Energy Efficiency Metrics for Storage

Herb Tanzer, Chuck Paridon
Hewlett Packard Company

Outline

- Energy Efficiency Basics
  - Tutorial; (work done/ power consumed)
- Storage Performance Council
  - Summary & sample HP data
- SNIA EMERALD™
  - Summary, predictions, and sample HP data
- EPA Energy Star™ for Data Center Storage
  - Expected requirements for V1.0
- TGG – DCsE (operational metrics)
  - Summarize TGG white paper
  - Use case: tiered storage COM
- Summary compare and key benefits of Metrics

Energy Efficiency Basics

- Energy efficiency
  (Useful work done / Power consumed)
- Storage useful work
  - Ready idle: Capacity
  - Active: IOs/sec, Throughput
- Storage energy efficiency metrics
  - GB/watt
  - IOPs/watt
  - MBPS/watt

Storage Performance Council

- SPC-1/E™

SPC Benchmark 1/Energy™ *(SPC-1/E™) is* an extension of the first industry standard storage benchmark, SPC Benchmark 1™

SPC-1™ is based on an exchange workload.

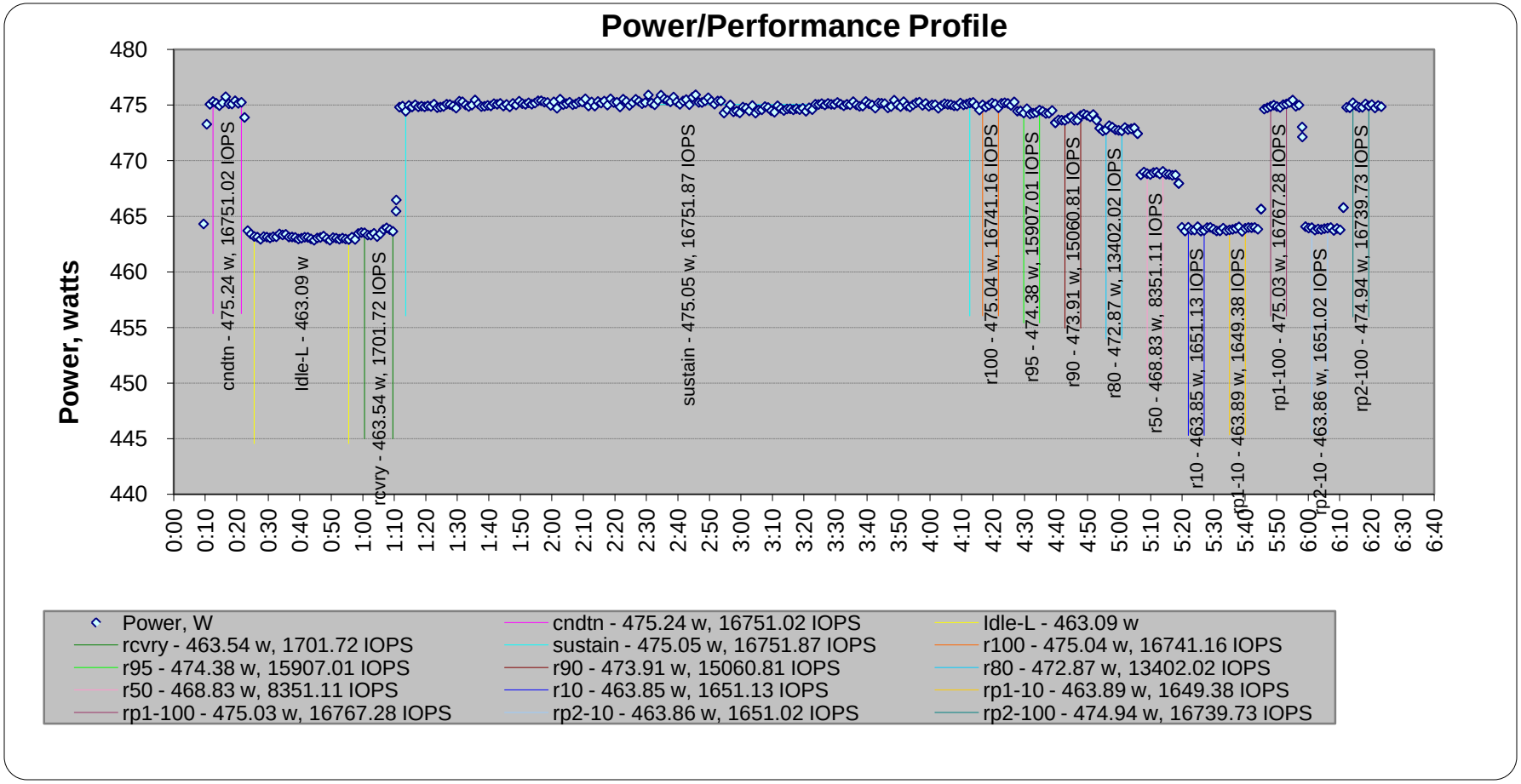
- SPC-2/E™

SPC-2/E™ incorporates the SPC Energy Extensions into the sequential workloads of SPC-2™

The SPC-2™ Workload is comprised of large block reads, writes and combinations thereof in addition to a Video on Demand Workload

Sample SPC data

HP EVA P6400 Power/Performance Profile



Sample SPC data (continued)

HP EVA P6400 Power/Performance Table

SPC-1/E Reported Data

The initial SPC-1/E energy extension temperature, recorded during the first one minute of the Idle Test was 81.50F. The final SPC-1/E energy extension temperature, recorded during the last one minute of the Primary Metrics Test was 82.64F.

Power Environment

Average RMS Voltage: 211.14 **Average Power Factor:** 0.809

Usage Profile

	Hours of Use per Day			Nominal	Nominal	Nominal	Nominal
	Heavy	Moderate	Idle	Power, W	Traffic, IOPS	IOPS/W	Heat, BTU/hr
Low Daily Usage:	0	8	16	227.85	3333.14	14.63	777.44
Medium Daily Usage:	4	14	6	231.49	8499.58	36.72	789.87
High Daily Usage:	18	6	0	235.84	14499.49	61.48	804.72
Composite Metrics:				231.73	8,777.40	37.88	

Annual Energy Use, kWh: 2,029.93

Energy Cost, $/kWh: $ 0.12

Annual Energy Cost, $: $ 243.59

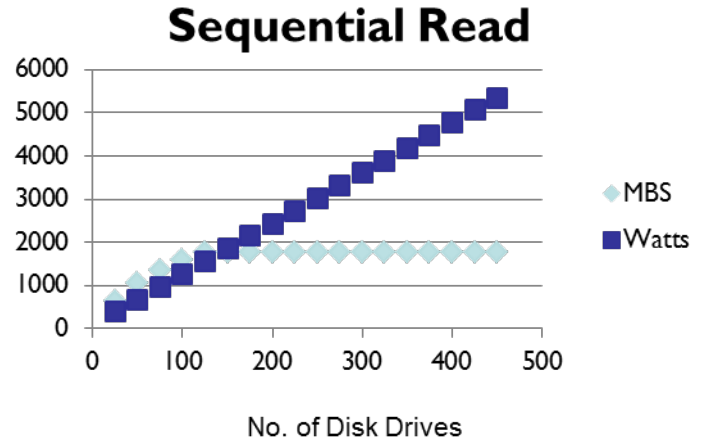
The above usage profile describes conditions in environments that respectively impose light ("low"), moderate ("medium"), and extensive ("high") demands on the Tested Storage Configuration (TSC).

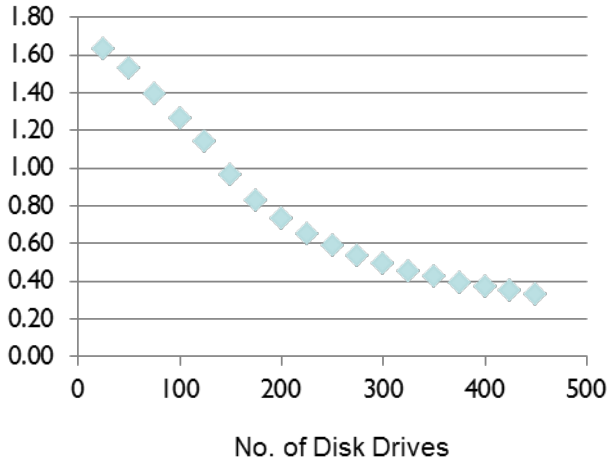
The SNIA EMERALD™ Program

- Provides a well-defined storage taxonomy and testing procedure, and public access to power efficiency data submissions
- Advocates using the Best Foot Forward (BFF) approach to identify and test optimized configurations that produce a set of peak power efficiency metrics for the test sequence
  - Random workloads [IOP/S/Watt]
  - Sequential throughput [MiB/S/Watt]
  - Idle Capacity [GB/Watt]
- Pending enhancements
  - Cache friendly / hotband workloads to replace uniform workloads
  - NAS filer workloads

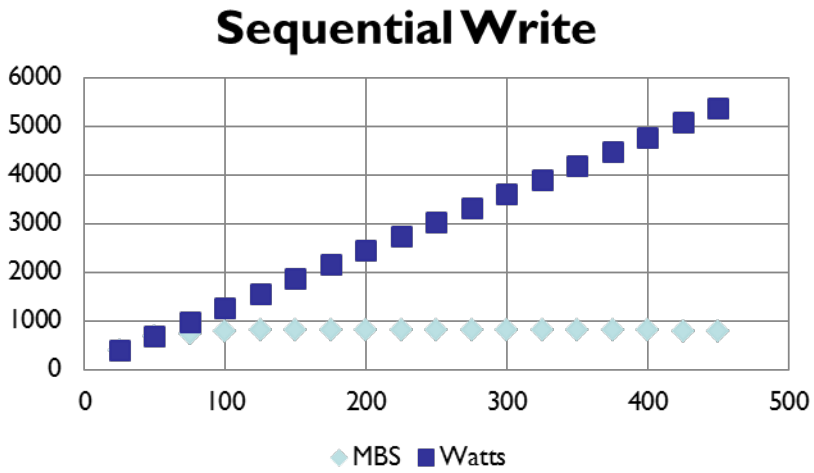
Example Predictions: 128K Sequential Read, Write

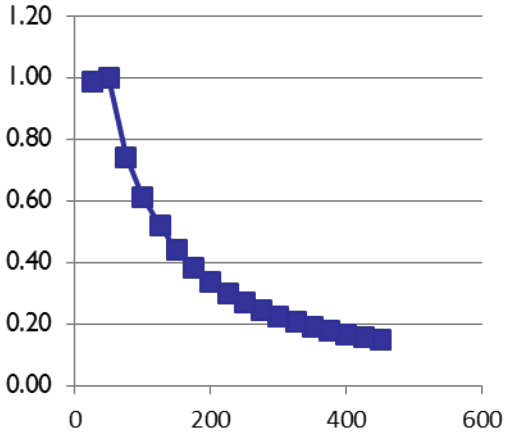
SFF 15K rpm, RAID 5





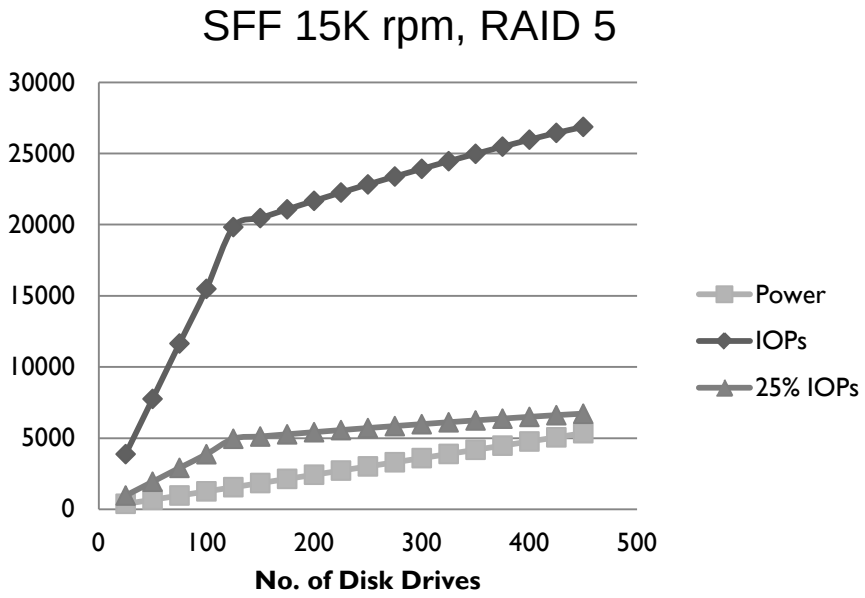
Peak Metric:
1.63 MB/S/Watt
and 25 drives

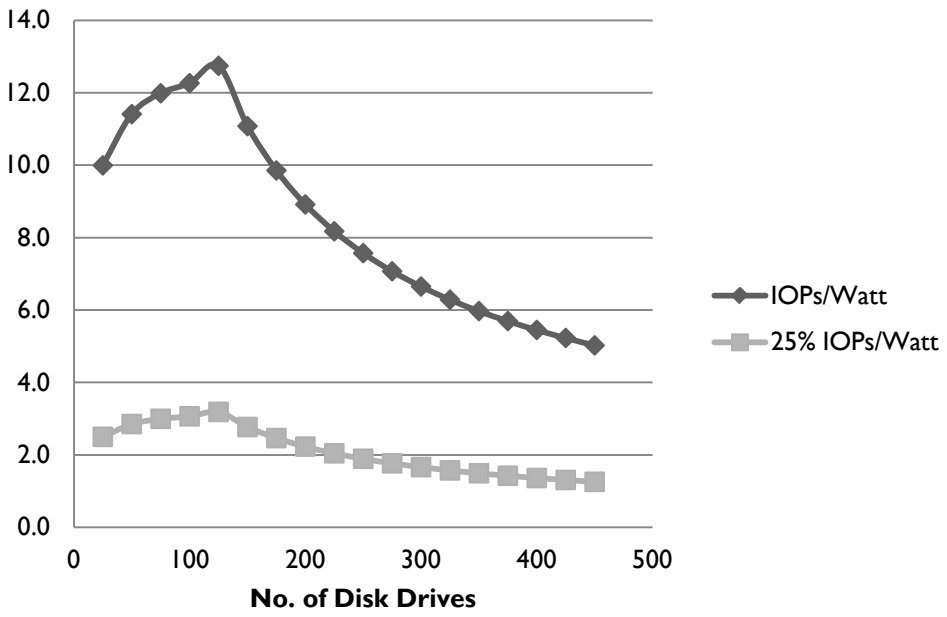




Peak Metric:
1.0 MB/S/Watt
and 50 drives

Example Predictions: Mixed workload 8K Random 70/30 R/W



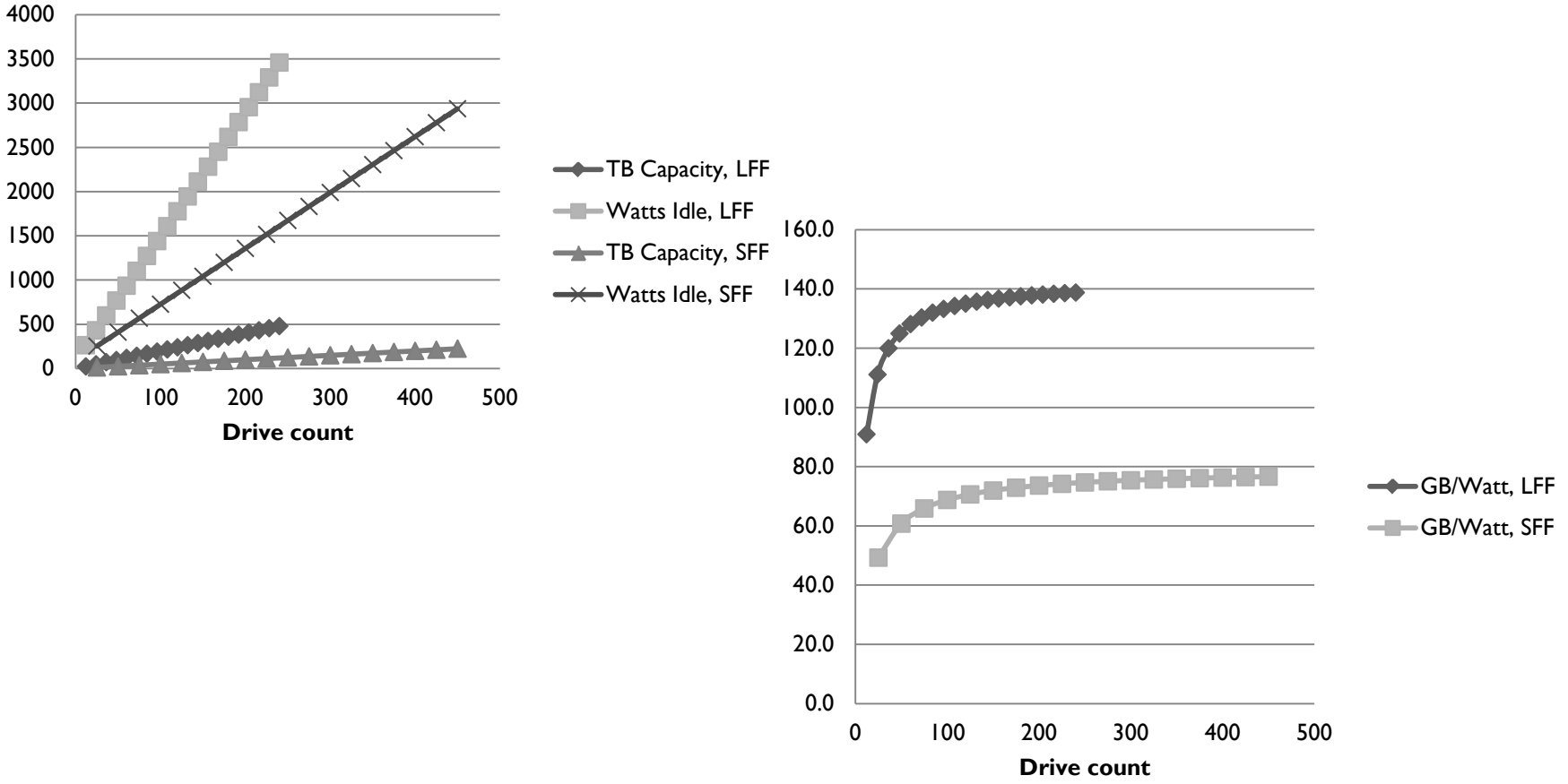


Peak metric = 12.7 IOPS/Watt at 125 drives

Changing the read/write mix changed the metric but not the drive count
60/40 r/w = 11.5 IOPS/W; 80/20 r/w = 14.9 IOPS/W

Example Predictions: Ready-Idle

LFF 2TB 7.2K rpm and SFF 500GB 7.2K rpm drives at Ready-Idle



Sample Emerald data submission (Online 3: HP P6500)

Operational Power

Idle power test

Average watts	592.692	W	
Raw capacity tested	7300	GB	GiB
EP_{RI}	12.317	GB/W	GiB/W

Standard idle metric: GB per Watt

Note: 1 GB = 10^9 bytes; 1 GiB = 2^30 bytes
a GiB is about 7.4% larger than a GB

Active power tests

EP_{RR}	17.255	run length (minutes)	30	Average latency	12 ms
Small random reads	I/Os per second per Watt				
EP_{RW}	9.155	run length (minutes)		Average latency	7 ms
Small random writes	I/Os per second per Watt				
EP_{SR}	3.01	run length (minutes)		Average latency	6 ms
Large sequential reads	MiB per second, per Watt				
EP_{SW}	1.22	run length (minutes)		Average latency	9 ms
Large sequential writes	MiB per second, per Watt				
EP_{MW1}	10.501	run length (minutes)		Average latency	13 ms
Mixed workload 1	I/Os per second per Watt				
70% random, 30% sequential, I/O intensity = 100					
EP_{MW2}	3.486	run length (minutes)		Average latency	4 ms
Mixed workload 2	I/Os per second per Watt				
70% random, 30% sequential, I/O intensity = 25					

NOTE: power-related numbers are required to be reported to three significant digits

SNIA Emerald™

The SNIA Emerald Test Data Report

Disclosure for storage systems and products

Capacity Optimizations	On during test?	Available in SUT?
Deduplication		NO
Compression		NO
Thin provisioning		YES
Parity RAID	YES	
Read-only delta snapshots		YES
Writeable delta snapshots		YES

Other mandatory disclosures, per spec

Test data provided is for a specific configuration that is tuned to achieve the best SeqRead and SeqWrite performance (the "sweet-spot"). The sweet spot data for alternate configurations that are tuned for the best Random and Idle metrics will be added in the near future.

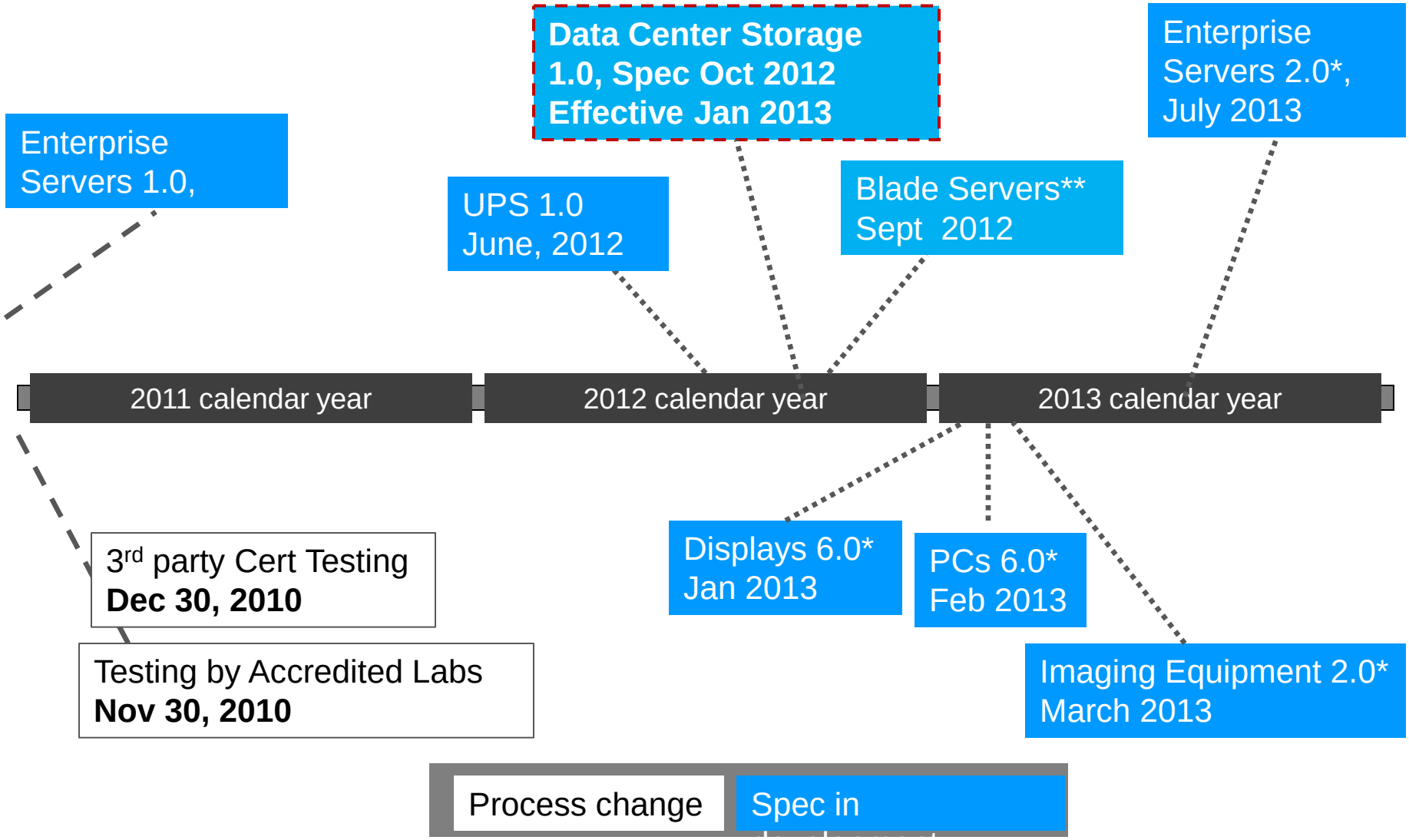
Key Point: This data submittal represents a configuration tuned for producing the best Sequential metrics

Emerald Enhancements: Data for cache friendly performance

HP Low & High End Array Products Comparison Chart

	Hot IOPS	Hot 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	1679	4.755 ms	844	9.47 ms	~26%	60%/24%
Large Array	16680	4.245 ms	7831	9.13 ms	~3%	60%/24%

ENERGY STAR™ Timeline for IT products



EPA Energy Star™ for Data Center Storage*

- Eligible product categories are low, mid, and high-end Online Arrays (Online 2, 3, 4)
- Requirements include checklist items (√) and data submissions
  - (√) Parity RAID, adaptive cooling, Silver-rated power supplies, power monitoring, x out of y Capacity Optimization Method (COM) features, online power modeler (optional)
  - Data for ready idle, sequential, and random "BFF" / optimal configuration. Eligible product SKU's are within -20% to +5% size range (drive count) Rounding to nearest full draw is allowed for ≥ 150 drive count
  - Drive combinations are permissible, based on % allocation of multiple optimal configurations
  - Storage device replacement (upgrade) if similar energy efficiency performance
- V2.0 will use the data from V1.0 to establish the top 25% cut line
  - Will also include: cache-friendly active workloads for Online, NAS filers, expanded product categories

* Expectations, based on latest review of Draft 3, V1.0 Spec

DCsE™(*)

- In contrast to the other “Acquisition” metrics, the DCsE metrics are “Operational”
  - The Capacity, Workload, and Throughput metrics are the same, but measurement and usage aspects differ.
  - For DCsE, measurements are made in the data center under normal operation
  - For DCsE, each storage system must consider it’s taxonomy category and it’s use case
- These represent key performance indicators (KPI) that show storage system efficiency over time

(*) Data Center storage Efficiency, per white paper by TGG and with SNIA collaboration

Case Study: Tiered Storage COM

- Description of the workload chosen for this case study
- Base-lining the performance using a single tier
- Re-running on the optimized tier structure
- Power/Performance benefit assessment

Case Study: Tiered Storage COM

The SPC-1 workload was chosen for the case study as it satisfies 3 important criteria:

- The asymmetry of IO demand
  - Asymmetric IO Demand A.K.A. “hot spots” within subregions of the physical address space not application space
  - The location of these hot spots must persist long enough to monitor, move and most importantly, execute after migration
  - The relative demand of these regions must be sufficiently higher than that of the remainder of the space
- The benchmark configuration is precisely defined in terms of capacity requirements and the amount of physical space utilization (no short stroking)
- The stimulus is also precisely defined by the specification

Smart Tier Environment (HP P9500)

P9500 Single Tier Configuration

1 x DKC Module 0
2 x Processor Blade
2 x 16-port 8Gbps FC CHA
1 x SAS DKA
2 x Cache Memory Adapter
6 x 16GB Cache Memory Module
2 x Express Switch Adapter
128 x 146GB SAS 15K 2.5" HDD

P9500 Dual Tier Configuration

1 x DKC Module 0
2 x Processor Blade
2 x 16-port 8Gbps FC CHA
1 x SAS DKA
2 x Cache Memory Adapter
6 x 16GB Cache Memory Module
2 x Express Switch Adapter
104 x 146GB SAS 15K 2.5" HDD
8 x 400 GB SAS SSD

Tiered Storage COM Benefit

Power/Performance Comparison Points

	Single Tier	Dual Tier	Delta
HDDs	128 x 146GB SAS 15K rpm SFF (18688 GB total)	104 x 146GB SAS 15k rpm SFF + 8 x 400GB SSD (18384 GB total)	14 fewer drives (~ same)
IOPS	~35,000*	~61,000*	+74%
Power Consumption (W)	2453	2346	- 4.3%
Power Efficiency (IOPS/W)	14.27	26.0	+82%

*These are reference IOPS and may NOT be referred to as SPC-1 IOPS

Tiered Storage Impact on DCsE

	Single Tier	Dual Tier
DCsE, cap (idle capacity)	18688/1960* = 9.53 GB/W	18384/1870* = 9.83 GB/W
DCsE, io (active workload)	35000/2453 = 14.27 IOPS	61000/2346 = 26.0 IOPS

- The active workload metric is the relevant KPI for the tiered storage COM implementation

* Idle power derived from Power Calculator

Summary Metrics Comparison

	Effective	Strengths – Key Benefits
SPC 1/E, 2/E	2010, 2012	Industry Standard - accepted workload & benchmark Precise, tightly controlled run rules (audited for accuracy) Benchmark produces both power/performance metrics & annual operating cost
Emerald	Oct 2011	Defined storage Taxonomy Documented test procedures/specification Database for energy-conscious Users Being leveraged by Energy Star
Energy Star	Jan 2013	World-wide recognized brand Will become a US government lockout on RFP's The Europeans may require it more broadly, e.g., required as part of EPEAT
DCsE	Future	Potential for smart Data Center

References

Storage Performance Council

http://www.storageperformance.org/home/

SNIA Emerald Program

http://sniaemerald.com/

EPA Energy Star Program

http://www.energystar.gov/index.cfm?c=new_specs.enterprise_storage

The Green Grid

http://www.thegreengrid.org/home

Thank You

Questions?

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