INTERNATIONAL STANDARD

ISO/IEC 30134-5

First edition
2017-11

Information technology — Data centres — Key performance indicators —

Part 5: IT Equipment Utilization for servers (ITEUsv)

Technologies de l'information — Centres de traitement de données — Indicateurs de performance clés —

Partie 5: Utilisation des appareils de technologies de l'information (TI) se rapportant aux serveurs

Reference number
ISO/IEC 30134-5:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/JTC 1, *Information technology*, Subcommittee SC 39, *Sustainability for and by Information Technology*.

A list of all parts in the ISO 30134 series can be found on the ISO website.

Introduction

The global economy is now reliant on information and communication technologies and the associated generation, transmission, dissemination, computation and storage of digital data. All markets have experienced exponential growth in that data for social, educational and business sectors and, while the Internet backbone carries the traffic, there are a wide variety of data centres at nodes and hubs within both private enterprise and shared/collocation facilities.

The historical data generation growth rate exceeds the capacity growth rate of the information and communications technology hardware and, with less than half of the world's population having access to an Internet connection as of 2014, that growth in data can only accelerate. In addition, with many governments having “digital agendas" to provide both citizens and businesses with faster broadband access, the very increase in network speed and capacity will, by itself, generate more usage (Jevon's Paradox). Data generation and the consequential increase in data manipulation and storage are directly linked to increasing power consumption.

With this background, it is clear that data centre growth, and power consumption in particular, is an inevitable consequence and that growth will demand increasing power consumption despite the most stringent energy efficiency strategies. This makes the need for Key Performance Indicators (KPIs) that cover the effective use of resources (including, but not limited to, energy) and the reduction of CO_2 emissions essential.

In order to determine the overall resource effectiveness or efficiency of a data centre, a holistic suite of metrics is required. For the resource effectiveness or efficiency of data centre infrastructures, power usage effectiveness (PUE) was defined in ISO/IEC 30134-2. PUE will be utilized to measure and improve energy efficiency of data centre infrastructures, such as cooling systems and power supply systems. For data centres that own not only infrastructure but also IT equipment, it is also necessary to measure and improve energy effectiveness or efficiency of their IT equipment. A data centre, which provides only infrastructure to the customer, is called a co-location data centre or housing service provider. For these data centres, PUE is essential. On the other hand, a data centre which owns and provides server, storage and network equipment is called a hosting or cloud service provider. These service providers can manage IT equipment and improve the energy effectiveness of a data centre by improving the energy effectiveness of the IT equipment which they own.

This document specifies the IT Equipment Utilization for servers in a Data Centre (ITEUsv) which specifies a method for measuring the average utilization of server equipment in a data centre. A data centre which owns servers can use this document to determine the average utilization of their current server equipment and to improve energy effectiveness by increasing the average utilization of servers. IT equipment installed in a data centre consists of servers, storage systems and network equipment. But it is difficult to calculate the summarized utilization among different types of IT equipment since the metrics for measuring their performance are different and a simple addition or average is not an appropriate method for summarizing. ITEUsv defines the method to obtain the average utilization of servers.

One approach to improve the resource effectiveness or efficiency of servers in a data centre is to

— improve the utilization ratio of servers by using such technologies as virtualization and server consolidation for sharing use of servers, and

— reduce the number of servers to achieve the same level of information processing.

ITEUsv is a KPI used to quantify the effects of this approach.

This document is part of a series of International Standards for such KPIs and has been produced in accordance with ISO/IEC 30134-1, which defines common requirements for a holistic suite of KPIs for data centre resource usage effectiveness or efficiency.

The ISO/IEC 30134 series does not specify limits or targets for any KPI and does not describe or imply, unless specifically stated, any form of aggregation of individual KPIs into a combined nor an overall KPI for data centre resource usage effectiveness or efficiency.

Information technology — Data centres — Key performance indicators —

Part 5: IT Equipment Utilization for servers (ITEUsv)

1 Scope

This document specifies the IT Equipment Utilization for servers (ITEUsv) as a Key Performance Indicator (KPI) to quantify the utilization of servers in a data centre. This document is intended as a KPI for improving the aggregate energy efficiency of servers in a given data centre.

This document

a) describes the purposes of ITEUsv,

b) defines ITEUsv in a conceptual manner,

c) describes how to use ITEUsv, and

d) describes reporting of ITEUsv.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 30134-1, *Information technology — Data centres — Key performance indicators — Part 1: Overview and general requirements*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 30134-1 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

3.1.1
Information Technology Equipment Utilization for servers
ITEUsv
average utilization ratio of all servers or a group of servers in a data centre

3.1.2
server
functional unit that provides services to workstations, to personal computers or to other functional units in a computer network

Note 1 to entry: Services may be dedicated services or shared services.

[SOURCE: IEC 60050-732:2010, 732-01-12, modified, Note 2 deleted.]

3.2 Abbreviated terms

CUS	CPU Utilization ratio of Server
CPU	Central Processing Unit
ITEEsv	IT Equipment Energy Efficiency for Servers
ITEUsv	IT Equipment Utilization for Servers

4 Relevance of ITEUsv

ITEUsv is a KPI which describes the utilization of the server equipment in the data centre in operational conditions. ITEUsv is developed with the knowledge that server energy efficiency tends to be optimal with higher utilization level. However, 100 % utilization may not necessarily be the optimal energy efficiency level. With some CPU architectures, high utilization, not peak utilization is more efficient.

Server energy effectiveness is a combination of

a) the capacity to do work per unit of energy (capability),

b) the amount of time the server is actually doing work (utilization), and

c) the ability of the server to reduce the energy consumption when the server is not fully loaded (power management).

ITEUsv accounts for utilization [see b)] and power management [see c)] aspects and is used to quantify the impact of one or both of the following:

— improving utilization ratio of servers by using such technologies as virtualization and server consolidation for sharing use of servers;

— reducing the number of servers to achieve the same level of information processing.

Capability [see a)] is addressed in ISO/IEC 30134-4.

NOTE ITEUsv does not involve data centre infrastructure such as air conditioning. IT equipment in the data centre includes servers, storage systems, and network equipment. ITEUsv only addresses utilization and power management of servers.

ITEUsv is a KPI intended for self-improvement of a given data centre and not for comparison among different data centres.

5 Determination of *ITEUsv(t)*

$CUS_i(t)$ is the average CPU utilization for server i at time t, measured by use of a performance monitoring tool provided by a server operating system. For performance monitoring tools, see A.2. When there are

N numbers of servers running in a data centre, the average CPU utilization of these servers at time t can be described by Formula (1).

$$ITEUsv(t) = \frac{\sum_{i=1}^{N} CUSi(t)}{N} \quad (1)$$

where

$CUS_i(t)$ is the CPU utilization ratio of server i at time t (%);

$ITEUsv(t)$ is the average CPU utilization of all servers or a group of servers in a data centre at time t;

N is the number of servers in a data centre or in a group running at time t.

NOTE When a server has multicore CPU, $CUS_i(t)$ is the average utilization of each core. When a server has multiple CPUs, $CUS_i(t)$ is the average utilization of each CPU.

When some of the servers are working at very low load, a data centre owner/operator can improve *ITEUsv(t)* by reducing the number of operating servers with virtualization techniques, while maintaining the whole workload of the data centre. If a data centre owner/operator utilizes server "power OFF" function to save energy consumption of servers in idle state, *ITEUsv(t)* shall be calculated by subtracting servers at "power OFF" state. The *ITEUsv(t)* value is high when most servers are running at high utilization status at time t. Energy effectiveness or efficiency of data centres can be promoted through encouraging larger *ITEUsv(t)* values.

6 Measurement of ITEUsv

The data centre shall monitor *ITEUsv(t)* for a year and collect all data. The data centre shall then calculate the annual average of *ITEUsv(t)*, which is the ITEUsv. ITEUsv indicates the average CPU utilization in a year for all servers or a group of servers. During the measurement, the maximum value of *ITEUsv(t)* and its time shall be recorded. The maximum value of *ITEUsv(t)*, or peak *ITEUsv(t)*, is useful to set a target value of ITEUsv. The definitional formula for ITEUsv is as shown by Formula (2):

$$\text{ITEUsv} = \frac{1}{a}\sum_{i=1}^{a}\left[ITEUsv\left(t_0 + e \times i\right)\right] \quad (2)$$

where

a is the number of *ITEUsv(t)* measurements intervals over a year (all intervals should be the same length);

t_0 is the starting time of measurement;

e is the interval of measurement, where $e \times a$ = one year.

The interval of measurement shall be between 1 min and 1 h. The data centre shall decide the interval based upon server operating conditions such as provisioning cycle, speed of change of the server load, change of server load in a day and characteristics of application.

In case the measurement was made in a partial section of the data centre or a group of servers in a data centre, then the measured area shall be stated. This group shall be identical and designated to a measurement group of ITEEsv. ITEUsv shall not be summed among these groups. Furthermore, when there is no CPU monitoring tool applicable to some servers, these servers shall not be included in these groups.

When monitoring the CPU utilization ratio, it is also common to record "duration of peak of CPU utilization". The server operator records how many seconds the *ITEUsv(t)* reaches its peak during a day

and calculates the percentage of the time in a day the *ITEUsv(t)* reaches its peak. This data is important to design a plan to improve ITEUsv.

ITEUsv relies on using CPU utilization as a proxy to the overall server utilization. A number of factors affects the effectiveness and/or accuracy of the use of CPU utilization as a proxy and should be taken into account. These factors include:

a) changes in the workload and/or its characteristics from original assessment;

b) memory utilization;

c) I/O utilization;

d) measurement differences between different CPU and OS architectures.

7 Application of ITEUsv

The use of this KPI allows the data centre owner/operator to establish an operational policy with effective operation of servers, in order to improve server energy efficiency. ITEUsv can be improved either by reducing the number of operating servers while maintaining the whole workload of the data centre or by maintaining the number of operating servers while increasing the whole workload of the data centre. Also, it can be used to check the improvement of data centre operational efficiency compared with the previous year. The KPI can assist in highlighting areas of improvement in data centre operation efficiency. By applying this KPI, the data centre owner/operator can improve the energy efficiency or effectiveness of the data centre.

The data centre operators should consider the implications of utilization peaks (>80 %) when using ITEUsv. If changes are made to increase average utilization an unacceptable number of peaks can occur which could impact application performance.

8 Reporting of ITEUsv

8.1 Requirements

For both calculating and reporting of ITEUsv, the data centre owner/operator shall disclose which measurement method (see Annex A) has been used, along with the values that serve as the basis of calculation, in order to secure reproducibility of calculated values.

The report shall include at least the following information:

a) the ITEUsv value;

b) the peak value of *ITEUsv(t)* and duration of peak of *ITEUsv(t)* (average minutes per day and per cent of the day being at peak utilization);

c) the measurement method of $CUS_i(t)$ (name of performance monitoring tool and measurement frequency or interval);

d) the start and end date of the one year measurement period;

e) the data centre identification (e.g. name, address, owner/operator).

8.2 Recommendations

The report should include the following information:

a) the purpose of reporting;

b) a historically recorded data of ITEUsv;

c) a reference to this document, i.e. ISO/IEC 30134-5.