, VTL) and tape
- The number of tape drives and use of robotics increases power, but enables very large libraries



Check out SNIA Tutorial:
Introduction to Data Protection: Backup to Tape, Disk, and Beyond by J. Iehl

Deduplication



= new unique data

= repeat data

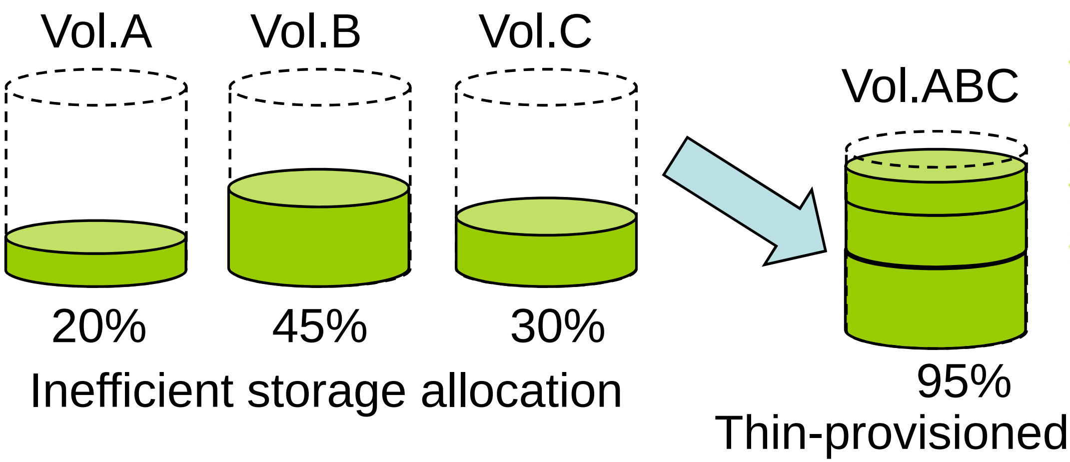
= pointer to unique data segment

- evaluate data
- identify redundancy
- create or update reference info
- store or transmit unique data once
- recall (read) or reproduce data

- Store (& transmit) new and unique data only (at sub-file level); update tracking information; further compression is possible
- Applicable to backup, some primary storage, disaster recovery, archive and wide area optimizations
- Must consider restore speeds (latency, bandwidth, data type); results will be application and data type dependent

Check out SNIA Tutorial:
Deduplication *Methods for Achieving Data Efficiency* by M. Brisse, G. Senderov

Thin Provisioning

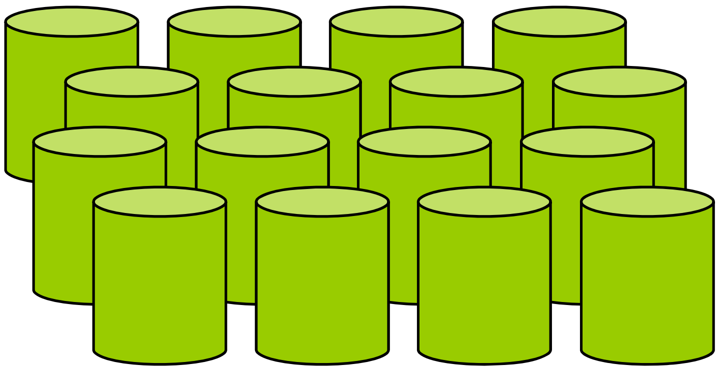


- user demands vary over time
- over-allocation anticipates future needs
- low capacity utilization
- low performance utilization

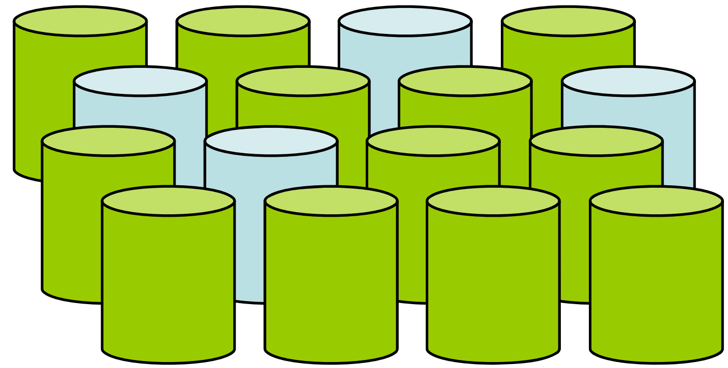
- Consolidate storage space to
  - improve utilization efficiency
  - reduce overall data center footprint
- Helpful to have good forecast of workloads
- Need dynamic provisioning to achieve best results

Check out SNIA Tutorial: **Software Technologies for Green Storage** by Alan Yoder

MAID



All disks spinning full-speed; high performance but no power saving



25% disks spun down; up to 25% power saving but some performance penalty

- MAID – Massive Array of Idle Disks
  - Spin down a fraction of idling drives – requires predictable workload
  - Lowers power, increases time-to-data
- Multiple MAID levels (10%, 25%, 50%) can optimize performance and power for different scenarios and applications
- Maximum power must be considered for design points

Outline

Education SNIA

- Current state of affairs and industry trends
- Power measurement
  - storage subsystems
  - idle and active modes
  - power supply loading / efficiencies
  - power measurement & monitoring equipment
- Green metrics and taxonomy
  - measuring green-ness
  - storage product categories
- Technologies for efficiency & power reduction
  - hard drives, solid state, tape
  - de-duplication, thin provisioning, MAID
- SNIA green storage efforts
  - unplugged fests, green standards, white papers / workshops, alliances

SNIA Green Efforts

- SNIA Green Storage Initiative (GSI) and SNIA Green Storage Technical Work Group (TWG)
  - on-going efforts to develop data-driven green standards & metrics
  - power measurements at multi-vendor “unplugged” fests
  - alliances with other active green organizations
    - (The Green Grid, 80PLUS/Climate Savers, DMTF, SPEC, SPC)
  - collaboration with EPA on the ENERGY STAR for Storage program
- Whitepapers / workshops
  - four tutorials at SNW; online tutorials available (www.snia.org/education/tutorials)
  - white papers from GSI

towards energy-efficient storage networking solutions

Summary

Education SNIA

- "Green is good" – for multiple reasons
  - a great engineering problem – doing more with less
  - saves money – great investment payback
  - helps save the planet – significant leverage
- Get involved with SNIA Green efforts
  - weekly discussions, regular face-to-face & "unplugged" fests (TWG)
    - Upcoming face-to-face meeting at SNIA TSG on May 12 & 13, 2009 in Chicago
    - Get a power meter and try the measurement spec on your own systems (!)
  - education and promotion (GSI)
  - promote these industry-wide efforts within your company
- Learn about wider green technology and opportunities
  - online resources; workshops by SNIA, EPA, The Green Grid
- Share your experience / knowledge

References

- SNIA Green Storage Initiative – www.snia.org/green
- The Green Grid – www.greengrid.org
- EPA ENERGY STAR
  (Data Center Energy Efficiency Initiatives –
  www.energystar.gov/index.cfm?c=prod_development.server_efficiency
- DOE Federal Energy Management Program –
  eere.energy.gov/femp
- Power calculators at various vendor sites
- SNIA Green Storage Outreach
  - www.snia.org/forums/green

Q&A / Feedback

- Thank you for your attention!
- Please send any questions or comments on this presentation to SNIA: trackstorage@snia.org and trackgreenstorage@snia.org

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

SNIA Education Committee

Patrick Chu	SNIA Green Storage Initiative members
Matthew Brisse	SNIA Green Storage TWG members
David Reinsel	
Edgar St.Pierre	
Willis Whittington	
Alan Yoder	
Wayne Adams	