Overview of SNIA Emerald™ Specification 3.0.1

Herb Tanzer, Chuck Paridon
SNIA Green TWG

SNIA Emerald™ Training

SNIA Emerald™ Power Efficiency Measurement Specification

Version 3.0

February-March 2018

The Emerald™ Specification

The SNIA Emerald™ Power Efficiency Measurement Specification identifies metrics by which energy consumption and efficiency of storage networking products can be measured for the purposes of new product development, end-user customer evaluation, and regulatory standards development. All versions of the Measurement Specification are available for download.

https://www.snia.org/tech_activities/standards/curr_standards/emerald

History of the Emerald™ Specification

Release	Date	Key changes, besides editorial / clarity
v1.0	Aug 23, 2011	Initial release of SNIA technical position (block storage only)
v2.0.2	Aug 12, 2013	Added definitions for Hot Band workload and auto-tiering; changed response times, revised Vdbench scripts
v2.1.1	Dec 2, 2015	Specified test data sets & revised test steps for COMs
v3.0.1	Sept 2017	Add tests for file access devices of Online and Near-Online taxonomy categories. Eliminate tests for Parity RAID as a COM. Improve tests for Data Deduplication and Compression COMs. Eliminate taxonomy categories Adjunct Product and Interconnect Element. Add Stable Storage. Align power requirements with ENERGY STAR Data Center Storage. Add temperature meters to Annex A.

Note: Planning to submit ISO version of v3.0.1 to ISO by ~ 2H18

Highlight of Emerald v3.0.1

The SNIA Emerald™ Specification V3.0.1 is a major extension of the V2.1.1 specification that:

- Adds a test methodology for file access storage devices based on the SPEC® SFS2014 benchmark, substantially expanding the range of testable storage devices;
- Retains all the major elements of V2.1.1 including the storage taxonomy and the block IO test methodology
- The ENERGY STAR Data Center Storage Program has been updated to V1.1, and now includes file access devices with testing based on Emerald v3.0.1

Emerald Specification v3.0.1 Contents Outline

- Clause 1 – Overview
- Clause 2 – Normative References
- Clause 3 – Scope
- Clause 4 – Definitions, Symbols, Abbreviations, and Conventions
- Clause 5 – Taxonomy
- Clause 6 – Capacity Optimization
- Clause 7 - Test Definition and Execution Rule
- Clause 8 – Metrics
- Annex A - J

Stable Storage has been added to v3.0.1

SNIA GSI | GREEN STORAGE

- Definition: Storage that retains its content over power failures

- Annex C: Stable storage is storage that retains data, for a minimum of 72 hours without external power, in the case of:
  1. Repeated power failures, including cascading power failures;
  2. Hardware failures (of any board, power supply, etc.);
  3. Repeated software crashes, including reboot cycle.

- Taxonomy: Stable Storage becomes a required attribute for Online 3 through 6

Data Center Storage Taxonomy Categories

(more detail in a later presentation)

SNIA GSI | GREEN STORAGE

- Online
- Near Online
- Removable Media Library
- Virtual Media Library

Attribute	Category			
	Online	**Near-Online**	**Removable Media Library**	**Virtual Media Library**
Access Pattern	Random/ Sequential	Random/ Sequential	Sequential	Sequential
MaxTTFD (t) [a]	t < 80 ms	t > 80 ms	t > 80 ms t < 5 min	t < 80 ms

- Block and File fall under the Online and Near-Online Categories
- Removed Adjunct Product and Interconnect Element *(change from 2.1.1)*
- Removed User Accessible Data attribute *(change from v2.1.1)*

Capacity Optimization Methods (COMs)

(more detail in another presentation)

SNIA GSI | GREEN STORAGE

- Are techniques that reduce the consumption of space required to store a data set
- COMs defined as:
  - Delta Snapshots
  - Thin Provisioning
  - Data Deduplication
  - Compression
  - Parity RAID
- Qualitative heuristics are simple pass/fail tests, intended only to verify the presence and activation of a particular capacity optimization method.
  Note: Are applied to both Block and File access
- For parity RAID, tests for presence are no longer defined *(change from 2.1.1)*

Block and File Access Tests

(File is new for v3.0.1)

SNIA GSI | GREEN STORAGE

- Vdbench, developed by Oracle, is used for Block workloads
- SPEC SFS 2014 is a file-based I/O benchmark, developed by the Storage Performance Evaluation Corporation (SPEC)
- There are four file system workloads based on SPEC SFS 2014 benchmark
  - Video Data Acquisition (VDA) Workload
    - Generally simulates applications that store data acquired from a temporally volatile source (e.g., surveillance cameras).
  - Database (DATABASE) Workload
    - Represents the typical behavior of a database.
  - Virtual Desktop Infrastructure (VDI) Workload
    - Simulates a steady-state high-intensity knowledge worker in a VDI environment
  - Software Build (SW Build) Workload
    - Classic meta-data intensive build workload derived from analysis of software builds.

Comparison of Block vs. File Access Workloads

Block Access Workloads

- Hot Band (mixed Streams w Hot Spots)
  - Measurements: IO/s/W
- Random Writes (100% write, 8k block)
  - Measurements: IO/s/W
- Random Reads (100% read, 8k block)
  - Measurements: IO/s/W
- Seq. Writes (100% write, 256k block)
  - Measurements: MiB/s /W
- Seq. Reads (100% read, 256k block)
  - Measurements: MiB/s/W

File Access Workloads

- Video Data Acquisition (VDA) Workload
  - Measurements: MiB/s/W*
- Data Base (DATABASE) Workload
  - Measurements: MiB/s/W*
- Virt. Desktop Infra. (VDI) Workload
  - Measurements: MiB/s/W*
- Software Build (SW Build) Workload
  - Measurements: MiB/s/W*

* these metrics are network through-put

File Access Workload Details

File Access IO Demand Intensity

Workload	Business Metric (LOAD parameter)
DATABASE	DATABASES
SWBUILD	BUILDS
VDA	STREAMS
VDI	DESKTOPS

At least 10 uniformly spaced business metric load values shall be specified for valid test execution

.

File Access Storage Capacity Requirements

DATABASE = 24 GB per DATABASE

SWBUILD = 5 GB per BUILD

VDA = 24 GB per STREAM

VDI = 12 GB per DESKTOP

Comparison of Block vs. File Access Execution

Block Access Test Execution

1. Pre-fill test, which puts data on the product under test;
2. Conditioning test, which assures accurate and reproducible measurements;
3. Active test, the basis for the active metrics;
4. Ready idle test, the basis of the ready idle metric;
5. Capacity optimization test (if defined), the basis of the secondary, capacity optimization metrics.

File Access Test Execution

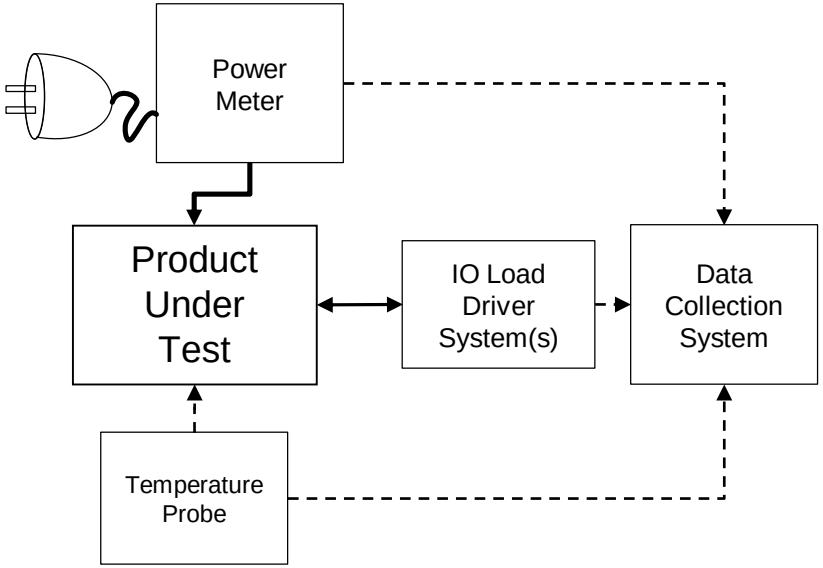
1. Execution of the four workloads, the basis for the active metrics;
2. Ready idle test, the basis of the ready idle metric;
3. Capacity optimization test (if defined), the basis of the secondary, capacity optimization metrics.

Using alternate IO measurement boundaries

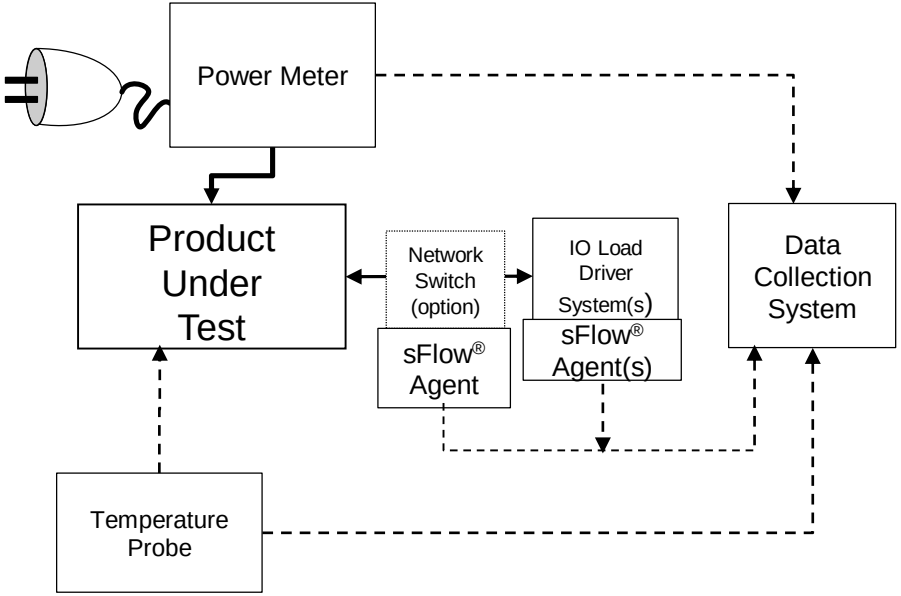
- Block IO accesses are non-buffered.
  - The IO is issued from the driver and the acknowledgement is returned after the IO is committed to stable storage within the storage device.
  - Thus a measure of the rate and response time of the storage subsystem
- File system accesses are buffered several times in the stack.
  - The IO is issued from the driver and the acknowledgement is returned after the IO is written in the client buffer cache.
  - From application space, the response time of the IO may have little to do with the storage subsystem
  - As a result, all IO rate metrics are derived from the Network IO driver or the interconnecting IP switch

Comparison of Block vs. File Access Schematic

Block Access Test Schematic



File Access Test Schematic



Comparative Data Collection Summary

SNIA GSI | GREEN STORAGE

Online Block Access Data Collection

Test	Collection Interval (seconds)		Workload Generator Data Collection		Minimum Duration (minutes)
	Power Meter	Temp Meter	Metric	sample (secs)	
Cond.	60	10	Average Response Time RTA_{sc} (milli-seconds)	60	720
Active	60	10	1) Operations Rate O_i (IO/s or MiB/s) 2) Average Response Time RTA_{sc} (milli-seconds)	60	40
Idle	60	10	N/A	N/A	120

Online File Access Data Collection

Test	Collection Interval (seconds)		Data Collection		Minimum Test Duration (minutes)
	Power Meter	Temp Meter	Metric	sample (secs)	
INIT	60	10	Operations Rate O_i (MiB/s)	10	N/A
Warm-up – per load point	60	10	Operations Rate O_i (MiB/s)	10	5
Active – per load point	60	10	Operations Rate O_i (MiB/s)	10	5
Ready Idle	60	10	N/A	N/A	120

Questions?