TECHNICAL REPORT

ISO/IEC TR 20913

First edition
2016-11-15

Information technology — Data centres — Guidelines on holistic investigation methodology for data centre key performance indicators

Technologies de l'information — Centres de données — Lignes directrices relatives à la méthodologie de recherche holistique pour les indicateurs de performance clé du centre de données

Reference number
ISO/IEC TR 20913:2016(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 meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information.

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 39, *Sustainability for and by Information Technology*.

Introduction

The ISO/IEC 30134 series defines key performance indicators (KPIs) for data centre resource effectiveness. There are many aspects to be considered in order to improve data centre resource effectiveness. As for resources, it may include not only energy, but also water and other natural resources. As for data centre components, they include air conditioning, power supply, servers, storages, and network equipment. However, it is difficult to include all aspects into one KPI, so multiple KPIs are under development, which measure each aspects of resource effectiveness improvement. Resource effectiveness improvement in each aspect will be performed by measuring each KPI. On the other hand, there is a need to observe the state and trend of data centre as a whole, or holistically, by monitoring multiple KPIs in a single view. Analysis of the KPIs from the overall perspective is also referred to as a holistic investigation method. This document describes a spider web chart-based method and control chart method extending the functionality of the conventional spider web chart for viewing and analysing KPIs for data centre resource effectiveness. It also investigates considerations for applying holistic investigation methods to resource effectiveness evaluation of multiple data centre KPIs. The usefulness and applicability of holistic methods are discussed using a SWOT analysis. The methods described in this document are intended for analysis and continuous improvement of a specific data centre and not for comparing different data centres.

Information technology — Data centres — Guidelines on holistic investigation methodology for data centre key performance indicators

1 Scope

This document describes backgrounds, motivation, and general concept of holistic methodology for data centre key performance indicators (KPIs) to investigate the status of KPIs. It discusses the usefulness of holistic investigation methodology in terms of aggregating a KPI across different contexts, aggregation of two or more KPIs within a single context, aggregation of two or more KPIs across multiple contexts, and aggregation of the multiple KPIs into a single indicator. This document presents a conventional spider web chart-based data centre KPIs status observation method and a control chart method including upper bound and lower bound of the operational status of KPIs. This document presents SWOT analysis results for both methodologies. The methods described in this document are aimed at the self-monitoring of a data centre, not comparison among data centres.

Specifically, this document

a) describes backgrounds, motivation, and general concept of holistic investigation methodology for data centre KPIs,

b) analyses the usefulness of holistic investigation methodology for aggregating KPIs,

c) describes a spider web chart-based KPIs status observation method and a control chart extending spider web chart to observe the operational status of KPIs,

d) describes alternative and/or additional methods of representing dissimilar KPIs to track holistic resource effectiveness of the data centre, and

e) presents SWOT analysis results for holistic investigation methods described in this document.

2 Normative references

There are no normative references in this document.

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions 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
holistic investigation method
data centre resource effectiveness investigation method considering multiple key performance indicators

3.1.2
spider web chart
chart that consists of multiple performance indicators which are set in a circle like a spider web

3.2 Abbreviated terms

IT	Information Technology
ITEEsv	IT Equipment Energy Efficiency for Servers
ITEUsv	IT Equipment Utilization for Servers
KPI	Key Performance Indicator
PUE	Power Usage Effectiveness
REF	Renewable Energy Factor
SWOT	Strength Weakness Opportunity Threat

4 Background and motivation

4.1 General concept of holistic investigation method

Improving the resource effectiveness and carbon footprint of a data centre requires the monitoring and analysis of multiple KPIs. ISO/IEC JTC 1/SC 39 has determined that it is impractical to aggregate multiple KPIs to determine the overall energy effectiveness of a data centre. There is a need to observe the state and trend of multiple KPIs in a single view.

With any performance indicator, it is necessary to understand the expected upper and lower limits and general behaviour of the performance indicator. There are typically two approaches that are applicable to holistic investigation of data centre KPIs:

— Engineering/modeling method: This method has been used to establish baseline performance. This methodology requires the development of an optimized economic and engineering model based on creating an idealized benchmark specific to each utility — incorporating the topology, demand patterns, and population density of the service territory. Typical limitations of this approach are as follows: the engineering models that support it can be very complicated, and the structure of the underlying components relationships can be obscured through a set of assumed coefficients used in the optimization process.

— Performance benchmarking method: This method includes a set of specific performance measurement indicators, such as volume billed per worker, consumed energy per product, quality of service (continuity, water quality, complaints), coverage, and key financial data. Usually, these indicators are presented in ratio form to control the scale of operations. These partial measures are generally available and provide the simplest way to perform comparisons: trends direct attention to potential problem areas.

Among the methods mentioned above, the performance benchmarking method is useful for evaluating the resource efficiency of data centres because ISO/IEC JTC 1/SC 39 is offering a selection of energy effectiveness KPIs. The performance benchmarking method may be further categorized into two types: performance indicator-based methods and chart-based methods.

— Performance indicator-based methods: In this category, the performance of the target is evaluated by developing performance indicators for the target. For example, Hz for CPU and bytes for storage are typical performance indicators. This category allows accurate performance evaluation and comparison among targets, if the performance indicators are defined. Typical limitation of this approach is that it is difficult to compare the evaluation results if performance indicators belong to different dimensions with different units.

— Chart-based methods: This category depicts the target's performance by using chart methods, such as pie, bar, line, and spider web, etc. This category is useful for evaluating performance by displaying multiple performance indicators, making analysis easier.

Since the chart-based approach supports multiple performance indicators simultaneously, it is appropriate for a holistic method. The spider web chart in particular is well suited for the display and analysis of multiple KPIs. A spider web chart is useful for displaying multiple KPIs in a single chart. It is also useful for displaying multiple measurement values of several KPIs in a single chart, for example, temporal measurement values of several KPIs. Thus, this document focuses on the spider web chart-based holistic KPI investigation methods. It is noted that the chart-based approach, especially spider web chart, has typical issues for applying a KPI investigation, such as scaling and normalization of KPI values, KPIs with different dimensions, ordering of KPIs in the chart, graphical interpretation of the chart, and so on. These typical issues are discussed in Clause 6 in detail.

4.2 Usefulness of spider web chart methods for visualizing data centre KPIs

The spider web chart consists of a bundle of performance indicators which are set in a circle. The indicators are usually normalized from zero to one, one indicating the highest possible performance, but unnormalized indicators may be utilized. Individual axes may need to be inverted in order for the different indicators to correlate. It is clear that the quality of the spider web charts depends on the validity, reliability, and comprehensiveness of the performance indicators. It is known that the spider web chart has strength on visualizing the status of performance indicators.

Regarding visualization capability, spider web charts provide a synoptic description of multiple performance measures and make trade-offs between performance measures visible. Figure 1 shows a spider web chart consisting of three sets of performance measurements and five performance indices. In the figure, the values of each index are originally measured and unnormalized ones, and the farther from centre of the chart implies the better. Each green, blue, and red polygon connecting measurement values of five index shows a single observation of the five indices, respectively. Using the chart, it is possible to visually compare the performance achievement among multiple performance measurements and indicators.

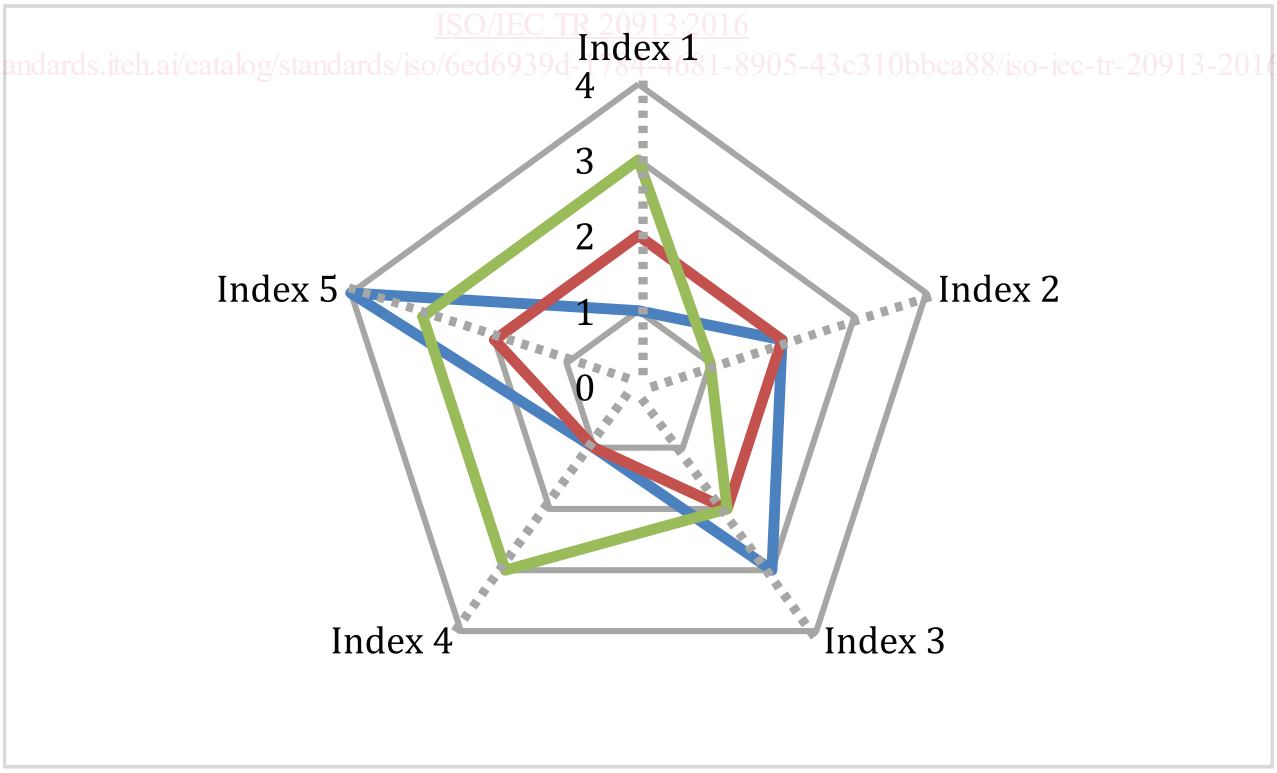

Figure 1 — Example of a spider web chart consisting three performance measurements (green, blue, and red)

Due to the advantages, spider web charts are popularly used to assess the performance of various evaluation objectives and to present a visual comparison of performance in various fields, especially business management. As discussed in this clause, the visualization capability of a spider web chart can help data centre administrators to monitor the specified performance KPIs of the data centre and their changes so that they can improve the efficiency of the data centre. For example, by regularly constructing the spider web chart showing the state of each KPI, the data centre administrator can effectively monitor the temporal behaviour of the KPIs. Further, the spider web chart has general advantages for assessing data centre KPIs rather than conventional charts such as bar chart when assessing the multiple measurement values of multiple KPIs. The advantages of a spider web chart are discussed in Clause 6 in detail.

4.3 Usefulness of aggregating data centre KPIs

The key objective of aggregating multiple performance indicators into a single indicator is to represent the overall achievement of each indicator as a single and integrated output. However, there are well-known problems with aggregating heterogeneous performance indicators. Each indicator can have a different dimension and scale, so aggregating multiple indicators by normalizing their original values can lose the characteristics of each indicator. Additionally, the aggregation process can cause a serious problem to be masked or a minor issue can be overstated depending on how the individual indicators are scaled. Further, depending on the dimension of indicators, it may be inappropriate to aggregate multiple indicators into a single indicator.

However, if all indicators are measured with the same dimension, aggregating multiple indicators into a single indicator may be useful. For example, if an indicator measures the operational achievement ratio of a KPI and its operational target value, the achievement ratio explains whether the data centre is operated effectively according to the operational target value. The operational target value of a KPI indicates the intended threshold for the KPI. Assume that a KPI measures the utilization ratio of IT server equipment and the administrator of the IT server sets upper bound and lower bound of the KPI as the operational target values. If the measured KPI value exceeds the upper bound of the KPI, the administrator may consider to install more IT servers in order to reduce the utilization ratio of IT servers. Whereas the administrator may consider to consolidate underutilized IT servers if the measured KPI value is below the lower bound of the KPI. Thus, by integrating the achievement ratios of each KPI into a single value, the data centre administrator can easily determine whether the data centre is operated as planned. It should be noted that observing the aggregated number of the measured KPI values may overlook the detailed characteristics. For example, by looking at the aggregate number of measured KPI values for server utilization, the overuse of one server may be masked by the underuse of another. Careful review of the individual server data for such events should be conducted to avoid data masking issues that may occur during KPI aggregation. However, even in this case, the relative importance (e.g. weighting) of each indicator is obscured. Thus, the aggregated overall operational achievement helps with management of the temporal changes in data centre operational efficiency. For example, let us assume that a data centre regularly examines the values of the overall operational achievement. At some time, if the overall operational achievement of the data centre is below the threshold for the data centre, the statuses of each KPI will be investigated and KPIs of which achievement is less than threshold could be managed by an administrator. Once the overall operational achievement value of the data centre exceeds the threshold, only the overall operational achievement value may be used to regularly manage the data centre.

5 Spider web chart-based KPIs status observation method

A holistic approach enables awareness of the effect of changes made to the data centre specific from an overall viewpoint by use of various efficiency metrics. A holistic approach helps the operator keep in mind the effects on all metrics simultaneously.