As the world's leading developer and provider of information infrastructure technologies, services, and solutions that enable people and organizations to transform the way they create value from their information. EMC Corporation appreciates the opportunity to respond to the ENERGY STAR Enterprise Storage Draft Specification Framework, issued June 4, 2009. The development of an appropriate specification is important to EMC, and we are committed to active participation in the process.

EMC's response is structured to reflect the overall format of the original framework document, with extended or detailed discussion appearing in addenda to improve readability. We have provided comments on all Building Block sections, and have omitted questions or subsections on which we had no recommendation or feedback.

Overview

There is a broad range of storage products intended for use in enterprise data centers, and the industry generally categorizes these systems using a hierarchical or tiered description that differentiates Storage Products by performance and the criticality of the data they contain. The term "Enterprise Storage" is typically used in the industry to refer to products in the two highest performing tiers, which correspond roughly to the Online 4 and Online 5 categories of SNIA's taxonomy. If EPA is interested in having the specification address the range of Storage Products present in data centers, the term that more accurately reflects the intended range of products that would be in scope would be "Data Center Storage". This term has the advantage of also meeting the industry definition of these types of products. (Note that the term "tier" is used to refer to the hierarchical layer of the storage, not a version number.)

While the storage industry continues to explore system performance characterization, no attempts thus far to provide a single workload benchmark that reflects all, or even typical, end-user applications have proven to reflect the wide variety of workloads applications present to storage products. The breadth of connectivity, capacity and performance scalability, the wide range of possible configurations, and the diversity of architectures and even connectivity methods (IP, Fibre Channel, etc.) further complicate the prospect of a single metric identifying the most energy efficient solution.

A further characteristic of most Storage Products found in enterprise data centers is that they standardly include a high level of redundancy given the enterprise need for timely and reliable access to its data. Further, since a single data center storage system may be shared by several different application servers, redundancy is also present in systems that are not handling mission-critical data so that a component failure will not interrupt the operation of multiple applications. These systems regularly include n+n power and redundant cooling to allow for component failures – some even include integrated backup power, while others leverage only external UPS to ensure continued availability through a power disruption. This can be true regardless of storage tier. Many also include logic to allow the unit to continue to service requests even if a controller fails. These factors alter the baseline energy profiles of Storage Products used in enterprise data centers, and are present because the target customer requires this level of storage availability.

Building Block #1: Definitions

Many of the definitions contained in this section do not reflect generally accepted storage industry definitions, while some other terms are not in common use in the industry. Addendum #1 provides the industry definitions or alternative terms that have industry acceptance.

With respect to the questions for discussion:

1. How are Active, Idle, Maximum and Full Load states defined in the industry?
   It is common practice in the industry for background services (e.g. maintenance functions, reliability-enhancing functions and capacity-enhancing functions) to be scheduled to run when client-initiated operations have ceased. This is not actually an idle system – the functions being performed are desired and expected by the customer, who does not want them to interfere with his production use of the Storage Product; in other words, these background services are seen as necessary work by the purchaser, and the system is not seen to be in an inactive state.
   A storage product is considered to be Active when it is servicing client-initiated requests, but is actually performing work in all states.

   The terms Maximum and Full Load are not used as such in the industry. The characteristics of Storage Products result in different maxima for each type of I/O operations (reads/writes, random/sequential) qualified by some performance or size characteristic, and so there is never a discussion of a single Maximum operating state. Similarly, the concept of Full Load is highly ambiguous, and can refer to one of many situations where system performance is bottlenecked in some dimension such as throughput, port load or response time, but this is not a specific condition or state. Occasionally, Full Load can refer to the I/O load at which a Storage Product demonstrates peak power use (see further discussion in Building Block #3, point 4).

2. Storage Products are defined to be Idle from a customer demand perspective according to the definition supplied in the framework. This is not a truly idle state as defined in #3 below.

3. Storage Products engage in a number of background services during the times when there is no client-generated I/O load. These services include maintenance functions performed by both the operating software in the product and the individual hardware components in the product, reliability enhancement functions performed by the operating software, and capacity enhancement functions performed by the operating software. These functions are necessary and expected by customers, and they are specifically expected to be performed in such a way that they do not impact the performance of the customers' applications, thus they are performed when customers consider the system to be "idle". EMC would recommend that either an additional state be defined that recognizes this background operation and differentiates it from "Idle" as defined in the framework, or that the "hardware idle state discussed in Addendum 1 be used to resolve the confusion about Storage Product state.

5. The SNIA taxonomy and glossary should continue to be a reference for Storage Product definitions.

6. The power supply definitions for servers are not completely appropriate for Storage Products, due to the differences between the products. Cooling in storage systems is generally not supplied in the rack, but is integrated with most PSUs, leading to a need to change the way measurement occurs. Data center products also generally contain redundant PSUs to ensure data availability since the storage is shared across multiple servers. Some PSUs also include integrated battery/standby power supplies – these require power to recharge and generate heat and consume power even when idle. Finally, most products support capacity scaling after installation, which results in a need to supply power to disks as they are added. The optimum number of incremental disks will vary as a function of the underlying architecture of each product. This means that any Storage Product contains multiple PSUs driven by a combination of base and capacity requirements. These requirements on the power system are not analogous to those seen for data center servers.

7. EMC does not believe that DC-powered storage systems are prevalent in most geographies.

Building Block #2: Eligible Product Categories

The Storage Industry uses multiple approaches to categorizing products. One is by access time, capacity, and availability, similar to the SNIA Taxonomy, and many EMC customers would agree that this is a useful approach. The industry also categorizes storage according to a hierarchy based on the criticality of data to an enterprise and the business purposes to which this data is put. Most EMC customers typically employ this methodology in defining their approach to matching storage products to business needs. EMC would welcome the opportunity to broker conversations with some of our customers to allow the EPA to understand these views first-hand. A discussion of this hierarchical approach is provided in Addendum #2. Finally, interface access methods to storage products (such as Fibre Channel, iSCSI, etc.) can have a significant impact on how they are architected, on the internal software that operates them and in their operation in a data center.

These differences can make comparisons across categories of Storage Products difficult, regardless of the categorization method used. In addition, due to the fact that products in one category generally do not substitute for those in another, customers may not find cross-category comparisons will greatly influence their buying decisions.

Building Block #3: Energy Efficiency Criteria and Test Procedures

As previously discussed, Storage Product power supplies can differ from those found in servers (see Building Block #1, point 6). As a result, it is recommended that fan power be excluded from measurements made of both Multi-Output and Single-Output supplies. It is further recommended that, since most data centers supply 230V to their equipment, both Multi- and Single-Output supplies be tested at 230V only.

In reference to the Questions for Discussion:

1. Given the earlier discussions of Idle (Building Block #1, point 3) and Maximum (Building Block #1, point 1), EMC recommends the use of Hardware Idle (Addendum #1, point e.1) and Peak Power as defined in point 4 of this section as the operational state in which Storage Product energy consumption should be measured. These measurement points are architecture and access interface neutral and allow meaningful comparisons between products.

4. There is no single recognized workload benchmark for all ranges of customer workloads, configurations or all categories of Storage Products. As a result, the existing benchmarks are either reflective of specific workloads or are designed to advantage specific architectures or interface categories. To avoid these biases, EMC would recommend that active power measurement be conducted by measuring the peak (maximum) power usage of each product while in an active state. Partners would have to identify and describe the workload used.

5. The SPEC sfs 2008 benchmark is not applicable to Storage Products other than NAS systems. In addition, the benchmark favors some system designs and configurations over others. In particular, some of the best-performing configurations may be amongst the least energy-efficient, resulting in the benchmark driving the wrong behavior for ENERGY STAR.

6. A significant fraction of Storage Arrays are made up of many assemblies which utilize their own redundant power supplies. As a result, a method will need to be developed for combining these assemblies to present an aggregate power measurement back to the data center user. In addition, some products integrate Network Switches, Hubs, “off the shelf” UPS modules or standby power, or other low power assemblies into the system. Since there are no ENERGY STAR requirements in place today for these products, it is recommended that they should be excluded from the power supply efficiency and thermal monitoring requirements identified in this framework if their total contribution to the system load is less than 10% of the total, and the total power consumption of the system exceeds 1 kW.

7. Given the need for highly available Storage Products, power supplies for these systems are designed to be run in redundant configurations. As a result, they typically run with loads ranging from 20% to 45%, unless one of the supplies in a system has failed.

8. There is very little experience with the Net Power Loss methodology for assessing power supply efficiency in Storage Products. For this reason, it is preferable to use the proposed benchtest methods with appropriate compensation for system redundancies.

9. The VAR sales channel has a significant role in the Data Center Storage market. Without an agreed definition of product families, it is difficult to determine the impact of the effort and expense represented by the requirements on the third-party sales channels.

Building Block #4: Information and Management Requirements

Responses to Questions for Discussion:

1. Storage Products found on the market today typically do not include the ability to measure or report input power measurement or inlet temperature data identified in the framework, unlike the situation found in servers. Incorporation of these capabilities will require changes to system design and system software to conform to the specification, at an added cost to the consumer. Typical product cycles in the storage industry are on the order of 18-24 months, which would result in very few qualifying products being available at the specification's effective date.

   Measuring and reporting capacity utilization is typically available on demand via the management interface(s) to a Storage Product. It is not, however, typical to provide performance utilization information as part of default product capabilities, and it may not be a supported feature on some products. Inclusion of these unsupported or optional features as part of the required feature set for ENERGY STAR labeled products will require changes to both product content and business practices, which should be considered in the decision-making process.

2. A key element of energy consumption in Storage Products is driven by the type (Enterprise Flash Drive, Fibre Channel, SATA, e.g), number and spindle speed (7k, 10k, 15k rpm) of disks included in any given configuration. This means that it is not sufficient to specify the capacity at which measurements were taken, but the number, capacity, spindle speed and type of each drive class used to achieve the measurement must be listed.

   A further important element that should be considered for inclusion in the data reported for Storage systems would be the enumeration of the types of background services that execute during the "idle" state of the system. This will provide potential customers of systems under consideration the ability to better understand the probable source of differences between systems. It is these software differences that lead us to recommend the use of the "hardware idle" state for measurements of energy consumed by "idle systems", providing a basis of comparison between hardware designs.

3. Capacity utilization is routinely discussed in the industry as a fraction of either Raw Capacity or Consumable Capacity as defined in Addendum #1. Performance capacity is measured in a variety of ways depending on the category of Storage Product and its architecture and configuration. There is relatively little commonality to the way this class of data is reported. Purchasing decisions, either of capacity expansion or additional Storage Product systems, are driven as a result of specific understanding of the application environment supported by the storage, the expected capacity

growth that will be experienced when compared to the actual consumption experienced, the performance each application is experiencing (as measured by the application rather than the storage system), and the backup, archiving and data reduction policies of the business. Within any given data center the criteria typically vary between application environments, and require information beyond that which the Storage Product can provide.

4. At this point in time, there are no industry-standard protocols for measuring temperature, utilization, etc.

Addendum #1: Definitions

a. Storage Hardware
    a. Storage Media: For consistency, this should refer to tape media, not tape drives, and optical media, not optical drives.
    b. Storage Product: The final sentence of this definition is inconsistent with the apparent scope of this specification, and should be stricken. We would further recommend that the definition be amended to read “"A system composed of integrated storage controllers, storage media, optional embedded network elements, and software that provides data storage services to one or more Computer Servers and/or other devices.”
    c. Storage Controller: We would recommend using the SNIA definition of an intelligent controller.
    d. Storage Product Family: Given the way products are structured and configuration options vary with respect to interface type (Fibre Channel, iSCSI, etc.) and speed (1GB/s, 4 Gb/s, etc.) as well as capacity options (Fibre Channel, SATA, EFD, etc.) and granularity this definition may require refinement.
    8. I/O device: We would recommend adding the following to this definition. “I/O devices do not initiate user I/O; they are used for the purpose of transmitting data and/or management information.”
b. Storage Characteristics
    1. Capacity: The definition offered is what is known in the industry as “Raw Capacity”. The industry and users distinguish between Raw Capacity and Consumable Capacity, which is the storage capacity available after formatting, system overhead, RAID parity and sparing has been deducted. We would recommend the use of these terms to provide clarity in measurement methodologies.
c. Other Data Center Hardware:
    a. Network Equipment: We would suggest the first sentence be modified to read “"A product whose primary function is to provide data connectivity among *an arbitrary combination of* devices connected to its several ports." to differentiate data center network equipment from embedded switch components.

d. Power Supplies:
   1. Given the differences between server and storage power supplies, it would be more appropriate to define the secondary outputs as a percentage of the primary rather than in absolute terms.

e. Operational States:
   1. Idle: While this does represent a valid state from the perspective of client-initiated workload, Storage Products are generally not in an unproductive state at this time. The only time in which Storage Products are not active is when they achieve the state of "hardware idle", i.e. when all background services have been completed and the system is exclusively in a wait state.

Addendum #2: Eligible Product Categories – Hierarchical Storage

Energy efficient storage practices depend on more than energy efficient storage products. The use of the storage products and practices designed for specific tiers in the information hierarchy provide data center operators with the ability to conserve overall energy by moving information from high-performance (and more energy consumptive) products to lower performing (and less consumptive) products as that information progresses through its lifecycle. This results in energy savings in two ways: first, it frees up capacity on high-performing storage, eliminating the need to bring more of it online, avoiding increasing energy spend, and second, migration to the proper tier allows greater use of energy-saving techniques such as removable media or disk drive spin-down.

The hierarchy of storage tiers reflects the criticality of information and its availability to a business, and is generally defined as:

- Tiers 0,1 – Data access is required in <80 ms, and must be continuously available, with access interruptions lasting less than minutes or seconds annually. Business applications using this data demonstrate dynamic workloads with the highest transaction loads of any applications in an enterprise. If a service interruption occurs, the business has a goal of zero to seconds of data loss for this data. Architectures in these tiers have no single points of failure (SPOFs) and provide for non-disruptive serviceability. Customer configurations in these tiers frequently include the use of 15K rpm disk drives, EFDs (Enterprise Flash Drives), and high-speed front-end connectivity modules (e.g. 4 Gb/s Fibre Channel, 1Gb/s iSCSI), contributing to the systems' energy profiles.
- Tier 2 – Data access is required in <80 ms, and must be very highly available, with access interruptions lasting less than hours or minutes annually. Business applications using this data demonstrate constant, heavy loads with high performance demands. If a service interruption occurs, the business has a goal of seconds to minutes of data loss for this data. Architectures in this tier usually have no SPOFs and generally provide for non-disruptive serviceability. Many of the same configuration options found in Tiers 0 and 1 also appear in Tier 2 configurations and affect these system's energy profiles.

- Tier 3 – Data access is required in <80 ms, and has an availability target that supports access interruptions lasting up to a few hours annually. Business applications using this data demonstrate primarily read access to information stored in this tier with moderate performance requirements. If a service interruption occurs, the business can tolerate the loss several minutes of data stored in this tier. Architectures in this tier may have SPOFs or require disruptive service.
- Tier 4 – Data access can take >80 ms, and has an availability target that supports access interruptions lasting up to a few hours annually. Business applications using this data demonstrate primarily read access to information stored in this tier, with moderate performance requirements, and may tolerate “internet-like” times to first access. If a service interruption occurs, the business can tolerate the loss of up to the last 24 hours data.