# User Guide
## for the *SNIA Emerald*<sup>TM</sup> *Power Efficiency Measurement Specification V2.1.1*

**Version 2.0 Revision 7**

**SNIA Emerald™**

**Sept 30, 2015**

## About the SNIA

The Storage Networking Industry Association is a not-for-profit global organization, made up of member companies spanning the global storage market. SNIA's mission is to lead the storage industry worldwide in developing and promoting standards, technologies, and educational services to empower organizations in the management of information. To this end, the SNIA is uniquely committed to delivering standards, education, and services that will propel open storage networking solutions into the broader market. For more information, visit www.snia.org.

## About the SNIA Green Storage Initiative

SNIA's Green Storage Initiative (GSI) is dedicated to advancing energy efficiency and conservation in all networked storage technologies in an effort to minimize the environmental impact of data storage operations. SNIA's Green Storage activities take place in two separate working bodies, the SNIA Green Storage Technical Working Group (TWG) and the Green Storage Initiative. The TWG is focused on developing test metrics by which energy consumption and efficiency can be measured. The Green Storage Initiative is focused on creating and publicizing best practices for energy efficient storage networking, educating the IT community, and promoting storage–centric applications that reduce storage footprint and associated power requirements.

## About the SNIA Emerald™ Program

The SNIA Emerald™ Program is a vendor-neutral, public service to the storage industry and end users that is sponsored and operated by the SNIA GSI. The program supports the use and evolution of the SNIA Emerald Power Efficiency Measurement Specification and the publication and use of product test data based on the Specification. The measurement procedure and test metrics are documented in the SNIA Emerald™ Power Efficiency Specification, which was developed, released, and is maintained by the Green Storage TWG under the guidance of the GSI.

The program provides a standardized way of reporting vendor-performed test results that characterize the various aspects of storage system energy usage and efficiency. Those test results—which are based on performance under different workload metrics—are combined with other power-related information about storage systems and their components. The

Program offers a common repository, accessible to the public, where test results are presented in a consistent format that can be used to help determine facility power usage and efficiency for data storage solutions, as well as the costs to operate such systems.

The EPA ENERGY STAR® Data Center Storage Program is based on the methodology defined in the Specification and offers another vehicle for publication of product test results created in accordance with the Specification.

Copyright © 2011, 2012, 2013, 2014, 2015 Storage Networking Industry Association.

The information contained in this publication is subject to change without notice. This guide represents a "best effort" attempt by the SNIA Green Storage Technical Working Group to provide guidance to those implementing the *SNIA Emerald™ Power Efficiency Measurement Specification*, and the guide may be updated or replaced at any time. The SNIA shall not be liable for errors contained herein.

Suggestions for revisions to this guide and questions concerning implementation of the *SNIA Emerald™ Power Efficiency Measurement Specification Version* can be directed (via email) to greentwg-chair@snia.org.

# Contents

# 1  Introduction

This document is intended to be used as an informative guide in connection with the *SNIA Emerald™ Power Efficiency Measurement Specification v2.1* (referred to within this document as simply the *Measurement Specification*), developed as part of the *SNIA Emerald™ Program*. The *SNIA Emerald™ Program* was set up to provide a consistent and credible way for storage system vendors to demonstrate product power efficiency. In order to facilitate this, the SNIA Green Storage Initiative (GSI) and Green Technical Working Group (GTWG) have developed a standard method to measure storage system efficiency along with a mechanism via the *SNIA Emerald™ Program* to store results accessible for information and comparison. This method is not intended to demonstrate the true power efficiency of a storage system at a customer site, but instead provide a general and comparable understanding of expected power efficiency.

The *Measurement Specification* has also been chosen by the EPA for its ENERGYSTAR® for Data Center Storage program and related test specification. Methods and advisory notes listed in this document will also be helpful in providing results for this program.

Any conflict between this document and the *Measurement Specification* shall defer to the *Measurement Specification*.

## 1.1  Audience

The target audience of this document includes organizations and individuals planning for and testing in accordance with the *Measurement Specification*. An organization or individual performing such testing is referred to as a test sponsor.

## 1.2  References

This guide is designed to be used with the following documents available at the *SNIA Emerald™* website http://www.sniaemerald.com/download:

- *Measurement Specification*
- Vdbench tool website link
- Related workload scripts
- COM data generation tool
- This guide
- Training materials
- Test Data Report Template

Additional information about the *SNIA Emerald™* Program, associated *SNIA Emerald™ Power Efficiency Measurement Specification*, and relation to the EPA ENERGYSTAR® for Data Center Storage program is available at these websites:

- http://www.sniaemerald.com, the SNIA Emerald™ Program website
- http://www.snia.org/forums/green, the SNIA Green Storage Initiative website
- http://www.energystar.gov/products/certified-products/detail/data-center-storage

# 2 Scope

## 2.1 General

SNIA developed the *Measurement Specification* and its related *Emerald™* program so that vendors and consumers of storage systems would have a reliable and consistent way to observe and compare storage power efficiency among different storage solutions. The complexity of these systems is reflected in the many details associated with the *Measurement Specification* and test implementation. This document is intended to help with the understanding and effective execution of the *Measurement Specification* test methods.

There are several aspects to the *Measurement Specification*. First, there is a taxonomy table that helps differentiate product types by their basic functionality. However, in order to facilitate testing, it is necessary to differentiate products further. This document addresses this need via product-family descriptions. It is also important to select appropriate metric test points based on selected configurations, both to provide valuable data and reasonably limit test time and expense. A so-called Best-Foot-Forward (a.k.a. optimal) metric point is defined to help facilitate this along with advice on how to find such metric test points for selected test configurations.

The *Measurement Specification* further defines and provides the means for determining values of several proxy power efficiency metrics. The IO/s/Watt and MiB/s/Watt metrics are described as active metrics, as they represent the power efficiency associated with moving data between the storage system and host(s). The GB/Watt metric is described as an idle metric and represents the power efficiency of storing data on the storage system.

There are also tests for determining the existence and basic operation of Capacity Optimization Methods (COMs) such as de-duplication and compression. The complexity of these functions made it difficult in many cases to include them in active or idle tests.

The main function of the *Measurement Specification* is to define the test configuration, instrumentation, benchmark driver requirements and IO test profiles, test execution methodology and sequence, and metric calculation methods. There are separate sections for each main taxonomy category.

## 2.2 Using this Guide

This document provides *Measurement Specification* supplemental advice on how to:

- develop a product family definition from a selected taxonomy category
- determine appropriate measurement configurations and test points
- understand site and instrumentation requirements
- use Vdbench and related associated scripts for workload generation
- set up and complete the measurement sequence including metric values and validity
- avoid problems
- submit results

Both the *Measurement Specification* and this guide refer to the configuration being tested as the *System Under Test (SUT)*.

While this document provides basic information related to all listed *Measurement Specification* taxonomy categories, its primary focus is on Online and Near Online systems. However, there is selected commentary on the Removable Media and Virtual Media Library categories.

# 3 Taxonomy Comments

Due to the wide spectrum of storage-oriented products, a taxonomy structure was created and placed in the *Measurement Specification*. This taxonomy presently has storage categories including Online, Near Online, Removable Media Library, and Virtual Media Library. Each category is then divided into classifications based on selected characteristics. Since each of these categories provides its own unique testing criteria, it is critical for valid measurement to correctly identify the SUT category and classification.

The Online taxonomy category and classifications deal with storage systems that can retrieve first data of a data block within 80ms. These systems are generally disk- and/or flash-based. Categories range from consumer/component to large systems supporting hundreds of storage devices.

To meet a particular Online classification, a system must support at least the maximum number of storage devices listed in the classification table. For example, an Online 3 storage product must physically support at least 12 storage devices. The storage product may be sold or tested with fewer storage devices.

Near Online category storage systems are those that may not be able to satisfy the 80ms time to first data requirement. However, these systems can support random and sequential IO requests. Like Online, Near Online system classifications also have lower boundary maximum supported configuration requirements.

Because of the similar nature of Online and Near Online, the *Measurement Specification* uses the same configuration and test sequence for both but with differences noted.

The Removable Media Library category is for tape libraries and optical juke boxes. These systems require more than 80ms to reach first data and can only support streaming IO requests. To keep the system-only power efficiency in scope for the SNIA Emerald™ program, the storage device (tape) must be within the library, hence the need for a maximum time to data of 5 minutes. Like other categories, each classification has a lower boundary maximum supported configuration requirement. Note that there is no Removable 4 classification in order to maintain numbering consistency across all categories.

The Virtual Media Library category is one that can meet a 80ms time to first data requirement. These systems tend to be disk-based designed for sequential I/O requests. This category also has a lower boundary maximum supported configuration requirement.

# 4  Identifying the Product Family

While the taxonomy is useful in differentiating storage systems, vendors may still have a wide range of products within a single category and classification. To further aid in test development, this document also includes advice on differentiating products and families along with suggestions on limiting test configurations and defining test points to minimize test efforts and costs. Even the smallest systems may have a significant number of configuration options, with each configuration requiring significant execution effort.

## 4.1  Overview and Goals

Several aspects come into play when considering which storage system configurations to test for efficiency. In particular, customers want a clear and reasonably complete method to gauge and evaluate efficiencies of particular product acquisition candidates. At the same time, storage system vendors wish to minimize efficiency measurement test variations for lowest cost and widest coverage from a potentially large set of product configurations and use cases.

## 4.2  Product/Family Definition

The *Measurement Specification* includes a taxonomy that divides storage products into relatively coarse categories and classifications. Once a product is aligned with a taxonomy category/classification, the question remains of how the product and its possible variations are actually measured per the goals listed in Section 4.1.

The concept of products and product families is presented here to help further define actual storage system test configurations. While vendors vary in the manner in which they define and sell their products, this product/family approach is believed to be generally applicable.

A product has different aspects depending on the observer. To the customer, a product represents a particular purchased and installed configuration. To the vendor, it can be a base (possibly entry) unit with a bounded set of configuration options. A product family can also have many interpretations.

In this document, a product and product family are arbitrarily defined as follows:

- A *product* represents a fundamental performance capability space that separates it from any other potentially related products.

- A *product family* represents the full *range* space of configuration variables and options for a particular *product.*

The terms *family* and *range* are used interchangeably within this section and may include such aspects as number and type of storage device (spinning or solid state drive), availability levels, etc.

Figure 1 depicts a simplified but possible product/family (range) differentiation depiction. Note that this figure could apply to any storage system architecture, e.g., monolithic, scale-up, or scale-out (with *scale-up* generally referring to a system of a limited number of controllers with scalable back-end storage and *scale-out* referring to systems constructed of interconnected compute-storage nodes, real or virtual).
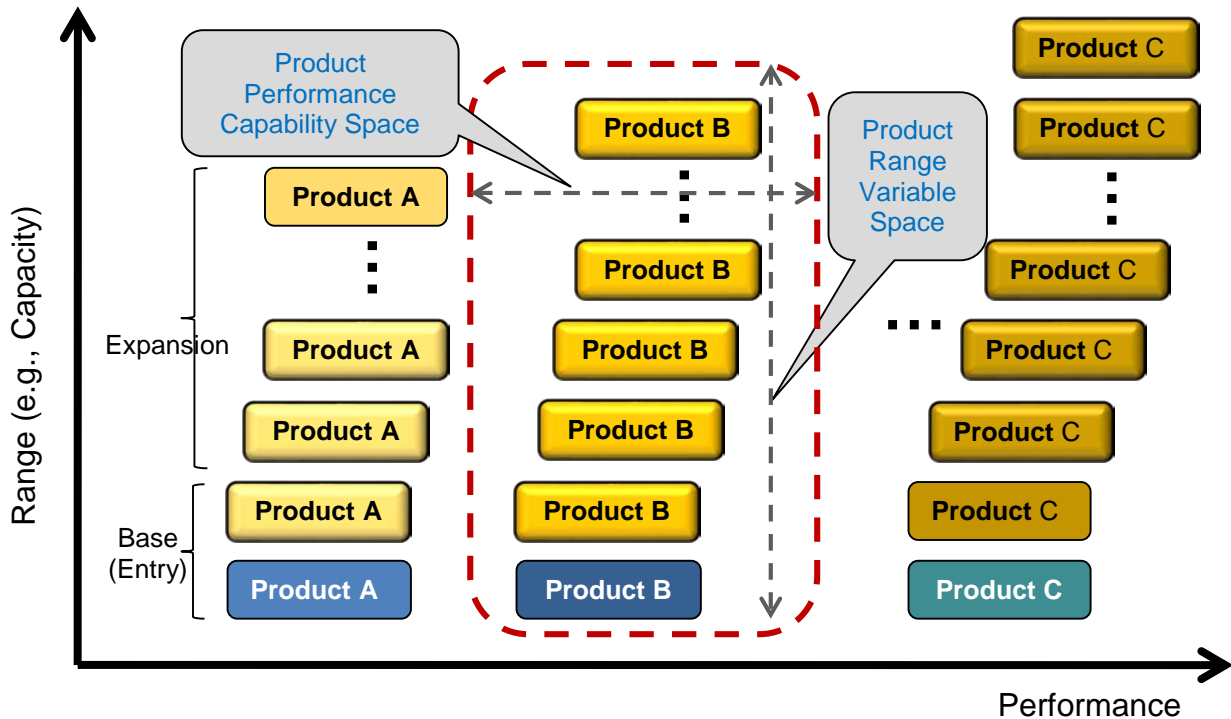
**Figure 1   Possible Product/Family (Range) Depiction**

The range variable space shown in Figure 1 focuses on capacity but can also imply storage device type or other variables. Note that some products illustrated may increase performance with added capacity and some may not, e.g., roll off, after a certain capacity/variable point.

## 4.3   Range Variable Discussion

As noted in the product/family discussion in Section 4.2, a full family range encompasses many variables both in number and type, of which SNIA has defined at least 25. The following list highlights those considered to have the highest potential energy consumption impact:

- Controller or related compute element - Typically defines the product performance space.

- Cache functions - May not always be aligned with the controller but not considered part of the user addressable space.

- Number and types of persistent storage devices - Define the user addressable space consisting of hard disk drives (HDD), solid state drives (SSD), etc.

- RAS items - Energy consuming functions necessary to meet requirements for reliability, availability, and serviceability.

- Capacity optimization - Functionality (usually software) that more effectively utilizes physical storage space, such as compression, deduplication, and thin provisioning.

Other items such as power supplies, IO (input/output) ports, cooling components, interconnect ports, etc., are not being ignored but are considered to be aligned and scale with performance and the items defined in this section.

Reduction of the variable space to the five items listed in this section still leaves vendors with a potentially very large set of test requirements and cases, each with significant set-up and execution times. Even configurations in which the number and type of HDDs and SSDs are the only variables can be too difficult to support. Maximum system size tests are expensive and cumbersome to manage. Customers would have similar issues in attempting to parse through a large number of test results and make effective vendor product comparisons. Rather than attempt to reduce this variable set further, a different method is proposed, the "best foot forward" (a.k.a. "sweet spot" or "optimal point") approach defined in Section 4.4.

## 4.4   Best Foot Forward Test Methodology

The Best Foot Forward (BFF) approach looks at a storage system product holistically. It allows the storage vendor to select and test one or more specific product/family configurations at operating points determined to be at or near *Measurement Specification* metric peak values, i.e., the "sweet spots." This results in a reduced test result set representative of the entire product family, which is both easier and less expensive for vendors to test and produces results simpler to understand and therefore more useful to customers.

The approach is based on the idea that the *Measurement Specification* active metrics have "peak" value points located within smaller— and hence more easily measurable —product/family configurations. The vendor selects one or more appropriately representative configurations and locates these *Measurement Specification* metric peak points. Key to this method is the avoidance of maximum configuration testing and other complex methods such as extrapolation and interpolation. (Note that in some cases of smaller systems, the maximum configuration may in fact be the BFF).

The diagram in Figure 2 shows an example of a hypothetical storage system in which system scaling is by capacity and performance tends to roll off at scale.
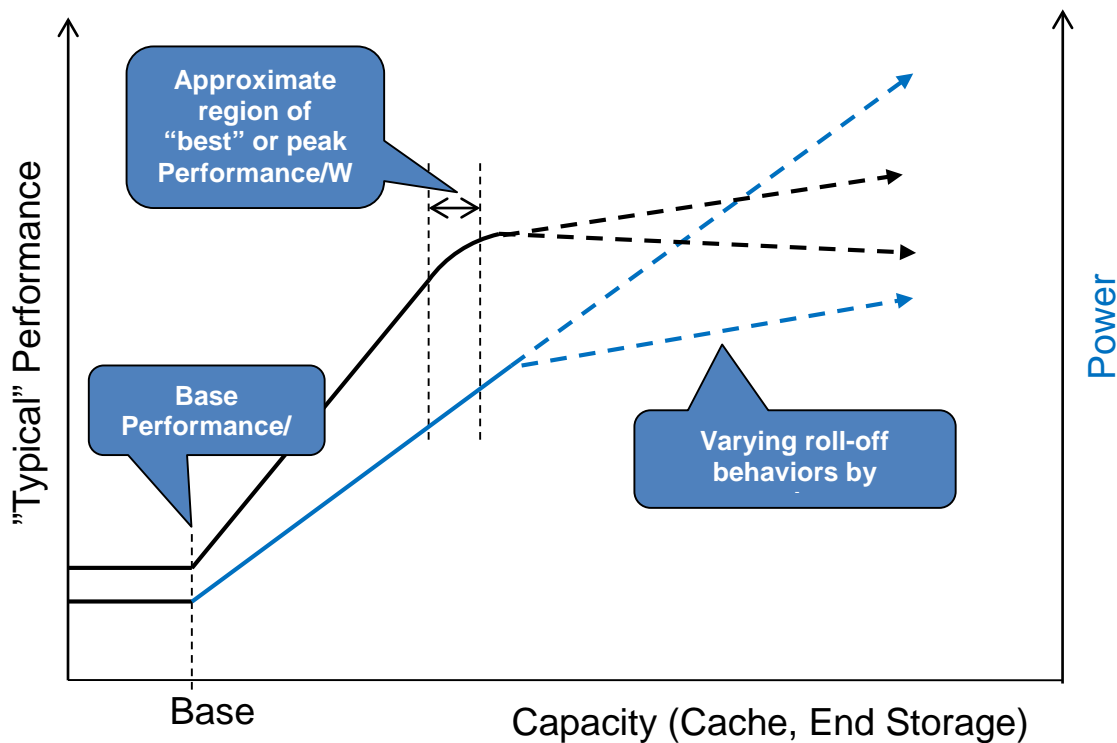
**Figure 2  Hypothetical Storage System Performance/Power Example**

The lines in Figure 2 represent highly simplified pictorial approximations and will vary with real systems (in scale-out systems the performance line may not roll off as extensively). Regardless, the example attempts to depict how a smaller representative system can be selected and tested at its vendor-determined peak *Measurement Specification* metric value points. One could also test at the base (entry point). However, it is usually not necessary to test beyond the peak points. In fact, many systems already scale capacity using high GB/Watt, high capacity HDDs. Similarly, scale-out systems can scale performance and capacity by step-and-repeat instantiation of the same devices as those tested per the BFF method.

# 5  Finding the Best Foot Forward

## 5.1  Overview and Goal

The Best Foot Forward (a.k.a." sweet spot") as a methodology for testing product/family configurations at the peak values of the power efficiency metrics was introduced in Section 4.4. The stated benefit of this approach is to reduce the testable sets from a large variable range to fewer in number (potentially just one) with the test results representative of the entire product family.

This section describes one method for finding the Best Foot Forward configuration by using prediction tools; it also provides characteristics of the approach. By using the described tools, a

large range of configuration variables can be evaluated and the predicted sweet spots arrived at relatively quickly.

## 5.2  A Step-wise Approach

To determine the Best Foot Forward, a vendor can follow these steps:

1.  Start with a product offering that fits within a taxonomy definition. If the product can be configured to fit into several taxonomies, then the vendor should consider a separate data submission for each applicable taxonomy category and classification.

2.  Considering all possible (and valid, i.e., saleable) product SKU's (Stock Keeping Unit), identify the optimized configurations that will give the peak power efficiency metrics.

    Since there are six different SNIA Emerald™ Program test profiles (five active and one idle), it is expected that there can be up to six different optimized (tuned) configurations that achieve peak metrics:

    *   1 x Hot band [IO/s/Watt], 2 x Random [IO/s/Watt], 2 x Sequential [MiB/s/Watt], and 1 x Ready-Idle [raw capacity, GB/Watt]

3.  Use estimator tools to predict the peak metrics. The alternative is to develop educated-guess derivations, which could potentially lead to a significant amount of labor- and resource-intensive testing. As long as the simulated results are reasonably accurate, the physical configuration selected to identify (by measurement) the peak value can be reasonable in size or range.

4.  Set up, test, and measure the peak metric values for your first sweet-spot:
    *   Run through the complete sequence of SNIA Emerald™ test profiles.
    *   Test, validate and data correlate the predicted results.

5.  Re-configure and re-test for each additional sweet spot of interest.

For each sweet spot, there is a tuned configuration that will produce a peak metric for a specific test profile. However, a single tuned configuration may, in fact, generate multiple peak metrics for related workloads (i.e., random or sequential). When submitting sweet-spot data, it may be advantageous to identify the SUT as *optimized to perform best at specific test profile "X."*

## 5.3  Discussion of Estimator/Simulation Tools

 When faced with the task of finding the peak metric values of full product/family range of configurations, estimator tools can be an invaluable aid. Storage vendors may have a variety of power calculator and performance estimator tools for their storage products. Some may even have tools that can predict a limited set of power efficiency metrics. These tools can be based on complex simulation methods and/or grounded on some data points with interpolation and extrapolation. The accuracy of prediction is always in question, and thus the predicted results will always need to identify completed data correlations before accuracy claims can be made. The *Test Data Report* (TDR) may contain spec sheet data that allows customers to perform these calculations, as well.

## *5.4 Example Exercises*

Using power calculator and performance estimator tools for a representative Online-3 array, some characteristic plots of performance, power and the power efficiency metrics were generated for the SNIA Emerald[TM] Program test profiles. The array controller performance options were fixed to a high level, and then the configuration variables in the drive type and drive count were evaluated (Note: SSDs were not included in this set of exercises). The maximum configuration size of this array is 240 large form factor (LFF) drives and/or 450 small form factor (SFF) drives. The objective of the prediction exercises is to find the peak metrics for power efficiency for each test profile. Several illustrative plots are shown in Figure 3, Figure 4, and Figure 5.



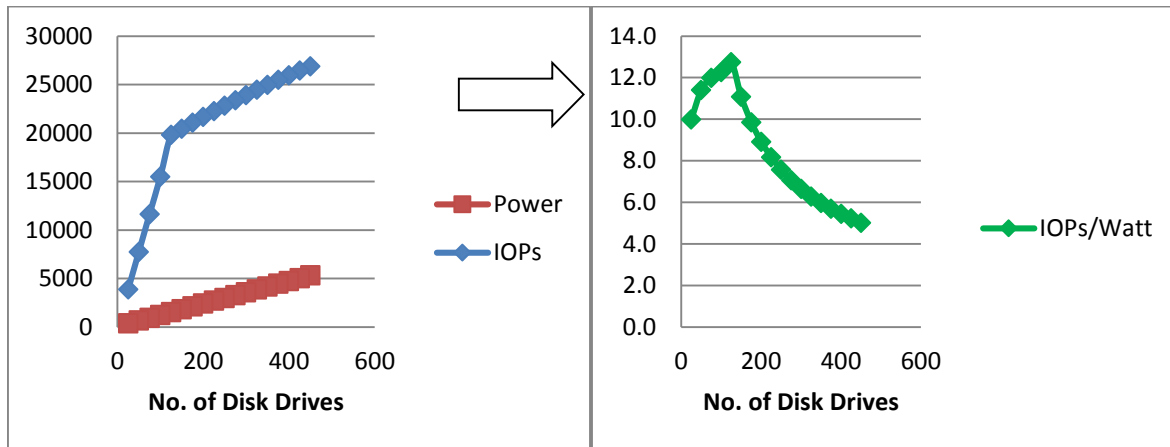**Figure 3   Performance, Power, and Power Efficiency Metric vs. Drive Count [Random workload, SFF 15K rpm SAS drives]**
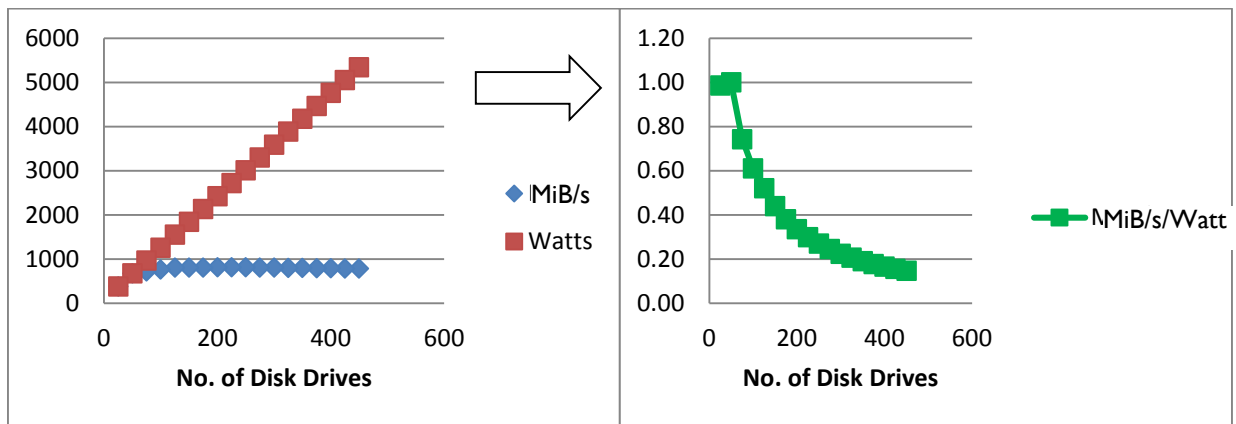


**Figure 4   Performance, Power, and Power Efficiency Metric vs. Drive Count [Sequential workload, SFF 15K rpm SAS drives]**
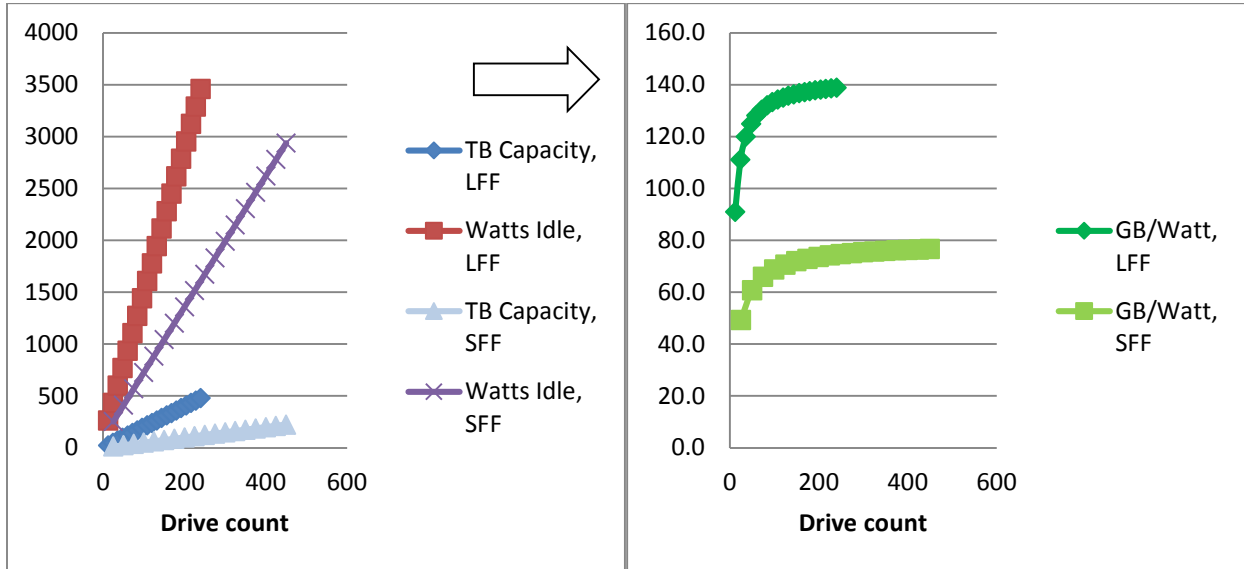
**Figure 5  Idle Capacity, Idle Power, and Idle Efficiency Metric vs. Drive Count
[LFF 2TB 7.2K rpm and SFF 500GB 7.2K rpm SAS drives]**

Obviously, based on the controller performance, bandwidth, and hardware efficiency, the slopes and shapes of these curves will vary. However, these observations can be made from this example:

- For all cases, the power steadily and regularly increases as the configuration size increases.

- For all active cases, the performance reaches a peak at a configuration considerably smaller than the largest drive count; then it levels out or goes down slightly.

- For all active cases, the peak metric [performance/power] is also reached at relatively low drive count configurations.

- For random and for sequential workloads, the peak metrics were achieved with the SFF, 15K rpm spinning drive.

- For the ready idle case, the peak metric continues to rise with drive count (as the controller electronics power is amortized over increasing numbers of drives).

# 6   Setting Up and Running the Measurement Tests

The *Measurement Specification* includes procedures used to derive the storage power efficiency metric values for Online, Near Online, Removable Media Library, and Virtual Media Library categories. While all procedures follow the same basic flow, each has variations due to their inherent characteristics.

This section will focus on the Online and Near Online categories. Their test procedures are similar enough that the *Measurement Specification* utilizes a single procedure with differences indicated. Aspects of Removable Media Library and Virtual Media Library categories are also noted.

Sections 6.1 through 6.5 provide detail on test configuration aspects, benchmark driver and power meter requirements plus test procedures and metric calculations.

## 6.1   *Configuration Set-Up*

### 6.1.1   Test Configuration

The storage system power efficiency metric measurements are intended to take place in a location indicative of a data center environment. The input power source for the storage system must meet the voltage requirements listed in the *Measurement Specification*. Environmental aspects such as temperature and humidity shall follow the *ASHRAE 2012 Thermal Guidelines Class A1* specifications.

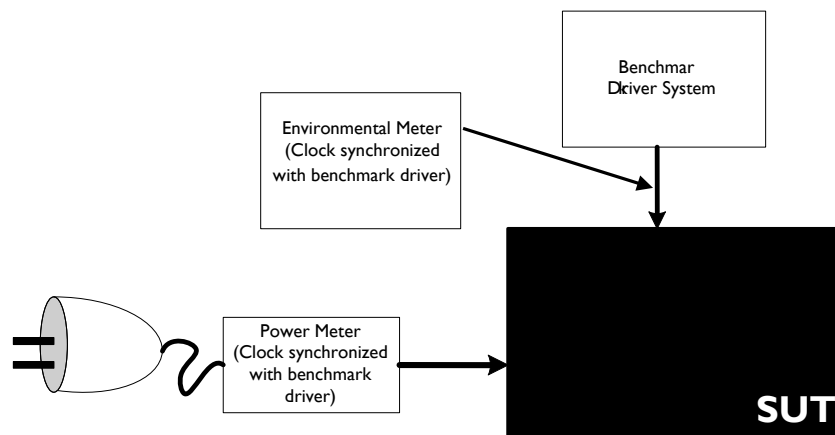The general setup is shown in Figure 6.



**Figure 6   Basic Measurement Setup**

### 6.1.2   Benchmark Driver Requirements

The Vdbench driver is used to provide workload generation and data collection. This driver plus an associated set of configuration scripts provide the generation of required workloads listed in the *Measurement Specification*. Both the specified Vdbench driver and script must be

used during testing. The scripts also specify variables settable by the test sponsor for determining optimal (e.g., BFF) results for particular test configurations. Access to Vdbench and associated scripts is via the Emerald website listed in Section 1.2 *References*.

The Measurement Specification defines five (5) active test workload IO profiles realized in the scripts:

- Hot Band
- Random Read
- Random Write
- Sequential Read
- Sequential Write

The Hot Band test is series of 13 random and sequential workloads designed to demonstrate the effectiveness of read caching. The remaining four (4) tests are collectively well known in the industry as 4-corner tests. The Online and Near Online test procedure utilizes all five workloads. The Removable Media Library and Virtual Media Library procedures only utilize sequential read and sequential write workloads.

Some recommendations before running the Vdbench scripts include:

- If you are not knowledgeable about Vdbench, it is strongly suggested that you read the user guide included with the Vdbench package.

- Use any available tools that may help you to configure your system for optimal use of energy based on your storage needs.

- Attempt, as much as possible, to do a single host run. It will make completing the measurement much easier.

### 6.1.3  Power Meter Requirements

The power meter is required to take accurate SUT power samples during selected tests in sync with the benchmark driver. Its requirements are listed in Table 1.

**Table 1: Power Meter Requirements**

| Power Consumption (p) | Minimum Accuracy |
|---|---|
| p ≤ 10 W | ± 0.01 W |
| 10 < p ≤ 100 W | ± 0.1 W |
| p > 100 W | ± 1.0 W |

### 6.1.4  Recommended Power Meters

A list of recommended power meters is located in appendix A of the *Measurement Specification*.

## *6.2   Test Procedures*

### 6.2.1   Online and Near Online

The *Measurement Specification* Online and Near Online test procedure follows this flow:

1.  Pre-fill Test to fill the SUT with random data.

2.  SUT Conditioning Test to get the SUT into a known state.

3.  Active Test phases that collect data for the active metrics, each with a method to assure metric value validity.

4.  Ready Idle Test that collects data for the capacity metric.

5.  Selected COM tests that demonstrate the storage system's ability to perform defined capacity optimization methods.

Tests are required to be run in an uninterrupted sequence, with the exception of the COM Test(s). The operation of COMs functions during active or idle tests is at the discretion of the vendor and test sponsor.

The Vdbench workload generator is used to provide the workload set to the SUT during the Pre-fill Test, SUT Conditioning Test, and Active Test phases. The Ready Idle Test requires no external IOs (IOs from or to a host), but the system shall be connected to the network or host and ready to support an IO request.

The COM tests have their own execution methods that in some cases require the use of particular data sets generated by a program referenced on the *SNIA Emerald™* website. See Section 6.3 *Capacity Optimization Method (COM) Tests*.

To generate a correct representation of the power efficiency of a SUT, it must be properly pre-filled with data and pre-conditioned. This is the goal of the Pre-fill and SUT Conditioning Tests, which are designed to get the system to a known state before active test phase measurements commence. The Pre-fill phase will fill up a percentage of the SUT with data that is two-to-one compressible. This fill percentage is specified in *Measurement Specification*.

The SUT Conditioning Test utilizes the Hot Band IO profile to demonstrate the SUT's ability to satisfy IO requests, ensure that the storage devices of the system are fully operational, achieve operational temperature, and place the storage system into a known state. The minimum SUT Conditioning Test time period is specified by *Measurement Specification* but may be increased in length as necessary. Online systems must have an average response time of less than 20ms for the last four hours of the SUT Conditioning Test. Near Online systems do not have an average response time requirement

Each Active Test phase requires determination of a validated, continuous 30-minute (1800 second) measurement interval. Data collected during this interval is used for calculating primary metrics (*Measurement Specification* Section 8). This entire interval must satisfy certain validity criteria including response time and metric stability. Any continuous 30-minute interval during an Active Test phase that meets all acceptance criteria may be used for calculating primary metrics.

Active Test phase data is gathered over consecutive 1-minute (60 second) intervals. As detailed in the *Measurement Specification*, performance data is collected and averaged over each 1-minute interval [O(60)]. The power meter collects and averages 5 second power measurement samples over the same 1-minute interval [PA(60). A performance per watt value is then calculated from the corresponding aligned average performance and average power values [EPP(60)].

For a continuous 30-minute measurement interval to be valid:

- The average response time of each 1-minute interval [RTA(60)] must meet the 80ms criteria set by the Active Test. This only applies to Hot band, Random Read and Random Write response times;
- The average response time of the entire 30-minute interval [RTA(1800)] must meet the 20ms criteria set by the Active Test. This only applies to Hot band, Random Read and Random Write response times;
- The 30 consecutive EPP(60) values must be deemed stable.

Stability is defined as the "flatness" of the EPP(60) values over a candidate span of 30 consecutive 1-minute intervals. Two methods are utilized to determine flatness. Both methods are required to pass in order to achieve stability:

a) Maximum allowed slope of a linear approximation of the candidate 30 EPP(60) values;
b) A moving average smoothing filter applied to the 30 candidate EPP(60) values and compared against a specified baseline.

Stability is determined for (a) by applying a least squares linear fit to the 30 candidate values. The absolute value of the resulting slope is not to be greater than the value specified in the *Measurement Specification*.

Stability is determined for (b) by applying the specified smoothing function to the 30 candidate values. The resulting values are compared against a defined baseline with deviation of all points limited to a validity range specified in the *Measurement Specification*.

The Ready Idle Test must last at least two hours. The last two hours are utilized for the metric.

### 6.2.2 Removable Media Library

The Removable Media Library test procedure has no Pre-fill or COM tests. There are also no response time requirements for the Conditioning or Active Tests. The Active Test only uses Sequential Read and Sequential Write workloads.

The *Measurement Specification* calls for the power efficiency measurement of these systems to be within 80% of the published data throughput due to their sequential IO nature. The published throughput of a system may have to be calculated by determining the throughput average of the various media devices.

### 6.2.3   Virtual Media Library

The Virtual Media Library test procedure has no Pre-fill or COM tests. There are also no response time requirements for the Conditioning or Active Tests. The Active Test only uses Sequential Read and Sequential Write workloads.

The *Measurement Specification* calls for the power efficiency measurement of these systems to be within 90% of the published data throughput due to their sequential IO nature.

### 6.2.4   Data Collection Summary

The data collection requirements for the taxonomy categories are listed in Table 2, which are taken directly from the *Measurement Specification*.

**Table 2: Data Collection Summary**

| Test | Collection Interval (seconds) | | Minimum Benchmark Driver Data Collection | | Minimum Test Duration (minutes) | |
|---|---|---|---|---|---|---|
| | Power Meter | Temp Meter | Online/ Near Online | Removable/ Virtual | Online/ Near Online | Removable/ Virtual |
| Conditioning | 5 | 60 | Response Time (per 1-minute interval) | Throughput (MiB/s) | 720 | 7 |
| Active | 5 | 60 | Response Time (per 1-minute interval) | Throughput (MiB/s) | 40 | 30 |
| Idle | 5 | 60 | N/A | N/A | 120 | 120 |

## 6.3   Capacity Optimization Method (COM) Tests

The COM tests are effectively existence tests and are only performed on Online and Near Online category systems.

COMs represent a class of particular (and potentially significant) storage efficiency capabilities otherwise difficult to acknowledge via Active Test phases. In order to provide a method of credit, the *Measurement Specification* utilizes a set of heuristic tests to ascertain the existence and active state of COMs. The goal of these tests is simply to provide a method for an independent third party to verify that the system under test is indeed capable of supporting selected COMs.

COM types include (but in the future are not limited to):

1) Delta snapshots (read and write);
2) Thin provisioning;
3) Data de-duplication;
4) Data compression;
5) Parity RAID.

The heuristic tests are meant to determine COM existence and not judge effectiveness. Hence, each is a basic yes/no test. Test sponsors may choose which heuristic tests to run (and receive credit). All tests are relatively straight forward but vary depending on the individual COM. The actual COM function determines what type of test is run.

Vendors must follow the given test steps for each COM they wish to be awarded credit on a given SUT. During a test sequence no media may be added or removed, changed in state (taken on- or off-line, made a spare, or incorporated, etc.), or RAID groupings changed. In the event of a disk failure and subsequent automated RAID rebuild at any time during a test, the test must be restarted after the rebuild is completed and the failed disk replaced per manufacturer guidelines.

Some COM tests require particular data sets to demonstrate existence. These data sets are generated by the COM Test Data Set Generator C program available at www.sniaemerald.com/download. This program is compiled and loaded on the test host prior to testing. Operational instructions are contained an associated readme file.

Three different data sets are generated each approximately 2GB in size:

- Completely irreducible: Cannot be significantly reduced in size by either compression or de-duplication methods;
- Dedupable but not easily compressible: Can be significantly reduced by de-duplication but not easily by compression methods;
- Compressible but not dedupable: Can be significantly reduced by compression but not by de-duplication methods.

The exclusive nature of the data sets supports systems with multiple active COMs, i.e. those that the SUT may be unable to individually disable.

### 6.3.1   Delta Snapshot Test

Delta snapshot heuristics utilize a before-after free space method to demonstrate basic COM functionality. A "container" is defined comprised of allocated space and a test data set.

Container free space is determined at the start of the test. The COM is tested on the data set and free space again determined and the end of the test. The test then specifies an existence threshold based on the two free space values. The process consists of the following steps:

a) Create a container on the SUT and query its amount of free space;

b) Write a 2GB irreducible data set into the container and create a read or write snapshot;

c) Read something from or write a few characters to the snapshot (depending on type);

d) Query the amount of container free space to determine whether significant additional storage space has been used.

Read-only and writeable delta snapshots are treated separately so that systems that only do read-only snapshots may get credit for them.

### 6.3.2   Thin Provisioning Test

The test for thin provisioning is simple:

- If thin provisioning is disabled, the SUT should not be allowed to allocate more than the available usable space;
- If thin provisioning is enabled, the SUT should be allowed to allocate more than the available usable space.

### 6.3.3   Parity RAID Test

Parity RAID existence is established via a simple equation, as compared with RAID 1, given that RAID group sizes and parity requirements are simple and well known.

### 6.3.4   De-Duplication and Compression Tests

Data de-duplication, and compression heuristics utilize a before-after free space method to demonstrate basic COM functionality. A "container" is defined comprised of allocated space and a test data set. Container free space is determined at the start of the test. The COM is run on the data set and free space again determined and the end of the test. Each test then specifies an existence threshold based on the two free space values.

These tests utilize all three generated data sets. As such, it is important that the data sets possess attributes conducive to testing different de-duplication and compression implementations.

SUTs may have minimum size thresholds before a de-duplication or compression function is executed. Hence, a data set greater than 2GB may be required. In this case, it is possible to construct larger data sets using combinations of the defined 2GB data sets. For instance, suppose a particular storage array does not enable de-duplication unless a data set size is at least 8GB.

A necessary data set can be constructed by concatenating the defined dedupable 2GB data set along with at least 3 irreducible 2GB data sets. However, the 3 irreducible must not collectively contribute to the de-duplication test. Hence, each irreducible data set is generated using a different "salt" value such that the combination of irreducible data sets is collectively irreducible. It is suggested that the user attempt to use prime numbers for each salt value to assure this uniqueness. The same construct can be applied to a compression data set as required.

The deduplication dataset contains many duplicated files of various sizes and many duplicated blocks aligned on block boundaries. It also contains duplicated blocks of variable lengths that are not aligned on block boundaries. This allows detection of block-based schemes, variable-

length schemes, and SIS schemes when used in place. To better understand deduplication, refer to "Understanding Data Deduplication ratios" -- DDSR SIG, located at this website:

http://www.snia.org/sites/default/files/Understanding_Data_Deduplication_Ratios-20080718.pdf

Another concern of compression existence testing is making sure that the related data set covers various compression methods. The existence test is again intended to make no judgment of how or what type of data is compressed, only that a compression method exists. To that end, the data set generation tool provides for both bit-level and pattern-oriented compression methods.

## 6.4 Avoiding Potential Pitfalls while Taking Measurements

With all the complexities of storage systems, not all potential issues associated with taking measurements on the system can addressed by the *Measurement Specification*. This section lists certain issues that may need to be addressed by the test sponsor when taking measurements. This is not intended to be all inclusive, but rather a list of general items that should be considered and/or addressed.

- Ensure that the measurement includes enough writing on the sequential write test to have enough written data for a stable sequential read test.

- All disclosed RAS features must be activated during measurement procedures. Certain tasks such as charging batteries should be completed before measurements start as this is not a typical operational function.

- Timing between the workload generator and the power/temperature meter should match. Any offset will cause the metric generation to be off. The time settings should be within one second of each other.

- The host providing the workload to the storage system should not be the bottleneck. The *Measurement Spec* does not prescribe client set-up except that Vdbench and the script (with user settable variables) are required. Use Vdbench per its own requirements. Proper client sizing is necessary for best results.

SNIA has identified and listed in the *Measurement Specification* several Reliability/Availability/Serviceability (RAS) features of storage systems with significant impacts on power consumption. These RAS features are requirements of contemporary highly available and serviceable storage systems. The issue with such functions is that their existence may contribute to power draw but have no direct positive benefit on performance and hence may have a detrimental impact on certain *Measurement Specification* metrics.

## 6.5 Reported Metrics

Section 8 of the *Measurement Spec* details the calculation of final power efficiency and COM metrics, segregated by taxonomy category. Active test values are calculated over the valid data 30-minute (1800 second) measurement interval. The Ready Idle result is calculated over its 2 hour (7200 second) measurement interval. Final COM metrics are represented by a simple true(1) or false(0) value depending on whether the COM test satisfied its associated heuristic.

# 7 Notes on Submitting Data

The resulting SUT metrics and configuration information are combined for submittal to the SNIA Emerald™ Program and/or the EPA ENERGY STAR® Data Center Storage Program.

The SNIA Emerald™ Program utilizes its own Test Data Report (TDR) for submission (see Section 1.2 *References*). This report provides entries for basic system information, test setup, and the metric results for each of the test phases defined for the taxonomy category. Average response time values must be submitted but may or may not be publicly published as indicated by selecting *yes* or *no* to the TDR publish average latency data question. Once the TDR is completed and verified for accuracy, it can be submitted to the SNIA Emerald™ Program website. The EPA uses a different but similar mechanism for data submittal to the ENERGY STAR® Data Center Storage Program. Details can be found at the EPA website (see Section 1.2 *References)*.

Per Section 4 *Identifying the Product Family* and Section 5 *Finding the Best Foot Forward*, there are tradeoffs between capacity, performance, and power. These tradeoffs need to be evaluated by storage system vendors when promoting their products to specific markets. It has not been possible to define a single storage power efficiency metric proxy for all capable system setups. As such, it is in the best interest of vendors to submit multiple system configurations as appropriate to the *SNIA Emerald™* and/or ENERGY STAR® Programs to demonstrate overall storage power efficiency.