Improvements in Storage System Energy Efficiency via Storage Subsystem Cache and Tiering

Chuck Paridon, H-P Master Storage Performance Architect
Herb Tanzer, H-P Storage Hardware Architect

SNIA Storage Developers Conference
September 16-19, 2013

Outline

SNIA SDC

- Storage Energy Efficiency Basics
  - Power / Performance Metrics, optimal power performance
  - Industry standards and status
- Methods to Improve Energy Efficiency
  - Capacity Optimization Methods (COMs) description and potential benefits
  - Product features
- Benefits of Caching and Smart Tiering
  - Increase IO performance via cache hits
  - Increased IOPs and decreased power consumption via tiering

Storage Energy Efficiency: Basics

SNIA SDC

- TGG definition
  - Efficiency is the % of time that IT equipment is utilized providing primary services
  - Productivity is useful work during operation in a data center, divided by total data center energy consumed producing this work
- SNIA, SPC definition
  - [Useful storage Work / Power consumed]
    - Useful Storage Work is:
      Ready idle – Capacity, Active – Workload, Throughput
    - Power consumed in Watts
  - Storage Energy Efficiency Metrics
    - GB/Watt
    - IOPs/Watt (IOs/sec/Watt)
    - MBPS/Watt (MB/sec/Watt)

Storage Energy Efficiency: Industry Standards

SNIA SDC

- Storage Performance Council
  - SPC Benchmarks SPC-1/E™ and SPC-2/E™ are energy extensions of the SPC-1™ (random workload) and SPC-2™ (sequential workloads)
- SNIA Emerald Program
  - SNIA Emerald™ Power Efficiency Measurement Specification has a well-defined Storage Taxonomy and testing procedure
- EPA Energy Star for Data Center Storage
  - V1.0 Spec being released Aug'13, and will become effective by late November
  - Requires a set of checklist items and data collection based upon the Emerald Specification. Data submittals are centered about an Optimal point and allowable extensions that determine the qualified range or product SKU's

Capacity Optimization Methods

SNIA SDC

- COMs are a set of techniques which reduce the consumption of space required to store a data set….this may result in less energy usage for a given task.
  - Compression, Data De-duplication, Thin Provision, and Delta snapshots.
  - Also include: Parity RAID as compared to mirroring (Raid 1)
- Including COMs in energy efficiency metrics
  - SPC is considering to allow COMs (thin provision, compression, de-duplication); will need to disclose
  - SNIA Emerald leaves this up to user discretion; will need to disclose
  - Energy Star requires a minimum number of COMs be made available as a selectable feature and that they be validation tested, but they are to be disabled during actual metric testing

Energy Efficient Storage: SNIA Emerald Program Specific COM Examples (70/30 R/W random)

SNIA SDC

COM Type	Raid Parity	Data Compression
COM Method	Interleaved Data and Parity (R5)	Grouped Data Patterns
Baseline Config	Raid 1 Mirrored and Striped	Uncompressed
Baseline Capacity (GB)	36,000	36,000
Baseline Power (W)	4046	4046
Baseline (GB/W)	8.90	8.90
Baseline Perf. (IOPs)	57,750	57,750
Baseline (IO/W)	14.27	14.27
Optimized Capacity (GB)	45,000	72,000
Optimized (GB/W)	11.12	17.79
Delta (GB/W)	+25%	+100%
Opt. Power (W)	4046	4046
Opt. Perf (IOPs)	37,310	115,500
Optimized (IO/W)	9.22	28.55
Delta (IO/W)	(35%)	+100%

Based on modeled data

Product Features

SNIA SDC

- COMs can improve the idle capacity efficiency metric, but when subjected to an active workload the resultant power/performance metric may or may not improve
- Caching & auto-tiering are features which under the right conditions are capable of improving the power/performance metric
  - SPC-1 has non-uniform access patterns (hot-spots)
  - Emerald test phases include workloads containing cache-friendly hot-band IO profiles, and allow for an optional migration phase to allow data relocation into appropriate tiers
  - Energy Star uses the Emerald test procedure, thus allows taking advantage of these features

Workloads Amenable to Caching and Tiering Optimization

SNIA SDC

- Typical workloads are comprised of several different IO streams, some of which contain hot spots, or regions of more intense IO demand.
- This results in varying degrees of logical block address re-referencing in certain regions of the overall IO space that can be either contained with the cache of an array, or can be placed on storage devices that deliver a higher IO rate.

Hot Band IO Patterns

SNIA SDC

Hot Band IO Profile *

IO Profile	% of workload	Read/Write Percentage	IO Size (KiB)	Access Pattern	Usable Address Range
Write Stream 1	5	0/100	See Table 12	Sequential	0-100%
Write Stream 2	5	0/100	See Table 12	Sequential	0-100%
Write Stream 3	5	0/100	See Table 12	Sequential	0-100%
Read Stream 1	5	100/0	See Table 12	Sequential	0-100%
Read Stream 2	5	100/0	See Table 12	Sequential	0-100%
Read Stream 3	5	100/0	See Table 12	Sequential	0-100%
Read Stream 4	5	100/0	See Table 12	Sequential	0-100%
Read Stream 5	5	100/0	See Table 12	Sequential	0-100%
Uniform Random	6	50/50	See Table 12	Random	0-100%
Hot Band 1	28	70/30	See Table 12	Random	10 -18%
Hot Band 2	14	70/30	See Table 12	Random	32-40 %
Hot Band 3	7	70/30	See Table 12	Random	55-63 %
Hot Band 4	5	70/30	See Table 12	Random	80-88 %

Hot Bands concentrate 54% of the IO in 32% of the space

* - invented and contributed to SNIA by Steven Johnson – Oracle Corp.

Energy Efficiency Benefits

SNIA SDC

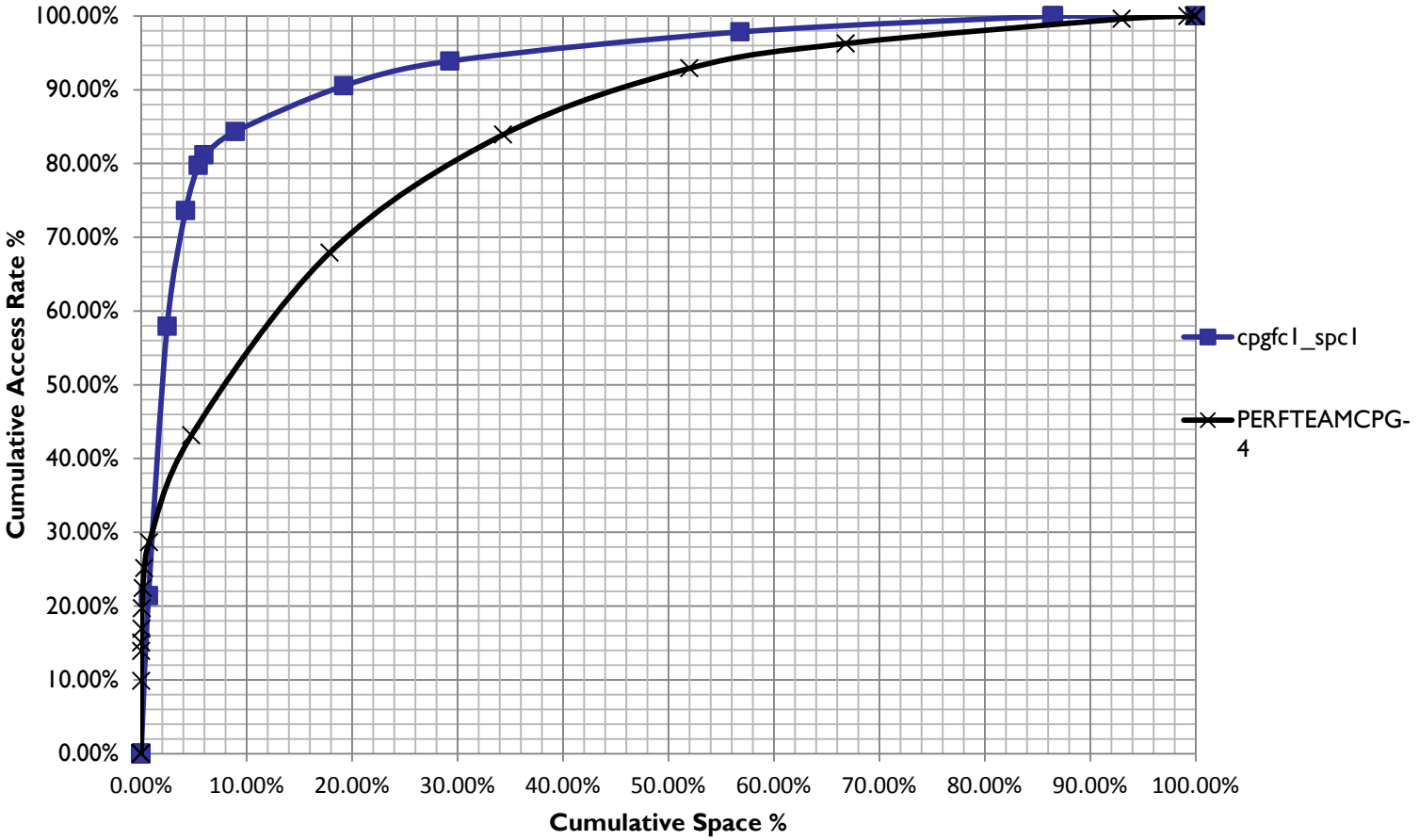
- These workloads, when run on High End Storage demonstrate the power/performance advantage of two product features
  - Array Based Cache
  - Storage Tiering
- Although the primary benefit of using this workload was solely cache focused, there is also a benefit of implementing faster tiers in the product such as HP P9500 Smart Tiering

Energy Efficient Performance Comparison Chart
Cache and Tiering Speedup on HP Storage Products

SNIA SDC

	Typical IOPS	Typical RT	Rnd IOPS	Rnd RT	C/WS ratio	Cache Hit Ratio
Small Array	4330	32.8 ms	4130	33.4 ms	<<	N/A
Large Array Cache Assist only	39900	8.97 ms	18410	22.59 ms	~3%	60%/24%
Large Array (Tiered)	42870	5.77 ms	N/A	N/A	~3%	60%/24%

Result of hot band workload on array back end after filtering through Storage Subsystem Cache



Blue line is the SPC-1 workload IO density data. Black is Hot Band. The steeper the curve and the sharper the knee, the more the IO rate is enhanced. SPC speedup ~50% Hot Band ~30%
So total speedup due to both cache and tiering ~ 100 + (40 *0.3) = 132%

- Derivation of Optimal Tiered Storage Configuration requires 2 processes :

- The Workload Analysis Process Consists of Two Steps

1. Create a single pool large enough to hold the desired working set
2. Run the workload on that pool and use the analysis tools to produce a report guiding the composition of the tier(s)

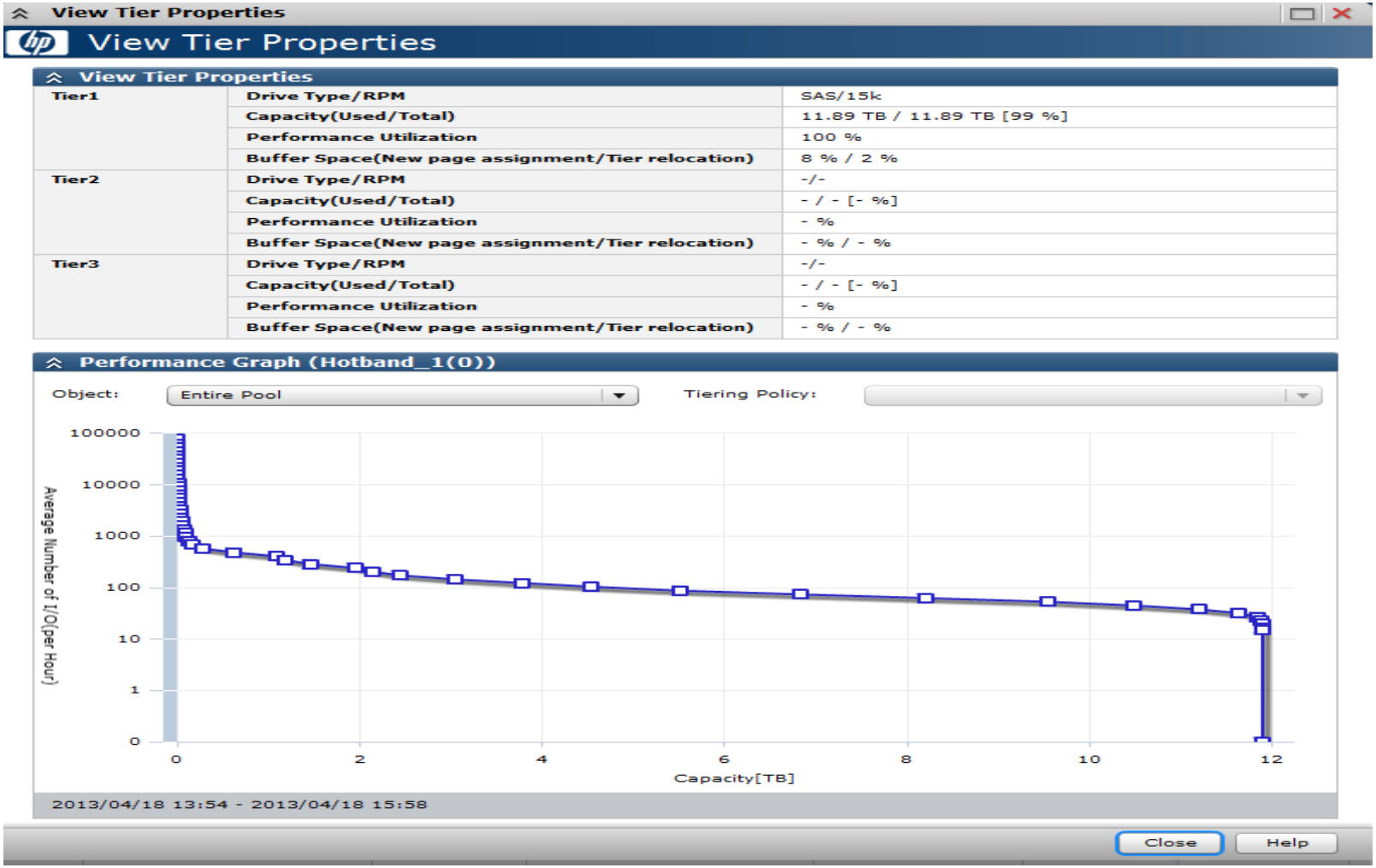
- The tier construction process involves 2 decision points.

1. Which technology to deploy
2. The capacities of the tier(s)

- The following slides illustrate an example of these activities

Energy Efficient Tiered Storage: Initial Tier Property Analysis

SNIA SDC



Energy Efficient Tiered Storage: Final Tier Property Analysis

SNIA SDC

View Tier Properties

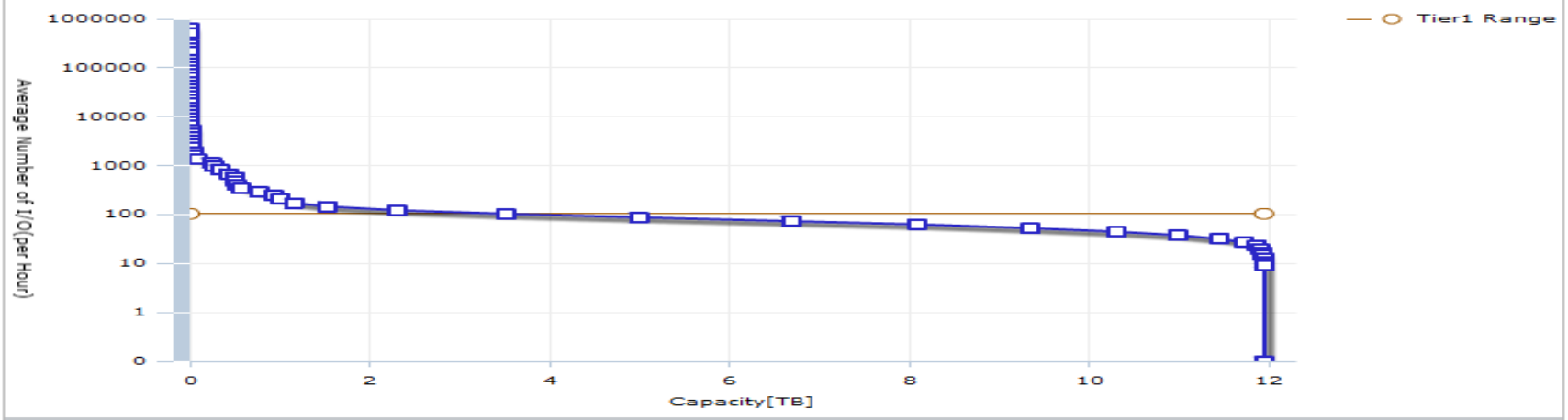
View Tier Properties

View Tier Properties

Tier1	Drive Type/RPM	SSD/-
	Capacity(Used/Total)	2.50 TB / 2.50 TB [100 %]
	Performance Utilization	74 %
	Buffer Space(New page assignment/Tier relocation)	0 % / 2 %
Tier2	Drive Type/RPM	SAS/15k
	Capacity(Used/Total)	9.43 TB / 10.06 TB [93 %]
	Performance Utilization	59 %
	Buffer Space(New page assignment/Tier relocation)	8 % / 2 %
Tier3	Drive Type/RPM	-/-
	Capacity(Used/Total)	- / - [- %]
	Performance Utilization	- %
	Buffer Space(New page assignment/Tier relocation)	- % / - %

Performance Graph (hot_band(0))

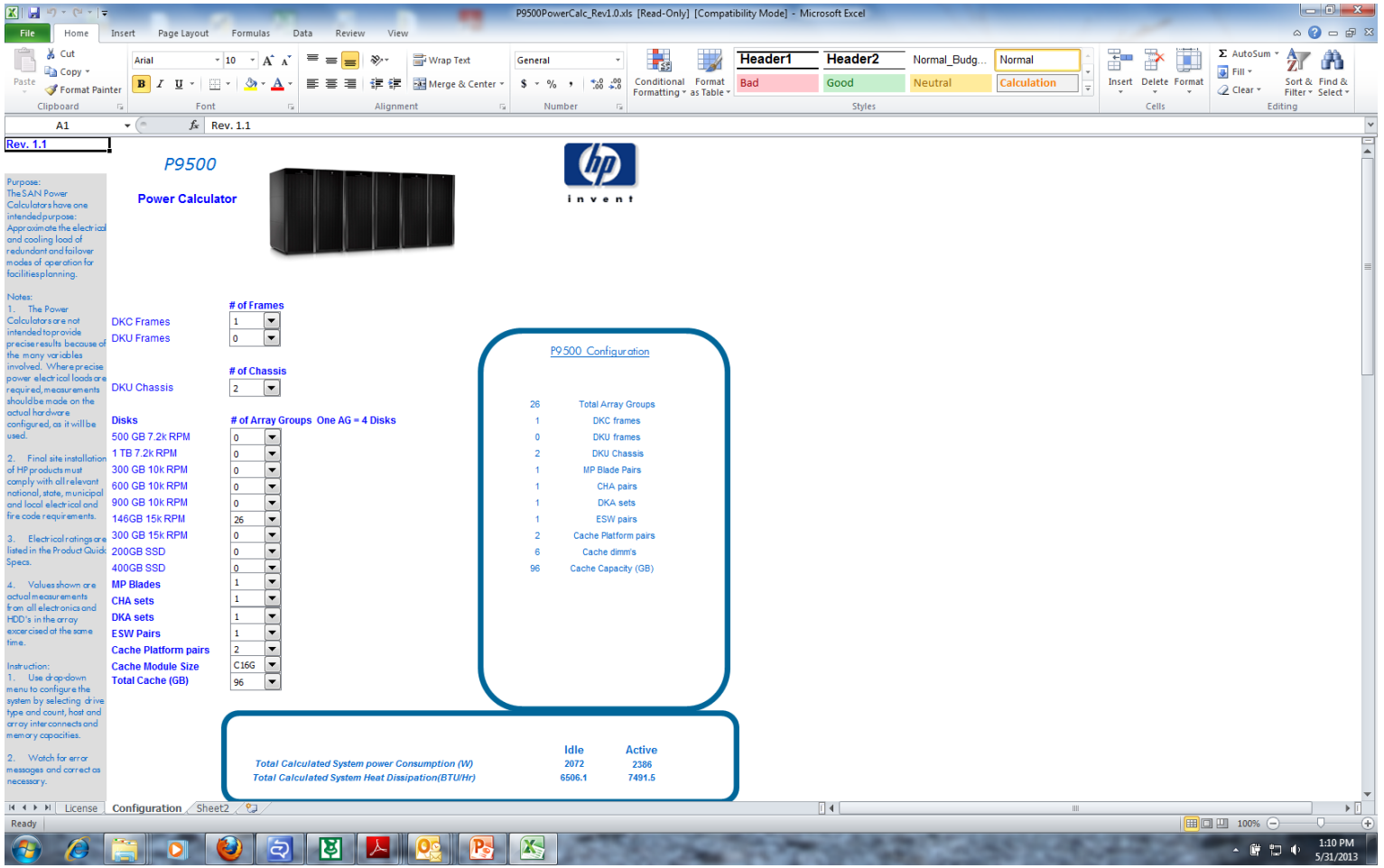
Object: Entire Pool Tiering Policy:



2013/04/26 09:54 - 2013/04/26 11:56

Close Help

The HP P9500 Power Calculator



Energy Efficient Tiered Storage: SNIA Emerald Program Primary Metric Comparison

SNIA SDC

Configuration	Tier Type	Power Consumption (Modeled)	IOPS	IOPS/Watt
Large Array (Initial)	15k RPM	7491 Watts	18410	2.457
Large Array (Cache Assist)	15k RPM	7491 Watts	39900	5.326
Large Array (Tiered)	15k + SSD	7283 Watts	42870	5.886

Energy Efficient Storage: Conclusions and Observations

SNIA SDC

- The Hot Band workload is amenable to performance optimization by both storage subsystem cache and the proper deployment of tiered storage.
- As a result of high cache hit rates (~60%) the overall performance contribution of tiered storage is limited.
- In addition to the increase in IO rate (132%), there is also a corresponding decrease in power consumption from the substitution of SSDs in the configuration
- The net effect of these two parameter changes is a 140% improvement in the SNIA primary active metric (IOPS/Watt) of

Thank You

SNIA SDC

- Questions?

Contact: chuck.paridon@hp.com
herb.tanzer@hp.com