ENERGY STAR®
Data Center Storage
Version 1.0

Preliminary Data Analysis &
Draft 1 Discussion

15 April 2010

Agenda

- Data Collection Summary
- Preliminary Data Analysis
  - SNIA
  - ENERGY STAR
- Draft 1 Specification Overview
- Timeline & Next Steps
- Q&A

Data Collection

- Data collection from 12-2009 to 3-2010.
- Focus on high-value taxonomy categories:
  - Online, Near-online, Removable Media Library, and Virtual Media Library
  - Groups 2, 3, and 4
- Defined Generic System Configurations (GSC)
  - GSC-1: Performance Non-HA* Configuration
  - GSC-2: Performance HA Configuration
  - GSC-3: Capacity Non-HA Configuration
  - GSC-4: Capacity HA Configuration

** HA = High Availability*

Objectives

- Understand relationship between hardware/software configuration and energy performance
- Evaluate both Active and Idle state energy performance
- Conduct sensitivity analysis on single-variable configuration changes
  - Hard Disk Drive type (e.g. capacity vs. performance)
  - RAS (High-availability) features (e.g. redundant controllers)
  - SFF and SSD storage technologies

Data Collection Procedure

- Developed with substantial stakeholder input over several months
- Intended to exercise a variety of performance states, as relevant to and supported by various system types
  - Active state
    - Sequential Read / Sequential Write
    - Random Read / Random Write
    - Ramp Random 70/30 R/W
  - Idle state
  - Deep Idle state

Data Submissions

- ENERGY STAR Data Set
  - 10 Online systems (*26 configurations*)
  - 1 Near-online system (*1 configuration*)
  - 4 Removable Media Libraries (*13 configurations*)
  - 2 Virtual Media Libraries (*2 configurations*)
- Subcategories
  - Two hybrid systems (HDD/SSD, HDD/Optical)
  - Two server-based storage devices (capable of running applications on the controller)

EPA

Submitted Data

# Systems (# Configs)	Online	Near-online	Removable Media Library	Virtual Media Library
Group 2	--	--	1 (2)	2 (2)
Group 3	5 (7)	--	2 (10)	--
Group 4	5 (19)	1 (1)	1 (1)	--

Single-variable Changes

- Online:
  - 7200 RPM to 15000 RPM drives
  - Full HDD to Hybrid HDD/SSD
  - RAID-5 to RAID-50 configuration
  - 50% to 100% & 75% to 100% loaded disk shelf
  - 2X & 4X number of disk shelves
  - 2X number of controllers
  - 2X controller cache
- Archive:
  - N to N+1 to N+2 drives
  - Single to Redundant PSU
  - 2X & 4X capacity

System Performance Assessment

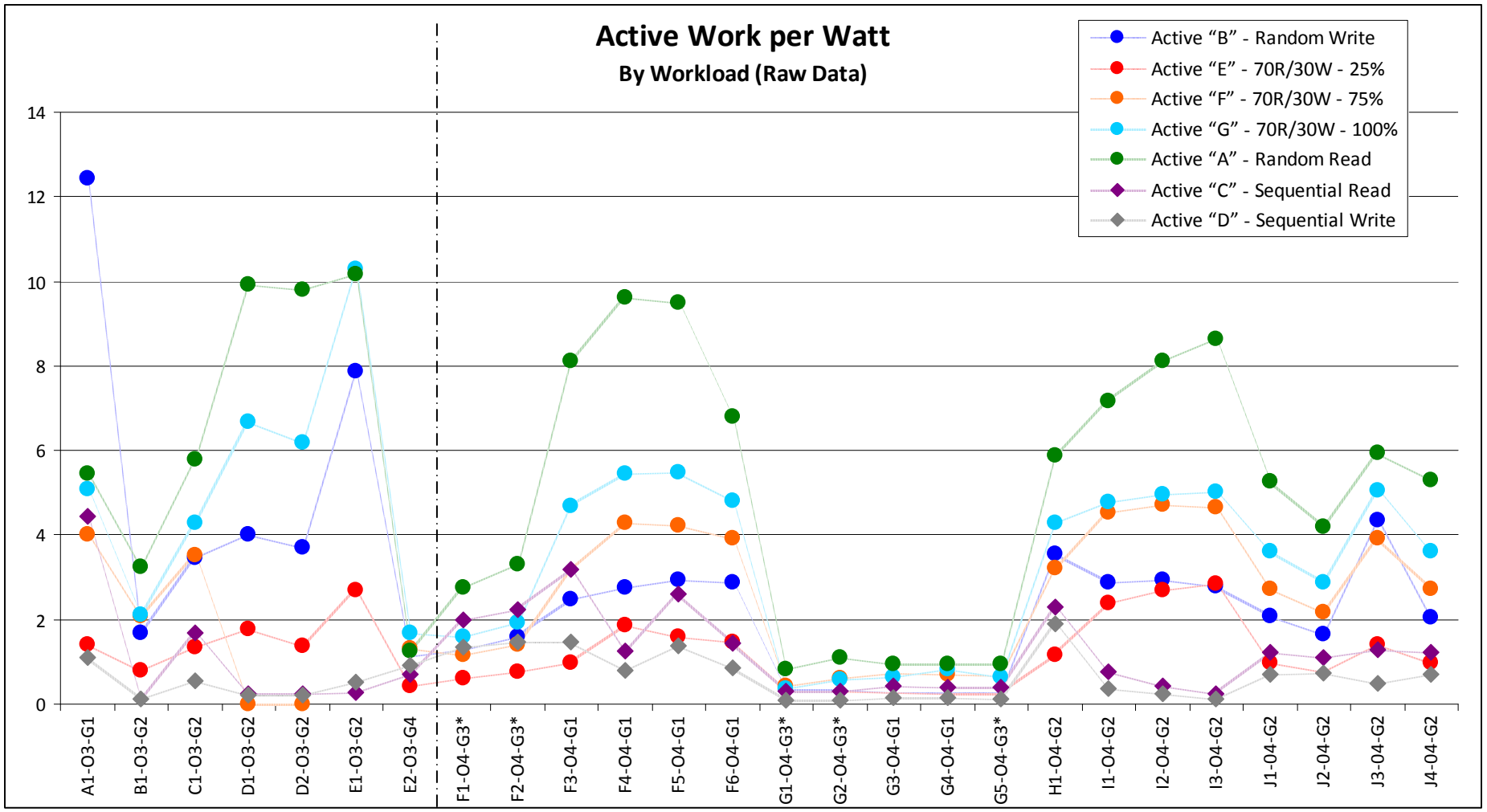
Items to Evaluate

- Performance variations as a function of taxonomy category
- PSU loading
- HDD type
- Raw vs. Effective (accounting for RAS features) capacity
- Other considerations
  - Hybrid systems
  - Server-based storage

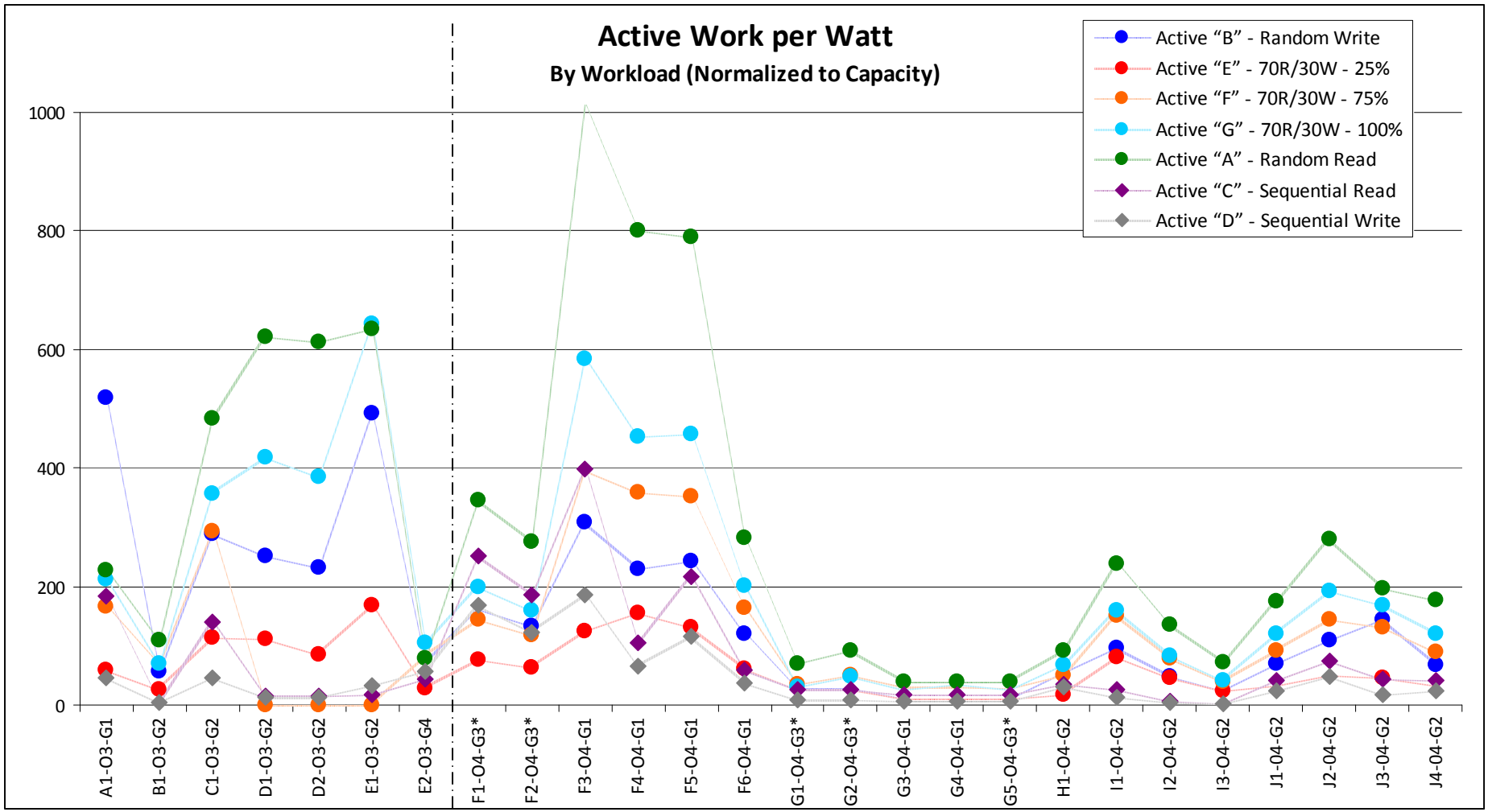
Online: Assessment

- Systems demonstrated individual strengths, though there was no clear leader in every category
  - Random or Sequential
  - Reading or Writing
- No strong correlation evident between Group 3 and Group 4 systems
- Additional RAS / Scalability showed little impact

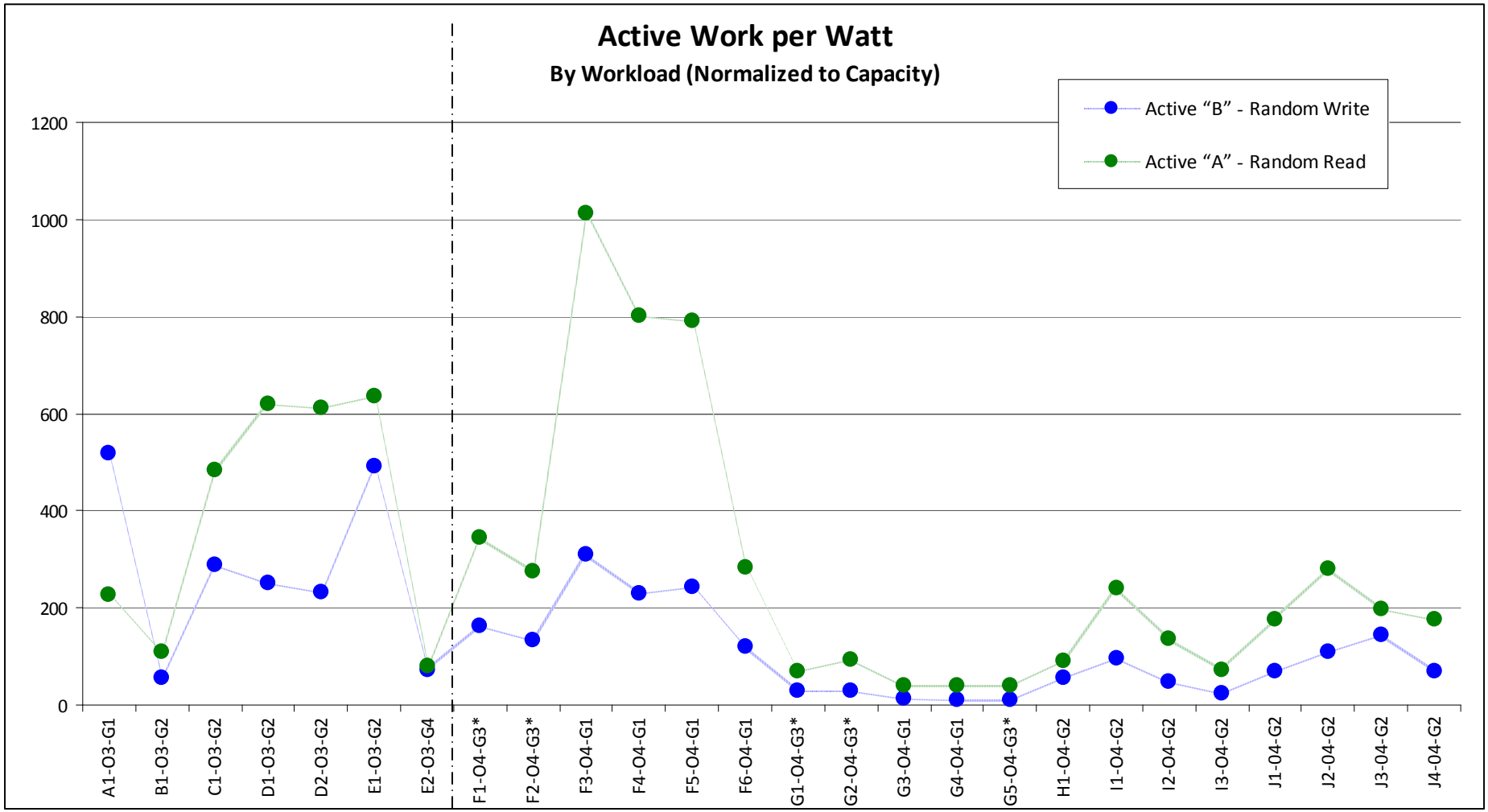
Online: Assessment



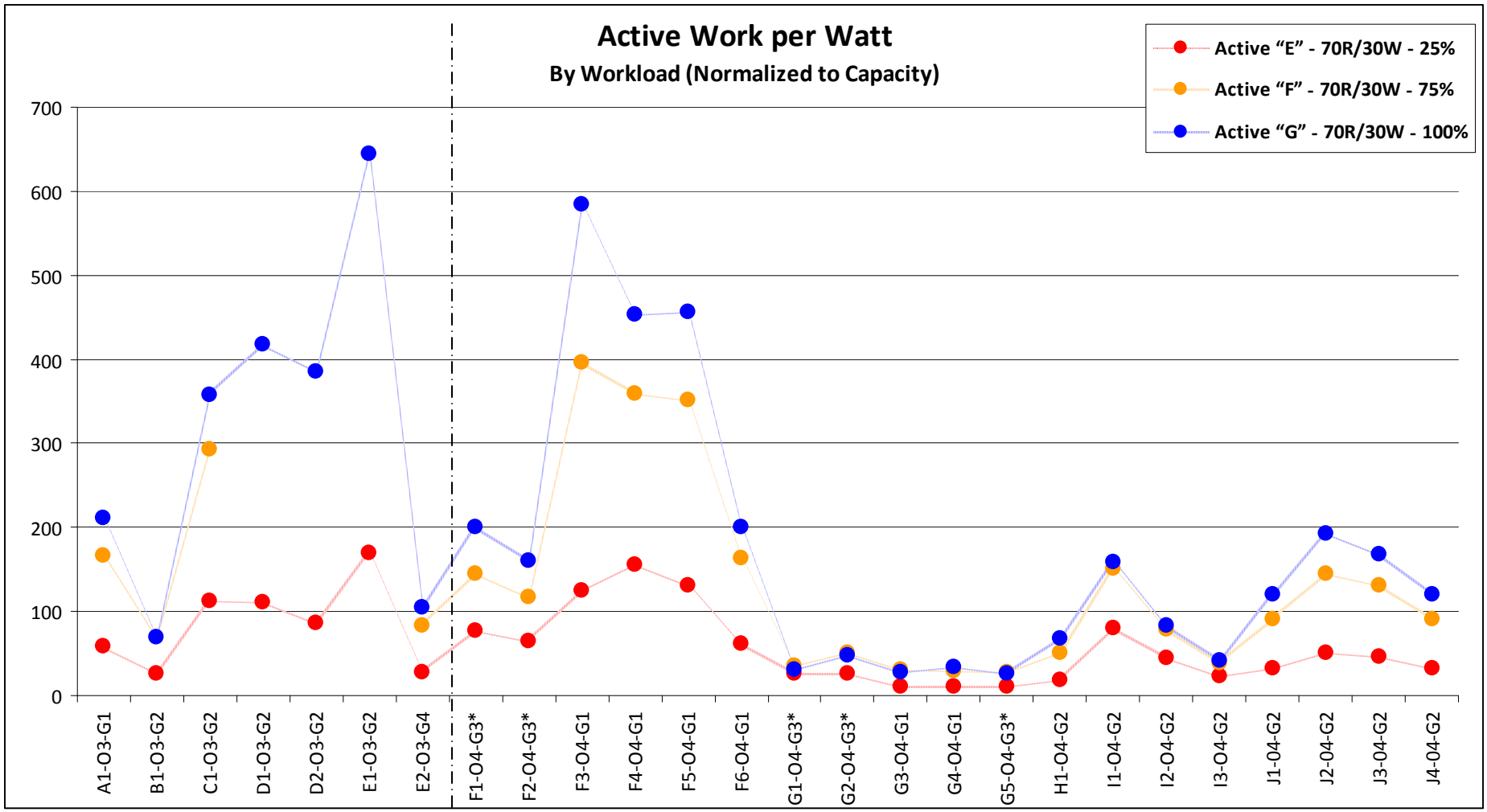
Online: Assessment



Online: Assessment



Online: Assessment



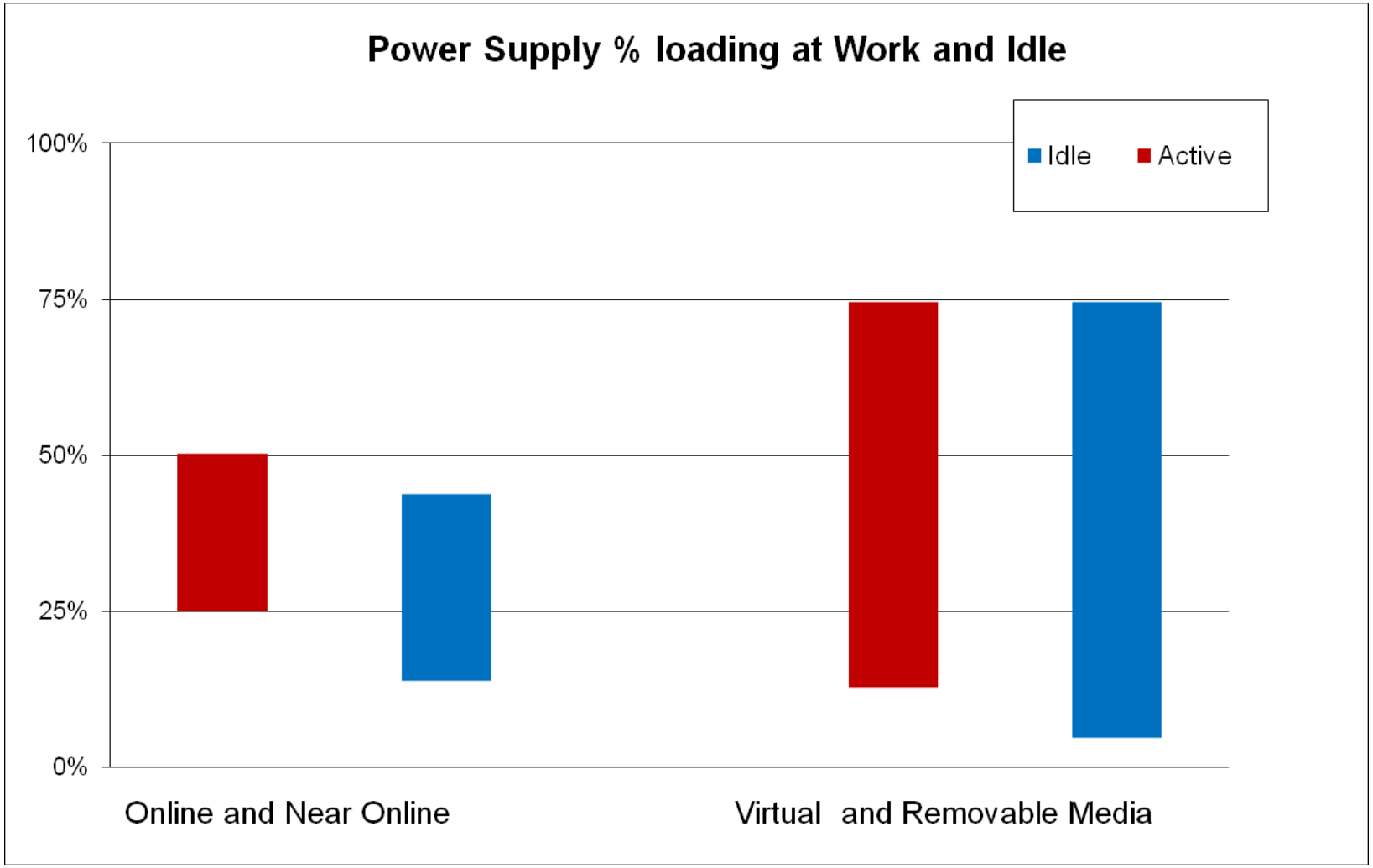
Online: Next Steps

- Test systems optimized for different workloads
- Merge Group 3 & 4? Would this extend to Group 2? Group 5?
- Conduct additional RAS testing?
  - Group 3 vs. Group 4 from same vendor
  - Same HDD speed / capacity, different interface

PSU Loading: Assessment

- Load as a function of taxonomy category:
  - Online & Near-online: 25-50% loading
  - Removable & Virtual: 13-74% loading
    - Several feature non-redundant PSUs
- Delta from Active to Idle state:
  - Online & Near-online: 7% average, 14% max improvement
  - Removable & Virtual: 0-50% improvement

PSU Loading: Assessment



PSU Loading: Next Steps

- Does the data support ENERGY STAR focus on efficiency in specific load ranges?
- Will data trend hold for Group 2 systems?
- Further investigate significance of the wide range of PSU loading for Archive systems?
- Revise test procedure & data collection sheet
  - Identify Controller vs. Drawer PSU
  - Isolate Controller vs. Drawer measurements?
  - Better accommodation for hybrid systems?

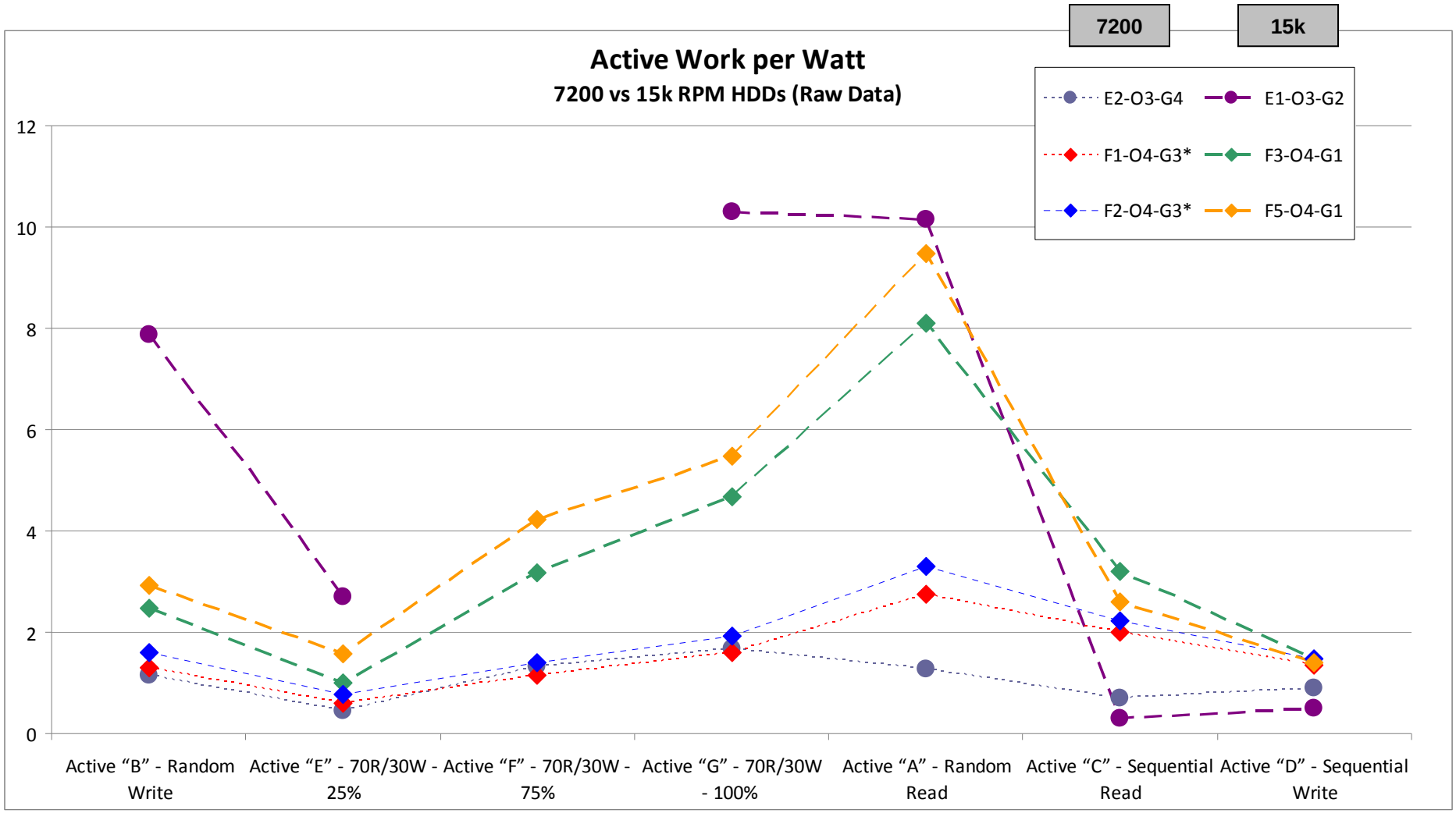
EPA

HDD Type: Assessment

- Workload dependent
  - Random workloads favor 15K HDDs
  - Sequential workloads do not highlight performance differences. Considering Work / Watt may favor high capacity HDDs
- No SSD-only systems were submitted for consideration
  - SPC results indicate a strong SSD performance advantage

EPA

HDD Type: Assessment



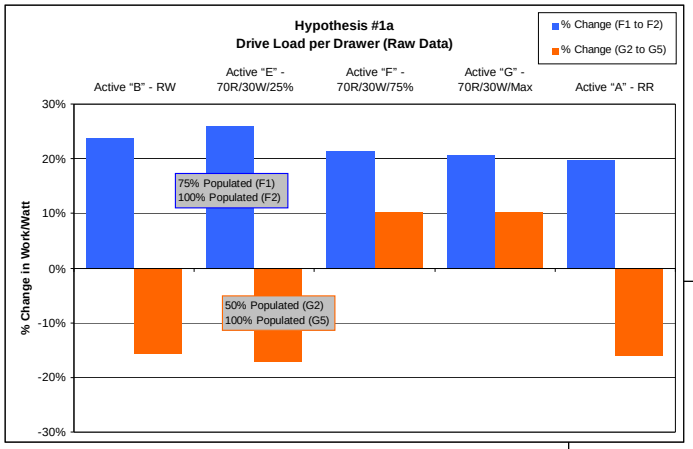
HDD Type: Next Steps

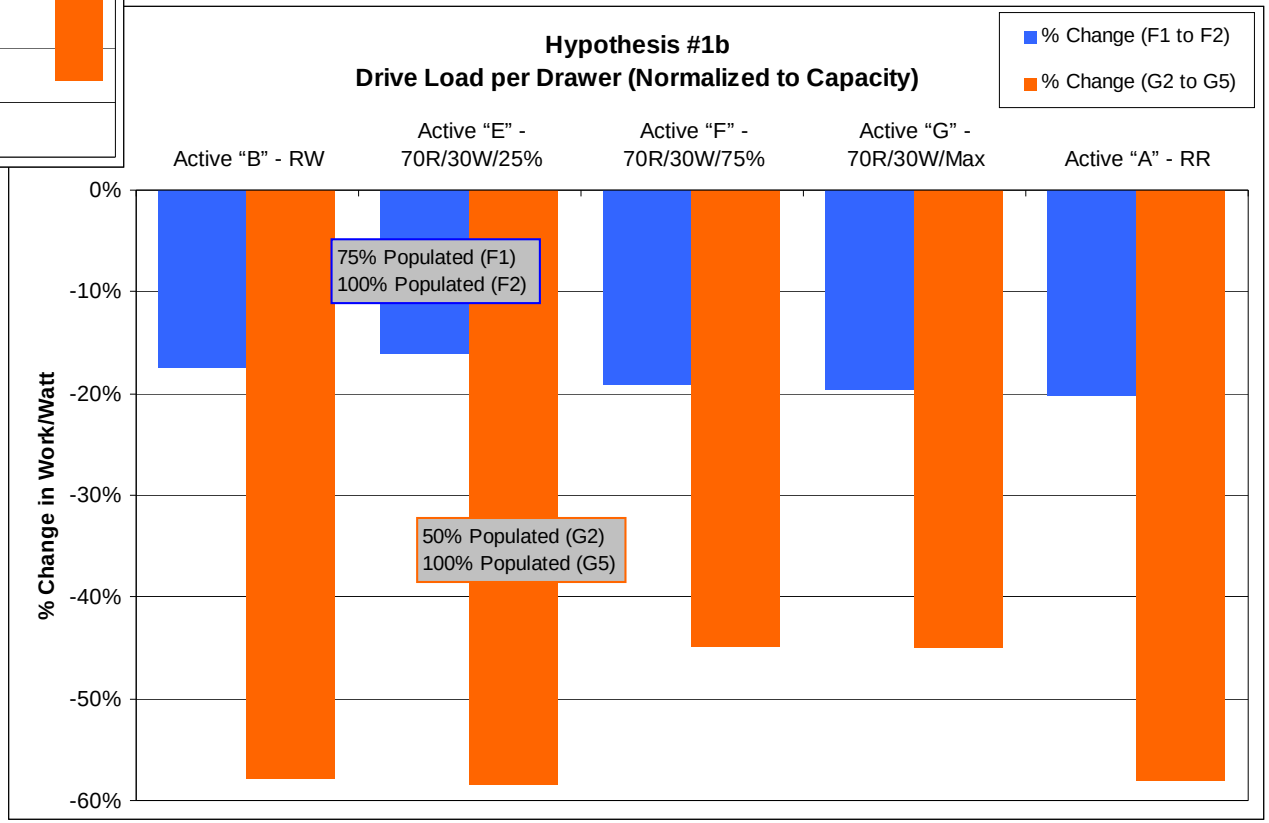
- HDD type is very relevant to test results
  - Selection addresses different end user goals – possibly a key factor for efficiency metrics
- Testing SFF and SDD?
  - One SFF system submitted
- Gather end-user insight regarding HDD selection vs. actual end-use applications?

Drawer Load: Assessment

- Partially- vs. Fully-loaded drawer
  - F1 / F2: 75% → 100%
  - G2 / G5: 50% → 100%
- No observed trend in Work / Watt
- Normalized view shows trend:
  - A 25% decrease drives per drawer = 20% increase in work/watt/capacity.
  - 50% decrease drives per drawer = 50% increase in work/watt/capacity.

Drawer Load: Assessment

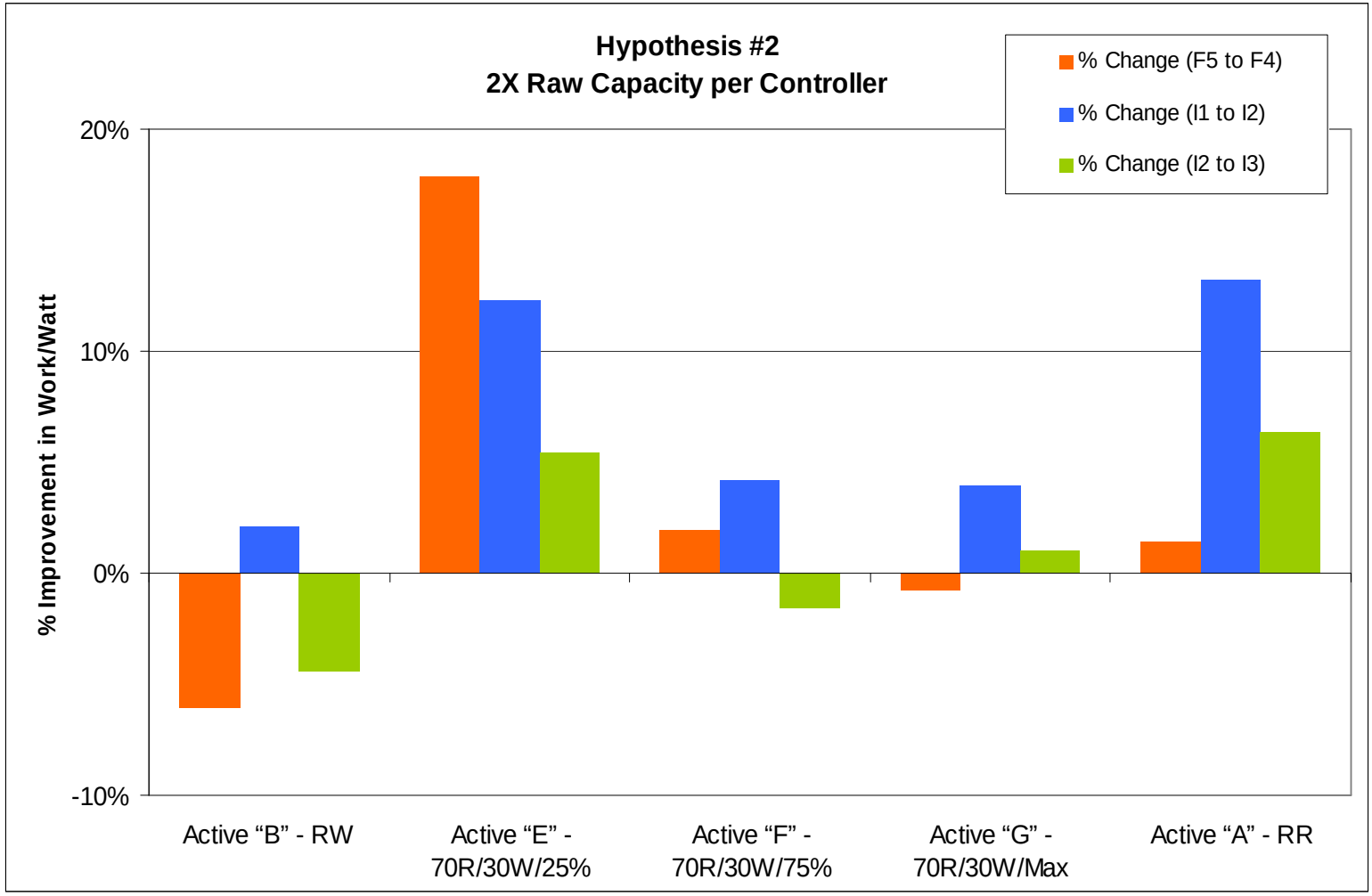




Controllers vs. Capacity

- Adding Drawers (HDDs) vs. Number of Controllers
  - F5 / F6: 12 → 24 HDDs; I1 / I2: 30 → 60 HDDs; I2 / I3: 60 → 120 HDDs
- Positive response to Read workloads, Negative response to Write workloads
- Responses are nonlinear
- Conclusions / Next Steps
  - Reinforces varied design points for storage: Optimized Workload, Optimal Capacity
  - Need to standardize testing configurations?

Controllers vs. Capacity

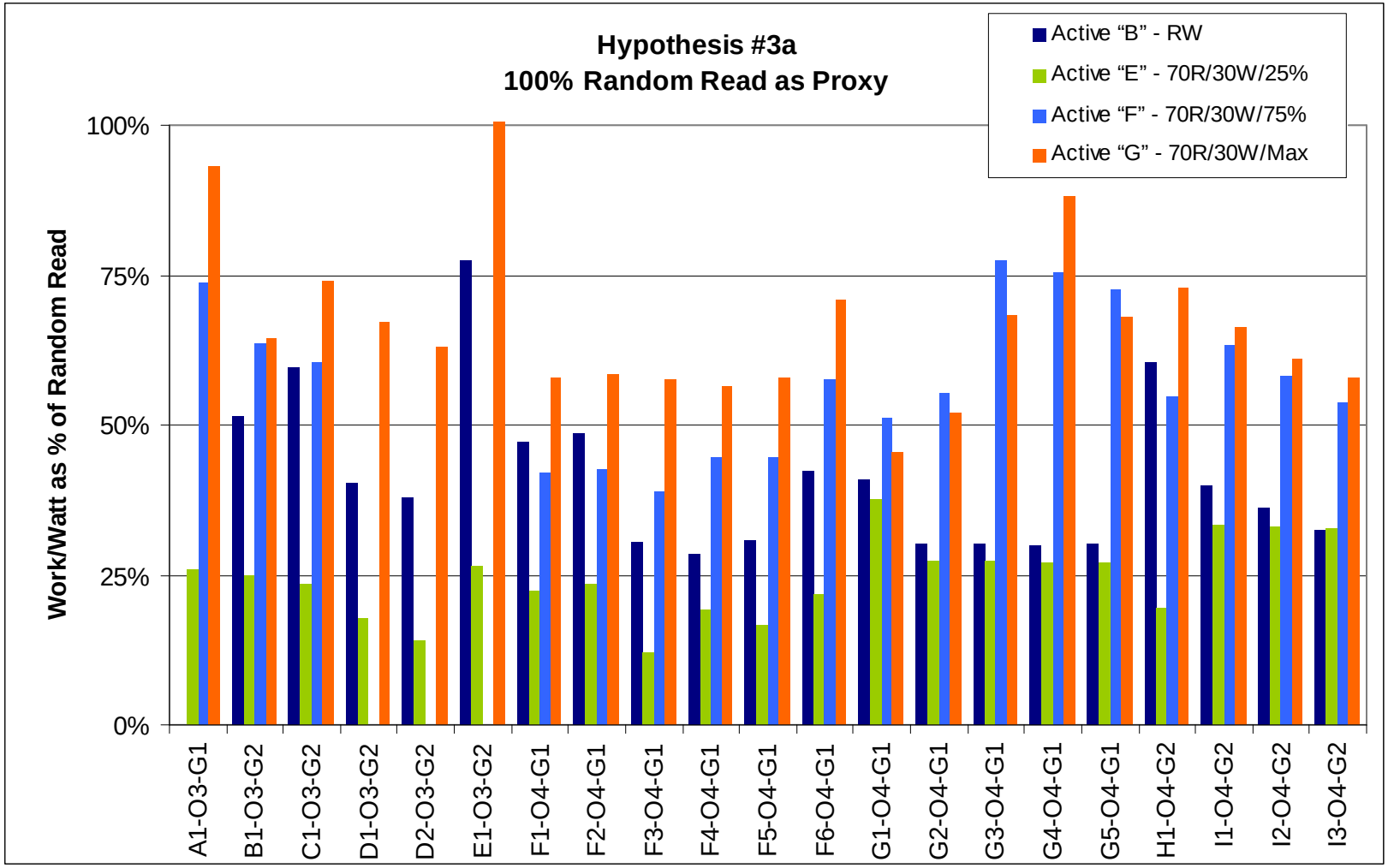


Test Procedure Assessment

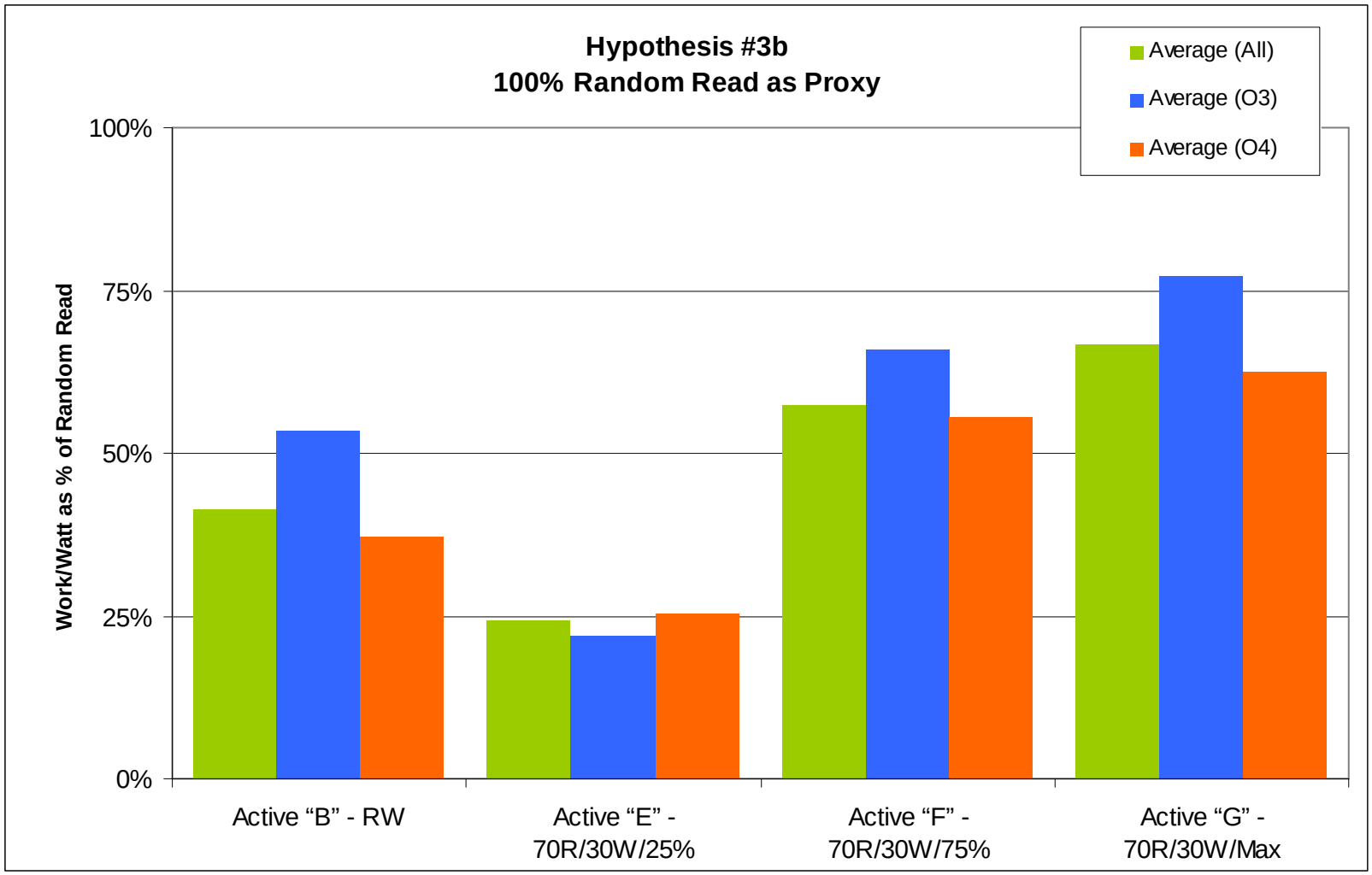
Random Read as Proxy

- In most cases, Online 3 systems are more “efficient” in random write and 70R/30W workloads than Online 4 systems.
- Appears that the 70/30 tests run systems into performance limits:
  - 25% test results in approximately 25% of the work/watt, but
  - 75% and 100% tests do not result in a linear increase in performance.

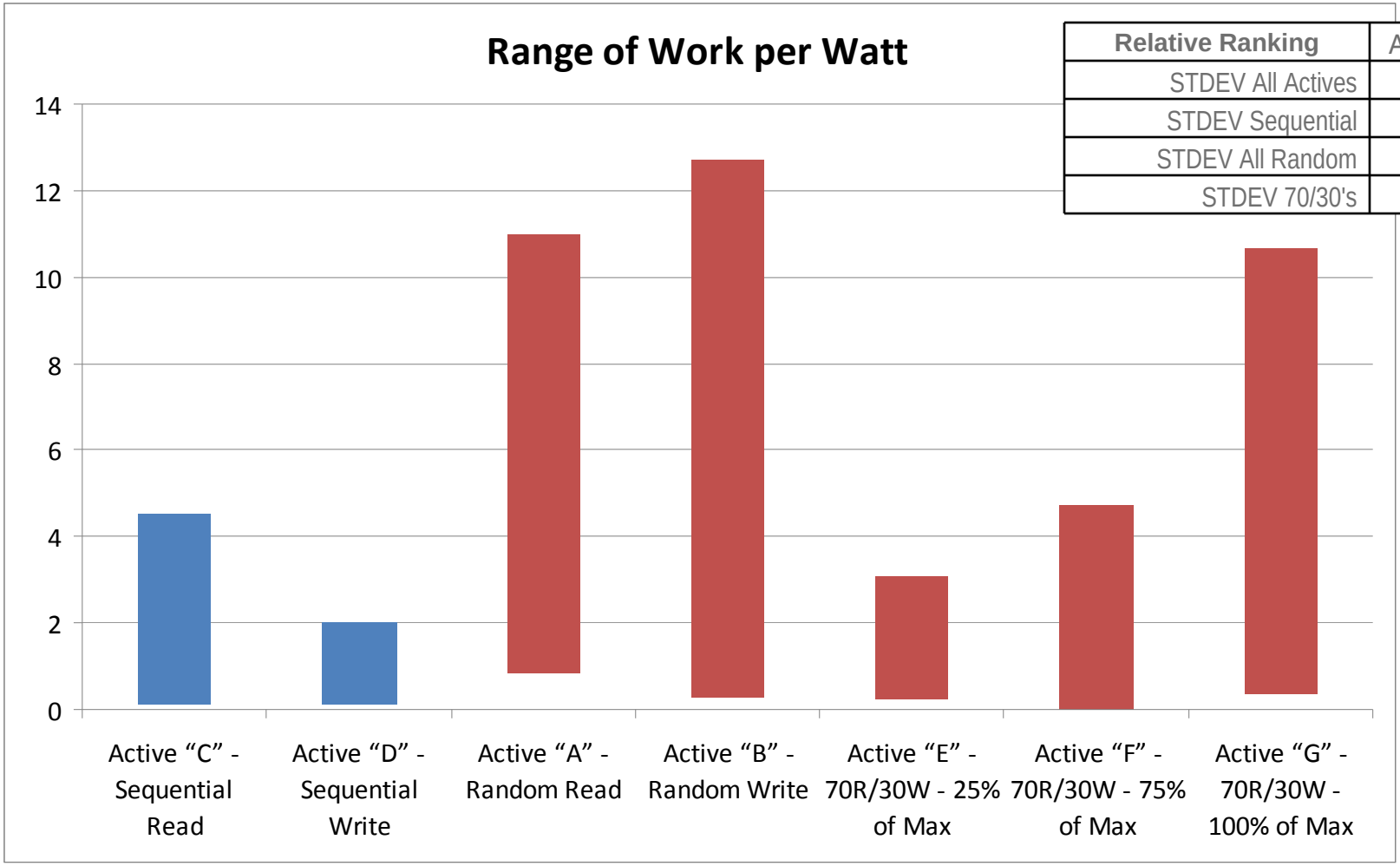
Random Read as Proxy



Random Read as Proxy



Value of Individual Test Phases



Relative Ranking	Average SD
STDEV All Actives	12
STDEV Sequential	12
STDEV All Random	4
STDEV 70/30's	2

Value of Test Phases

- Observations
  - Individual tests highlighted strengths of different systems. Standard deviation of rankings from 2 to 11.
- Conversation
  - Is each test applicable to specific end use cases?
  - Does each test or combination of tests provide valuable insight for end user?
- Conclusions / Next Steps
  - Additional investigation of a smaller set of proxy tests

Other Topics

- Server Based storage
  - Storage which is able to execute productivity applications on the storage controller
  - How accommodate for additional energy usage of server portion of storage
- Hybrid storage
  - How to showcase benefit of hybrid storage during active evaluation

Next Steps: Testing

Supplemental Data Collection

- Data set will be posted to an online collaboration forum for continued analysis by stakeholders. Email storage@energystar.gov for access.
- Further guidance for additional data collection to be provided. Will continue to focus on single-variable changes and on rounding out the data set. For consideration:
  - Online - Supplement Existing: 7200 RPM to 15000 RPM drives, Full HDD to Hybrid HDD/SSD, RAID-5 to RAID-50 configuration, 50% to 100% & 75% to 100% loaded disk shelf, 2X & 4X number of disk shelves, 2X number of controllers, 2X controller cache
  - Archive - Supplement Existing: N to N+1 to N+2 drives, Single to Redundant PSU, 2X & 4X capacity
  - File I/O? Additional hybrid systems? Additional Near-online and VTL data points? Others?

Draft 1 Specification Overview

Industry Collaboration

- Received feedback from individual organizations and larger industry groups such as the Storage Networking Industry Association (SNIA) and The Green Grid
- The sum of the contributions has been more valuable than the individual parts
- EPA continues to welcome feedback from all interested parties
- Goal is to encourage dialog about the specification as it pertains to a vendor- and organization-agnostic program

Draft 1: Goals

- Expose energy performance data for purchasers and operators. The act of reporting in a consistent format will benefit the industry
- Develop a system-level efficiency approach, with minimal component or subsystem requirements (PSUs, etc.)
- Develop a model for quantifying & recognizing the efficiency benefits of various software implementations

Draft 1: Partner Commitments

- Partner Commitments will be expanded in the final specification to include new testing requirements and program changes per the ENERGY STAR Enhanced Program Plan. There is a separate stakeholder process for these changes
- Please visit www.energystar.gov/mou for additional information

Draft 1: Product Family

- The concept of “product family” is included in Draft 1
- The family concept may be well suited to storage due to the high degree of customization and configurability of products
- EPA may or may not choose to define product families based once the initial data collection process and data analysis is complete. These definitions will be developed or deleted, accordingly

Draft 1: Taxonomy

- Objective: Define requirements consistent across taxonomy categories, in order to simplify
- Find appropriate solutions (both test procedures and specification criteria) for “hybrid” systems
- Stakeholders noted risks of qualifying products that use 3rd-party components
  - Specifying requirements at the system level allows maximum flexibility to source components and subsystems that will allow their products to earn the ENERGY STAR
  - Primary exception is for PSUs, which represent a simple and effective path to improved efficiency across a product’s full range of operation

Draft 1: PSUs

- Because storage PSUs are often installed in redundant configurations, EPA intends to focus its efforts on encouraging;
  - the use of PSUs that are most efficient at low loads (<40%) for use in redundant active/active configurations,
  - right-sizing of PSUs to application requirements, and
  - novel approaches to redundant PSU installations to allow for greater overall system efficiencies
- Require that all storage PSUs meet ENERGY STAR qualification criteria
- It may be necessary to define unique efficiency requirements for various PSU types, especially given the long lifetimes and long refresh cycles for PSUs
- Continue to track Climate Savers and 80 PLUS® efforts

Draft 1: Other Terms

- Many definitions have been aligned with the Fall 2009 edition of the SNIA Dictionary
- Definitions for “I/O Adapter” and “I/O Port” have been removed from Draft 1 since there are no direct references
- Several other terms have been suggested but have not yet been added, including: “Storage Protection,” “Small Computer System Interface (SCSI),” “Count Key Data (CKD),” and “Fixed Block Architecture (FBA)”

Draft 1: Definitions

***Storage Product**: A fully-functional storage system that supplies data storage services to clients and devices attached directly or through a network. Components and subsystems that are an integral part of the storage product architecture (e.g., to provide internal communications between controllers and disks) are considered to be part of the storage product. In contrast, components that are normally associated with a storage environment at the data center level (e.g., devices required for operation of an external SAN) are not considered to be part of the storage product. A storage product composed of integrated storage controllers, storage media, embedded network elements, software, and other devices. For purposes of this specification, a storage product is a unique configuration of one or more SKUs prepared for sale to an end user.*

- A concise and unambiguous definition for “Storage Product” is critical. This definition is a work in progress
- It is the storage product that will ultimately be subject to ENERGY STAR qualification. Subsystems and components will not be eligible for qualification

Draft 1: Software

- Software is an important contributor to system energy efficiency
- EPA supports the efforts of the SNIA Capacity Optimization Subgroup (COS) to develop a single “data storage efficiency” metric
- Intend to evaluate software metrics for inclusion either as reporting or qualification requirements

Draft 1: Active State

***Active State**: The state in which a storage product is processing external I/O requests.*

- SNIA defines Active State as, *"An operational state in which a Storage Product can consistently maintain a level of service implied by its taxonomy rating. This generally means that most or all of its storage elements are fully-powered and active, and that background data cleansing and other operations are as active as I/O load allows."*
- EPA proposed a simplified definition for this specification

Draft 1: Idle State

***Idle State**: An operational state in which the Storage Product is capable of completing I/O transactions, but no active I/Os are requested or pending. The system may, however, be servicing self-initiated I/Os from background data protection and cleansing, and other operations.*

> ***Ready Idle**: The state in which a storage product is able to respond to I/O requests within the MaxTTD limits for its taxonomy category, but is not receiving external I/O requests. The storage product may perform routine housekeeping tasks during Ready Idle, provided such operations do not compromise the product's ability to meet MaxTTD requirements.*
>
> ***Deep Idle**: A state in which one or more storage product components or subsystems have been placed into a low-power state for purpose of conserving energy. A storage product in Deep Idle may not be able to respond to I/O requests within the MaxTTD limits for its taxonomy category, and may need to perform a managed 'wake-up' function in order to return to a Ready Idle or Active state. Deep Idle capability must be a user-selected, optional feature of the Storage Product.*

- A definition for "Hardware Idle" was proposed but is not included at this time, since EPA intends to focus on energy efficiency of systems in active use

Draft 1: Efficiency Metrics

- Active state efficiency requirements will be developed in subsequent draft specifications
  - Goal is to develop a simple, easy-to-understand energy performance metric for data center storage products
  - Data analysis will ideally result in one or more metrics (such as GB/Watt or IOPS/Watt) applicable across all storage products within a taxonomy category
- EPA understands that true idle conditions occur only infrequently in most storage applications
  - Data analysis will try to determine whether idle state energy efficiency can be a suitable proxy for active state energy efficiency

Draft 1: Reporting

- Power & Performance Data Sheet
  - Review the latest proposed Servers data sheet and send comments regarding data fields that should be included in a PPDS for data center storage products
  - EPA plans to distribute a first draft of the storage PPDS for review along with the Draft 2 specification
- EPA would like to assess the relative costs and benefits of power, temperature, and other data reporting at the rack, shelf, or component (PSU, disk) level

Next Steps: Program

Development Schedule

- *6/4/09 Framework distributed*
- *7/20/09 Stakeholder meeting (San Jose)*
- *10/15/09 Test Procedure Workshop (Phoenix)*
- *12/28/09 Start 1st round data collection*
- *2/2/10 Stakeholder meeting (San Jose)*
- *3/1/10 Complete 1st round data collection*
- *4/9/10 Draft 1 distribution*
- *4/15/10 Stakeholder meeting (Orlando)*
- *5/21/10 Draft 1 comments due to EPA*
- *May-Jun Supplemental data collection*

Acknowledgments

EPA would like to thank the Storage Networking Industry Association (SNIA) for hosting today's event

Contact Information

- *Andrew Fanara (US EPA)*
  - fanara.andrew@epa.gov // 206.553.6377
- *Steve Pantano (ICF)*
  - spantano@icfi.com // 202.862.1551
- *Al Thomason (TBWC)*
  - thomasonw@gmail.com

More Info:
http://www.energystar.gov/NewSpecs