July 10, 2009

Dear EPA,

Please find enclosed Intel comments and feedback on the Specification Framework Document for the Version 1.0 Energy Star for Enterprise Storage Specification, dated June 4, 2009.

Intel remains committed and supportive of the US EPA's efforts to define energy efficiency goals and targets across the spectrum of computer products including the development of Energy Star for Enterprise Storage, v1.0. We hope our input and work with industry stakeholders continue to be beneficial in development specification. Intel is actively involved with storage industry stakeholders on determining evaluation methods of energy efficient performance across the multiple tiers* of data storage architectures. Please feel free to contact us if there are aspects where we may be able to improve these interactions.

We continue to work extensively with our industry colleagues in Storage Network Information Association (SNIA), The Green Grid (TGG), Climate Savers Computing Initiative (CSCI), IT Information Council (ITI), and the Alliance for Telecommunications Industry Solutions (ATIS), in addition to supporting the Energy Star for Enterprise Storage program to deliver increasing energy efficiency.

If you have any questions please feel free to contact myself or Henry L Wong, henry.l.wong@intel.com.

Sincerely,

Lorie Wigle
General Manager
Eco-Technology Program Office

* Tier-ed Data Storage Hierarchy, see Figure 1 attached.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Intel appreciates the opportunity to provide the EPA with the comments and recommendations for the framework of the Energy Star Program Requirements for Enterprise Storage v1.0 specification. We hope the EPA can take these items under consideration before issuing the initial draft of the specification.

Summary

The Definition section will provide end users helpful system identification with inclusion of storage system Tier architecture definitions. Additionally, Intel has supplied clarity for many of the other terms in the Definition section.

Operational States: The Operational States are not defined in terms that align with current storage system architectures. Tier 1 storage systems are never truly "idle" (i.e. inactive). Most of Tier 2 systems are very seldom idle due to the 'background' or storage management functions these systems provide. Tier 3 storage systems can clearly be classified with idle and active states. Intel recommends addressing these 'state' definitions for the initial draft of the specification to allow storage vendors to better focus resources on energy efficient performance metric(s) in Building Blocks #3 and #4 sections.

Energy Efficiency Improvements: Intel has provided several recommendations which can significantly impact energy consumption for storage in data centers. These are listed in the Energy Efficiency Criteria and Test Procedures question section. The product and procedural changes recommended include: utilizing larger capacity disks, massive arrays of idle disk (MAID) systems, thin provisioning, utilizing data center storage tiers and data de-duplication.

Power Measurements in Different Operational States: Storage systems perform a diverse set of tasks, from caching to indexing, to read and writes, to storage volume and system management. Intel recommends evaluating the storage power consumption by measuring watts per usable terabyte.

Information and Management Requirements: Storage system management is very diverse amongst vendors. Many functions are provided via proprietary solutions, others are standards based. Intel recommends contacting the Storage Vendors for standards identification and classifications of the non-proprietary means for system thermal and power management of data access and storage systems.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Per Section Response

For reference, this response provides recommended copy to provide clarity that businesses and individuals will commonly understand in relation to storage architecture and products. Recommended copy is provided on the labeled sections of the Framework document issued June 4, 2009.

Building Block #1: Definitions

c. Preliminary List of Definitions

c.0 Enterprise Storage:

Data has many purposes within a business, and these purposes change over time. Availability, performance and value of data changes as the data gets older. As such data is stored on different types of systems and in different types of tiers or arrangements within a data center. Most commercial data can be classified into:

- **Mission-critical:** Most valuable to an enterprise such as billing and manufacturing systems, needs high performance, high availability storage with near zero downtime.
- **Business-critical data:** important to the enterprise such as inventory systems, needs reasonable performance, good availability storage with monthly downtime no more than a few minutes.
- **Accessible online data:** typically work process or data held for compliance of regulations, such as research and email. This data requires low access, high availability storage with monthly downtime no more than 1-2 hours.
- **Nearline data:** large volumes of low accessed data, this may include test and development records, in addition to customer/supplier records. Automated storage retrieval with less than one hour access time.
- **Offline data:** archived data, backup or compliance information. Used in remote locations for disaster recovery. Storage with 72 hour seek time is sufficient.

Data integrity, accuracy and security are key activity drivers for these storage systems. IT professionals quickly realize the performance requirements of their storage systems are beyond simple capacity given the

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

number of changes, the coalescing of updated information, and the multiple access points that require security. In addition, data requirements include compliance to regulations and standards such as Sarbanes-Oxley, medical record retention, state and federal legal data retention requirements, and financial data retention to name a few. Many of these functions and efficiency methods in handling such information occur without direct end user interface, and rarely gets recognized as the useful work of these systems.

Functions of a storage system are designed to support these business-critical access times, while reducing the risk of lost data, reducing incremental labor time and operational effort to transfer data in addition to minimizing the impact to host machines in moving and migrating data. Storage systems need to be designed and provisioned to support the lifecycle of data, applications and the business activities of that operation or data center.

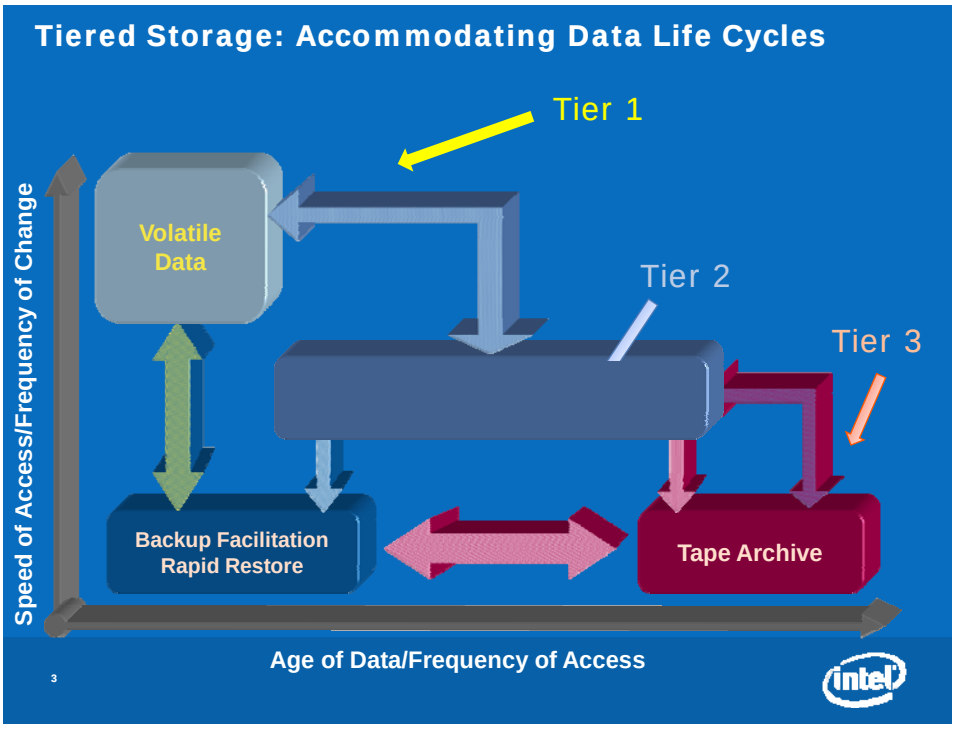
Some of the 'background' storage functions while there are no data requests or changes (i.e. 'idle') include:

- Disk management: provisioning disks, de-fragmentation, spreading data across multiple disks for performance and protection
- Data protection services, which includes: encryption, backup, data monitoring, including multiple copies (snap copies, volume copies, remote asynchronous/synchronous copies) for different access or locations and snap copy management, and finally data recovery.
- Storage pool management: data monitoring for lifecycle aging, presentation of different LUN sizes for better mapping of storage resources for optimization of utilization of allocated storage resources.

Energy efficient storage in a data center is divided into 'tiers'. The business or organization's requirement for speed of access to the data and the frequency of change of the data determines data placement within the storage tiers. Tier-ed storage is not simply storing data on various storage systems with different system architectures. Multi-tiered storage is constructed of a single-image pool of storage, with integration by segmented storage tiers, all controlled by a unified data storage architecture. Data is promoted or demoted seamlessly through the tiers. Tier-ed data storage hierarchy provides structure to safeguard data, make it available for transport as well as securely saving data for government regulations and

archiving; all with the most energy efficiency. The diagram below illustrates a tier-ed data storage hierarchy in a data center.

Figure 1:Tier-ed Data Storage Hierarchy within a data center



Tier 1 Storage systems are used for dynamic data, frequent changes and typically with high business value for the data (e.g. any downtime would be very costly to the business). Tier 1 storage systems are never idle. Tier 2 storage systems are used for frequently used data that can tolerate a few seconds of downtime. Tier 2 systems are very rarely idle. Tier 3 storage systems are used for archiving data; these systems can tolerate minutes of downtime without severely affecting the business. A Tier 3 storage system is primarily used for archival of rarely accessed information. These types of systems may be comprised of a robotic mechanism which will mount (insert) and dismount removable mass storage media into a storage device according to the system's demands (or their logical equivalents). When a Tier 3 storage system becomes idle, it will dismount the volumes of the mass storage media. Tier 3 storage examples include tape libraries and optical jukeboxes.

The storage systems within this EPA Storage framework encompass systems in Tiers 1 and 2. Note that Tier 1 storage systems should not have an idle requirement, since that level of inactivity could jeopardize the activity

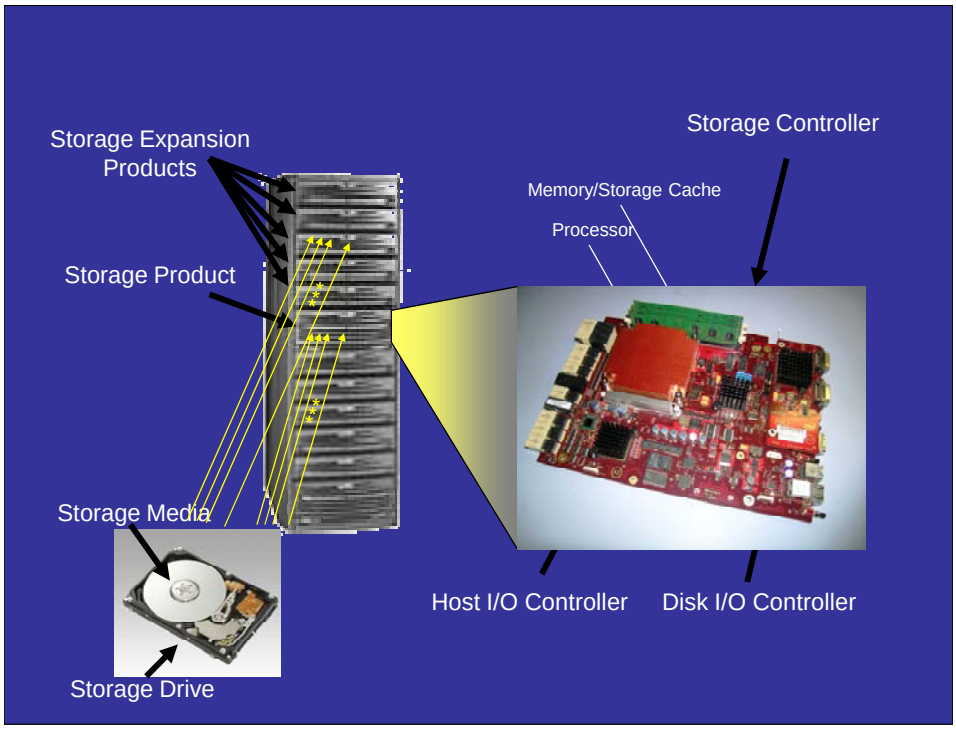
designed for this tier of storage. To better describe the devices and functions, we recommend using the following changes to the terms and definitions in building block 1 of the framework document.

1. **Storage Hardware/Device**: A collective term for disks, tapes, disk arrays, tape arrays, and any other mechanisms capable of non-volatile data storage. Any device designed and built primarily for the purpose of persistent data storage and delivery. This definition includes I/O processing with disk drives, tape drives, RAID array subsystems, robotic tape libraries, filers, file servers, and any other types of storage devices. (Source: SNIA Dictionary v.2008.1.ENG, published 2008)
2. **Storage Media:** The physical material that stores digital information/data. Storage Media used in Enterprise Storage applications may be silicon based (e.g. solid state drive), magnetic (hard disk platter, tape), or optical (optical disc). Solid state disks (SSDs) are a combination of silicon chips where the storage capability is provided by solid-state, non-volatile random access memory.
3. **Storage Drive:** A subsystem consisting of storage media and associated control mechanisms that supports access to storage media via an I/O interface.
4. **Storage Product:** *Recommend this definition be removed, as it is duplicate with Storage Hardware/Device and may be confusing with standard storage definitions.*
5. **Storage Controller:** An intelligent controller to which storage devices are attached. Provides I/O interface to Computer Servers, interface to storage device with storage media, and processing elements the execute software which manages the storage device's and services I/O requests from the servers. (Source: SNIA Dictionary v.2008.1.ENG, published 2008)
6. **Blade System:** A blade server is a server chassis housing multiple thin, modular electronic circuit boards, known as server blades. Either self-standing or rack mounted, the chassis provides the power supply, and each blade has its own CPU, memory and hard disk. Redundant power supplies may be an option. Each blade typically comes with one or two local SATA or SAS drives. For additional storage, blade servers can connect to a storage pool facilitated by a network-attached storage (NAS), Fiber Channel, or iSCSI storage-area network (SAN). In this definition, a Blade System contains its own storage or is

connected to storage systems, versus a Storage Blade which is a high density networked storage architecture.

7. **Blade Chassis:** *(recommend moving this from this section to 'c. Other Data Center Hardware' section as this is not about specific storage definitions.*
8. ~~*Blade Storage*~~ *(reverse order- should read)* **Storage Blade:** A storage-specific element that relies on shared resources (e.g., power supply, cooling, etc.) for operation. Blade Storage units are designed to be installed in a Blade Chassis, are hot-swappable and are incapable of operating independent of the chassis.
9. **I/O ~~Device~~ Subsystem:** A collective term for the set of devices and software components that operate together to provide data transfer services. A storage subsystem is one type of I/O subsystem. Many I/O subsystems are characterized by the types of ports they support, such as Ethernet, Fibre Channel, InfiniBand or SAS.

Figure 2: Illustration of key parts within a storage system



Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

b. Storage Characteristics:

1. **Capacity: ‘**Raw Capacity**’**-The sum of the raw unformatted, uncompressed
   capacity of all Storage Media installed in a Storage System. ‘Storage capacity’- is the ‘user consumable’ portion of the storage media, which does not include space required to enable storage on the media, such as the file system and tables, storage management functions, formatting, RAID parity and sparing.
2. **Direct-connected:** Direct-attached storage (DAS) is a dedicated storage device that is directly attached to one server or more servers. Acronym DAS. (Source: SNIA Dictionary v.2008.1.ENG, published 2008)
3. **~~Network-connected~~:** *This term should be Network Attached Storage*
4. **Network-Attached Storage (NAS):** NAS stands for Network Attached Storage. NAS allows users to access file storage across an Ethernet network, using TCP/IP. A host system that uses network attached storage uses a file system device driver to access data using file access protocols such as NFS or CIFS.
5. **Storage Area Network (SAN)**: A storage system consisting of storage elements, storage devices, computer systems, and/or appliances, plus all control software, communicating over a network. Using the SCSI (Small Computer System Interface) protocol, SAN allows for storage disks on separate servers to be shared. The term SAN is usually (but not necessarily) identified with block I/O services rather than file access services. NAS elements may be attached to any type of net work.

c. Other Data Center Hardware:

1. **Computer Server:** A computer that provides services and manages networked resources for client devices, e.g., desktop computers, notebook computers, thin clients, wireless devices, PDAs, IP telephones, other Computer Servers and other networked devices. Computer Servers are designed to respond to requests and are primarily accessed via network connections, and not through direct user input devices such as a keyboard, mouse, etc.

2. **Blade Chassis** *(recommend moving this from Building Block Definitions #1 to 'c. Other Data Center Hardware' section as Blade Chassis are Other Data Center Hardware:*

d. Power Supplies: No comments. Please refer to Climate Savers Computing Initiative (CSCI) for additional comments.

e. Operational States:

1. **Idle:** A system state in which the operating system and other software have completed loading, but no active workload transactions are requested or pending by the system. The tiers of disk drives in the system are set into standby mode. No activity occurs on any primary storage controller. In standby mode, the mechanical head assembly is parked within the disk drive, and the spindle is stopped. Fan usage is automatically shut-down in idle state. When the tier of disks is in the standby mode, the drive logic continues to listen for new commands. When disk access commands are received by the drive logic, the idle tier of drives will be returned to active mode. Idle state occurs on single storage pool servers for small and medium business and also in Tier 2 and Tier 3 storage systems in a data center.

2. **Active:** A system state where all drives are spinning, the drive logic is issuing disk access and write commands and storage/system maintenance functions are being performed**.** When active, a storage system's 'background' operations are being performed to support: regulation compliance, data accessibility, security and disaster recovery.

   Storage 'background' maintenance functions include: data redundancy for disk failure protection, snapshots for immediate data protection, duplicate copies of the transaction logs, local and remote replication for disaster recovery protection, incremental daily backups, weekly full backups, monthly full backups. Other active storage functions include hot swapping disks for online capacity expansion and online RAID level migration. Hot swapping is the function for disks to be physically added or removed; OCE is the function so a data array can be expanded across new disks; and ORLM allows the type of RAID array to be changed to a new type on the fly. These background system functions occur during active system states to allow modification of the

storage system without interrupting operations, such as dynamically resizing flexible volumes. The active state occurs in both single pool storage systems and in Tiered storage systems

3. **Maximum:** The storage system state where all disks are made available to all datasets through a common pool of storage, both performance and capacity utilization are maximized. When disk space is no longer needed by a particular application, it is returned to the free pool and made available to other applications as their storage needs grow
4. **Full Load:** *Same as Maximum*

Building Block #2: Eligible Product Categories

C. Questions for Discussion:

1. What are some additional means of segmenting the storage market? *Answer: Storage Market Segmentations are DAS (direct-attached storage), NAS (network- attached storage) and SAN (storage area network – both Fibre Channel and iSCSI). Storage systems can range from a few hundred dollars to over $500,000 in each of the three storage market segments.*
2. Are there any upcoming technologies or product types in development which are not included on the list provided in section (b) and should be considered for inclusion in ENERGY STAR? *Answer: Solid-State Storage Devices*
3. What portion of the enterprise storage market has the greatest need for an ENERGY STAR label to help customers identify the most efficient products? *Answer: in all three segments (DAS, NAS, SAN).* In what market segments can EPA expect to garner the most energy savings? *Answer: typically, larger/more expensive systems operate with hundreds of disks. Spinning disks consume the most energy within a storage system, and are the main source of heat generation – which requires more energy for system and data center cooling. Using fewer, higher capacity drives can significantly alter the storage power consumption. Typical SATA disk drives consume about 50% less power per tera-byte than equivalent capacity Fibre Channel drives.*

They also offer the highest available storage density per drive, further helping minimize power consumption.

Building Block #3: Energy Efficiency Criteria and Test Procedures

d. **Questions for Discussion:**

1. Which operational modes (e.g., Idle, Active, Full Load, Maximum) should EPA address in the specification? In which mode(s) might the highest energy savings be achieved? *Answer: Active state in which all disk drives are spinning. This is where the most power is consumed and this is where architectural and technology improvements can have the biggest impact on energy demand reduction.*

   There are a few practices/technologies which can reduce the need for constant spinning disks. These include: utilizing larger capacity disks, massive arrays of idle disk (MAID) systems, thin provisioning, utilizing data center storage tiers and data de-duplication.

   There is a direct correlation between the number and speed of disk drives and the electricity these devices consume. The more disks there are spinning, the more energy is required to keep spindle motors turning and to remove waste heat byproduct. Reducing the energy requirements for disk systems is relatively straightforward: reduce the number of spinning disks or spin the disks at a slower rotational rate. These techniques work well to store infrequently accessed data such as archival or backup data, but the same techniques may not meet the business needs of OLTP applications.

   Starting with simple approaches first, migrating data to larger capacity disk drives can reduce power consumption by 5 to 20%. For example, if backups today are written to 250GB SATA disks, write those backups to 500GB SATA disks. If this is performed, then half the number of disks will be needed and energy and floor space will be conserved. Storage-system software and infrastructure architecture optimized for SATA drive are more energy efficient than Fibre Channel.

   MAID storage systems detect when a disk drive hasn't been accessed for several minutes. Once the preset period of inactivity is reached,

the drives are rotated at a slower speed which saves energy. When a request to access data on the disk is received, the drive increases to its normal speed. This is perfect for archival data, hierarchical storage implementations or backup systems which are seldom if ever accessed. Data centers using MAID storage systems report a 10% to 20% reduction in storage power consumption.

Storage growth (and hence power use) can be reduced using a thin provisioning architecture. Thin provisioning is a technique that allocates disk blocks to a given application only when the blocks are actually written rather than at initial provisioning and partitioning. By employing thin provisioning along with good storage management software, not only is disk consumption reduced, but the storage management system itself can accurately calculate and project how much storage will be needed based on use history, instead of implementing over-capacity (and hence high power consumption) as a safe-guard.

Another solution for storage energy efficiency is appropriate use data center storage tiers to include use of optical and tape systems. The power and cooling requirements for tape systems are much less than similar capacity disk systems. Tape cartridges and optical media, when stored on a shelf, require no electricity and are ideal for archiving and system back-ups.

The most logical way, however, to save power is to store less data. Less data requires fewer operational cycles performing the data functions mentioned previously to ensure accuracy, availability, and security. Data centers using the best practice and/or deploying state-of-the-art storage systems use software that recognizes and eliminates duplication (sometimes referred to as de-duplication software). Storage growth almost always occurs because of stored redundant data. Enterprises should identify data that isn't accessed actively and move it to secondary or "near-line" storage such as optical or tape storage to maximize control of storage growth. Data center managers report data-reduction rates as high as 10:1 or even 20:1. Storing less data requires fewer disks and saves energy.

A final point: data de-duplication has significant benefit for heavily virtualized environments. Each virtual instance redeploys the operating system files repeatedly. Data de-duplication not only reduces the storage space required by virtualization, but it also increases the likelihood that when data is needed, it's already in filer's cache. Both go a long way toward improving the energy efficiency and performance of virtualized storage.

2. For various types or classes of Storage Products, what is the typical breakdown of energy consumption across operational states? E.g. an xx
   TB Storage Product typically consumes between xx and xx watts, and over the course of a year, xx % of the power is consumed in Idle state and
   xx% in Active state.
   Answer: given the diverse tasks of a storage system, from caching to indexing to read and writes, a good method of evaluating storage power consumption is to measure watts per usable terabyte. This can be expressed with the following formula:

$$\frac{\textbf{Watts to power system}}{\textbf{Total system TB} \times \textbf{\% System Utilization}} = \textbf{Watts per Usable TB}$$

   To calculate the power efficiency of a particular storage system, divide the total
   watts per system by the total number of TBs in that system times the system
   utilization. System utilization is equal to that percent of disks actually available
   for use.

3. Are there any additional power consumption or efficiency test procedures that should be considered for reference in the ENERGY STAR specification? *Answer: yes, the one given above.*

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124

Building Block #4: Information and Management Requirements

Recommend you consult with manufacturers on several of the non proprietary means for system thermal and power management of data access and storage systems.

Intel Corporation
2111 NE 25th Avenue
Hillsboro, OR 97124