Draft ETSI EN 303 804 V0.0.9 (2025-01)

Environmental Engineering (EE);
Energy efficiency metrics and measurement methods for data storage equipment

Reference
DEN/EE-EEPS44

Keywords
data storage, data storage equipment, energy efficiency, environment, measurement, metrics, storage

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI IPR online database.

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This draft European Standard (EN) has been produced by ETSI Technical Committee Environmental Engineering (EE), and is now submitted for the combined Public Enquiry and Vote phase of the ETSI EN Approval Procedure (ENAP).

Proposed national transposition dates	
Date of latest announcement of this EN (doa):	3 months after ETSI publication
Date of latest publication of new National Standard or endorsement of this EN (dop/e):	6 months after doa
Date of withdrawal of any conflicting National Standard (dow):	6 months after doa

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The present document specifies:

1) A metric for the assessment of energy efficiency of Data Storage Equipment (DSE).

The metric for energy efficiency of DSE is different from server's. For DSE it is important to evaluate the performance of data operation in the unit of Input/Output operations or data throughput per second, while for servers it is mainly to evaluate the performance of computing.

For comparison, evaluations should be conducted across similar types of categories of DSE. The present document categorizes DSE to address applicability, configuration groupings to represent a family of DSE to address the broad range of custom configurations possible within each family, and tool revision control to ensure comparability and consistency of the resulting metric value.

2) Test and evaluation methodologies for the assessment of energy efficiency of DSE.

The present document formalizes the tools, conditions and calculations used to generate a single figure of merit of DSE. The present document formalizes the methodology for evaluating energy saving level of DSE from the perspective of supported energy saving feature.

The present document considers some benchmark documents on energy efficiency of DSE.

1 Scope

The present document is based upon Energy Efficiency Benchmark for Data storage products.

The present document specifies:

1) an active state metric, test conditions and product family configuration for the assessment of energy efficiency of DSE using reliable, accurate and reproducible measurement methods;
2) an idle state metric and the calculation of the idle state power;
3) a measurement method of the active state power;
4) a measurement method of the idle state power;
5) the measurement and calculation of the maximum power;
6) the measurement and calculation of the operating condition class, the ASHRAE validation, using reliable, accurate and reproducible measurement methods, which take into account the recognized state of the art;
7) requirements for equipment to perform the measurements and analysis;
8) documentation and reporting requirements;
9) evaluation methodology for energy saving level from the perspective of supported energy saving features.

The present document addresses DSE.

The present document is applicable at the energy efficiency of:

- online storage;
- nearline storage.

The present document defines metric for the assessment of energy efficiency of DSE and related testing methodology considering data storage equipment HW and system.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found in the ETSI docbox.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] ANSI INCITS 400-2004: "Information Technology - SCSI Object-Based Storage Device Commands (OSD)".

[2] ANSI INCITS 458-2011: "Information Technology - SCSI Object Based Storage Device Commands- 2 (OSD 2)".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] "SNIA Emerald™ Power Efficiency Measurement Specification", Version 4.0.0.

[i.2] BenchDEE.

[i.3] "SPECstorage® Solution 2020".

[i.4] "CTS Lite Device Test - Power Efficiency Measurement".

[i.5] ASHRAE TC 9.9 Reference Card: "Equipment Thermal Guidelines for Data Processing Environments".

3 Definition of terms, symbols, abbreviations and conversions

3.1 Terms

For the purposes of the present document, the following terms apply:

a.c.-d.c. power supply unit: power supply unit that converts line-voltage alternating current (a.c.) input power into one or more direct current (d.c.) power outputs

active state: operational state of a DSE (as opposed to the idle state) in which the DSE is carrying out work in response to prior or concurrent external requests (e.g. instruction over the network), mainly data operation work

auto-tiering: policy-based system that automatically places and moves data across tiers to optimize performance service levels, cost targets, and overall energy consumption

NOTE: Each storage tier may comprise different storage technologies, offering varying performance, cost, and energy consumption characteristics.

blade storage: storage device that is designed for use in a blade chassis and that is dependent upon shared blade chassis resources (e.g. power supplies, cooling) for operation

cache: temporary data storage, not directly addressable by end-user applications, used to store data for expedited access to or from slower media

committed data: data that has been written to stable storage

compression: process of encoding data to reduce its size

controller system: computer or computer server that manages a benchmark evaluation process

Count-Key-Data (CKD): disk data organization model in which the disk is assumed to consist of a fixed number of tracks, each having a maximum data capacity

NOTE: The CKD architecture derives its name from the record format, which consists of a field containing the number of bytes of data and a record address, an optional key field by which particular records can be easily recognized, and the data itself.

data deduplication: replacement of multiple copies of data at variable levels of granularity with references to a shared copy in order to save storage space and/or bandwidth

Data Storage Equipment (DSE): collective term for disk drives, solid state drives and modules, tape cartridges, and any other mechanisms providing non-volatile data storage

data storage server: enterprise storage device which contains the same components as a computer server together with ≥ 10 storage devices and software (vendor or 3rd party) that supports storage system connectivity, capacity optimization management, virtualized storage environment and software defined storage

NOTE: Supporting features are described by the product's datasheet description and are either accompanied with vendor specific utilities and/or commercially available software supporting these functions.

data storage system: fully-functional system that supplies data storage services to clients and devices attached directly or through a network

NOTE 1: A storage system may be composed of integrated storage controllers, storage devices, embedded network elements, software, and other devices.

NOTE 2: While storage system may contain one or more embedded processors, these processors do not execute user-supplied software applications but may execute data-specific applications (e.g. data replication, backup utilities, data compression, install agents).

NOTE 3: 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 system.

NOTE 4: Components that are normally associated with a storage environment (e.g. devices required for operation of an external Storage Area Network) are not considered to be part of the storage system.

d.c.-d.c. power supply unit: power supply unit that converts line-voltage direct current (d.c.) input power to one or more d.c. outputs

NOTE: For purposes of the present document, a d.c-d.c. converter that is internal to a data storage system and is used to convert a low voltage d.c. (e.g. 12 VDC) into other d.c. power outputs for use by components inside is not considered a d.c-d.c. power supply unit.

dedupable: property that a collection of data is said to possess if the needed storage capacity for the data is reduced significantly by data deduplication

delta snapshot: type of point in time copy that preserves the state of data at an instant in time, by storing only those blocks that are different from an already existing full copy of the data

direct-connected: storage designed to be under the control of a single host, or multiple hosts in a non-shared environment

efficiency: ratio of workload output to the resource input to the system

energy efficiency: ratio of workload output to the power input to the system

file: abstract data object made up of:

a) an ordered sequence of data bytes stored on a disk or tape;

b) a symbolic name by which the object can be uniquely identified; and

c) a set of properties, such as ownership and access permissions;

that allow the object to be managed by a file system or backup manager

file system: software component that imposes structure on the address space of one or more physical or virtual disks so that applications may deal more conveniently with abstract named data objects of variable size (files)

Fixed Block Architecture (FBA): model of disks in which storage space is organized as linear, dense address spaces of blocks of a fixed size

NOTE: Fixed block architecture is the disk model on which SCSI is predicated.

formatted capacity: total number of bytes available to be written after a system or device has been formatted for use, e.g. by an object store, file system or block services manager

NOTE: Formatted capacity is less than or equal to raw capacity. It does not include areas set aside for system use, spares, RAID parity areas, checksum space, host- or file system-level remapping, "right sizing" of disks, disk labelling, caches, file system metadata, and so on. However, it may include areas that are normally reserved such as snapshot set-asides if they can alternatively be configured for ordinary data storage by the storage administrator.

free space: amount of additional irreducible data that can be written to the product under test as configured

hot band: simulation of naturally occurring areas of storage address space accessed more frequently than is typical

hybrid Solid State Storage and magnetic disk system: storage system whose formatted capacity is provided by a combination of Solid State Storage and magnetic disk storage devices

hyper converged system: software-defined set of resources with at least two of the following: compute, storage, networking and/or virtualization

ICT equipment: equipment providing data storage, processing and transport services

NOTE: A combination of Information Technology Equipment and Network Telecommunications Equipment.

idle state: operational state of a DSE (as opposed to the active state) in which the DSE is not performing any useful work

NOTE 1: The DSE is capable of completing workload transactions, but no active workload transactions are requested or pending by the system.

NOTE 2: For systems where ACPI standards are applicable, idle state correlates only to ACPI System Level S0.

I/O device: device which provides data input and output capability between a DSE and other devices

NOTE: An I/O device may be integral to the DSE motherboard or may be connected to the motherboard via expansion slots.

I/O intensity: measure of the number of IOPS requested by a load generator

NOTE: IO intensity is phrased as a percentage of selected maximum IOPS level that satisfies the timing requirement(s) for a taxonomy category.

I/O port: physical circuitry within an I/O device where an independent I/O session can be established

NOTE: A port is not the same as a connector receptacle; it is possible that a single connector receptacle can service multiple ports of the same interface.

irreducible data: data that is neither compressible nor dedupable

Just a Bunch Of Disks (JBOD): shelf that does not provide storage protection

NOTE: A JBOD is a "simple" "just a bunch of disks," an architecture that does not provide storage protection capabilities. The storage devices of a JBOD can function as individual volumes or can be combined to form a single logical volume by a simple controller. A JBOD provides no redundancy or resilience, so failure of a single storage device amounts to failure of a whole logical volume. A JBOD is a single shelf.

load generator: hardware and software environment executing the workload generator to drive the product under test during measurements

Logical Unit (LU): entity within a SCSI target that executes IO commands

Logical Unit Number (LUN): synonym for logical unit

maximum power: peak sustained or root means square power consumption value while operating the worst case functions

Maximum Time to First Data (MaxTTFD): maximum time required to start receiving data from a storage system to satisfy a read request for arbitrary data

memory: server component external to the processor in which information is stored for immediate use by the processor

motherboard: main circuit board of the DSE typically accommodating the processor, memory, expansion slots and enabling the attachment of additional circuit boards

network-connected: storage designed to be connected to a host via a network protocol (e.g. TCP/IP, IB, and FC)

non-disruptive serviceability: support for continued availability of data during all service operations

NOTE 1: Examples of non-disruptive serviceability are FRU replacement, code patches, software/firmware upgrades, configuration changes, data migrations, and system expansion done during production time.

NOTE 2: Service operations may result in performance impacts to data availability, but will not result in a loss of access.

non-volatile: property of retaining data in the event of the loss of power

normalized performance: relative performance values calibrated to a baseline common to the set of equipment being evaluated

parity RAID: collective term used to refer to Berkeley RAID Levels 3, 4, 5 and 6

permanent storage: data storage media which can retain data indefinitely without a power source

power supply unit: self-contained device, physically separable from the motherboard of the computer server, that converts a.c. or d.c. input power to one or more d.c. power outputs for powering the computer server via a removable or hard-wired electrical connection

raw capacity: sum of the raw, unformatted, uncompressed capacity of each of the storage device in the product under test

ready idle: operational state in which a system is capable of satisfying an arbitrary IO request within the response time and MaxTTFD constraints of its selected taxonomy category, but no user-initiated IO requests are being submitted to the system

NOTE: In the ready idle state, background I/O activity, autonomously initiated by the solution under test, may take place.

sequential read: IO load consisting of consecutively issued read requests to logically adjacent data

sequential write: IO load consisting of consecutively issued write requests to logically adjacent data

shelf: modular enclosure suitable for installation in a rack

NOTE: A shelf typically houses storage devices, storage controllers, power supplies, and cooling devices. A shelf typically has a pre-wired backplane that carries power and I/O interconnect signals to the housed components.

Single Point of Failure (SPOF): one component or path in a system, the failure of which would make the system inoperable or data inaccessible

solid state drive: storage device that uses memory chips instead of rotating magnetic platters for data storage

solid state storage: storage capability built from non-volatile solid state electronic devices

storage controller: device for handling storage requests that includes a processor or sequencer programmed to autonomously process a substantial portion of IO requests directed to storage devices

NOTE: This definition is specifically intended to exclude aggregating storage elements such as RAID array subsystems, robotic tape libraries, filers, and file servers. Also excluded are storage devices which are not directly accessible by end-user application programs, and are instead employed as a form of internal cache.

storage protection: any combination of hardware and software (e.g. RAID, NVRAM, disk sparing and background disk scrubbing or media scan) that assures that all completed IO operations will be preserved in the event of power loss or storage device failure

system crash: hardware or software failure which causes data to be temporarily unavailable and which requires a reboot of one or more hardware components and/or re-initialization of one or more software components in order for data access to be restored

Uninterruptible Power Supply (UPS): combination of convertors, switches, and energy storage devices (such as batteries) constituting a power system for maintaining continuity of load power in case of input power failure

virtual drive: removable media storage device, e.g. tape drive, that is emulated using other storage devices

weighted geometric mean: geometric mean calculated using a predetermined factor for each of the elements prior to aggregation

worklet: synthetic software routine, using real application functions focused on a particular type of data operation activity, which stresses a particular characteristic of the system

workload generator: software used in the load generator to drive the product under test during measurement

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

a.c., AC	Alternating Current
ACPI	Advanced Configuration and Power Interface
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BIOS	Basic Input/Output System
CPU	Central Processor Unit
d.c., DC	Direct Current
DSE	Data Storage Equipment
ESC	Energy Saving Cooling
ETSI	European Telecommunications Standards Institute
EUT	Equipment Under Test
FBA	Fixed Block Architecture
FRU	Field-Replaceable Unit
GB	GigaByte
HDD	Hard Disk Drive
I/O	Input/Output
ICT	Information and Communication Technology
ID	IDentification
INCITS	InterNational Committee for Information Technology Standards
IOPS	Input/Output operation Per Second
JBOD	Just a Bunch Of Disks
MaxTTFD	Maximum Time To First Data
NVRAM	Non-Volatile Random Access Memory
NVSS	Non-Volatile Solid State
OS	Operating System
OSD	Object-based Storage Device
PDU	Power Distribution Unit
PSU	Power Supply Unit
RAID	Redundant Array of Independent Disks
RAS	Reliability, Availability, and Serviceability
RMS	Root Mean Square
RVML	Removable and Virtual Media Library
SCSI	Small Computer System Interface
SPOF	Single Point of Failure
SSD	Solid State Drive
UPS	Uninterruptible Power Supply
VDA	Video Data Acquisition

VDC Volts (DC)

3.4 Conversions

For the purposes of the present document, the following conversions apply:

A kilobyte (KB) is equal to 10^3 B.
A megabyte (MB) is equal to 10^6 B.
A gigabyte (GB) is equal to 10^9 B.
A terabyte (TB) is equal to 10^{12} B
A petabyte (PB) is equal to 10^{15} B.
An exabyte (EB) is equal to 10^{18} B.
A kibibyte (KiB) is equal to 2^{10} B.
A mebibyte (MiB) is equal to 2^{20} B.
A gibibyte (GiB) is equal to 2^{30} B.
A tebibyte (TiB) is equal to 2^{40} B.
A pebibyte (PiB) is equal to 2^{50} B.
An exbibyte (EiB) is equal to 2^{60} B.

4 Taxonomy

4.1 Categories

The Taxonomy of DSE is specified in order to define suitable energy efficiency measurement methods for different types of DSE. Performance and energy efficiency comparison is suggested to conducted among same categories of DSEs.

Categories of DSE are defined based on attributes including media type and access paradigm.

Media types include magnetic disk, magnetic tape, optical disk, solid state disk and hybrid disk. A DSE with hybrid disk has both magnetic disks and solid state disks.

Access paradigm is one or more of block access, file access, object access, and memory access. A DSE with block access paradigm shall provide fixed-block or CKD access to data, e.g. via SCSI block access for disks/tapes. A DSE with file access paradigm shall provide file access to data, e.g. via the NFS protocol. A DSE with object access paradigm shall provide object access to data, e.g. via the SCSI OSD protocol (ANSI INCITS 400-2004 [1] and ANSI INCITS 458-2011 [2]). A DSE with memory access paradigm shall provide access to data via CPU memory access instructions, e.g. load and store.

In conclusion, DSEs are classified into the following categories:

- Online disk
- Near-Online disk
- Removable Media Library
- Virtual Media Library
- Disk Access NVSS
- Memory Access NVSS

Attributes of different categories of DSE are listed in Table 1.

Table 1: Taxonomy categories overview

Category	Disk		RVML		NVSS	
	Online disk	Near-online disk	Removable Media Library	Virtual Media Library	Disk access NVSS	Memory access NVSS
Media Type	Magnetic disk	Magnetic disk	Magnetic tape, optical disk	Magnetic disk, Solid State Storage	Solid State Storage + optional magnetic diska	Solid State Storage
Access Paradigm	Block, File, Object	Block, File, Object	Block	Block	Block, File, Object	memory

4.2 Classifications

4.2.1 Attributes

4.2.1.1 Access Paradigm

Access Paradigm is one or more of Block, File, Object, and Memory.

A storage product having the Block Access Paradigm attribute shall provide fixed-block or CKD access to data, e.g. via SCSI block access for disks/tapes.

A storage product having the File Access Paradigm attribute shall provide file access to data, e.g. via the NFS protocol.

A storage product having the Object Access Paradigm attribute shall provide object access to data, e.g. via the SCSI OSD protocol [1] and [2].

A storage product having the Memory Paradigm attribute shall provide access to data via CPU memory access instructions, e.g. load and store.

4.2.1.2 Access Pattern

Access Pattern is one or more of Random and Sequential.

A storage product having the Random Access Pattern attribute shall provide roughly equal access time to any stored data.

A storage product having the Sequential Access Pattern attribute may provide faster access to data at the address following the last accessed data than to other data.

4.2.1.3 Consumer/Component

A consumer product is any of a wide array of storage products which are used primarily for personal, family, household, or small-business purposes.

Components are devices (e.g. a stand-alone tape drive or an individual hard disk drive) that are a part of a data-center storage product.

A storage product having the Consumer/Component attribute shall have one of the above properties.

4.2.1.4 FBA/CKD Support

A storage product having the FBA/CKD Support attribute shall provide count-key-data as defined by IBM Corporation.

4.2.1.5 MaxTTFD

A storage product meeting the requirements of a given Classification/Category shall offer a MaxTTFD within the range specified in the associated Classification table.

4.2.1.6 Media Type

Media Type identifies the type of storage media providing the formatted capacity of a product. A storage product meeting the Media Type attribute shall include the type(s) of media specified in the associated Classification table.

The NVSS Set Disk Access Category Media Type "Solid State Storage + optional magnetic disk" encompasses either a purely Solid State Storage system or a hybrid Solid State Storage and magnetic disk system.

4.2.1.7 Multi-host Shareability

Multi-host Shareability identifies the number of hosts that can share the storage product. A storage product meeting the Multi-host Shareability attribute shall support the number of concurrently connected hosts given in the associated Classification table.

4.2.1.8 No SPOF

No SPOF is a storage product attribute. A storage product meeting the No SPOF attribute shall not have any component or path whose failure makes the system inoperable or data inaccessible.

4.2.1.9 Non-Disruptive Serviceability

Non-Disruptive Serviceability is a storage product attribute. A storage product meeting the Non-Disruptive Serviceability attribute shall support continued availability of data during all service operations, including FRU replacement, code patches, software/firmware upgrades, configuration changes, data migrations, and system expansion.

4.2.1.10 Robotics

A storage product having the Robotics attribute shall include an automatic mechanical system for loading and unloading storage media.

4.2.1.11 Stable Storage Support

A storage product having the Stable Storage Support attribute shall provide stable storage.

4.2.1.12 Storage Controller

A storage product having the Storage Controller attribute shall have a storage controller.

4.2.1.13 Storage Protection

A storage product having the Storage Protection attribute shall assure that all completed IO operations will be preserved in the event of power loss or storage device failure. This assurance can be provided by a combination of hardware and/or software, e.g. RAID, NVRAM, disk sparing, background disk scrubbing, and/or background media scan.

4.2.1.14 System Capacity

System Capacity is a metric used to help discriminate between systems of different sizes.

Disk Set and RVML Set Classifications: For the Disk Set Online, Disk Set Near-Online, RVML Set Removable Media Library, and RVML Set Virtual Media Library Classifications, System Capacity is the number of storage devices supported by a storage product. A storage product meeting the System Capacity attribute shall support the number of storage devices within the specified range.

NVSS Set Disk Access Classification: For the NVSS Set Disk Access Classification, System Capacity is specified qualitatively as one of:

- Consumer: A system marketed primarily for personal, family, household, or small-business purposes.
- JBOD: A single enclosure providing no storage protection.

- Very small: An entry level system with relatively small capacity and limited expandability.
- Small: An entry level system with relatively higher capacity and more expandability than the very small system.
- Medium: A higher capacity system that may have higher expandability than the small capacity system and with significantly more capabilities than the small systems.
- Large: A system with very high capacity and greater expandability compared to the medium and small systems intended for large Enterprises.
- Mainframe: A large system intended for Mainframe applications.

NVSS Set Memory Access Classification: For the NVSS Set Memory Access Classification, System Capacity is specified qualitatively as one of:

- Consumer: A system marketed primarily for personal, family, household, or small-business purposes.
- JBOD: A single enclosure providing no storage protection.
- Very small: An entry level system with relatively small capacity.
- Small: An entry level system with relatively higher capacity than the very small system
- Medium: A higher capacity system than the small capacity system and with significantly more capabilities than the small systems.
- Large: A system with very high capacity compared to the medium and small systems intended for large Enterprises.
- Mainframe: A large system intended for Mainframe applications.

4.2.2 Classifications overview

Classifications define combinations of settings or values for the attributes within a Category.

Taxonomy Classifications discriminate between different systems. The goal of the Classifications is to differentiate between systems of differing size, complexity, and target market. The Classifications are in general the following:

- Consumer/Component (<classification> 1);
- JBOD (<classification> 1,5);
- Low-end (<classification> 2);
- Mid-range (<classification> 3) and (<classification> 4);
- High-end (<classification> 5); and
- Mainframe (<classification> 6).

Consumer (<classification> 1) systems are any of a wide array of manufactured goods which are purchased primarily for personal, family, household, or small-business purposes. Components (<classification> 1) are components (e.g. a stand-alone tape drive or an individual hard disk drive) that are a part of a data-center storage product.

JBOD (<classification> 1,5) is a "simple" JBOD, an architecture that does not have storage protection capabilities built in. The storage devices in a JBOD can function as individual volumes or can be connected to form a single logical volume by a simple, integrated controller providing no redundancy. A JBOD provides no redundancy or resilience, so failure of a single storage device amounts to failure of a whole logical volume. A JBOD is a single shelf.

Low-end systems (<classification> 2) typically are a single shelf with one or two integrated controllers and limited expandability.

Mid-range (<classification> 3 and <classification> 4) systems are sub-divided into small and large midrange systems. Small Mid-range systems (<classification> 3) generally have limited expandability and robustness compared to Large Mid-range systems (<classification> 4).

High-end (<classification> 5) systems are very large systems that can have very high capacity and offer a high level of robustness.

Mainframe (<classification> 6) is differentiated from High End (<classification> 5) in that FBA/CKD support is required.

Table 2 provides an overview of the taxonomy classifications. Attributes of each classification are described in detail in clauses 4.2.3 to 4.2.8.

Table 2: Taxonomy overview

Categories	Data Storage Device					
	Disk		RVML		NVSS	
	Online disk	**Near-online disk**	**Removable Media Library**	**Virtual Media Library**	**Disk access NVSS**	**Memory access NVSS**
Consumer/Component (<classification> 1)	Online disk 1	Near-online disk 1	Removable Media Library 1	Virtual Media Library 1	Disk access NVSS 1	Memory access NVSS 1
JBOD (<classification> 1,5)	Online disk 1,5	Near-online disk 1,5	Removable Media Library 1,5	Virtual Media Library 1,5	Disk access NVSS 1,5	Memory access NVSS 1,5
Low-end (<classification> 2)	Online disk 2	Near-online disk 2	Removable Media Library 2	Virtual Media Library 2	Disk access NVSS 2	Memory access NVSS 2
Mid-range (<classification> 3)	Online disk 3	Near-online disk 3	Removable Media Library 3	Virtual Media Library 3	Disk access NVSS 3	Memory access NVSS 3
Mid-range (<classification> 4)	Online disk 4	Near-online disk 4	Removable Media Library 4	Virtual Media Library 4	Disk access NVSS 4	Memory access NVSS 4
High-end (<classification> 5)	Online disk 5	Near-online disk 5	Removable Media Library 5	Virtual Media Library 5	Disk access NVSS 5	Memory access NVSS 5
Mainframe (<classification> 6).	Online disk 6	Near-online disk 6	Removable Media Library 6	Virtual Media Library 6	Disk access NVSS 6	Memory access NVSS 6

4.2.3 Online disk

Table 3: Online disk classifications

Attribute	Online disk 1	Online disk 1,5	Online disk 2	Online disk 3	Online disk 4	Online disk 5	Online disk 6
Access Pattern	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential
MaxTTFD	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms
Media Type	Magnetic disk	Magnetic disk	Magnetic disk	Magnetic disk	Magnetic disk	Magnetic disk	Magnetic disk
Access Paradigm	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object
Multi-host Shareability	Not Specified	Ability to share with 1 or more hosts	Ability to share with 1 or more hosts	Ability to share with 1 or more hosts	Ability to share with 1 or more hosts	Ability to share with 1 or more hosts	Ability to share with 1 or more hosts
Consumer/ Component	Yes	No	No	No	No	No	No
Storage Controller	Optional	Optional	Required	Required	Required	Required	Required
Storage Protection	Optional	No	Required	Required	Required	Required	Required
No SPOF	Optional	Optional	Optional	Optional	Required	Required	Required
Stable Storage Support	Optional, unless Required by protocol	Optional, unless Required by protocol	Optional, unless Required by protocol	Required	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Optional	Required	Required
FBA/CKD Support	Optional	No	Optional	Optional	Optional	Optional	Required
System Capacity	≥ 1	≥ 4	≥ 4	≥ 12	> 100	> 400	> 400

4.2.4 Near-online disk

Table 4

Attribute	Near-online disk 1	Near-online disk 1,5	Near-online disk 2	Near-online disk 3	Near-online disk 4	Near-online disk 5	Near-online disk 6
Access Pattern	Random/ Sequential	-	Random/Sequential	Random/Sequential	-	Random/Sequential	Random/Sequential
MaxTTFD	> 80 ms	-	> 80 ms	> 80 ms	-	> 80 ms	> 80 ms
Media Type	Magnetic disk	-	Magnetic dis	Magnetic dis	-	Magnetic dis	Magnetic dis
Access Paradigm	Block, File, Object	-	Block, File, Object	Block, File, Object	-	Block, File, Object	Block, File, Object
Multi-host Shareability	Not Specified	-	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	-	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/ Component	Yes	-	No	No	-	No	No
Storage Controller	Optional	-	Optional	Required	-	Required	Required
Storage Protection	Optional	-	Optional	Required	-	Required	Required
No SPOF	Optional	-	Optional	Optional	-	Required	Required
Stable Storage Support	Optional, unless Required by protocol	-	Optional, unless Required by protocol	Required	-	Required	Required
Non-Disruptive Serviceability	Optional	-	Optional	Optional	-	Required	Required
FBA/CKD Support	Optional	-	Optional	Optional	-	Optional	Required
System Capacity	≥ 1	-	≥ 4	≥ 12	-	> 100	> 1 000

4.2.5 Removable Media Library

Table 5

Attribute	Removable Media Library 1	Removable Media Library 1,5	Removable Media Library 2	Removable Media Library 3	Removable Media Library 4	Removable Media Library 5	Removable Media Library 6
Access Pattern	Sequential	-	Sequential	Sequential	-	Sequential	Sequential
MaxTTFD	≤ 5 min	-	≤ 5 min	≤ 5 min	-	≤ 5 min	≤ 5 min
Media Type	Magnetic tape, optical disk	-	Magnetic tape, optical disk	Magnetic tape, optical disk	-	Magnetic tape, optical disk	Magnetic tape, optical disk
Access Paradigm	Block	-	Block	Block	-	Block	Block
Robotics	Prohibited	-	Required	Required	-	Required	Required
No SPOF	Optional	-	Optional	Optional	-	Optional	Required
Non-Disruptive Serviceability	Optional	-	Optional	Optional	-	Optional	Required
System Capacity	Not Specified	-	≤ 4	≥ 5[b]	-	≥ 25[b]	≥ 25[b]

4.2.6 Virtual Media Library

Table 6

Attribute	Virtual Media Library 1	Virtual Media Library 1,5	Virtual Media Library 2	Virtual Media Library 3	Virtual Media Library 4	Virtual Media Library 5	Virtual Media Library 6
Access Pattern	Sequential	-	Sequential	Sequential	-	Sequential	Sequential
MaxTTFD	≤ 80 ms	-	≤ 80 ms	≤ 80 ms	-	≤ 80 ms	≤ 80 ms
Media Type	Magnetic disk, Solid State Storage	-	Magnetic disk, Solid State Storage	Magnetic disk, Solid State Storage	-	Magnetic disk, Solid State Storage	Magnetic disk, Solid State Storage
Access Paradigm	Block	-	Block	Block	-	Block	Block
Robotics	Optional	-	Optional	Required	-	Required	Required
No SPOF	Optional	-	Optional	Optional	-	Optional	Required
Non-Disruptive Serviceability	Optional	-	Optional	Optional	-	Optional	Required
System Capacity	≤ 12	-	> 12[b]	> 48[b]	-	> 96[b]	> 96

4.2.7 Disk access NVSS

Table 7

Attribute	Disk access NVSS 1	Disk access NVSS 1,5	Disk access NVSS 2	Disk access NVSS 3	Disk access NVSS 4	Disk access NVSS 5	Disk access NVSS 6
Access Pattern	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential	Random/Sequential
MaxTTFD	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms
Media Type	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]	Solid State Storage + optional magnetic disk[b]
Access Paradigm	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object	Block, File, Object
Multi-host Shareability	Not Specified	Ability to share with 1 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts	Ability to share with 2 or more hosts
Consumer/ Component	Yes	No	No	No	No	No	No
Storage Controller	Optional	Optional	Required	Required	Required	Required	Required
Storage Protection	Optional	Not integrated	Required	Required	Required	Required	Required
No SPOF	Optional	Optional	Optional	Optional	Required	Required	Required
Stable Storage Support	Optional, unless Required by protocol	Not integrated	Optional, unless Required by protocol	Required	Required	Required	Required
Non-Disruptive Serviceability	Optional	Optional	Optional	Optional	Optional	Required	Required
FBA/CKD Support	Optional	Not integrated	Optional	Optional	Optional	Optional	Required
System Capacity	Consumer	JBOD	Very small	Small	Medium	Large	Mainframe

4.2.8 Memory access NVSS

Table 8

Attribute	Memory access NVSS 1	Memory access NVSS 1,5	Memory access NVSS 2	Memory access NVSS 3	Memory access NVSS 4	Memory access NVSS 5	Memory access NVSS 6
Access Pattern	Random	Random	Random	Random	Random	Random	Random
MaxTTFD	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms	≤ 80 ms
Media Type	Solid State Storage	Solid State Storage	Solid State Storage	Solid State Storage	Solid State Storage	Solid State Storage	Solid State Storage
Access Paradigm	Memory	Memory	Memory	Memory	Memory	Memory	Memory
Multi-host Shareability	Not Specified	TBD	TBD	TBD	TBD	TBD	TBD
Consumer/ Component	Yes	No	No	No	No	No	No
Storage Protection	Optional	TBD	Optional	Required	Required	Required	Required
No SPOF	Optional	TBD	Optional	Optional	Required	Required	Required
Stable Storage Support	Optional, unless Required by protocol	TBD	Optional, unless Required by protocol	Required	Required	Required	Required
Non-Disruptive Serviceability	Optional	TBD	Optional	Optional	Optional	Required	Required
System Capacity	Consumer	JBOD	Very small	Small	Medium	Large	Mainfram

5 Metrics

5.1 Performance metric

- Block access(Random): random access (transactional) of the data per second, Input Output Per Second (IOPS)
- Block access(Sequential): sequential access (streaming) of the data per second, Megabyte per second (MB/s)
- File access: performance metric is MB/s
- Storage Capacity, Gigabyte (GB)

5.2 Power metric

- Power of the EUT, Watt (W)

5.3 Energy efficiency metric

5.3.1 Active state energy efficiency metric

- Block access (Random): random access (transactional) of the data per unit of power, Input Output per second per watt (IOPS/W)
- Block access (Sequential): sequential access (streaming) of the data per unit of power, Megabyte per second per Watt (MB/s/W)
- File access: energy efficiency metric is MB/s/W

- Storage Capacity per unit of power(maximum power), Gigabyte per Watt (GB/W)

5.3.2 Idle state energy efficiency metric

- Idle state energy efficiency, Storage Capacity per unit of power at idle state, Gigabyte per Watt (GB/W)

6 Test definition and execution rules

6.1 Test setup

6.1.1 Configuration

Figure 1 is provided as guidelines for the test configuration. Test sponsors are free to modify their configuration to suit their particular needs and equipment, provided no other requirement of the present document is violated.

Power meter shall be the only power supply of the EUT. IO load drive system is the device that generates data operation applied on the EUT, and could be integrated with controller physically.

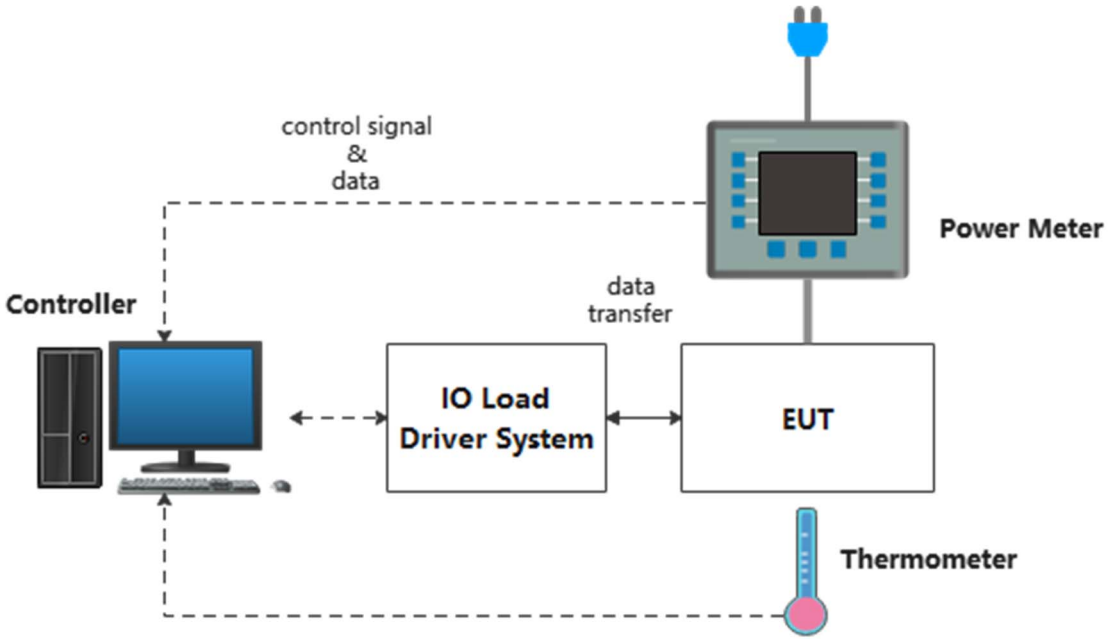

Figure 1: Test configuration

6.1.2 Environment

1) All measurements shall be conducted in a climate-controlled facility.

NOTE: The environmental conditions specified in in the following, satisfying ASHRAE Class A1 standards for data centers as described in ASHRAE TC 9.9 Reference Card [i.5].

2) Temperature Requirements: Ambient temperature shall be within the range 18 °C to 28 °C for the duration of the test.

3) Humidity Requirements: Humidity shall be within the range 15 % relative humidity to 80 % relative humidity for the duration of the test.

6.1.3 Power

Attributes of the power supply including supply voltage (AC RMS), phases, voltage tolerance, frequency tolerance maximum total harmonic distortion shall be evaluated and meet the requirement in Table 9.

Table 9: Input Power Requirements

Supply Voltage (AC RMS)	Phases	Voltage Tolerance	Frequency Tolerance	Maximum Total Harmonic Distortion
100 V, 115 V, 230 V	1	±1 % (Rated Power ≤ 1 500 W) ±5 % (Rated Power > 1 500 W)	±1 %	2 % (Rated Power ≤ 1 500 W) 5 % (Rated Power > 1 500 W)
200 V, 208 V, 400 V	3	±1 % (Rated Power ≤ 1 500 W) ±5 % (Rated Power > 1 500 W)	±1 %	2 % (Rated Power ≤ 1 500 W) 5 % (Rated Power > 1 500 W)

6.1.4 RAS

RAS features can have a significant impact on the power consumption of the product under test.

Typical RAS features are summarized as following:

- Dual Controller (No SPOF Controller)
- Mirroring (Local or remote, sync or async)
- RAID 1, 4/5, 6
- Snapshots (Full or Delta)
- Disk Scrubbing
- Multi-pathing
- Disk Sparing
- Dual Robotics
- Drive-level Maintenance
- Dual Power Supply
- Variable-speed Fans

Any RAS features required to satisfy the requirements of the EUT shall be enabled. The choice of what additional RAS features to enable in a product under test is left to the test sponsor.

If the product under test includes RAS features that are enabled for any test or test phase, then they shall be enabled for all tests and test phase, unless disabling of RAS features is explicitly allowed in the definition of a given test or test phase.

6.1.5 Measurement equipment requirement

Power meter and thermometer are required in the test and shall fulfil the following requirements:

1) The power meter shall be active throughout all tests and test phases.
2) The power meter shall record input voltage and power consumption of the EUT.
3) The accuracy of input power measurement shall be within 2 % at the 95 % confidence level.
4) The minimum resolution of power consumption measurement shall be within ±10 mW for $p \leq 10$ W, within ±100 mW for $10 < p \leq 100$ W and within ±1 W for $p > 100$ W.

5) The thermometer shall be placed to measure no more than 50 mm upwind of the primary air inlet port for the EUT.

6) The thermometer shall be active throughout all test and test phases.

7) The temperature is measured in degree Celsius.

8) The accuracy of temperature shall be within ±0,5 °C.

9) The temperature shall be recorded to durable media using a reading rate of one reading every 10 s.

6.1.6 EUT Consistency

The physical and logical configuration of the EUT, including its configuration and tuning parameters, shall not be changed between or during a test or test phase unless explicitly allowed in the definition of the test or test phase.

6.2 Equipment Under Test (EUT) Configuration

The configuration of the EUT shall be as specified in Table 10.

Table 10: Configuration of EUT

A)	As-shipped condition	Products shall be tested in their "as-shipped" condition, which includes both hardware revision and system settings, unless otherwise specified in this test method. Where relevant, all software options shall be set to their default condition.
B)	Measurement location	All power measurements shall be taken at a point between the a.c. power source and the EUT. Uninterruptible Power Supply (UPS) units shall not be connected between the power meter and the EUT. Internal batteries shall be at off state in which they do not supply or consume any energy from EUT. The power meter shall remain in place until all Idle and Active State efficiency power data are fully recorded. When testing a blade system, power shall be measured at the input of the blade chassis (i.e. at the power supplies that provide chassis distribution power). For a multi-node system, the power consumption per node of the EUT shall be measured in the fully-populated blade chassis configuration. All multi-node servers installed in the blade chassis shall be identical, sharing the same configuration.
C)	Air flow	Purposefully directing air in the vicinity of the measured equipment in a way that would be inconsistent with normal practices at the intended installation location is prohibited.
D)	Power supplies	All PSUs shall be connected and operational. For EUT with multiple PSUs: • all power supplies shall be connected to the a.c. power source and operational during the test; • if necessary, a Power Distribution Unit (PDU) may be used to connect multiple power supplies to a single source (if a PDU is used, any overhead electrical use from the PDU shall be included in the power measurement of the EUT). For blade system with half-populated chassis configurations, the power supplies for the unpopulated power domains can be disconnected.
E)	Power Management and Operating System	The as-shipped operating system or a representative operating system shall be installed. Products that are shipped without operating systems shall be tested with any compatible operating system installed. For all tests, the power management techniques and/or power saving features shall be left as-shipped. Any power management features which require the presence of an operating system (i.e. those that are not explicitly controlled by the Basic Input Output System (BIOS) or management controller) shall be tested using only those power management features enabled by the operating system by default.

6.3 Workload

6.3.1 General

According to different access paradigms of different categories of DSE, suitable workload shall be selected for energy efficiency test. Table 11 specifies workload definition for different categories of DSE.

Table 11: Workload definition

Categories	Data Storage Device					
	Disk		RVML		NVSS	
	Online disk	Near-online disk	Removable Media Library	Virtual Media Library	Disk access NVSS	Memory access NVSS
Access paradigms	Block, File, Object	Block, File, Object	Block	Block	Block, ObjectFile,	Memory
Workload	Block Access Workload File Access Workload	Block Access Workload File Access Workload	Block Access Workload	Block Access Workload	Block Access Workload File Access Workload	-

6.3.2 Block Access IO Profiles

The particular IO stimuli used to drive the product under test during a test or test phase are specified in terms of an IO profile (a.k.a. workload) made up of multiple attributes:

- Name: the name of the IO pattern for this stimulus. The identifier for the associated test phase is included parenthetically, when appropriate.
- IO Size: the number of bytes requested by a given read or write operation.
- Read/Write Percentage: the mixture of read/write IO requests within an IO profile.
- Transfer Alignment: Minimum granularity of IO transfer addresses. All transfer addresses within an IO stream shall be a multiple of this value.
- Access Pattern: either one or the other of the following two alternatives:
  - Random: Randomly distributed throughout the address space of the product under test;
  - Sequential: The first IO within an IO Stream with a sequential access pattern shall use an offset randomly distributed throughout the address range provided to the workload generator, and rounded down to satisfy the transfer alignment requirement.
- Data Pattern: compression ratio of 2:1.
- Hot band IO profile.

The goal of the hot band IO profile is to provide a workload that considers the contribution of auto- tiering mechanisms, e.g. read caching. This workload consists of a mix of different IO sizes and access patterns with a skewed access across a range of blocks. For example, this skewed access tends to hold data in cache and creates "cache hits" for improved throughput and reduced power consumption.

Typical block access IO profile is presented in detail in Annex E.

6.3.3 File Access IO Profiles

6.3.3.1 Overview of Workloads

Four workloads are specified to be executed in the File Access sequence of tests. These are five different file access workloads with different data operation characteristics. The warmup interval of these workloads may be adjusted by the test sponsor to ensure stable operation throughout the measurement period.

Typical file access IO profile is presented in detail in Annex F.

6.3.3.2 Software Building Workload

The software building workload is a classic meta-data intensive build workload. This workload was derived from analysis of software builds, and traces collected on systems in the software build arena. Conceptually, these tests are similar to running Unix "make' against several tens of thousands of files. The file attributes are checked (metadata operations) and if necessary, the file is read, compiled, then data is written back out to storage.

6.3.3.3 Video Data Acquisition Workload

The workload generally simulates applications that store data acquired from a temporally volatile source (e.g. surveillance cameras). A stream refers to an instance of the application storing data from a single source (e.g. one video feed). The storage admin is concerned primarily about maintaining a minimum fixed bit rate per stream and secondarily about maintaining the fidelity of the stream. The goal of the storage admin is to provide as many simultaneous streams as possible while meeting the bit rate and fidelity constraints.

The workload consists of two workload objects: VDA1 (data stream) and VDA2 (companion applications). Each stream corresponds to a roughly 36 Mb/s bit rate, which is in the upper range of high definition video.

6.3.3.4 Electronic Design Automation Workload

Electronic Design Automation Workload represents the typical behaviour of a mixture of EDA applications, represent very compute-heavy processes with high concurrency. The workload comprises of large numbers of small files with a low percent of large files. Mixed random and sequential IO of metadata operations representing two high level design phases: Frontend design and Backend design. The complete workload is a mixture of two subcomponents: the EDA_FRONTEND and EDA_BACKEND workloads. The EDA_FRONTEND workload is the EDA frontend processing applications, and EDA_BACKEND represents the EDA backend applications that generate the final output files.

6.3.3.5 AI Image Workload

AI Image Workload is representative of AI Tensorflow image processing environment. There are four subcomponents that make up the aggregate AI workload. The first two are the Data Preparation Phase, the second two make up the Training Phase:

- AI_SF: small (image) file ingest.
- AI_TF: tensor flow record creation.
- AI_TR: training consumption of tensor flow records.
- AI_CP: Represents the checkpointing functionality and may occur infrequently, if at all, during a typical run.

6.3.3.6 Genomics Workload

The Genomics workload models the entire pipeline of the Genomics workflow. The traces used to construct this workload came from commercial and research facilities that perform genetic analysis. The I/O behaviour was captured and is synthesized by the benchmark. The data has been sanitized so that it does not contain any of the original genome data.

6.3.4 Workload Generator

Workload Generator performing block and file access workload generation function shall be used as the workload generator. Block access workload IO profiles are specified in clause 6.3.2 and file access workload IO profiles are specified in clause 6.3.3. Test sponsors shall use a provided workload configuration scripts, containing user adjustable parameters that the Test Sponsor may adjust.

Hardware for workload generation should have enough performance that ensure it would not be the performance bottleneck of the test system. Workload generator software should meet the following requirements:

1) Generating either or both of block access and file access workload IO profiles specified in the present document.

2) Executing automated test during the whole test procedure.

3) Recording and replaying metrics versus time during the test. Recorded metrics include IOPS, throughput, response time, power, CPU usage, read/write mix percentage, outstanding IO.

4) Compatible with mainstream PCs and computer servers as workload generator hardware.

5) Compatible with mainstream data storage systems and data storage devices as EUT. Mainstream data storage systems include storage server, large scale distributed storage system. Mainstream data storage devices include single or group (up to 12, disk classification 3) of disk, SSD, etc.

The present document does not specify an energy efficiency benchmark tool. For the benchmark software that meets the requirements of the present document are provided in Annex A, Annex B and Annex C.

6.4 Test procedures

6.4.1 Block Access workload Test procedure

Block access work execution consists of the following sequence of tests:

1) Purge test, which clear the data and reset the EUT to ensure its status is close to unused status and contains no valid data.

   If purge is not supported on the EUT, this step can be skipped and the fact that purge is not supported/run should be documented in the report.

2) Pre-fill test, which puts data on the product under test.

   Apply Sequential Write workload with 256 KiB to the EUT for twice the user capacity. For HDD, prefill test step if optional. Whether pre-fill test is run should be documented in the report.

3) Conditioning test, which assures accurate and reproducible measurements.

   Apply the Hot Band Workload for a minimum of 5 Rounds of 30 minutes each.

The EUT shall achieve steady state wherein the linear least squares best-fit curve for five consecutive 30 minutes Rounds does not exceed a ± 10 % deviation from average IOPS of the five Rounds and the slope of linear least squares best-fit curve over the 5 Rounds does not exceed 10 %. This steady state determination procedure is referred to as the 5 Rounds Steady State Determination.

If the DUT is not in steady state, perform additional Rounds(up to a total of 25) until Steady State is achieved.

4) Hot Band.

This step includes a warm-up test for minimum interval of 10 minutes and a measurement interval of 30 minutes.

Hot Band workload is applied in this step.

5) Random Write.

This step includes a warm-up test for minimum interval of 10 minutes and a measurement interval of 30 minutes.

Random Write workload with date size of 8 KiB is applied in this step.

6) Random Read.

This step includes a warm-up test for minimum interval of 10 minutes and a measurement interval of 30 minutes.

Random Read workload with date size of 8 KiB is applied in this step.

7) Sequential Write.

This step includes a warm-up test for minimum interval of 10 minutes and a measurement interval of 30 minutes.

Sequential Write workload with date size of 256 KiB is applied in this step.

8) Sequential Read.

This step includes a warm-up test for minimum interval of 10 minutes and a measurement interval of 30 minutes.

Sequential Read workload with date size of 256 KiB is applied in this step.

9) Ready idle test, the basis of the ready idle metric. The minimum interval of ready idle test is 30 minutes.

During warm-up test of each active state test step, a self-optimizing sweep procedure is performed. Target workload is applied with different queue depth and thread count to find the max performance (IOPS) with response time less than 20 ms. Optimized queue depth and thread count should be documented in the report.

Details of IO profiles of block access workloads are in Annex E.

6.4.2 File Access workload Test procedure

File access work execution consists of the following sequence of tests:

1) Software Building Workload.

Software Building Workload with at least 10 load points is applied in this step.

Each load point includes a warm-up test for minimum interval of 5 minutes and a measurement interval of 5 minutes.

2) Video Data Acquisition Workload.

Video Data Acquisition with at least 10 load points is applied in this step.

Each load point includes a warm-up test for minimum interval of 5 minutes and a measurement interval of 5 minutes.

3) Electronic Design Automation Workload.

Electronic Design Automation Workload with at least 10 load points is applied in this step.

Each load point includes a warm-up test for minimum interval of 5 minutes and a measurement interval of 5 minutes.

4) AI Image Workload.

AI Image Workload with at least 10 load points is applied in this step.

Each load point includes a warm-up test for minimum interval of 5 minutes and a measurement interval of 5 minutes.

5) Genomics Workload.

Genomics Workload d with at least 10 load points is applied in this step.

Each load point includes a warm-up test for minimum interval of 5 minutes and a measurement interval of 5 minutes.

6) Ready idle test, the basis of the ready idle metric; The minimum interval of ready idle test is 30 minutes.

Details of IO profiles of file access workloads are in Annex E.

6.4.3 Test procedure for idle power at normal operating conditions

Two methods of measurement are possible:

- the test tool which automatically records and reports the idle power value; or
- manual collection of the idle power data from the power meter.

Below is the manual collection method for idle power:

1) Obtain a power measurement instrument which meets the requirements in clause 6.1.5.

2) Correctly configure the power analyser and measurement logging software (often distributed by the power analyser manufacturer) to log the EUT total energy consumption (volts and amps) at least one time per second. Measurement logging software shall be run on a separate system from the EUT to avoid affecting the measurement.

3) Configure the EUT as described in clause 6.2. If the default power configuration causes the EUT to enter a limited functionality low power state, such as sleep or hibernate, before the end of the measurement period, the setting to delay entering such states until after the measurement period, shall be set.

4) Boot the EUT and do not manually make any modifications to the EUT, such as opening or closing application or services.

5) Let the system sit idle state for between 5 and 6 min in an ambient temperature and a relative humidity as defined in clause 6.1.2.

6) Measure the EUT's energy consumption for 1 min, collecting at least one measurement per second.

7) Idle power for the EUT shall equal the arithmetic mean of power, in watts, during the 1-min measurement period.

7 Measurement

7.1 Measurement for active state efficiency

The measurement shall be in accordance with Table 12.

Table 12: Measurement of active state efficiency

A)	The EUT shall be re-booted. System caches and any stored information that may affect the active state energy efficiency metric shall be flushed.
B)	The test shall be started between 5 and 15 minutes after the completion of initial boot or log in.
C)	There shall be no manual intervention or optimization of the Controller System, EUT, or its internal and external environment during the execution of benchmarking tool.
D)	Once the execution of benchmarking tool is completed, the output files generated by the benchmarking tool shall be included with all testing results.

7.2 Measurement for idle state efficiency

In idle state, the useful work of the EUT is data reservation and there is no data operation executing. Performance metric of idle state is storage capacity in Gigabyte (GB). Storage capacity can be obtained from technical specification of the EUT. Power measurement for idle state shall be in accordance with Table 13.

Table 13: Measurement of idle state efficiency

A)	There shall be no manual intervention or optimization of the Controller System, EUT, or its internal and external environment during the execution of benchmarking tool or during manual collection interval.
B)	Once the execution of benchmarking tool is completed, the output files generated by the benchmarking tool shall be included with all testing results.

7.3 Measurement metrics and intervals

Table 14 specified the measurement metrics and intervals for the test.

During active state test, average power of each test phase should be collected from power meter, and temperature should be collected from thermometer. Workload data should be collected from workload benchmark tool or other throughput collection tool.

Table 14: Measurement metrics and intervals

Test	Test phase	Power recording interval(s)	Temperature recoding interval(s)	Workload Data Collection		Minimum Test Duration (minutes)
				Metric	Collection interval (seconds)	
Block Access	Conditioning	10	10	1) Average throughput for each drive (MiB/s) 2) Operations Rate (IOPS)	10	30
	Active	10	10	1) Average throughput for each drive (MiB/s) 2) Operations Rate (IOPS)	10	30
	Ready Idle	10	10	Capacity (GB)	-	30
File Access	Active per load point	10	10	Average throughput (MiB/s)	10	5
	Ready Idle	10	10	Capacity (GB)	-	30

7.4 Sensitivity analysis

Due to the manufacturing variance in components, number of significant power elements in the system, and run to run variations, re-test of an individual DSE, a DSE from the same family, or a product family will result in values that may vary from the initial testing.

Adjusting for a 90 % confidence interval, re-testing values that are within 10 % of the passing level shall be acceptable under re-test. Re-test or audit resulting in greater than 10 % error to the designated passing level shall be re-evaluated as a product family to determine compliance.

For audits and re-verification of Power Supply Units, re-testing of PSU efficiency shall not be lower than the declared value by more than 2 % and the power factor shall not be lower than the declared value by more than 10 %. Variance beyond these levels shall require re-evaluation of the Power Supply Unit.

7.5 Active state Periodic energy efficiency

The periodic energy efficiency for a test or test phase i, EEi(T), is the ratio of performance over a specified time interval T and the energy consumption during the same time interval T as illustrated by the following:

$$EE_i(T) = \frac{Perf_i(T)}{P_i(T)}$$

Where:

- $EE_i(T)$is the periodic energy efficiency during test or test phase i, taken over a time interval of T seconds.
- $Perf_i(T)$ is the total performance value during test or test phase i, taken over the same time interval of T seconds; For random block access workload test, $Perf_i(T)$ equals to the average operation rate in IOPS. For sequential block access workload test, $Perf_i(T)$ equals to the average data transfer rate in MiB/s. For file access workload test, $Perf_i(T)$ equals to the average performance matric over time interval T in MiB/s.

- $P_i(T)$ is the average power during test or test phase i, taken over the same time interval of T seconds.

7.6 Idle state energy efficiency

Idle state energy efficiency is calculated as following:

$$EE_{idle}(T) = \frac{Capacity}{P_{idle}(T)}$$

Where:

- $EE_{idle}(T)$ is the idle state energy efficiency during idle test interval.
- $Capacity$ is the total data storage capacity as the performance metric for idle state.
- $P_{idle}(T)$ is the average power during idle test interval.

7.7 Total energy efficiency score

Energy efficiency test of EUT is conducted under different workloads and loadlevels including active states and an idle state, and an energy efficiency score is generated from each test interval. A comprehensive energy efficiency score is defined from interval results to generally evaluate the energy efficiency of EUT.

Since the energy efficiency metrics are different between active states and idle state, normalization is conducted before calculating comprehensive score. Reference values of energy efficiency metrics are defined from a reference DSE and same for all tests.

Geometric mean function is used to calculate energy efficiency score of each workload from scores of all loadlevel intervals, and to calculate total energy efficiency score of the total test from scores of all workloads and idle state.

$$EE_{total} = \sqrt[n]{\prod_k EE_k}$$

Where:

- EE_{total} is the total energy efficiency.
- EE_k is the normalized period energy efficiency of phase k.
- k is test phase including all active state phases and active idle state.
- n are the numbers of phases including all active state phases and active idle state.

8 Test report

The following energy efficiency metrics shall be listed in the test report:

a) Performance metric of each test interval

b) Power metric

c) Active state energy efficiency metric;

d) Idle state energy efficiency metric.

e) Comprehensive active energy efficiency scores of all workloads

f) Comprehensive energy efficiency score of the total test

g) reported maximum power

It is accepted that the measurement accuracy of the metrics is ±10 %. Any subsequent assessment within this range shall be considered to be consistent with the quoted value.

Diagrams of key energy efficiency metrics shall be shown in the report.

Details of test results are listed in Table 15.

Table 15: Required test results in test report

Block access workload test								
	Metric	**Units**	**Hot Band**	**Random Write**	**Random Read**	**Sequential Write**	**Sequential Read**	**Ready Idle**
Key result	Energy Efficiency	IOPS/W	value	value	value	value	value	-
		MiB/s/W	value	value	value	value	value	-
		GB/W	-	-	-	-	-	value
Supporting data	Performance	IOPS	value	value	value	value	value	-
		MiB/s	value	value	value	value	value	-
		GB	-	-	-	-	-	value
	Power	W	value	value	value	value	value	value
	Thread count	-	value	value	value	value	value	-
	Queue depth	-	value	value	value	value	value	-
Other test parameters	Steady state	Yes/No	value					
	Purge	Yes/No	value					
	Pre-fill	Yes/No	value					
	Active LBA range	Percent	value					
	Data pattern	Random 2:1 compressible						
File access workload test								
	Metric	**Units**	**Software Building Workload (each load point)**	**Video Data Acquisition Workload (each load point)**	**Electronic Design Automation Workload (each load point)**	**AI Image Workload (each load point)**	**Genomics Workload (each load point)**	**Ready Idle**
Key result	Energy Efficiency	IOPS/W	value	value	value	value	value	-
		MiB/s/W	value	value	value	value	value	-
		GB/W	-	-	-	-	-	value
Supporting data	Performance	IOPS	value	value	value	value	value	-
		MiB/s	value	value	value	value	value	-
		GB	-	-	-	-	-	value
	Power	W	value	value	value	value	value	value

The following additional information shall also be reported/provided under the technical documentation:

1) Automatedly generated test report and supporting data from test tool.

2) Author, site, and date of the test.

3) EUT category and classification.

4) Configuration information of DSE, including:

   a) Product family with the low-end performance configuration and the high-end performance configuration.

   b) Manufacturer name.

   c) Product model.

   d) For data storage system, component manufacturer, product ID, number of units and component product type for CPU, memory, drive (HDD or SSD) shall be reported.

5) Power Supply Unit test information or reference to previously conducted test report.

6) The extremes of EUT inlet test temperature during the test.

7) Revision numbers for each of test software elements used.

8) Test tool information including name and version.

9) Software environment of test tool, including OS type, OS version, Java™ revision and source used.

10) Controller model.

11) Measurement equipment (power meter, thermal sensor/meter): manufacturer, model, ID, and calibration date.

9 Energy saving level evaluation

9.1 General

The purpose of defining energy saving level is to guide the development and support of energy saving features on DSE. Energy saving level is defined according to the supported energy saving features of a DSE and the level of each feature. DSE supporting more features has higher energy saving level score.

9.2 Energy saving feature

9.2.1 Energy saving features at idle and low loadlevel state

Energy saving features at idle state are in the following:

- PSU Active-Standby mode (A-S mode): The PSU active/standby mode is used in idle and light-load scenarios, and the PSU load balancing mode is used in medium- and high-load scenarios to improve power supply efficiency.

9.2.2 Energy saving features in active state

Energy saving features in active state are in the following:

- Power Capping: This technology is used to improve cabinet deployment density, improve power supply utilization, and ensure power supply security.
- Energy Saving Cooling (ESC): The fan speed adjustment technology adopts the PID control algorithm.

9.3 Evaluation methodology

9.3.1 Information obtaining

Supported energy saving features and details for determining feature level shall be obtained from formal manual of a DSE or from formal test report. Uncertain features shall be marked as not supported.

9.3.2 scoring rules

Table 16 specifies the scoring rule for each energy saving feature.

Table 16: Scoring rule for each energy saving feature

Supported	Not supported	Supported
energy saving level score	0	10

Total energy saving level score $Score_{ES}$ is calculated as following:

$$Score_{ES} = \sum_i Score_{ES,i}$$

where i represents feature i listed in clause 9.2 and $Score_{ES,i}$ represents score of feature $Score_{ES,i}$ according to Table 16.

Combining total energy efficiency test score EE_{total} (see clause 7.6) and total energy saving level score specified in this clause, a total energy efficiency score could be calculated:

$$EE'_{total} = EE_{total} + \alpha Score_{ES}$$

where EE'_{total} is a comprehensive EE score and $\alpha Score_{ES}$ is the extra score for energy saving feature. $\alpha = 0{,}1$ is the extra factor.

Annex A (informative): BenchDEE Benchmark

BenchDEE [i.2] is an energy efficiency benchmark developed by the China National Institute of Standardization (CNIS).

BenchDEE provides full auto test function including workload generation, power measurement, temperature measurement and energy efficiency assessment. The workloads of BenchDEE fulfil the requirement of both file access and block access test. Measurement of performance metrics of both block access and file access workloads is integrated in BenchDEE.

This document describes the use of BenchDEE Benchmark as the primary workload generator for the full testing.

Annex B (informative): SPECstorage® Solution 2020

SPECstorage® Solution 2020 [i.3] is a performance benchmark developed by the Standard Performance Evaluation Corporation (SPEC).

This document describes the use of SPECstorage® Solution 2020 as the alternative workload generator for file access testing.

Annex C (informative): CTS Lite Device

CTS Lite [i.4] is an IO performance test software product for data storage devices developed by Calypso Systems, Inc.

This document describes the use of CTS Lite as the alternative workload generator for file access testing.

Annex D (informative): SNIA Emerald™ Power Efficiency Measurement

SNIA Emerald™ Power Efficiency Measurement [i.1] is released by the Storage Networking Industry Association (SNIA).

This document 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.

This document describes a standardized method to assess the energy efficiency of commercial storage products in both active and idle states of operation.

Annex E (normative): Block Access workload IO Profile

IO profile presented in this clause applies to block access testing of online disk category, near-online disk category, and disk access NVSS category DSE.

Table E.1: Pre-fill test IO profile of block access workload

IO Profile	Read/Write Percentage	IO Intensity	Access Pattern	Data Pattern
Sequential Write	0/100	100	Sequential	2:1 compression

Table E.2: Conditioning test IO profile of block access workload

IO Profile	Read/Write Percentage	IO Intensity	Access Pattern
Hot Band Workload	See Table E.4 to Table E.6	100	See Table E.4 to Table E.6

Table E.3: Active test IO profile of block access workload

IO Profile	IO Size (KiB)	Read/Write Percentage	IO Intensity	Transfer Alignment (KiB)	Access Pattern
Hot Band Workload	See Table E.4 to Table E.6	See Table E.4 to Table E.6	100	See Table E.4 to Table E.6	See Table E.4 to Table E.6
Random Write	8	0/100	100	8	Random
Random Read	8	100/0	100	8	Random
Sequential Write	256	0/100	100	256	Sequential
Sequential Read	256	100/0	100	256	Sequential

Table E.4: Hot band IO profile

IO Profile	% of workload	Read/Write Percentage	Access Pattern	Usable Address Range
Write Stream 1	5	0/100	Sequential	0 % to 100 %
Write Stream 2	5	0/100	Sequential	0 % to 100 %
Write Stream 3	5	0/100	Sequential	0 % to 100 %
Read Stream 1	5	100/0	Sequential	0 % to 100 %
Read Stream 2	5	100/0	Sequential	0 % to 100 %
Read Stream 3	5	100/0	Sequential	0 % to 100 %
Read Stream 4	5	100/0	Sequential	0 % to 100 %
Read Stream 5	5	100/0	Sequential	0 % to 100 %
Uniform Random	6	50/50	Random	0 % to 100 %
Hot Band 1	28	70/30	Random	10 % to 18 %
Hot Band 2	14	70/30	Random	32 % to 40 %
Hot Band 3	7	70/30	Random	55 % to 63 %
Hot Band 4	5	70/30	Random	80 % to 88 %

Table E.5: IO Transfer size within the hot band IO profile for 512 Byte native devices

Xfer	Streaming Write	Streaming Read	Uniform	Hot Band
512 B			2 %	2 %
1 KiB			2 %	2 %
4 KiB	29 %	29 %	27 %	27 %
8 KiB	33 %	33 %	31 %	31 %
16 KiB	6 %	6 %	5 %	5 %
32 KiB	5 %	5 %	5 %	5 %
48 KiB			1 %	1 %
56 KiB			1 %	1 %
60 KiB			2 %	2 %
64 KiB	22 %	22 %	20 %	20 %
128 KiB	3 %	3 %	2 %	2 %
256 KiB	2 %	2 %	2 %	2 %

Table E.6: IO Transfer size within the hot band IO profile for 4k Byte native devices

Xfer	Streaming Write	Streaming Read	Uniform	Hot Band
512 B				
1 KiB				
4 KiB	29 %	29 %	31 %	31 %
8 KiB	33 %	33 %	31 %	31 %
16 KiB	6 %	6 %	5 %	5 %
32 KiB	5 %	5 %	5 %	5 %
48 KiB			1 %	1 %
56 KiB			1 %	1 %
60 KiB			2 %	2 %
64 KiB	22 %	22 %	20 %	20 %
128 KiB	3 %	3 %	2 %	2 %
256 KiB	2 %	2 %	2 %	2 %

Annex F (normative): File Access workload IO Profile

Table F.1 present typical characteristics of file access workload required in the test defined in the present document, describing the percentage of all data operations in each workload.

Table F.1: File operation distribution of file access workloads

Operation	Software Building Workload	Video Data Acquisition Workload		Electronic Design Automation Workload	
		VDA1	VDA2	EDA_FRONTEND	EDA_BACKEND
	Percentage (%)	Percentage (%)	Percentage (%)	Percentage (%)	Percentage (%)
read	0	0	5	0	50
mmap read	0	0	0	0	0
write	0	100	0	0	50
mmap write	0	0	0	0	0
rmw	0	0	2	0	0
mkdir	1	0	0	1	0
readdir	2	0	3	0	0
unlink	2	0	1	1	0
stat	70	0	2	39	0
rename	0	0	0	0	0
lock	0	0	0	0	0
statfs	0	0	0	0	0
trunc	0	0	0	0	0
read file	6	0	84	7	0
rand read	0	0	0	8	0
write file	7	0	0	10	0
rand write	0	0	0	15	0
append	0	0	0	0	0
rmdir	0	0	1	0	0
create	1	0	0	2	0
unlink2	0	0	2	1	0
access	6	0	0	15	0
copyfile	0	0	0	0	0
chmod	5	0	0	1	0
pathconf	0	0		0	0
neg_stat	0	0		0	0
Operation	**AI Image Workload**				**Genomics Workload**
	AI_SF	**AI_TF**	**AI_TR**	**AI_CP**	
	Percentage (%)	**Percentage (%)**	**Percentage (%)**	**Percentage (%)**	**Percentage (%)**
read	37	0	95	0	70
mmap read	0	0	0	0	0
write	0	100	0	100	8
mmap write	0	0	0	0	0
rmw	0	0	0	0	0
mkdir	0	0	0	0	0
readdir	0	0	0	0	0
unlink	0	0	0	0	1
stat	56	0	5	5	12
rename	0	0	0	0	0
lock	0	0	0	0	0
statfs	0	0	0	0	0
trunc	0	0	0	0	0
read file	0	0	0	0	0
rand read	0	0	0	0	2
write file	0	0	0	0	0
rand write	0	0	0	0	1
append	0	0	0	0	0
rmdir	0	0	0	0	0
create	0	0	0	0	1
unlink2	0	0	0	0	0
access	7	0	0	0	4

Operation	AI Image Workload				Genomics Workload
	AI_SF	AI_TF	AI_TR	AI_CP	
	Percentage (%)	Percentage (%)	Percentage (%)	Percentage (%)	Percentage (%)
copyfile	0	0	0	0	0
chmod	0	0	0	0	1
pathconf	0	0	0	0	0
neg_stat	0	0	0	0	0

Annex G (informative): Bibliography

- Commission Regulation (EU) 2019/424 of 15 March 2019 laying down ecodesign requirements for servers and data storage products pursuant to Directive 2009/125/EC of the European Parliament and of the Council and amending Commission Regulation (EU) No 617/2013.
- Directive 2009/125/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for the setting of ecodesign requirements for energy-related products.
- ETSI GS OEU 016 (V1.1.1): "Energy Consumption Measurement of Operational Information Technology Storage Units".
- ISO/IEC 24091:2019: "Information technology — Power efficiency measurement specification for data center storage".

History

Document history			
V0.0.9	January 2025	ENAP process	AP 20250430: 2025-01-30 to 2025-04-30