ETSI EN 303 470 V1.1.1 (2019-03)

EUROPEAN STANDARD

Environmental Engineering (EE); Energy Efficiency measurement methodology and metrics for servers

Reference
DEN/EE-EEPS24

Keywords
energy efficiency, ICT, server, sustainability

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 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

List of tables

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is 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 Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, 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.

Foreword

This European Standard (EN) has been produced by ETSI Technical Committee Environmental Engineering (EE).

National transposition dates	
Date of adoption of this EN:	11 March 2019
Date of latest announcement of this EN (doa):	30 June 2019
Date of latest publication of new National Standard or endorsement of this EN (dop/e):	31 December 2019
Date of withdrawal of any conflicting National Standard (dow):	31 December 2019

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 a metric for the assessment of energy efficiency of computer servers using reliable, accurate and reproducible measurement methods, which take into account the recognized state of the art.

The present document formalizes the tools, conditions and calculations used to generate a single figure of merit of a single computer server representing its relative efficiency and power impact. The metric is targeted for use as a tool in the selection process of servers to be provisioned.

For comparisons, evaluations should be conducted across similar server types or categories. The efficiency metric is targeted for use in a pass/fail selection process by differentiating the ability of servers to be provisioned for general purpose operations. The present document does not prescribe the levels or values for acceptance but prescribes a standard method of evaluation that energy efficiency programs would use to establish such criteria.

As there are many operational deployments of servers resulting in a range of specialized equipment and configurations for a single server product, a metric that evaluates provisioning impacts to general purpose operations may not be applicable. ICT equipment and servers in particular, are generally customized and commissioned on site for deployment. As with most IT equipment, new technologies are regularly introduced, which may require product level customization or an industry wide tool upgrade to more appropriately represent the efficiency of the servers. The present document categorizes servers to address applicability, configuration groupings to represent a family of servers to address the broad range of custom configurations possible within each server product family, and tool revision control to ensure comparability and consistency of the resulting metric value.

The present document is based upon the Server Efficiency Rating Tool™ (SERT™) of the Standard Performance Evaluation Corporation (SPEC) and takes into account:

- the Eco-design Technical Assistance Study on Standards for ErP Lot 9 Enterprise Servers and Enterprise Data Storage;
- activity related to the analysis of output of Server Efficiency Rating Tool (SERT™) measurements and deployed power by The Green Grid;
- ENERGY STAR® for Computer Servers [i.2].

The present document defines energy efficiency metrics and measurement methodology for server equipment under standardization mandate M/462 of the European Commission [i.3].

1 Scope

The present document specifies a metric using the Server Efficiency Rating Tool (SERT™), test conditions and product family configuration for the assessment of energy efficiency of computer servers using reliable, accurate and reproducible measurement methods. The metric applies to general purpose computer servers with up to four processor sockets and with their own dedicated power supply.

NOTE 1: The term "socket" also applies to design in which processors are installed without sockets (e.g. soldered products).

The metric applies to a computer server model and to a computer server product family, including type and count of CPU, memory, storage, power supplies, cooling (e.g. fans) and any other add-on hardware expected to be present when deployed.

The present document defines:

- an energy efficiency metric to support procurement or market entry requirements;
- requirements for equipment to perform the measurements and analysis;
- requirements for the measurement process;
- requirements for the management of the metric calculation;
- operation or run rules to configure, execute, and monitor the testing;
- documentation and reporting requirements;
- a validation process for the metric using the Deployed Power Assessment.

The present document is not applicable to:

- fully fault tolerant servers;
- High Performance Computing (HPC) systems;
- hyper-converged servers;
- large scale servers;
- servers with integrated APA(s);
- networking equipment including network servers;
- server appliances;
- storage device including blade storage and storage servers.

NOTE 2: Products whose feature set and intended operation are not addressed by active mode testing parameters are excluded from this evaluation method. The above list shows products for which SERT™ efficiency evaluations are not appropriate.

The present document does not address home servers and small servers that fall under the scope of mandate M/545 [i.8].

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 at https://docbox.etsi.org/Reference/.

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] CENELEC EN 62623:2013: "Desktop and notebook computers. Measurement of energy consumption".

[2] Standard Performance Evaluation Corporation (SPEC): "Server Efficiency Rating Tool (SERT) version 2 Run and Reporting Rules".

NOTE: Available at https://www.spec.org/sert2/SERT-runrules.pdf.

[3] Standard Performance Evaluation Corporation (SPEC): "Server Efficiency Rating Tool (SERT) version 2 User Guide".

NOTE: Available at https://www.spec.org/sert2/SERT-userguide.pdf.

[4] IEEE 802.3™: "IEEE Standard for Ethernet".

NOTE: Available at https://standards.ieee.org/findstds/standard/802.3-2015.html.

[5] IEEE 802.3az™: "Energy Efficient Ethernet".

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] CENELEC EN 60297 series: "Mechanical structures for electrical and electronic equipment. Dimensions of mechanical structures of the 482,6 mm (19 in) series".

[i.2] ENERGY STAR®: "Product Specification for Computer Servers".

[i.3] Standardization mandate M/462: "M/462 Standardisation mandate addressed to CEN, CENELEC and ETSI in the field of ICT to enable efficient energy use in fixed and mobile information and communication networks".

[i.4] ETSI EN 300 119 series: "Equipment Engineering (EE); European telecommunication standard for equipment practice".

[i.5] Standard Performance Evaluation Corporation (SPEC): "Server Efficiency Rating Tool (SERT) version 2 Design Document".

NOTE: Available at https://www.spec.org/sert2/SERT-designdocument.pdf.

[i.6] SERT Client Configurations (JVM Options).

NOTE: Available at https://www.spec.org/sert2/SERT-JVM_Options-2.0.html.

[i.7] SERT Result File Fields.

NOTE: Available at https://www.spec.org/sert2/SERT-resultfilefields.html.

[i.8] Mandate M/545: "Commission Implementing Decision of 6.1.2016 on a standardisation request to the European standardisation organisations as regards computers and computer servers, in support of the implementation of Commission Regulation (EU) No 617/2013 of 26 June 2013, implementing Directive 2009/125/EC of the European Parliament and of the Council with regard to ecodesign requirements for computers and computer servers".

3 Definition of terms, symbols and abbreviations

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 for powering a computer server

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

NOTE: The work includes, but is not restricted to, active processing and data seeking/retrieval from memory, cache, or internal/external storage while awaiting further input over the network.

Auxiliary Processing Accelerator (APA): additional compute device installed in the computer server that handles parallelized workloads

NOTE 1: This includes, but is not limited to, Graphical Processing Units (GPUs) or Field Programmable Gate Array chips which can be installed in a server either on Graphics or Extension add-in cards installed in general-purpose add-in expansion slots (e.g. GPGPUs, CPU accelerators, etc. installed in a PCI slot) or directly attached to a server component such as the motherboard.

NOTE 2: There are two specific types of APAs used in servers:

a) Expansion APA: An APA that is on an add-in card installed in an add-in expansion slot (e.g. GPGPUs, CPU accelerators, etc. installed in a PCI slot). An expansion APA add-in card may include one or more APAs.

b) Integrated APA: An APA that is integrated into the motherboard or CPU package or an expansion APA that has part of its subsystem, such as switches, included in the non-APA server configuration that would be used to run the energy efficiency test (SERT™ suite).

blade chassis: enclosure that contains shared resources for the operation of blade servers, blade storage, and other blade form-factor devices

NOTE: Shared resources provided by a chassis include, but are not restricted to, power supplies, data storage, and hardware for d.c. power distribution, thermal management, system management, and network services.

blade server: computer server, designed for use in a blade chassis, that is a high-density device and functions as an independent computer server and includes at least one processor and system memory, which is dependent upon shared blade chassis resources (e.g. power supplies, cooling) for operation

NOTE: A processor or memory module that is intended to scale up a standalone server is not considered a blade server.

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

blade system: blade chassis and one or more removable blade servers and/or other units (e.g. blade storage, blade networking equipment) which provide a scalable means for combining multiple blade server or storage units in a single enclosure

NOTE: A blade system is designed to allow service technicians to easily add or replace (hot-swap) blades in the field.

buffered Double Data Rate (DDR) channel: channel or memory port connecting a memory controller to a defined number of memory devices (e.g. dual in-line memory modules (DIMMs)) in a computer server

NOTE 1: A typical computer server may contain multiple memory controllers, which may in turn support one or more buffered DDR channels.

NOTE 2: Each buffered DDR channel serves only a fraction of the total addressable memory space in a computer server.

computer server: computer, sold through enterprise channels, that provides services and manages networked resources for client devices

NOTE 1: Client devices include, but are not restricted to desktop computers, notebook computers, thin clients, wireless devices, Personal Digital Assistants, IP telephones, other computer servers, or other network devices.

NOTE 2: A computer server is primarily accessed via network connections, versus directly-connected user input devices such as a keyboard or mouse.

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

data averaging interval: time period over which all samples captured by the high-speed sampling electronics of the power analyser are averaged to provide the measurement set

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 for powering a computer server

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

deployed power: average power level of the utilization applicable to the total number of servers provisioned to meet an aggregate peak load

direct current server: computer server that is designed solely to operate on a direct current (d.c.) power source

double-wide blade server: blade server requiring twice the width of a standard blade server bay

efficiency: defined workload output divided by the resource input to the system

fully fault tolerant server: computer server that is designed with complete hardware redundancy, in which every computing component is replicated between two nodes running identical and concurrent workloads (i.e. if one node fails or needs repair, the second node can run the workload alone to avoid downtime) and that uses two systems to simultaneously and repetitively run a single workload for continuous availability in a mission critical application

half-height blade server: blade server requiring one half the height of a standard blade server bay

hard disk drive: primary computer storage device which reads and writes to one or more rotating magnetic disk platters

High Performance Computing (HPC) system: computing system which is designed (or assembled), optimized, marketed and sold to execute highly parallel applications for higher performance computing applications

NOTE 1: HPC systems support applications including, but not restricted to, deep learning or artificial intelligence.

NOTE 2: HPC systems feature multiple clustered nodes to increase computational capability and often featuring high speed inter-processing interconnects as well as large memory capability and bandwidth.

NOTE 3: HPC systems may be comprised of multiple clusters of homogenous nodes, for which the clusters may be heterogeneous.

hyper-converged server: highly integrated enterprise device which contains the same components as a computer server in addition to the features of a network server and storage server

hypervisor: supervisory system level software that establishes and manages a virtualized environment which enables multiple operating systems to run on a single physical system at the same time

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 computer server in which the operating system and other software have completed loading but is not performing any useful work

NOTE 1: The computer server 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 computer server and other devices

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

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.

large scale server: resilient/scalable server which ships as a pre-integrated/pre-tested system housed in one or more full frames or racks and that includes a high connectivity input/output subsystem with a minimum of 32 dedicated input/output slots

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

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

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

multi-bay blade server: blade server requiring more than one bay for installation in a blade chassis

multi-output power supply unit: power supply unit designed to deliver the majority of its rated output power to more than one primary direct current (d.c.) output for the purpose of powering a computer server

NOTE 1: Multi-output power supply units may offer one or more standby outputs that remain active whenever connected to an input power source.

NOTE 2: The total rated power output from any additional power supply unit outputs, other than primary or standby outputs, is greater than or equal to 20 W.

network client (testing): computer or computer server that generates workload traffic for transmission to an Equipment Under Test connected via a network switch

network server: large network device which contains the same components as a computer server together with more than 11 ports, has a total line rate throughput of greater than or equal to 12 Gb/s and is designed to dynamically reconfigure ports and speed and to support a virtualized network environment, software defined networking

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.

networking equipment: device whose primary function is to pass data among various network interfaces, providing data connectivity among connected devices via the routing of data packets (e.g. routers and switches)

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

pedestal server: self-contained computer server that is designed with power supply units, cooling, input/output devices, and other resources necessary for stand-alone operation within a frame similar to that of a tower client computer

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

processor: central processing unit of the computer server comprising logic circuitry that responds to and processes the basic instructions that drive the server

product category: second-order classification or sub-type within a product group or form-factor that is based on product features and installed components

NOTE: Used in the present document to determine qualification and test requirements.

quarter-height blade server: blade server requiring one quarter the height of a standard server bay

rack-mounted server: computer server that is designed for deployment in a standard 19 inch ICT equipment rack as defined by CENELEC EN 60297 [i.1] or ETSI EN 300 119 [i.4]

NOTE: For the purposes of the present document, a blade server is considered under a separate product category and excluded from the rack-mounted product category.

reported maximum power: highest maximum power recorded on the eleven SERT™ worklet scores for the two tested configurations

resilient server: computer server designed with extensive Reliability, Availability, Serviceability (RAS) and scalability features integrated in the micro architecture of the system, Central Processor Unit (CPU) and chipset

NOTE: The requirements are listed in annex A.

server appliance: server that is not intended to execute user-supplied software, delivers services through one or more networks, is typically managed through a web or command line interface and is bundled with a pre-installed OS and application software that is used to perform a dedicated function or set of tightly coupled functions

server processor utilization: ratio of processor computing activity to full-load processor computing activity at a specified voltage and frequency, measured instantaneously or with a short term average of use over a set of active and/or idle cycles

server product family: group of servers sharing one chassis and motherboard combination that may contain multiple hardware and software configurations

single output power supply unit: power supply unit designed to deliver the majority of its rated output power to one primary direct current (d.c.) output for the purpose of powering a computer server

NOTE 1: Single-output power supply units may offer one or more standby outputs that remain active whenever connected to an input power source.

NOTE 2: The total rated power output from any additional power supply units outputs, other than primary and standby outputs, is not greater than 20 W.

NOTE 3: Power supply units that offer multiple outputs at the same voltage as the primary output are considered single-output power supply units unless those outputs are generated from separate converters or have separate output rectification stages, or have independent current limits.

single-wide blade server: blade server requiring the width of a standard blade server bay

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

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

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

NOTE 2: While storage products 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 product.

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 product.

storage server: enterprise storage device which contains the same components as a computer server together with $\geq$ 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.

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

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 computing activity, which stresses a particular characteristic of the system

NOTE: A floating point and integer performance stress code is an example of a CPU worklet.

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
APA	Auxiliary Processing Accelerator
BIOS	Basic Input/Output System
CENELEC	European Committee for Electrotechnical Standardization
CPU	Central Processor Unit
d.c., DC	Direct Current
DDR	Double Data Rate
DIMM	Dual In-line Memory Module
EEE	Energy Efficient Ethernet
EUT	Equipment Under Test
FPGA	Field Programmable Gate Array
GB	GigaByte
GPGPU	General-Purpose computing on Graphics Processing Units
HDD	Hard Disk Drive
HPC	Higher Performance Computing
I/O	Input/Output
ICT	Information and Communication Technology
ID	Identification
ISO/IEC	International Organization for Standardization/International Electrotechnical Committee

IT	Information Technology
JVM	Java™ Virtual Machine
LU	Lower-Upper
OS	Operating System
PCI	Peripheral Component Interconnect
PDU	Power Distribution Unit
PSU	Power Supply Unit
RAS	Reliability, Availability and Serviceability
rms	root mean square
SERT™	Server Efficiency Rating Tool
SOR	Successive Over-Relaxation
SPEC	Standard Performance Evaluation Corporation
SSD	Solid State Drive
SSJ	Server Side Java™
UPS	Uninterruptible Power Supply
VDC	Volts (DC)
W_{CPU}	Weighting of CPU
W_{Memory}	Weighting of Memory
$W_{Storage}$	Weighting of Storage

4 Server product categories and representative product family configurations

4.1 General

Servers are sold in different form factors, processor socket counts, resilience levels and configuration types and different groups of servers will have distinct performance capability and power demands.

Servers are categorized as defined in clause 4.4 by specific form factor and configuration parameters to enable the setting of appropriate idle power or active efficiency thresholds and assessment of like products with regards to procurement or market entry requirements.

To compare or evaluate systems, the evaluation or metric used shall only be made against like products. Like products are grouped into categories. Products of different categories shall not be compared using the metric of the present document.

Even though servers are classified in categories by the type of system, each server is customized by its configuration to best match the application for which they are being sold or purchased. As a result, a product is represented by a fixed set of configurations.

For an appropriate evaluation, the category shall be defined and the product family configurations itemized.

The product family configuration establishes a single group representing the efficiency of the product, as this then covers the range of configurations of that product that would be sold. Since the products sold are custom-configured, there should be a minimum of 2 configurations that would represent the family of configurations, bounded by the high-end configuration (see clause 4.5.2) and the low-end configuration (see clause 4.5.3).

4.2 Applications and metric applicability

Computer servers are architected in such a way as to be configurable to different groups of applications. The full configuration, including logical and physical elements, is optimized to deliver the most effective platform for operating those applications.

The software component of the metric is designed to execute typical real world applications and is designed to stress and assess the elements and associated functionality included in the server systems.

By stressing the elements in a fashion that replicates real world applications, a process for measuring workload output and its associated power demand is established.

Changes in application target, technology, configuration, or elements will impact the results and applicability of the metric. Therefore, it is necessary to establish categories of products which describe the elements which group similar products, determine representative configurations, and ensure applicability of the metric.

Categories or comparison groups of servers are formed by a combination of physical characteristics and limitations. The categories are separated based on computer architecture and physical differences that determine a different energy profile unique to that group. The server metric in the present document is targeted for general purpose servers and may not be applicable to certain categories of systems due to the elements described in clause 4.1.

4.3 Computer servers

4.3.1 General requirements

For the purposes of the present document a computer server shall meet all of the following criteria:

- be marketed and sold as a computer server;
- be designed for and listed as supporting one or more computer server operating systems (OS) and/or hypervisors;
- be designed such that all processors have access to shared system memory and are visible to a single OS or hypervisor;
- be targeted to run user-installed applications typically, but not exclusively, enterprise in nature;
- be packaged and sold with one or more a.c.-d.c. or d.c.-d.c. power supplies;
- provide support for error-correcting code and/or buffered memory (including both buffered dual in-line memory modules (DIMMs)) and buffered on board configurations).

4.3.2 Form-factors

For the purposes of the present document, three computer server form-factors are considered as defined in clause 3.1:

- blade;
- rack-mounted;
- pedestal.

4.4 Computer server product categories

The general purpose server metrics of the present document are applicable to the following server product categories as defined and uniquely characterized in clause 3.1:

- blade 1-socket server;
- blade 2-socket server;
- blade 2-socket resilient server;
- blade 4-socket resilient server;
- direct current server;
- pedestal 1-, 2- and 4-socket server;
- rack 1-socket server;
- rack 2-socket server;
- rack 4-socket server;

- rack 2-socket resilient server;
- rack 4-socket resilient server.

The requirements of a resilient server are specified in annex A.

4.5 Server product family configuration

4.5.1 General

A server product family configuration shall:

- be from the same model line or machine type;
- either share the same form-factor (i.e. rack-mounted, blade, pedestal) or share the same mechanical and electrical designs with only superficial mechanical differences to enable a design to support multiple form-factors;
- either share processors from a single defined processor series or share processors that plug into a common socket type;
- be dependent on one or more internal power supply unit(s).

For the purposes of defining the configurations of clauses 4.5.2 and 4.5.3, the following apply:

- Calculated Processor Capacity (dimensionless) = "the number of central processor units (CPUs)" x "the number of cores per CPU" x "the number of threads per core" x "CPU frequency (GHz)".

EXAMPLE 1: 2 CPUs, each CPU has 4 cores and 2 threads per core, frequency of 2,2

Processor capacity = 35,2.

- Calculated Memory Capacity (GB) = "the number of central processor units (CPUs)" x "the number of cores per CPU" x "the number of threads per core" which is subsequently rounded up to "the number of memory channels" x "the lowest capacity DIMM available for the product family".

EXAMPLE 2: 2 CPUs, each CPU has 4 cores and 2 threads per core

Memory capacity = 16 GB

If the product has 4 memory channels and the lowest capacity DIMM is 2 GB, then 8 x 2 GB_DIMMs (assuming 2 slots per channel).

4.5.2 "High-end" performance configuration

The "high-end" configuration is comprised of a minimum of two solid state drives (SSDs), with a minimum of 3 times the Calculated Memory Capacity of the EUT and with the Calculated Processor Capacity which represents the highest performance product model within the server product family.

All memory channels shall be populated with the same DIMM raw card design and capacity.

4.5.3 "Low-end" performance configuration

The "low-end" configuration is comprised of a minimum of two 10 000 rpm 3,5" HDDs, with a minimum of 1 time the Calculated Memory Capacity of the EUT and with the Calculated Processor Capacity which represents the lowest performance product model within the server product family.

All memory channels shall be populated with the same DIMM raw card design and capacity.

5 Metrics

5.1 Active state metric

5.1.1 Worklets

The SERT™ tool reports performance and power data for:

- a CPU metric comprising:
    - 6 CPU worklets i.e. Compress, LU, CryptoAES, SOR, Sort and SHA256;

NOTE 1: XML_validate is represented in the capacity metric and is therefore excluded from the CPU workload calculation.

    - 1 hybrid worklet: Hybrid SSJ.

NOTE 2: The Hybrid SSJ worklet is also considered as a CPU workload for the purposes of creating a single combined efficiency metric.

- 2 memory worklets: Flood3 and Capacity3;
- 2 storage worklets: Sequential and Random.

For each worklet, data is reported for a set of proportional performance intervals and associated, measured power values along with other test measurements.

The set of individually measured Performance and Power values with their associated efficiency value is termed "interval data".

5.1.2 Formulae

5.1.2.1 General

The geometric mean function is used to combine the interval data to produce a worklet efficiency score, the worklet efficiency scores to create workload (CPU, memory, storage) efficiency scores and the workload efficiency scores to create a single efficiency metric. Using the geometric mean prevents any single performance, power, worklet or workload efficiency score from unduly influencing the single metric.

In order to create a single energy efficiency metric for a server it is necessary to combine the interval Efficiency values for all the different worklets using the following general procedure:

a) combine the interval Efficiency values for the individual worklets using the geometric mean to obtain individual worklet Efficiency values for the worklet;

b) combining worklet Efficiency scores using the geometric mean function by workload type (CPU, Memory, Storage) to obtain a workload type value;

c) combining the three workload types using a weighted geometric mean function to obtain a single, total server Efficiency value.

In order to facilitate the deployed power assessment for each server configuration (see annex B), the Performance and Power scores can be combined using the same process as that above.

5.1.2.2 Active State Metric definition

The Active State Metric is defined as:

$$Eff_{server} = \exp\left[W_{CPU} \times \ln\left(Eff_{CPU}\right) + W_{Memory} \times \ln\left(Eff_{Memory}\right) + W_{Storage} \times \ln\left(Eff_{Storage}\right)\right] \quad (1)$$

or

$$Eff_{server} = Eff_{CPU}^{W_{CPU}} \times Eff_{Memory}^{W_{Memory}} \times Eff_{Storage}^{W_{Storage}} \quad (1a)$$

where W_{CPU}, W_{Memory} and $W_{Storage}$ are the weightings applied to the CPU, Memory and Storage worklets respectively (see clause 5.1.3) and:

$$Eff_{CPU} = \left(\prod_{i=1}^{7} Eff_i\right)^{1/7} \quad (2)$$

where: i = 1 for normalized interval efficiency of worklet$_{Compress}$,

i = 2 for normalized interval efficiency of worklet$_{LU}$,

i = 3 for normalized interval efficiency of worklet$_{SOR}$,

i = 4 for normalized interval efficiency of worklet$_{Crypto}$,

i = 5 for normalized interval efficiency of worklet$_{Sort}$,

i = 6 for normalized interval efficiency of worklet$_{SHA256}$, and

i = 7 for normalized interval efficiency of worklet$_{Hybrid\ SSJ}$.

$$Eff_{Memory} = \left(\prod_{i=1}^{2} Eff_i\right)^{1/2} \quad (3)$$

where: i = 1 for normalized interval efficiency of worklet$_{Flood3}$,

i = 2 for normalized interval efficiency of worklet$_{Capacity3}$.

$$Eff_{Storage} = \left(\prod_{i=1}^{2} Eff_i\right)^{1/2} \quad (4)$$

where: i = 1 for normalized interval efficiency of worklet$_{Sequential}$,

i = 2 for normalized interval efficiency of worklet$_{Random}$,

and:

$$Eff_i = 1\,000\ \frac{Perf_i}{Pwr_i} \quad (5)$$

where: Perfi: Geometric mean of the normalized interval performance measurements.

Pwri: Geometric mean of the measured interval power values.

5.1.3 Weightings

In the present document:

W_{CPU} is the weighting assigned to the CPU worklets = 0,65

W_{Memory} is the weighting assigned to the Memory worklets = 0,30

$W_{Storage}$ is the weighting assigned to the Storage worklets = 0,05

5.2 Idle state metric

Idle power is the alternating current (a.c.) power of the device with no activity either before or after running the SERT™ worklets. Two methods of measurement are possible:

- manual collection of the idle power data from the power meter;

- the SERT™ test automatically records and reports the idle power value.

Idle power is reported by SERT™. Idle power of the Equipment Under Test (EUT) typically scales based on the features and performance added. In order to assess idle power as a performance metric for the server, it is necessary to adjust any specific idle power limit based on the components installed on the test configuration or EUT. The specific aspects to be considered when determining an idle limit for a server include:

a) a base idle power allowance which is typically specified for a given form factor and number of processor sockets and covers a minimum quantity of memory in GB, a specified number of storage devices(s) which may be zero, and embedded I/O ports. The base idle power allowance is specified in Watts;

b) idle power adders for components including:

 1) memory: in W/GB;

 2) storage devices: W/device by device type or group of device types;

 3) I/O devices: in Watt per port or device, segregated by connection speed and/or interface type again with the potential to group by devices with similar power profiles;

 4) auxiliary processing accelerators: W/device segregated by device type (e.g. GPGPU, FPGA) where appropriate;

 5) CPU performance: this adder is determined using the geometric mean of the 7 SERT™ CPU worklets at the 100 % interval performance values, referred to as $Perf_{CPU}$. It requires a system performance multiplier, which is applied to the $Perf_{CPU}$ value to get the system performance adder in Watt. When calculating CPU performance for the performance adder calculations equation (6) is evaluated using the performance values of the 7 CPU worklets at the 100 % utilization level.

$$Perf_{CPU} = \left(\prod_{i-1}^{7} Perf_i\right)^{\frac{1}{7}} \quad (6)$$

Where:

i = 1 for normalized performance of Compress worklet at 100 % utilization

i = 2 for normalized performance of LU worklet at 100 % utilization

i = 3 for normalized performance of SOR worklet at 100 % utilization

i = 4 for normalized performance of Crypto worklet at 100 % utilization

i = 5 for normalized performance of SORT worklet at 100 % utilization

i = 6 for normalized performance of SHA256 worklet at 100 % utilization

i = 7 for normalized performance of SSJ worklet at 100 % utilization

The idle limit for a given configuration would be the sum of the base idle allowance plus the sum of the appropriate idle power adder allowances, also in Watt.

6 Test setup

6.1 General

A single test setup shall be used to undertake the measurements to determine both the active state and idle state metrics.

A separate test setup shall be used to evaluate the power supply efficiency.

6.2 Input power

During the test, the input voltage tolerance to the EUT shall be as specified below:

a) ≤ 1,0 % if power consumption is ≤ 1 500 W;

b) ≤ 4,0 % if power consumption is > 1 500 W.

For a.c. input voltages, the frequency tolerance shall be ≤ 1,0 % and the total harmonic distortion shall be as specified below:

a) ≤ 2,0 % if power consumption is ≤ 1 500 W;

b) ≤ 5,0 % if power consumption is > 1 500 W.

6.3 Environmental conditions

6.3.1 Ambient temperature

Ambient temperature shall be 25 ±5 °C.

6.3.2 Relative humidity

Relative humidity shall be between 15 % and 80 %.

6.4 Power analyser

The following requirements apply to single-phase power.

The power analyser shall report true root mean square (r.m.s.) power and at least two of the following parameters: voltage, current, and power factor.

The power analyser shall:

a) have a valid calibration certificate or equivalent, to support its use at the time the tests are carried out;

b) feature an available current crest factor of 3 or more at its rated range value; or

c) for power analysers that do not specify the current crest factor, the power analyser shall be capable of measuring an amperage spike of at least 3 times the maximum amperage measured during any 1 s sample;

d) have a minimum frequency response of 3,0 kHz;

e) have a minimum resolution of:

- 0,01 W for measurement values less than 10 W;
- 0,1 W for measurement values from 10 W to 100 W; and
- 1,0 W for measurement values greater than 100 W;

f) have a power measurement accuracy of no greater than 1,0 %;

g) have a logging performance of:

- minimum reading rate: one set of measurements (power measurement in W) per second;
- data averaging interval equal to the reading rate.

6.5 Temperature sensor

The temperature sensor shall:

a) have a temperature measurement accuracy of no greater than ±0,5 °C when measured no more than 50 mm in front of (upwind of) the main airflow inlet of the EUT;

b) have a logging performance of minimum reading rate: four samples per minute.

6.6 Active state test tool

The active state test tool shall be that provided by the Standard Performance Evaluation Corporation® (SPEC) and specified in the version 2 of SERT™ User Guide [3].

6.7 Controller system

The controller system shall be capable of the following functions:

a) start and stop each segment (phase) of the performance benchmark;

b) control the workload demands of the performance benchmark;

c) start and stop data collection from the power analyser so that power and performance data from each phase can be correlated;

d) store log files containing benchmark power and performance information;

e) convert raw data into a suitable format for benchmark reporting, submission and validation;

f) collect and store environmental data, if automated for the benchmark.

The controller system may be a server, a desktop computer, or a laptop and shall be used to record power from the equipment specified in clause 6.4 and temperature data from the equipment specified in clause 6.5.

The controller system and the EUT shall be connected to each other via one port of an Ethernet network switch.

6.8 General SERT™ requirements

Any additional requirements specified in the two following supporting documents shall be followed, unless otherwise specified in the test method in the present document:

a) SERT™ User Guide [3];

b) SERT™ Run and Reporting Rules [2].

Other supporting documents from SPEC® are:

a) SERT™ Design Document [i.5];

b) SERT™ JVM Options [i.6];

c) SERT™ Result File Fields [i.7].

7 Equipment Under Test (EUT)

7.1 Configuration

The configuration of the EUT shall be as specified in table 1.

Table 1: Configuration of EUT

A)	As-shipped condition	Products shall be tested in their "as-shipped" condition, which includes both hardware configuration 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. The power meter shall remain in place until all Idle and Active State 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).
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 servers with half-populated chassis configurations, the power supplies for the unpopulated power domains can be disconnected (table 2, D) for more information).
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.
F)	Storage	Products shall be tested for qualification with at least two HDD or two SSD installed. Products that do not include pre-installed drives (HDD or SSD) shall be tested using a storage configuration used in an identical model for sale that does include preinstalled drives. Products that do not support installation of drives (HDD or SSD) and, instead, rely exclusively on external storage solutions (e.g. storage area network) shall be tested using external storage solutions.
G)	Blade System and Dual/Multi-Node Servers	A blade system or dual/multi-node server shall have identical configurations for each node or blade server including all hardware components and software/power management settings. These systems shall also be measured in a way that ensures all power from all tested nodes/blade servers is captured by the power meter during the entire test.
H)	Blade Chassis	The blade chassis, at a minimum, shall have power, cooling, and networking capabilities for all the blade servers. The blade chassis shall be populated as specified in table 2, D). All power measurements for blade systems shall be made at the input of the blade chassis.
I)	BIOS and EUT System Settings	All BIOS settings shall remain as-shipped unless otherwise specified in the test method.

J)	Input/Output (I/O) and Network Connection	The EUT shall have at least one port connected to an Ethernet network switch. The switch shall be capable of supporting the EUT's highest and lowest rated network speeds. The network connection shall be live during all tests, and, although the link shall be ready and able to transmit packets, no specific traffic is required over the connection during testing. For the purpose of testing ensure the EUT offers at least one Ethernet port (using a single add-in card only if no on-board Ethernet support is offered).
K)	Energy Efficient Ethernet (EEE)	Products shipped with support for Energy Efficient Ethernet (compliant with IEEE 802.3az [5]) shall be connected only to Energy Efficient Ethernet compliant networking equipment during testing. Appropriate measures shall be taken to enable EEE features on both ends of the network link during all tests.

7.2 Test procedure

The EUT test configuration shall be in accordance with table 2.

Table 2: Test configuration of EUT

A)	The EUT shall be tested with all processor sockets populated.
B)	The EUT shall be installed in a test rack or other static location and shall not be physically moved until testing is complete.
C)	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.
D)	For a blade system, the blade server power consumption of the EUT shall be measured in the half-populated blade chassis configuration with an additional option of testing the EUT in the fully populated blade chassis configuration with an additional option of testing the EUT in the fully populated blade chassis configuration. For blade systems, the blade chassis shall be populated as follows: 1. Individual blade server configuration: All blade servers installed in the blade chassis shall be identical, sharing the same configuration (homogeneous). 2. Half -chassis population (required): The number of blade servers required to populate half the number of single-wide blade server slots available in the blade chassis shall be calculated. For blade chassis having multiple power domains, the number of power domains shall be chosen that is closest to filling half of the blade chassis. In a case where there are two choices that are equally close to filling half of the blade chassis, test with the domain or combination of domains which utilize a higher number of blade servers. EXAMPLE 1: A blade chassis supports up to 7 single-wide blade servers on two power domains. One power domain supports 3 blade servers and the other supports 4 blade servers. In this example, the power domain which supports 4 blade servers would be fully populated during testing, while the other power domain would remain unpopulated. EXAMPLE 2: A blade chassis supports up to 16 single-wide blade servers on four power domains. Each of the four power domains supports 4 blade servers. In this example, two of the power domains would be fully populated during testing, while the other two power domains would remain unpopulated. All user manual or manufacturer recommendations shall be followed for partially populating the blade chassis, which may include disconnecting some of the power supplies and cooling fans for the unpopulated power domains. If user manual recommendations are not available or are incomplete, then the following guidance shall be followed: • Completely populate the power domains. • If possible, disconnect the power supplies and cooling fans for unpopulated power domains. • Fill all empty bays with blanking panels or an equivalent airflow restriction for the duration of testing. 3. Full-chassis population (optional): • All available blade chassis bays shall be populated. • All power supplies and cooling fans shall be connected. • Proceed with all required tests in the test procedure as specified in table 3.

E)	The EUT shall be connected to a live Ethernet (IEEE 802.3 [4]) network switch. The live connection shall be maintained for the duration of testing, except for brief lapses necessary for transitioning between link speeds.
F)	The Controller System required to provide SERT™ workload harness control, data acquisition, or other EUT testing support shall be connected to the same network switch as the EUT and satisfy all other EUT network requirements. Both the EUT and Controller System shall be configured to communicate via the network.
G)	The power meter shall be connected to an a.c. voltage source set to the appropriate voltage and frequency for the test, as specified in table 1.
H)	The EUT shall be connected to the measurement power outlet on the power meter following the guidelines in table 1, B).
I)	The data output interface of the power meter and the temperature sensor shall be connected to the appropriate inputs of the Controller System.
J)	It shall be verified that the EUT is configured in its as-shipped configuration.
K)	It shall be verified that the Controller System and EUT are connected on the same internal network via an Ethernet network switch.
L)	Using a normal ping command, It shall be verified that the Controller System and EUT can communicate with each other.
M)	The most current SERT™ shall be installed on the EUT and the Controller System as specified in the most current SERT™ version 2 User Guide [3].

8 Measurement

8.1 Measurement for active state

The measurement shall be in accordance with table 3.

Table 3: Measurement of active state

A)	The EUT shall be re-booted. System caches and any stored information that may affect the active state metric shall be flushed.
B)	Between 5 and 15 minutes after the completion of initial boot or log in, the most current SERT™ version 2 User Guide [3] shall be followed to engage SERT™.
C)	All steps outlined in the most current SERT™ version 2 User Guide [3] shall be followed to successfully run SERT™. There shall be no manual intervention or optimization of the Controller System, EUT, or its internal and external environment during the execution of SERT™.
D)	Once SERT™ is completed, the following output files shall be included with all testing results: 1. Results.xml 2. Results.html 3. Results.txt 4. All results-chart png files (e.g. results-chart0.png, results-chart1.png, etc.) 5. Results-details.html 6. Results-details.txt 7. All results-details-chart .png files (e.g. results-details-chart0.png, results-details-chart1.png, etc.)

8.2 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 server, a server from the same family, or a server family will result in values that may vary from the initial testing.

Adjusting for a 90 % confidence interval, re-testing values that are within 15 % of the passing level shall be acceptable under re-test. Re-test or audit resulting in greater than 15 % 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 be 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.

8.3 Measurement for power supply

8.3.1 Measurement for internal power supply

The measurement shall be in accordance with table 4.

Table 4: Measurement of internal power supply

A)	For all types of internal power supplies, the efficiency and the power factor shall be measured at 10 %, 20 %, 50 % and 100 % of the rated [nameplate] output power.
B)	Test setup, test conditions, and measurement instrument specifications shall comply with clause 6.3.
C)	This test procedure assumes that the internal power supply meets the following criteria: • Detailed input and output ratings are available on the name plate or in manufacturer's literature, specifying the maximum loads that can safely be placed on each individual d.c. output voltage bus and, where necessary, groupings of those voltage busses. • The power supply has connectors that allow the d.c. output voltage busses to be connected and disconnected from the powered product non-destructively. • The power supply can be easily detached from the housing of the product it powers, without causing harm to other circuits and components of the product.
D)	In the event the above criteria are not met, a test board (see clause 8.3.2) shall be provided to enable testing.
NOTE 1: The power supply can be easily detached from the housing of the product it powers, without causing harm to other circuits and components of the product. NOTE 2: Such data could already be available from the manufacturer of the power supply; in such cases, the manufacturer could decide to use them. However, where 3rd party test results are used, it is the responsibility of the manufacturer to assess the trustworthiness of the sources. NOTE 3: The EPRI/ECOVA Generalized Test Protocol [i.2] is an acceptable basis for providing the required data.	

8.3.2 Measurement for test board power supply

8.3.2.1 General

Tests specified in this clause shall be made on either:

- the power supply of the computer under test, after it has been disconnected from the powered parts and extracted from the housing; or alternatively;
- another unit, representative of the built-in power supply.

8.3.2.2 Test loads

Active loads such as electronic loads or passive loads such as rheostats may be used as d.c. test loads. They shall be able to maintain the required current loading set point for each output voltage within an accuracy of ±0,5 %.

8.3.2.3 Test leads and wiring

Appropriate wires shall be used to avoid excessive overheating and reduce voltage drop across the wires. If measurements are not taken directly at the connector pins, voltage drop against the additional wires shall be taken into account.

NOTE: If applicable, voltage drop across the input wires will be subtracted from the measured input voltage, and the voltage drop across the output wires will be added to the measured input voltage for the measurement of power efficiency.

8.3.2.4 Warm up time

Whereas internal temperature of the components could impact its efficiency, the power supply under test shall be loaded up to the test load for a period of at least 15 minutes or until the reading over two consecutive five-minute intervals does not change by more than ±0,2 %.

8.3.2.5 Power measurements

The true RMS wattmeter used to carry out a.c. input power measurements shall meet the requirements of clauses 5.7 and 5.8 of CENELEC EN 62623:2013 [1]. Input power shall be determined using an averaging technique over a minimum of 32 input cycles utilizing the measurement instrument averaging function.

For appliances connected to more than one phase, the power measurement instrument shall be equipped to measure the total power of all phases connected.

D.c. output power measurements shall be made either with a suitably calibrated voltmeter and ampere meter or with a suitably calibrated power meter.

9 Measurement report

The following metrics/measurements shall be listed in the measurement report:

a) the active state metric;

b) the idle state metric.

It is accepted that the measurement accuracy of the metrics is ±15 %. Any subsequent assessment within this range shall be considered to be consistent with the quoted value.

The following additional information shall also be reported/provided under the measurement report:

1) automated SERT™ test report and supporting data;

2) author, site, and date of the testing;

3) product category;

4) product family and configuration type;

5) server configuration including:

    a) system test model number;

    b) manufacturer name;

    c) server product ID;

    d) component manufacturer, product ID, number of units and component product type for CPU, memory, drive (HDD or SSD);

6) Power Supply Unit test information or reference to previously conducted test report;

7) the extremes of server inlet test temperature during the test;

8) revision numbers for each of test software elements used;

9) Java™ revision and source used;

10) controller product model ID;

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

12) SERT™ tool suite revisions;

13) SERT™ product configuration revision.

Annex A (normative): Resilient server requirements

A.1 Reliability, Availability and Serviceability (RAS) features

Reliability features support a server's ability to perform its intended function without interruption due to component failures (e.g. component selection, temperature and/or voltage de-rating, error detection and correction).

Availability features support a server's ability to maximize operation at normal capacity for a given duration of downtime (e.g. redundancy - both at micro- and macro-level).

Serviceability features support a server's ability to be serviced without interrupting operation of the server (e.g. hot plugging).

A.2 Reliability, Availability and Serviceability (RAS) requirements

In order to classify as a resilient server all of the features listed in table A.1 shall be provided.

Table A.1: Resilient Server requirements

A)	**Processor RAS:** The processor shall have capabilities to detect, correct, and contain data errors, as described by all of the following: 1. Error recovery by means of instruction retry for certain processor faults. 2. Error detection on L1 caches, directories and address translation buffers using parity protection. 3. Single bit error correction (or better) on caches that can contain modified data. Corrected data is delivered to the recipient as part of the request completion.
B)	**System Recovery and Resiliency:** No fewer than six of the following characteristics shall be present in the server: a) Error recovery and containment by means of (1) data poison indication (tagging) and propagation which includes mechanism to notify the OS or hypervisor to contain the error, thereby reducing the need for system reboots. (2) Containment of address/command errors by preventing possibly contaminated data from being committed to permanent storage. b) The processor technology used in resilient and scalable servers is designed to provide additional capability and functionality without additional chipsets, enabling them to be designed into systems with 4 or more processor sockets. c) Memory Mirroring: A portion of available memory can be proactively partitioned such that a duplicate set may be utilized upon non-correctable memory errors. This can be implemented at the granularity of DIMMs or logical memory blocks. d) Memory Sparing: A portion of available memory may be pre-allocated to a spare function such that data may be migrated to the spare upon a perceived impending failure. e) Support for making additional resources available without the need for a system restart. This may be achieved either by processor (cores, memory, I/O) on-lining support, or by dynamic allocation/deallocation of processor cores, memory and I/O to a partition. f) Support of redundant I/O devices (storage controllers, networking controllers). g) Has I/O adapters or storage devices that are hot-swappable. h) Identify failing processor-to-processor lane(s) and dynamically reduce the width of the link in order to use only non-failing lanes or provide a spare lane for failover without disruption. i) Capability to partition the system such that it enables running instances of the OS or hypervisor in separate partitions. Partition isolation is enforced by the platform and/or hypervisor and each partition is capable of independently booting. j) Uses memory buffers for connection of higher speed processor -memory links to DIMMs attached to lower speed DDR channels. Memory buffer can be a separate, standalone buffer chip which is integrated on the system board, or integrated on custom-built memory cards.
C)	**Power Supply RAS**
	All PSUs installed or shipped with the server shall be redundant and concurrently maintainable. The redundant and repairable components may also be housed within a single physical power supply, but shall be repairable without requiring the system to be powered down. Support shall be present to operate the system in degraded mode.
D)	**Thermal and Cooling RAS**
	All active cooling components shall be redundant and concurrently maintainable. The processor complex shall have mechanisms to allow it to be throttled under thermal emergencies. Support shall be present to operate the system in degraded mode when thermal emergencies are detected in system components.

Annex B (informative): Deployed Power Assessment

B.1 Overview

This clause provides an overview of the Deployed Power Assessment methodology for the validation of a proposed server efficiency metric.

In order to evaluate potential metrics, a series of assessments are made of the relative power impact of selecting the server to a definitive confirmation of a metric's ability to predict efficiency of the server as deployed.

The assessment comprises of a graphical correlation and rank comparison between the power impact of a number of servers deployed to execute a defined workload versus the aggregated efficiency metric for the server. The selection of servers mimics the provisioning method used by IT professionals when determining the data processing needs of that location in which the servers are to be deployed. This is termed the Deployed Power Assessment method.

The Deployed Power Assessment is based on determining the ability of provisioning a set of servers for a targeted workload that results in a minimum expenditure of energy across the various utilization levels. The metric validation is to ensure that a better efficiency score will result in a lower deployed server power demand to execute that work. The use of a deployed power calculation enables differentiation between the effectiveness of a low performance, low power server and a high performance, high power server, as it enables an assessment of the number of servers and their associated energy use required to deliver a given workload in a given location or operating environment.

To calculate the number of servers needed to perform a workload and their associated deployed power, it is necessary to select a workload level to use to calculate the number of deployed servers:

1) Target workload should represent the maximum composite of the work targeted for the intended location. Based on actual workloads deployed, the weights used are 65 % CPU, 30 % Memory and 5 % Storage. The performance of each sub-workload category, e.g. CPU, Memory, and Storage, is the geometric mean of the maximum performance of each worklet in the category.

2) A value of "100 × the maximum performance value of the group of tested servers" defines the target workload performance level for the evaluation. This value is large enough to avoid quantization effects.

3) The following power/utilization/workload types are used in our comparison analysis. The intent is to assess a combination of workload types and power consumption levels that a set of servers will experience in an operating environment:

    a) idle power as measured by the SERT™ tool;

    b) geometric mean of the power for all workloads at the 25 % utilization (light workload);

    c) geometric mean of the power for all workloads at the 50 % utilization (medium workload);

    d) geometric mean of the power for all workloads at the 100 % utilization (heavy workload).

Multiple workloads and power use scenarios were assessed to validate that the combined metric is balanced and representative of efficiency across the range of workloads that servers are expected to perform and to avoid the assertion that use of a single power/workload/utilization might be biased to a particular outcome.

B.2 Determining the number of deployed servers

B.2.1 General

Any attempt to create a single efficiency metric based upon the SERT™ tool uses some method to combine the individual worklet values to create a single value. The aggregation used to determine a single value is with the geometric mean combinatory method and the designated component (CPU, memory and storage) weightings to calculate the number of servers required.

In order to determine the number of servers required for any given server model it is necessary to determine both a performance target for the dataset and a performance capability for each individual server. The individual worklet performance values reported in the SERT™ tool are combined and divided into the performance target to determine the number of servers required to meet the performance target.

B.2.2 Establishing target performance

In order to minimize quantization issues, since deployed power is based on an integral number of servers, a 100 times the weighted performance of the highest performance server in the data set is used for the performance target. The number of servers required to meet a desired performance level is calculated according to equation (B.1).

$$Deployed_{Qty_n} = Roundup\left(\frac{100 * Perf_{max}}{Perf_{weighted_n}}\right) \quad \text{(B.1)}$$

where:

- $Deployed_{Qty_r}$ = the number of servers deployed in the data centre to meet the target performance/workload.
- $Perf_{max}$ = the performance values for all servers in the data set.
- $Perf_{weighted_n}$ = the calculated weighted performance of server n.

B.2.3 Weighting factors

A weighted geometric mean (CPU 65 %: Memory 30 %: Storage 5 %) of performance, designated as the weighted performance, is used.

Equation (B.2) aggregates the geometric mean of the normalized performance values to obtain a single performance number for each individual server. The weighted maximum performance number represents the capability of the server and what would be used to provision the servers in the intended location. The value calculated from this equation is used as the denominator in equation (B.1):

$$Perf_{weighted} = EXP[0,65 * \ln(Perf_{CPU}) + 0,30 * \ln(Perf_{Memory}) + 0,05 * \ln(Perf_{Storage})] \quad \text{(B.2)}$$

where:

$Perf_{weighted}$ represents the newly calculated weighted performance;
$Perf_{workload}$ represents the geometric mean performance score for each workload from the SERT™ data base.

and:

$$Perf_{CPU} = \left(\prod_{i=1}^{7} Perf_i\right)^{1/7} \quad \text{(B.3)}$$

where: i = 1 for normalized interval performance of worklet$_{Compress}$,

i = 2 for normalized interval performance of worklet$_{LU}$,

i = 3 for normalized interval performance of worklet$_{SOR}$,

i = 4 for normalized interval performance of worklet$_{Crypto}$,

i = 5 for normalized interval performance of worklet$_{Sort}$,

i = 6 for normalized interval performance of worklet$_{SHA256}$, and

i = 7 for normalized interval performance of worklet$_{Hybrid\ SSJ}$.

$$Perf_{Memory} = \left(\prod_{i=1}^{7} Perf_i\right)^{1/2} \quad \text{(B.4)}$$

where: i = 1 for normalized interval performance of worklet$_{Flood3}$,

i = 2 for normalized interval performance of worklet$_{Capacity3}$.

$$Perf_{Storage} = \left(\prod_{i=1}^{2} Perf_i\right)^{1/2} \quad \text{(B.5)}$$

where: i = 1 for normalized interval performance of worklet$_{Sequential}$,

i = 2 for normalized interval performance of worklet$_{Random}$.

Annex C (informative): Alternative calculation

This is an alternative method of calculating the efficiency of a server. This formula also contains the Performance and Power metric of the server.

$$Eff_{server} = 1\,000\ \frac{Perf_{server}}{Pwr_{server}} \tag{C.1}$$

$$Perf_{server} = \exp\left[W_{CPU} \times \ln\left(Perf_{CPU}\right) + W_{Memory} \times \ln\left(Perf_{Memory}\right) + W_{Storage} \times \ln\left(Perf_{Storage}\right)\right] \tag{C.2}$$

$$Pwr_{server} = \exp\left[W_{CPU} \times \ln\left(Pwr_{CPU}\right) + W_{Memory} \times \ln\left(Pwr_{Memory}\right) + W_{Storage} \times \ln\left(Pwr_{Storage}\right)\right] \tag{C.3}$$

with:

$$Perf_{CPU} = \left(\prod_{i=1}^{7} Perf_i\right)^{1/7} \text{ and } Pwr_{CPU} = \left(\prod_{i=1}^{7} Pwr_i\right)^{1/7} \qquad \text{(C.4) (C.5)}$$

where:

i = 1 for normalized interval performance of worklet$_{Compress}$,

i = 2 for normalized interval performance of worklet$_{LU}$,

i = 3 for normalized interval performance of worklet$_{SOR}$,

i = 4 for normalized interval performance of worklet$_{Crypto}$,

i = 5 for normalized interval performance of worklet$_{Sort}$,

i = 6 for normalized interval performance of worklet$_{SHA256,}$ and

i = 7 for normalized interval performance of worklet$_{Hybrid\ SSJ}$.

$$Perf_{Memory} = \left(\prod_{i=1}^{2} Perf_i\right)^{1/2} \text{ and } Pwr_{Memory} = \left(\prod_{i=1}^{2} Pwr_i\right)^{1/2} \qquad \text{(C.6) (C.7)}$$

where:

i = 1 for normalized interval performance of worklet$_{Flood3}$,

i = 2 for normalized interval performance of worklet$_{Capacity3}$.

$$Perf_{Storage} = \left(\prod_{i=1}^{2} Perf_i\right)^{1/2} \text{ and } Pwr_{Storage} = \left(\prod_{i=1}^{2} Pwr_i\right)^{1/2} \qquad \text{(C.8) (C.9)}$$

where:

i = 1 for normalized interval performance of worklet$_{Sequential}$,

i = 2 for normalized interval performance of worklet$_{Random}$.

Annex D (informative): Bibliography

- Final Report: "Ecodesign Technical Assistance Study on Standards for ErP Lot 9 Enterprise Servers and Enterprise Data Storage".
- "Server efficiency Metric from SERT™ Worklet Results" DIGITALEUROPE Proposal, 5 August 2016, Version 2.0.
- ISO/IEC 30134-4: "Information technology - ICT sites - Key Performance Indicators - Part 4: IT Equipment Energy Efficiency for Servers ($ITEE_{SV}$)".
- Deployed power: "The Green Grid".
- European Environmental Citizens Organisation for Standards.

NOTE: Available at www.ecostandard.org.

- SPEC®: "Power and Performance Methodology".

NOTE: Available at https://www.spec.org/power/docs/SPEC-Power_and_Performance_Methodology.pdf.

- SPEC®: "Power Measurement Setup Guide".

NOTE: Available at https://www.spec.org/power/docs/SPECpower-Measurement_Setup_Guide.pdf.

- SPEC®: "PTDaemon Design Document".

NOTE: Available at https://www.spec.org/power/docs/SPEC-PTDaemon_Design.pdf.

- SPEC®: "SERT™ – Power and Performance".

NOTE: Available at https://www.spec.org/power/docs/SPEC-Power_Analyzer_Acceptance_Process.pdf.

- EPRI/ECOVA: "Generalized Test Protocol for Calculating the Energy Efficiency of Internal A.c.-D.c. and D.c-D.c. Power Supplies Revision 6.7", March 2014.
- IEEE 1100™: "IEEE Recommended Practice for Powering and Grounding Electronic Equipment".

History

Document history				
V1.0.0	June 2018	EN Approval Procedure	AP 20180927:	2018-06-29 to 2018-09-27
V1.1.0	January 2019	Vote	V 20190311:	2019-01-10 to 2019-03-11
V1.1.1	March 2019	Publication		