

Product/Family, Best Foot Forward (BFF), Test Points and Qualification Ranges

Jim Espy – EMC Herb Tanzer - HP

SNIA Emerald[™] Training

SNIA Emerald Power Efficiency Measurement Specification, for use in EPA ENERGY STAR®

June 24-27, 2013









- Product/Family/Best Foot Forward (BFF) Descriptions
- BFF Methodology and Examples
- ◆ Emerald[™] and ENERGY STAR[®] Test Point and Qualification Range Definitions





Jim Espy - EMC

- Current Title: Senior Consulting Engineer
- Experience: 30+ years in storage and related industries

Herb Tanzer – HP

- Current Title: Storage Hardware Architect
- Experience: 30+ years in storage and related industries





Wide Spectrum of Storage-Oriented Products

- Created a taxonomy to narrow scope
- Categories: On-Line, Near-Line, etc.
- Classifications: Further granularity of each Category

Still too Broad in Scope

- Vendors may have multiple Category/Classification products
- Each product may have <u>many</u> configuration variables

Requirement/Challenge: Select Appropriate Test Configs

- Comprehensive and usable results for customer
- Minimized, lower cost, but effective testing methods for vendor







Product:

 Represents a fundamental performance capability space that separates it from any other potentially related products

Product Family:

 Represents the full range space of configuration variables and options for a particular product.

Term Usage:

• Terms *family* and *range* are used interchangeably and may include such aspects as number and type of storage devices (spinning or solid state drive), cache size, availability levels, etc.







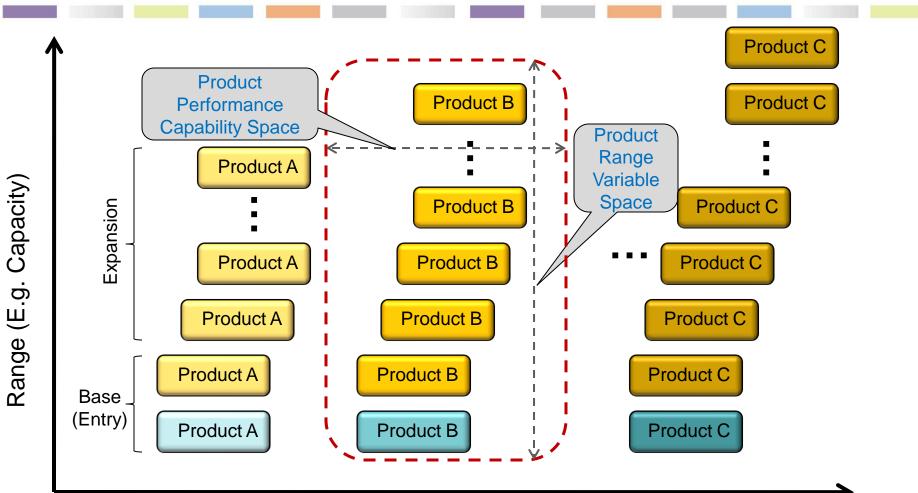
- Vendor Aligns Product(s) with SNIA Taxonomy Category
 - Hopefully straight forward Taxonomy will adapt over time
- Vendor Aligns Product(s) with Category Classification
 - Will be some boundary gray areas E.g. OL-3 or OL-4?
- Vendor Further Defines Product/Family Configurations
 - The really hard part...

Conceptual Representation

- Next slide depicts a possible product/family (range) differentiation
- Believed applicable to most storage system architectures



Simplified Product/Family Representations





Performance



Products Could be of Various Architectural Types

- Monolithic Little or no scaling but may still have family aspects
- Scale-up E.g. base controller + storage expansion
- Scale-out E.g. base compute/storage + compute/storage expansion
- Others TBD

Product Performance Typically Scales With Expansion

- Varying degrees
 - > Scale-up performance typically rolls off at varying degrees before max config
 - > Scale-out performance can be linear with increasing configurations
- Inter-product performance overlap driven by vendor's market positioning





Range Variables

- Example on previous product/family depiction focuses on capacity
- Could involve other variables

Range Variable Types

- Particular Items of highest potential energy consumption impact:
 - > Controller or related compute element Typically defines performance aspect
 - > Cache Also perf oriented Not considered part of the user-addressable space
 - > Number and type of persistent storage devices Defines user-addressable space
 - > RAS items As necessary for reliability, availability, serviceability requirements
 - Capacity optimization Functionality (typically software) that more effectively utilizes physical storage space, e.g. thin provisioning, compression, de-duplication
- Many other examples



> Power supplies, cooling, I/O, etc.

Approach to Range Variable Reduction



Range Variable Reduction is Difficult

- Even with the 5 listed items, still too many test cases
 - > Significant set-up and execution times
 - > Complex results sets
- Max system size testing is expensive and cumbersome to manage
- Need a simpler alternative...

"Best Foot Forward" (aka Sweet Spot) - BFF

- Find proxy family configuration(s)
 - > Intended to be reasonably representative of the all range variables
- Find test point(s) where Measurement Spec active metrics are best
 - > The "sweet spot"
 - Suitable for any architecture
 - > E.g. scale-up, scale-out, hybrid, ...



BFF Looks Holistically at Storage System Product/Family

- Allows vendor to select and test one product/family configuration
 - > Or more if desired
- At operating points near the Measurement Spec metric peak values
 - > I.e. the "sweet spot"
- Results reasonably representative of the entire family
 - > Easier and less expensive for the vendor
 - > Simple and understandable results for the potential customer

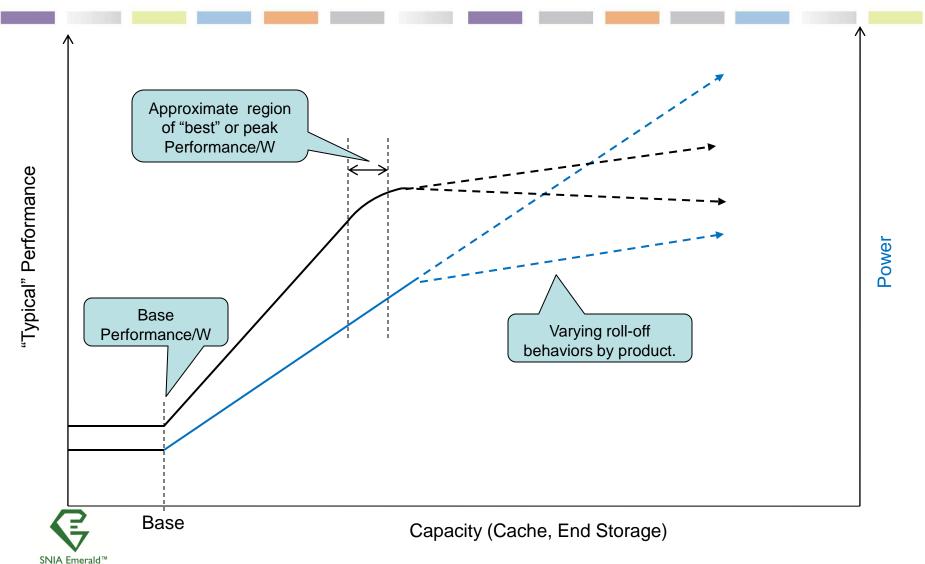
Scale Up Example on Following Slide

- Notion that Measurement Spec active metrics have peak values
- Peaks typically located at points below maximum configurations



Best Foot Forward Approach Scale-Up System







Previous Slide is a Rough Approximation

- Capacity increases are actually more stepwise
- Degree of performance roll-off can vary by product
 - Dashed lines attempt to show one (of possibly many) changes due to different storage technology tiers, e.g. scaling capacity w/large SATA drives
- Storage device enclosures shelves may impact
 - > Efficiencies of partial vs fully populated shelves
- Regardless, example depicts a smaller test configuration

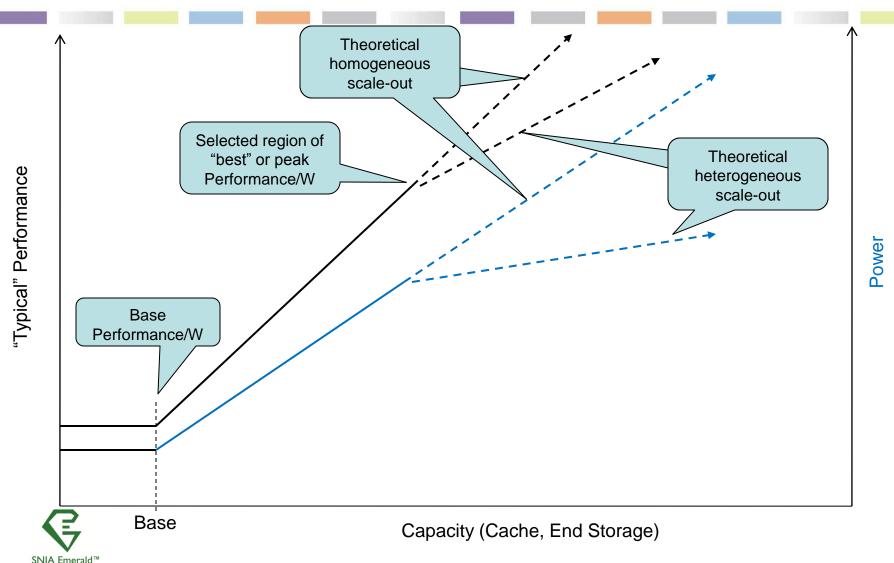
Scale Out Example on Following Slide

• What if there no clearly discernible peak for a reasonable test config?



Best Foot Forward Approach Scale-Out System







Again a Rough Approximation

- Capacity increases are actually more stepwise
- Dashed lines attempt to show one (of possibly many) changes due homogeneous vs heterogeneous scale-out configurations
- Can still select a smaller test configuration





Given Known Taxonomy Category and Classification

- Vendor determines one or more family representative configurations
- Vendor locates Measurement Spec active metric peak points
- Tests are performed on this reduced configuration (set)
 - > Note: For smaller systems, the BFF may in fact be the maximum configuration

Where is the Performance/W Peak?

- Depends on numerical increase of performance numerator vs power denominator with capacity
- If numerator initially increases more than denominator, a clear peak
- Else it becomes harder Just pick a point before it rolls off?





Next: BFF Methodology and Examples





- The benefit of Best Foot Forward (BFF) is to reduce the full range of variables of a product family to just a few test configurations. This reduced test set can be considered representative of the entire product family
- The BFF consists of the configurations that will produce a set of peak power efficiency metrics of a product family for the different test phases
 - Random [IO/s/Watt]
 - Sequential [MiB/s/Watt]
 - Idle Capacity [GiB/Watt]



A general approach for Emerald data submission



- Start by aligning your product family within a taxonomy definition
- Consider a baseline run to establish the test process "get a feel"
 - Start w/ available configuration; no particular "tuning" in affect
 - Identify any issues with conditioning, stability, response times, etc. (per the run rules), post-processing, reporting, etc.
- Consider all possible (and valid) product SKU's to identify configurations that will give the peak power efficiency metrics
- Using Estimator tools, identify the "best-foot- forward" or "sweetspot" relative to each specific test profile
- Set-up, test, and measure the peak metric values for your 1st BFF
 - Run through the complete sequence of test phases
 - Test validate and data correlate
- For each additional BFF of interest, re-configure and re-test



Candidate SUT: A shipping Online-3 SAN



- Two controller performance points, with variable cache and front-end interfaces
- The lower product class can support 120xLFF or 250xSFF and the higher product class can support 240xLFF or 450xSFF (6Gb SAS)

SFF

- > 146GB, 15K
- > 300GB, 10K
- > 450GB, 10K
- > 600GB, 10K
- > 500GB, 7.2K midline
- > 200GB SSD*
- > 400GB SSD*
- LFF
 - > 300GB, 15K
 - > 450GB, 15K
 - > 600GB, 15K
 - > 2TB, 7.2K midline





Test Phase IO Profiles for Online & Near Online



- The complete test is run in an un-interrupted sequence, and consists of Pre-fill, Conditioning, the Active tests, and Ready Idle
- Each Active IO test phase IO profile shall last a minimum of 40 minutes, comprised of a minimum of 10 minutes to establish stability followed by 30 minutes as the measurement interval

IO Profile	IO Size (KIB)	Read/Write Percentage	IO Intensity	Transfer Alignment (KIB)	Access Pattern
Hot Band Workload (I-HB)	See Table 11	See Table 11	100	See Table 11	See Table 11
Random Write (I-RW)	8	0/100	100	8	Random
Random Read (I=RR)	8	100/0	100	8	Random
Sequential Write (I=SW)	256	0/100	100	256	Sequential
Sequential Read (I-SR)	256	100/0	100	256	Sequential

A 2-hour ready idle test follows the Active tests



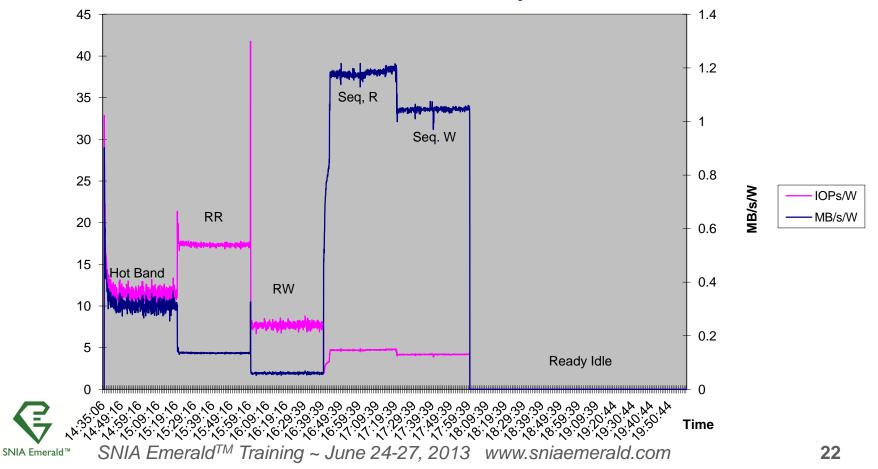
SNIA Emerald[™] Training ~ June 24-27, 2013 www.sniaemerald.com

Test Results for Candidate SUT (50SFF, 148GB 15K)



P6500 Power Performance Plot

Combined Vdbench /Power Analyzer Plot





- Peak transactional efficiency metric occurred during RR phase (~ 15 IO/s/W)
- Peak throughput efficiency metric [MB/s/W] occurred during SR phase (~ 1.2 MB/s/W)
- Power consumed for any workload varies only 13.4% (451.2W at idle to 512.7W during RR)





Since there are 6 different Emerald test profiles for Online we can expect up to 6 different BFF configurations

- 1 X Hotband (IOP/S/Watt]
- 2 x Random [IOP/S/Watt]
- 2 x Sequential [MiB/S/Watt]
- 1 x Ready-Idle [raw capacity, GiB/Watt]
- Recommend to use Estimator tools that combine power and performance to predict the peak metrics
 - The alternative is educated derivations and potentially a lot of testing that is very labor and resource intensive.
 - As long as the simulated results are reasonably accurate, the physical configuration selected to measure the peak value can be reduced in range





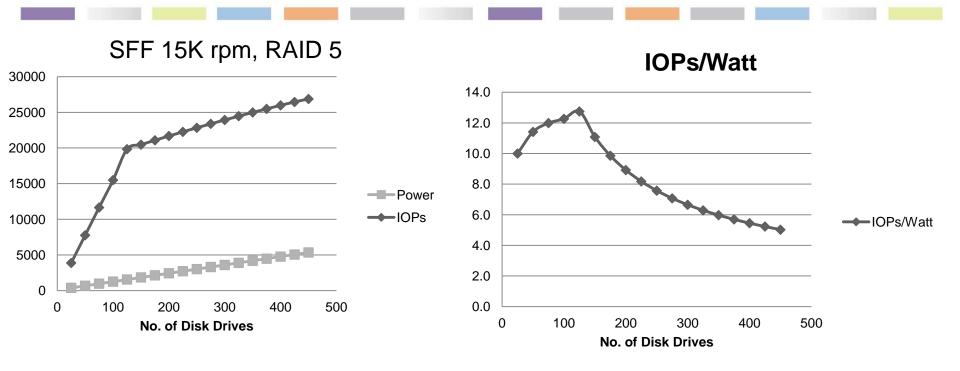
Exercise #	Prediction basis
1 1.5	Mixed Workload, Random 70/30 R/W Granular level, single drives *
2	Random Read (100/0 R/W) & Random Write (0/100 R/W)
3	Sequential Read (100/0 R/W) & Sequential Write (0/100 R/W)
4	Ready Idle
	*Note: Elsewhere, drive count is

incremented by full drive shelf



Exercise 1: Mixed Workload 8K Random 70/30 R/W





• Peak metric = 12.7 IOP/S/Watt at 125 drives

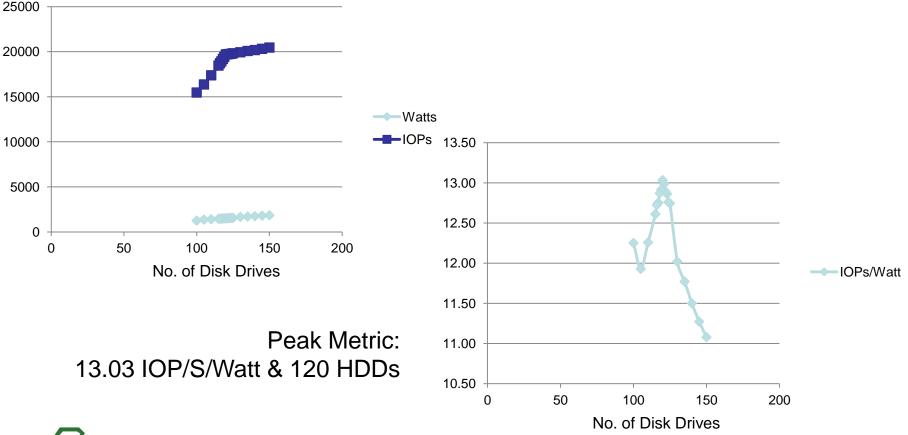
• Changing the read/write mix changed the metric but not the drive count 60/40 r/w = 11.5 IOP/S/W; 80/20 r/w = 14.9 IOP/S/W



Exercise 1.5: Granular drive counts (increment by single HDDs)



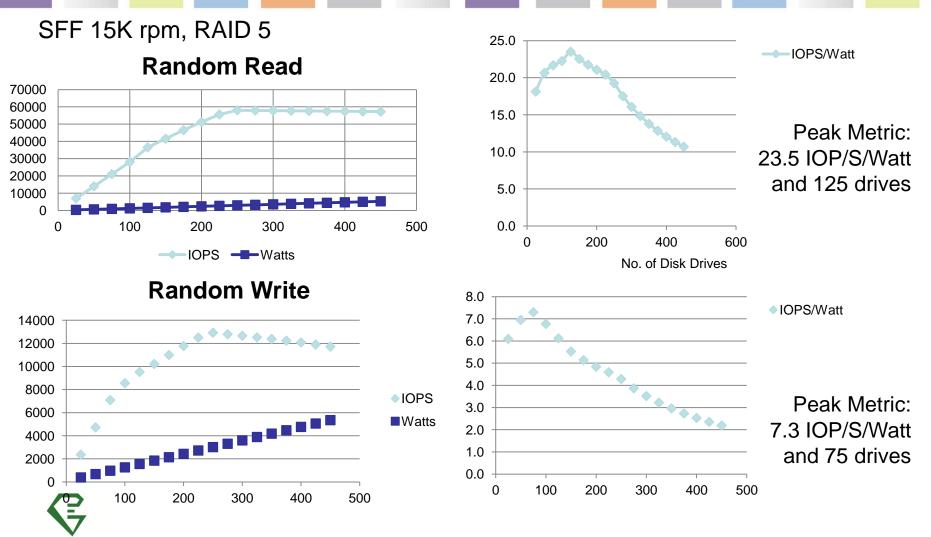
8K Random, 70/30 R/W, SFF 15K rpm, RAID 5





Exercise 2: 8K Random Read, Write

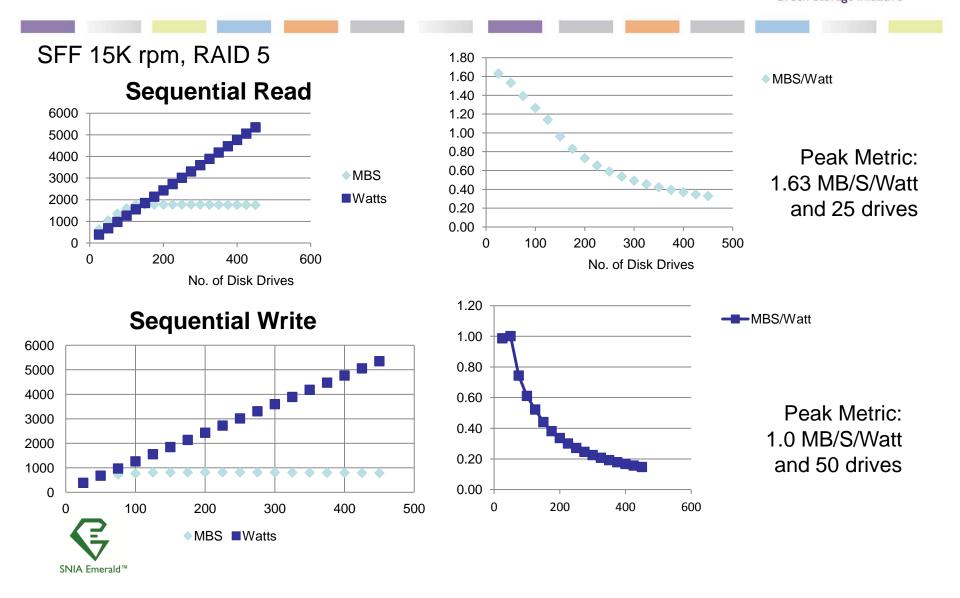




SNIA Emerald™

Exercise 3: 128K Sequential Read, Write

Green Storage Initiative

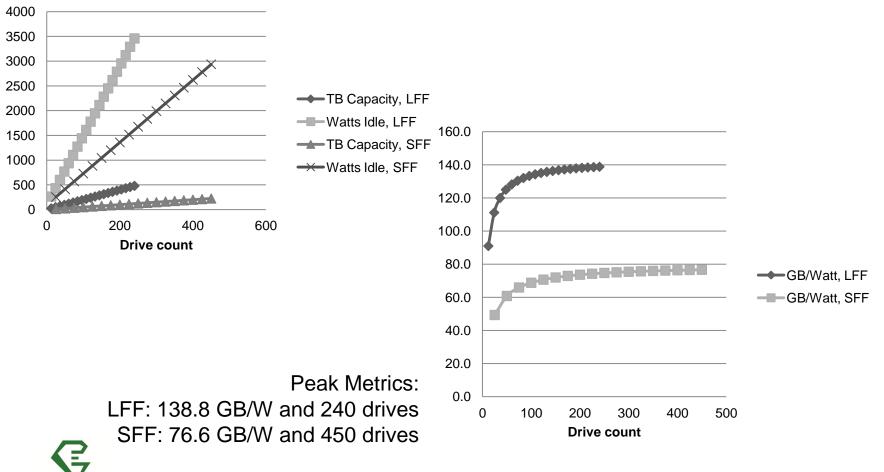


Exercise 4: Ready-Idle

SNIA Emerald[™]



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



General Observations for the Candidate SUT



- Active cases the Performance* reaches a highest rate relatively early (i.e., smaller drive count) before rolling off and leveling. The peak [Performance/Power] metric seems to coincide with the rate inflection
 - All peak predictions for Random are reached with the same drive type (15K, SFF) and close in drive count (125 or 75)
 - All peak predictions for Sequential reached with the same drive type (15K, SFF) and close in drive count (50 or 25)
- Ready-idle case the peak metric levels but continues to slowly rise with drive count (as the controller electronics power is amortized over increasing numbers of drives)

*Note: very dependent on specific Controller performance and bandwidth behavior











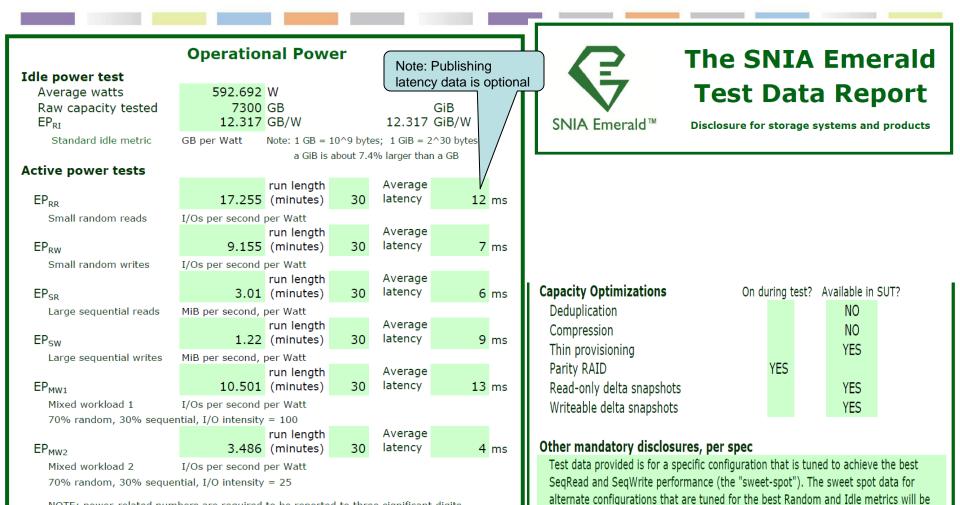
BFF Performance/Watt point

- For all vendor-selected product/storage device configurations
- Could Test at Other Capacity Points but not Required



Sample data submission (Online-3 SUT)





added in the near future.

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





BFF Performance/Watt point

For all vendor-selected product/storage device configurations

Two Additional Performance/Watt points

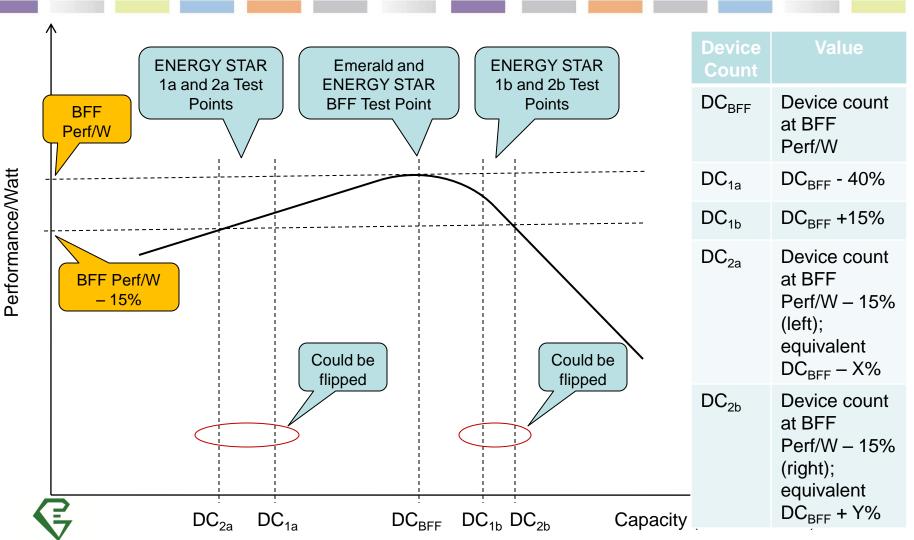
- For particular product configuration but *only for most commonly sold* and shipped storage device type
- Either:
 - 1) -40% and +15% of the BFF storage device count
 - -or-
 - 2) Storage device count points where Perf/W value is 15% < the BFF point

Whichever provides the wider Performance/Watt curve coverage



Test Point Definitions Scale-Up System Example





SNIA Emerald™





For Emerald

- Complete TDR(s)
 - > Product configuration/storage device type(s) and BFF data test point(s)
 - > Product configuration and storage device types/combinations you choose

For ENERGY STAR

- Submit Test Data
 - > Product configuration/storage device type(s) and all required data test point(s)
 - > Product configuration and storage device types/combinations you choose
- But What Defines an ENERGY STAR Qualified Product?



ENERGY STAR Qualification Range



Qualification Range Option 1

• If:

Goal is to select the option providing the widest qualified range.

- > Test points utilize -40% and +15% of the BFF storage device count
- Then:
 - > Qualified range extends from -20% to +5% of the BFF storage device count
- Further If:
 - > Perf/W values at -40% and (or) +15% device counts are > BFF Perf/W 15% level
- Then
 - > Qualified range widens to -40% and (or) +15% of the BFF storage device count

Qualification Range Option 2

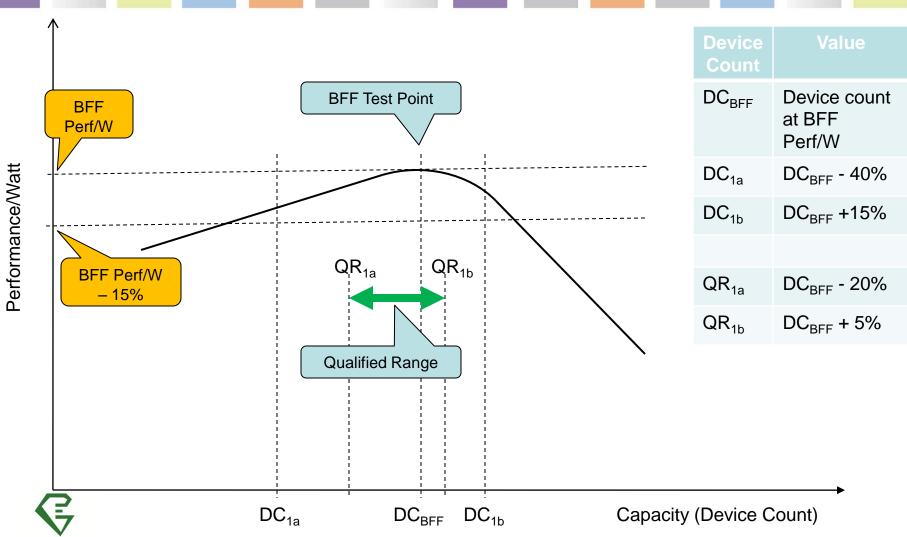
- If:
 - Test points utilize storage device count points where Perf/W values = BFF Perf/W value 15%; equivalent of -X% and +Y% of the BFF storage device count

Then:

SNIA Emerald > Qualified range extends from -X% to +Y% of the BFF storage device count

ENERGY STAR Qualification Range Option 1 Example

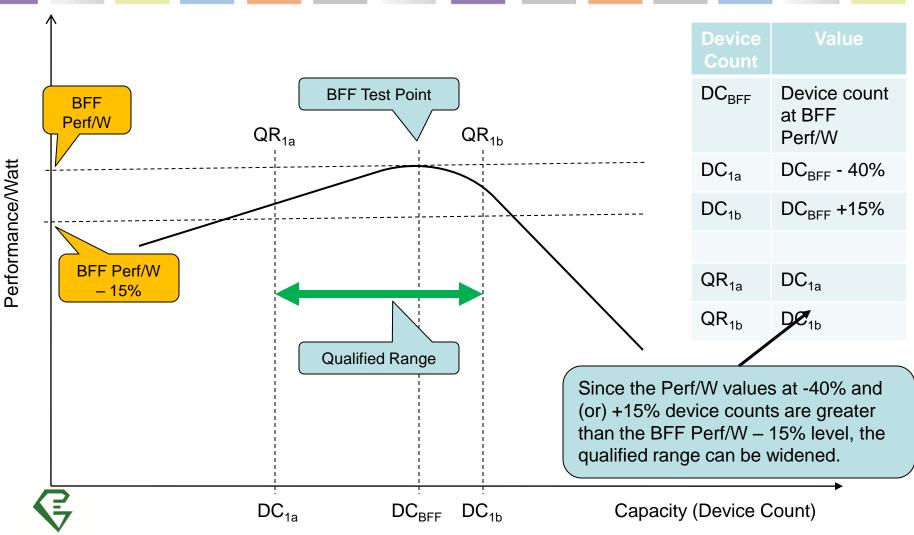




SNIA Emerald™

ENERGY STAR Qualification Range Option 1 Example Extended (Widened)

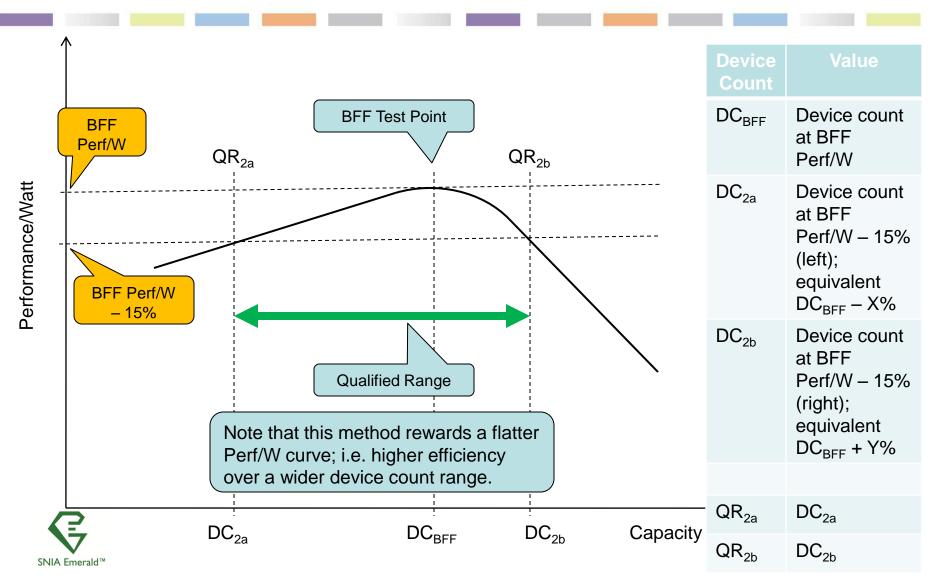




SNIA Emerald[™]

ENERGY STAR Qualification Range Option 2 Example: X=60%; Y=20%







- Test points exist for up to five Active + one Idle conditions
- Rounding to full drawers
- Combinations of single device type optimal configurations, based on percentage allocation of devices
- Storage device replacement (if similar or better)
- Multiple device type optimal configurations (must have auto-tiering technology)
 - Auto-tiering BFF or optimal configuration is not yet fully understood, but it is suggested that initial submissions are representative of actual selling systems.
 - Maintain ratios of device types for –X% / +Y% test points





Whichever Qualification Range Method is Determined

- Any extension to the qualified range based on testing of the most commonly sold and shipped storage device type may be applied to additional storage devices within the same product and workload type, without the need to physically retest the outer points
- BFF point still needs to be tested for each additional storage device)







- Questions? Can't Imagine What Questions There Could be? ③
- jim.espy@emc.com
- herb.tanzer@hp.com

