J Supercomput (2013) 63:639–656
DOI 10.1007/s11227-011-0704-3

Review of performance metrics for green data centers: a taxonomy study

Lizhe Wang · Samee U. Khan

Published online: 13 October 2011

Abstract Data centers now play an important role in modern IT infrastructures. Although much research effort has been made in the field of green data center computing, performance metrics for green data centers have been left ignored. This paper is devoted to categorization of green computing performance metrics in data centers, such as basic metrics like power metrics, thermal metrics and extended performance metrics i.e. multiple data center indicators. Based on a taxonomy of performance metrics, this paper summarizes features of currently available metrics and presents insights for the study on green data center computing.

Keywords Data center · Green computing · Performance metrics

1 Introduction

A data center is a home to computational power, storage, and applications necessary to support an enterprise business. A data center is central to modern IT infrastructure, as all content is sourced from or passes through it. Electricity usage is the most expensive portion of a data center's operational costs. In fact, the U.S. Environmental Protection Agency (EPA) reported that 61 billion KWh, 1.5% of US electricity consumption, is used for data center computing [31]. Additionally, the energy consumption in data centers doubled between 2000 and 2006. Continuing this trend, the EPA estimates that the energy usage will double again by 2011. It is reported that

L. Wang (✉)
Pervasive Technology Institute, Indiana University, 2719 E. 10th St., Bloomington, IN 47408, USA
e-mail: Lizhe.Wang@gmail.com

S.U. Khan
Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58108-6050, USA
e-mail: samee.khan@ndsu.edu

power and cooling cost are the most dominant costs in data centers [40]. Thus there is growing pressure from the general public for the IT society to offer "green" data center infrastructures and services [46].

There are two major and complementary methods [33] to build a "green" data center: (1) involve "green" elements in the design and building process of a data center, (2) "greenify" the process of running and operating a data center in everyday usage. Various research activities have been carried out to manage data centers in a "green" mode (the latter method), such as reducing data center temperature [28, 39, 42], increasing server utilization [21, 26, 37], and decreasing power consumption of computing resources [7, 8, 12, 20]. A fundamental research topic for the above study is how we can define performance metrics to identify how "green" a data center is. "Green" performance metrics, for data centers, are a set of measurements that can qualitatively or quantitatively evaluate the environmental effects by operating a data center. Nowadays, the term "green computing" requires clear scientific definitions and it is more or less a marketing concept. The research on developing green data center metrics brings solid specifications and definitions for green data centers in the following aspects:

- identify and specify clearly how "green" a data center is, for example by calculating its energy efficiency or greenhouse gas emission per time unit,
- evaluate data center products and compare similar data centers,
- track "green" performance to increase a data center's "green" efficiency, and
- provide guidance to engineers, manufacturers, and service providers to develop research and development on future green data center technologies.

Despite the importance of green metrics, there has been relatively little activity on organizing and summarizing performance metrics for green data centers compared with a large body of work that has been done on data center power management and optimization. This paper tries to fill this gap by reviewing related performance metrics for green data center computing and makes a taxonomy study on these metrics. The contribution of this paper is two fold: (1) it presents research on performance metrics definition and evaluation, and (2) it develops a taxonomy study on related data center green performance metrics for its daily operation and management.

The rest of this paper is organized as follows. First, Sect. 2 introduces how performance metrics are achieved via current available benchmarks in data centers. Section 3 presents a study to organizing performance metrics for green data centers. Sections 4–7 discuss basic performance metrics such as, greenhouse gases, humidity, and thermal & energy metrics. Sections 8–9 detail the extended green data center metrics of multiple indicators and total ownership cost. Section 10 concludes the paper.

2 Green data center practice and benchmarking

A performance metric is typically associated with a benchmark or a workload and typical execution environments. A benchmark is used to emulate real-world computing tasks. A data center benchmark is the act of running a set of programs, or other operations, in order to assess the relative performance of data center performance metrics. Typically data center benchmarking is carried out under a certain workload, which is artificially generated or produced in real life: real or artificial workload → benchmark → performance metrics. Although some benchmark workloads are available, for example, server-side Java under various loads for SPECpower [23], external sort for JouleSort [32, 33], data center benchmarking is generally taken under normal data center practice and workload.

As this paper is devoted to the research of green metrics for data centers, we are more interested in two types of benchmark: server level benchmarks and data center level benchmarks. Server level benchmarks, like GCPI, Green500 and JouleSort, normally measure energy consumption in a compute server or a cluster. Data center level benchmarks, such as, HVAC Effectiveness Index and Computer Power Consumption Index from LBNL, present a measurement for the entire data center.

2.1 Computer server level benchmark

This section introduces several performance metrics for computer server level performance evaluation: GCPI, SPECpower, Green500 project, and JouleSort.

SiCortex [36] proposes an inclusive metric for comparing energy efficiency in High-Productivity Computing segment,

The Green Computing Performance Index (GCPI) [13] is an inclusive metric for comparing energy efficiency in High-Productivity Computing segment proposed by SiCortex [36]. The GCPI analyzes computing performance-per-kWatt across a spectrum of industry-standard benchmarks, providing organizations with much-needed guidance in the era of out-of-control data center energy consumption. It is declared that the GCPI is the first index to measure, analyze and rank computers on a broad range of performance metrics relative to energy consumed.

In 2006, the SPEC community started to establish SPECpower [23], an initiative to augment existing SPEC benchmarks with a power/energy measurements. SPECpower's workload is a Java application that generates and completes a mix of transactions; the reported throughput is the number of transactions completed per second over a fixed period. SPECpower reports the energy efficiency in terms of overall operations per watt. This metric represents the sum of the performance measured at each target load level divided by the sum of the average power (in watts) at each target load.

The Green500 project [11] aims at increasing awareness about the environmental impact and long-term sustainability of high-end supercomputing by providing a ranking of the most energy-efficient supercomputers in the world. The Green500 list uses "performance per watt (PPW)" as its metric to rank the energy efficiency of supercomputers. The performance is defined as the achieved maximal performance, GFLOPS (Giga FLoating-point OPerations per Second), by the Linpack [10] benchmark on the entire system. The power is defined as the average system power consumption during the execution of Linpack with a problem size that delivers.

JouleSort [32] is an I/O-centric benchmark that measures the energy efficiency of systems at peak use. It is an extension of the sort benchmark, which is used to measure the performance and cost-performance of computer systems. The JouleSort

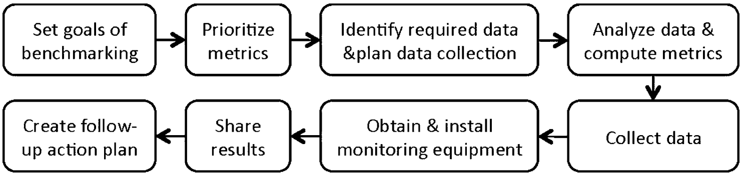


Fig. 1 Data center benchmarking process [24]

benchmark measures the cost for doing some amount of work, which reflects some measure of power use, e.g. average power, peak power, total energy, and energy-delay.

2.2 Data center level benchmarking

Data center level benchmarks are used to evaluate how "green" a data center is and is also used to compare similar data centers. Lawrence Berkeley National Laboratory (LBNL) releases a set of efforts and practices for data center benchmarking [15]. These practices include improved air management, emphasizing control and isolation of hot and cold air streams; rightsizing central plants and ventilation systems to operate efficiently both at inception and as the data center load increases over time; optimized central chiller plants, designed and controlled to maximize overall cooling plant efficiency, central air-handling units, in lieu of distributed units. Figure 1 shows the formal process for benchmarking a data center [24]. A set of software is developed, such as Calarch, Comis, DoE-2, EnergyPlus, Genopt, home energy saver, for benchmarking and designing energy efficient buildings and data centers.

3 A study on green performance metrics

Developing data center green performance metrics must meet the following criteria:

– Technically sound: The metrics should be adaptable to various data center types, including internal data centers and outsourced data centers, multi-tier data centers, and data centers with redundancy.
– Holistic and balanced system evaluation: A green performance metric should evaluate the whole data center system, without overemphasizing or overlooking any components of a data center.
– All-featured evaluation: Green performance metrics should be able to capture a data center's behavior at typical workloads and operation modes.
– Low cost to implement: Complex and expensive implementation of performance metrics may circumscribe its wide adoption.
– Ease of use: Data center management staff and users would be glad to adapt to some metrics, which are easy to use and intuitive to understand.

Typically, a green performance metric for data centers is defined with the following semantic:

$$\text{Performance metric} = \frac{\text{Useful work}}{\text{Environmental cost}} \tag{1}$$

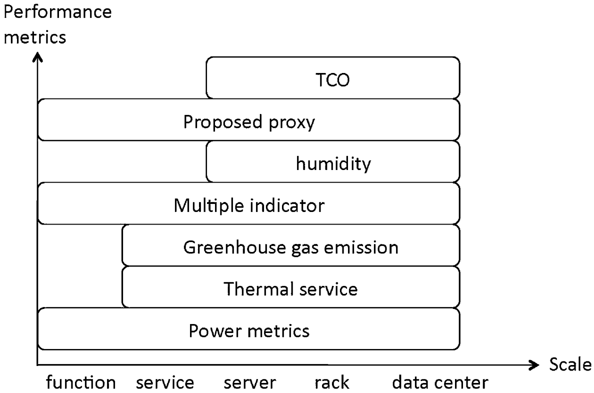


Fig. 2 Green computing performance metrics vs. grain

There are several research challenges for defining green performance metrics:

– There are multiple roles, such as data center owners, data center administrators and users (who join the management), and operation & usage of data centers. It is thus difficult to define "useful" work for all participants of a data center.
– A data center contains multiple hardware/software subsystems and components. It requires sophisticated efforts to define "useful work" to characterize various data center subsystems.
– The term of environmental cost is difficult to define. Normally, several green related parameters are used, such as temperature, power or energy usage, humidity and CO_2 emission.

We also classify data center performance metrics based on (1) the measure grain: function, service and application, server, rack, or the entire data center, and (2) viewpoints of data center users or owners. Data center users are interested in the environmental impacts of their applications in different scales, like per function or per service. Data center owners and administrators, on the other side, are more interested how "green" the resources in a data center are, like per server, per rack, or the whole data center. Different performance metrics can be obtained in various grains. For example, power metrics can be measured or estimated in a wide range of per bit, per instruction, per service, or per rack. CO_2 and thermal metrics can be measured in coarser grain, for example, per service, per server, or per rack. Other metrics, like Total Cost of Ownership (TCO) and humidity, must be measured in a coarse grain, for example, in a server level, a rack level, or the whole data center. Figure 2 shows the measure grains for performance metrics.

In this paper, performance metrics for green data center computing are divided into two categories: basic metrics and extended metrics. The basic metrics, such as greenhouse gas, humidity, thermal metrics, and power metrics, deliver an intuitive image of "how green a data center is". The extended metrics are functions of basic metrics, such as data center indicators and TCO, which give an in-depth look on a data center. Figure 3 overviews the taxonomy of green data center performance metrics and Sects. 4–9 discuss them in detail.

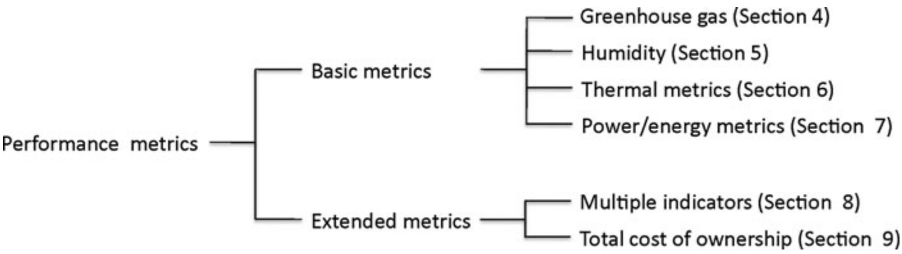


Fig. 3 Green computing performance metrics taxonomy

4 Greenhouse gas emission

Greenhouse gases are gases in the atmosphere that absorb and emit radiation within the thermal infrared range. This process is the fundamental cause of the greenhouse effect. Carbon dioxide (chemical formula CO_2) is an important greenhouse gas, which contributes 9%–26% of the greenhouse effect [22]. A data center consumes significant amount of power and a mass of greenhouse gas is produced in the process of power generation. Therefore, the energy consumption for data center computing can produce huge amounts of greenhouse gas.

It is hard to measure the CO_2 emission that is produced by data center computing due to the following reasons [41]:

- Data centers are complex with shared infrastructures by a number of users and manage support. A fine-grain calculation of CO_2 produced by an application, a user, or a computing server is difficult.
- CO_2 produced is affected by multiple factors, for example, data center efficiency and energy sources.

Wang et al. [47, 48] report a coarse estimation of CO_2 emission in a data center, New York State University at Buffalo:

1. Total power consumption of a university data center is measured.
2. Electrical power source break down of the data center is analyzed. For example, Table 1 shows the supplied power distribution and CO_2 emission to generate one unit of power (kWh) for the data center in New York State University at Buffalo.
3. Then the total CO_2 emission per hour can be calculated.
4. The CO_2 emission for some specific service can be estimated if its power consumption can be measured.

5 Humidity

Humidity is a measurement of moisture content in the air [9]. High humidity in a data center can cause more hardware failures and can also increase a data center's cooling costs. Thus, humidity control is important for physical media, like tape storage, and is generally not critical for rest of the data center equipment [24].

Table 1 Breakdown of power sources in New York State and CO_2 generation

Power source	CO_2 emission (g/kWh)	Power source break down
Oil	881	16%
Coal	963	14%
Natural gas	569	22%
Nuclear	6	29%
Hydroelectric	4	17%
Wind Power	3–22	≤2%

The traditional way of measuring humidity in the data center is to look at relative humidity. Relative humidity is given as a percentage and measures the amount of water in the air at a given temperature compared to the maximum amount of water that air can hold. American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) [3] recommends the relative humidity be within the 40% to 55% range [30] and the range between 20% and 80% is “acceptable”.

Relative Humidity Difference (RHD) is used to characterize the difference of the return and supply air relative humidity in the data center:

$$\mathrm{RHD} = \mathrm{Rhumidity} - \mathrm{Shumidity} \tag{2}$$

where Shumidity is the supply air relative humidity and Rhumidity is the return air relative humidity. A small relative humidity difference range suggests opportunities to reduce energy use, especially if there is active humidification and dehumidification [24].

6 Thermal metrics

It is reported that cooling costs can be up to 50% of the total energy cost in a data center [34]. Therefore thermal metrics are quite important for operating a green data center. Figure 4 shows a taxonomy of thermal metrics: data center temperature, British Thermal Unit (BTU), Airflow performance index, and Cooling system efficiency metrics.

6.1 Data center temperature

Operating computing systems for extended periods of time at high temperatures greatly reduces reliability, longevity of components and will likely decrease QoS. An ambient temperature range of 68°F to 75°F (20°C to 24°C) is optimal for system reliability [30]. It is recommended that expensive IT equipment should not be operated in a computer room or data center where the ambient room temperature has exceeded 85°F (30°C) [1].

The difference between the supply and return air temperature in a data center is also of interests. It is proposed by the ASHRAE guidelines [4] that: (1) a temperature range of 64°F–80°F is recommended; (2) a temperature range of 59°F–90°F is allowed. A low supply air temperature, and a small temperature differential between

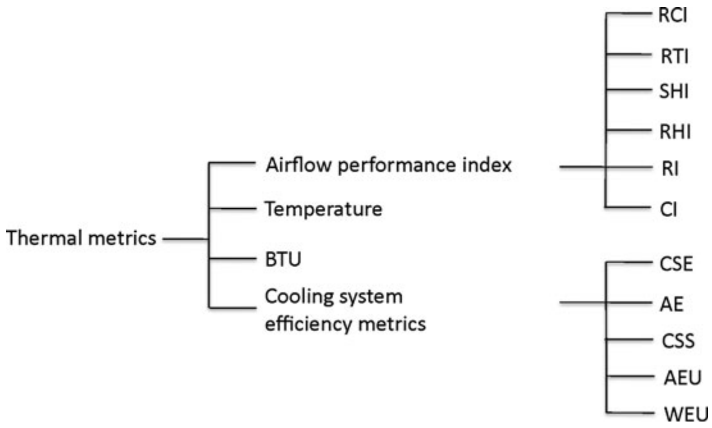


Fig. 4 Thermal metrics taxonomy

supply and return air temperature, typically indicates the opportunity to improve air management, raise supply air temperature, and thereby reduce energy use.

6.2 British Thermal Unit

British Thermal Unit (BTU) [49] is used to manage solutions for cooling a server room or a data center. A BTU is the amount of energy required to raise the temperature of a pound of water 1°F. There are several factors that contribute to how much cooling power a specific room or data center requires. For example, to calculate the BTU for cooling power for a data center, there are several rules of thumbs:

– The size of a server room has a lot to do with how many BTUs are required to cool down a server room or datacenter. The space contribution for cooling a data center is $\text{BTU} = 330 \times \text{Length} \times \text{Width}$.
– Another factor for cooling your data center is the equipment. A data center, or a server room, demands BTU for the equipment: $\text{BTU} = 3.5 \times$ total wattage running equipments.
– Lighting generates a noticeable amount of heat in a data center. Total BTU for lighting: $\text{BTU} = 4.25 \times$ wattage of lighting.

6.3 Airflow performance index

Airflow performance and cooling system efficiency are used to label the efficiency of a data center's thermal aspect. This section introduces how to evaluate the performance of airflow in data center and the next section discusses how to calculate the cooling system efficiency. For more information of the data center cooling system and airflow, please refer to [25, 29].

6.3.1 Rack Cooling Index

The Rack Cooling Index (RCI) is designed to measure how effectively equipment racks are cooled and maintained within industry thermal guidelines and standards [17]. There are two RCI metrics:

- RCI_{HI}
  The RCI_{HI} is a measure of the absence of over-temperatures, which characterizes the equipment health at the high end of the temperature range. 100% means that no over-temperatures exist. The lower the percentage the greater risk the equipment experiences excessive intake temperatures.
- RCI_{LO}
  The RCI_{LO} complements the RCI_{HI} when the supply conditions are below the minimum recommended temperature. The significance of a low RCI_{LO} is the potential for harmful relative humidity levels and that the equipment may not be qualified at such temperatures.

6.3.2 Return Temperature Index

The Return Temperature Index (RTI) is a measure of the energy performance of the air management [18], whose objective is to isolate hot and cold airstreams. The RTI takes both the supply and return temperatures, and maximizes the difference between them while keeping the inlet temperatures within ASHRAE recommendations. It also allows reduction of the system air flow rate. This strategy allows the HVAC equipment to operate more efficiently. The RTI is ideal at 100% wherein the return air temperature is the same as the temperature leaving the IT equipment and the supply air temperature is the same as the rack inlet temperature.

6.3.3 Supply Heat Index and Return Heat Index

The metrics of the Supply Heat Index (SHI) and Return Heat Index (RHI) [35] are introduced to assess the magnitude of recirculation and mixing of hot and cold streams. The SHI and RHI are based on non-dimensional parameters that describe the geometry of the data center and they measure the degree of separation between cool supply and warm return streams. RHI measures the extent to which cool supply air mixes with the warm return stream, while SHI measures the extent to which warm return air mixes with cool supply air. The values of RHI and SHI are between 0 and 1. High RHI and low SHI values denote good separation of air streams.

6.3.4 Recirculation Index

The Recirculation Index (RI) [43] is an airflow-based cooling performance index for racks for analyzing raised-floor cold-aisle clusters. The RI represents the ratio of a rack's inlet airflow that did not come directly from the perforated tiles.

6.3.5 Capture Index

The Capture Index (CI) [44] is proposed as a cooling performance metric based solely on the airflow patterns associated with the supply of cool air to, or the removal of hot air from, a rack. The cold-aisle CI is defined as the ratio of air ingested by the rack that originates from local cooling resources (e.g., perforated floor tiles or local coolers). The hot-aisle CI is defined as the ratio of air exhausted by a rack that is captured by local extracts (e.g., local coolers or return vents). The value of CI is between 0 and 1. A high CI value indicates a better cooling performance.

6.4 Cooling system efficiency metrics

6.4.1 Data Center Cooling System Efficiency

The Data Center Cooling System Efficiency (CSE) characterizes the overall efficiency of the cooling system (including chillers, pumps, and cooling towers) in terms of energy input per unit of cooling output. It is an average value depicting the average power of the cooling system with respect to the cooling load in the data center:

$$\text{CSE} = \frac{\text{Average cooling system power usage}}{\text{Average cooling load}} \tag{3}$$

Based on data from the LBNL practice [24], 0.8 kW/ton could be considered as good practice benchmark and 0.6 kW/ton as a better practice benchmark.

6.4.2 Airflow Efficiency

Airflow Efficiency (AE) characterizes overall airflow efficiency in terms of the total fan power required per unit of airflow. This metric provides an overall measure of how efficiently air is moved through the data center, from the supply to the return, and takes into account low pressure drop design as well as fan system efficiency:

$$\text{AE} = 1000 \times \frac{\text{total fan power}}{\text{Total fan airflow}} \tag{4}$$

There are limited data on airflow efficiency in data centers. The data from the LBNL study [24] suggest that 0.5 W/cfm might be considered a threshold of better practice. A high value of AE indicates that the fan system (motors, belts, drives) is inefficient and the pressure drops airflow distribution system need to be reduced. Improving the design of the duct work can significantly reduce the pressure drop in the system.

6.4.3 Cooling System Sizing

Cooling System Sizing (CSS) [24] is the ratio of the installed cooling capacity to the peak cooling load:

$$\text{CSS} = \frac{\text{Installed Chiller Capacity}}{\text{Peak Chiller Load}} \tag{5}$$

A high CSS value indicates a good potential and scalability of the cooling system.

6.4.4 *Air Economizer Utilization*

Air Economizer Utilization (AEU) [24] characterizes the extent to which air-side economizer system is being used to provide "free cooling". It is defined as the percentage of hours in a year that the economizer system can be in either full or complete operation (i.e. without any cooling being provided by the chiller plant):

$$\text{AEU} = \frac{\text{Air economizer hours}}{24 \times 365} \tag{6}$$

A low value for AEU indicates potential for increasing energy savings from using an air-side economizer system [24].

6.4.5 *Water Economizer Utilization*

Water Economizer Utilization (WEU) is the percentage hours in a year that the water side economizer system meets the entire cooling load of the data center.

$$\text{WEU} = \frac{\text{Water economizer hours}}{24 \times 365} \tag{7}$$

WEU provides information on the energy savings from using a water-side economizer system. Increasing the chilled water temperatures to the chiller plant increases the hours of economizer use [24].

7 Power/energy metrics

As discussed in the first section, current data centers consume tremendous amounts of power. Power/energy metrics of various scales and various components for data center computing have been proposed and applied in data center computing practice. Figure 5 shows a summary of popular power related performance metrics.

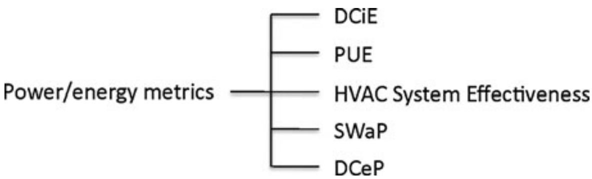
7.1 DCiE: Data Center Infrastructure Efficiency

The Data Center Infrastructure Efficiency, DCiE or DCE [5, 45], is an industry accepted metric. It is defined by

$$\text{DCiE} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}} \tag{8}$$

where we have the following.

Fig. 5 Power consumption metrics taxonomy



– IT Equipment Power
  This includes the load associated with all of the IT equipment, i.e. computation, storage and network equipment, along with supplemental equipment i.e. KVM switches, monitors, and workstations/laptops used to monitor or otherwise control the data center.
– Total Facility Power
  This includes all IT Equipment power as described above plus everything that supports the IT equipment load such as:
  · Power delivery components i.e. UPS, switch gear, generators, PDUs, batteries and distribution losses external to the IT equipment
  · Cooling system components i.e. chillers, computer room air conditioning units, direct expansion air handler units, pumps, and cooling towers
  · Other miscellaneous component loads such as data center lighting

The LBNL practice [24] suggests that a DCiE value of about 0.5 is considered typical practice and 0.7 and above is better practice. Some data centers are capable of achieving 0.9 or higher [14].

7.2 PUE: Power Usage Effectiveness

An organization for green computing provides the following standard metric for calculating energy efficiency of the data centers: Power Usage Effectiveness (PUE) [5]:

$$\text{PUE} = \frac{1}{\text{DCiE}} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}} \tag{9}$$

PUE shows the relation between the energy used by IT equipments and energy used by other facilities, such as cooling needed for operating the IT equipment. For example, a PUE of 2.0 indicates that for every watt of IT power, an additional watt is consumed to cool and distribute power to the IT equipment. At the present time, the PUE of a typical enterprise data center is around between 1 to 3.

7.3 HVAC System Effectiveness

The HVAC (Heating, Ventilation, and Air Conditioning) system of a data center typically includes the computer room air conditioning and ventilation, a large central cooling plant, and lighting & other minor loads. The HVAC System Effectiveness is the fraction of the IT equipment energy to the HVAC system energy. The HVAC system energy is the sum of the electrical energy for cooling, fan movement, and any other HVAC energy use like steam or chilled water. The HVAC System Effectiveness is calculated as follows:

$$\text{HVAC Effectiveness} = \frac{\text{IT}}{\text{HVAC} + (\text{Fuel} + \text{Steam} + \text{Chilled Water}) \times 293} \tag{10}$$

where IT: Annual IT Electrical Energy Use. HVAC: Annual HVAC Electrical Energy Use. Fuel: Annual Fuel Energy Use. Steam: Annual District Steam Energy Use. Chilled Water:Annual District Chilled Water Energy Use.

The HVAC System Effectiveness denotes the overall efficiency potential for HVAC systems. A higher value of this metric, compared to the peers, means higher potential to reduce HVAC energy use [24].

7.4 SWaP: Space, Watts and Performance

The metric of SWaP (Space, Watts and Performance) [38] characterizes a data center's energy efficiency by introducing three parameters of space, energy and performance together. The SWaP is calculated as follows:

$$\mathrm{SWaP} = \frac{\mathrm{Performance}}{\mathrm{Space} \times \mathrm{Power\ Consumption}} \tag{11}$$

where Performance: Using industry-standard benchmarks. Space: Measuring the height of the server in rack units (RUs). Power: Determining the watts consumed by the system, using data from actual benchmark runs or vendor site planning guides.

The SWaP metric gives users an effective cross-comparison and total view of a server's overall efficiency. Users are able to accurately compare the performance of different servers and determine which ones deliver the optimum performance for their needs. The SWaP can help users better plan for current and future needs and control their data center costs.

7.5 DCeP: Data Center energy Productivity

The DCeP [2, 16] is proposed to characterize the resource consumed is energy for useful work in a data center. The DCeP is calculated as follows:

$$\mathrm{DCeP} = \frac{\mathrm{Useful\ work\ produced}}{\mathrm{Total\ energy\ consumed\ to\ produce\ that\ work}} \tag{12}$$

Useful work is the tasks performed by the hardware within an assessment window. The calculation of total energy consumed is the kWh of the hardware times the PUE of the facility.

8 Multiple indicators

8.1 Data center indicator

Data center indicators contain several performance metrics, which render an overview image on how "green" a data center is from multiple aspects simultaneously. As shown in Fig. 6, we summarize several key data center indicators [6] as follows:

– DCiE is discussed in Sect. 7.1.
– Server utilization counts the activity of the processor relative to its maximum ability in the highest frequency state.
– Data center utilization measures the amount of power consumed by IT equipments relative to the actual capacity of the data center.

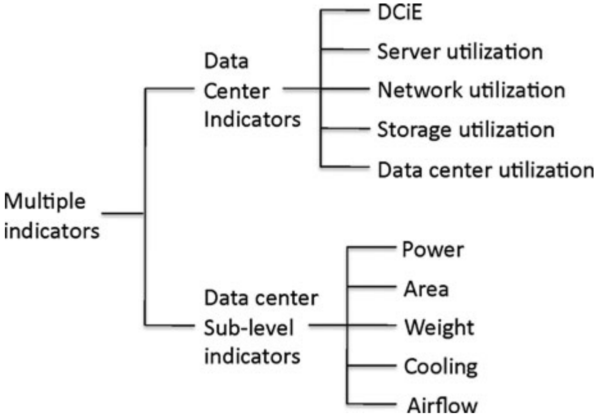


Fig. 6 Multiple indicators taxonomy

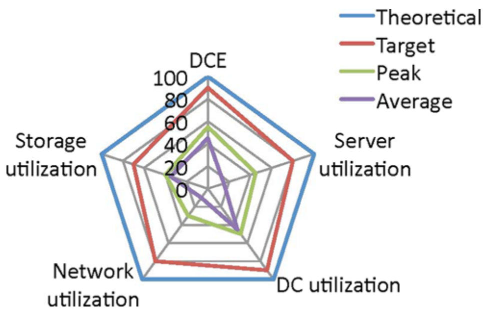


Fig. 7 Data center indicators [6]

– Storage utilization is the percentage of storage used relative to the overall storage capacity within a data center.
– Network utilization is the ratio of bandwidth used relative to the bandwidth capacity in a data center.

Figure 7 shows multiple evaluation levels of a data center with data center indicators.

8.2 Data center sub-level indicators

For above data center indicators, we perform the same type of analysis by developing multiple sub-indicators to locate an in-depth look at the limitations and overall utilization of each axis topic [6]. For example, for the data center utilization, five sub-indicators are often studied: power, cooling, airflow, weight, and area. Figure 8 shows a sample evaluation for data center sub-level indicators.

9 TCO: Total Cost of Ownership

Total Cost of Ownership (TCO) represents the cost to the owner to purchase/build, operate, and maintain a data center [27]. Typically a data center TCO can be split into two parts: the capital expenses and the operational expenses. The capital cost

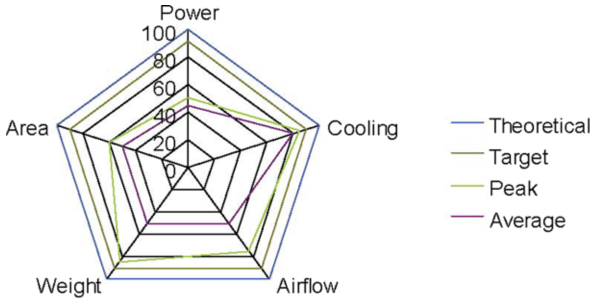


Fig. 8 Data center sub-level indicators [6]

is the initial investments for purchasing and building a data center. Since main subsystems of a data center, such as power, cooling and space, roughly scale linearly with energy consumed, it is therefore meaningful to characterize the capital cost in terms of dollars per watt for a data center. As a rule of thumb, building a large data center typically costs around 12–15$/Watt and small data center building costs more. In general, 80% of capital cost goes to power and cooling, and constructing a data center costs the remaining 20% [19].

Operational cost refers to the recurring monthly costs of actually running a data center [19]. Many factors can affect the operational cost, for example, implementation techniques, climate, management cost. Typically in USA to operate large data centers cost 0.02–0.08 $/Watt excluding electricity cost.

10 Summary and discussion

Performance metrics are important for designing, building and evaluating a data center. It also encourages the use of new technologies for next generation data centers. This paper has discussed a taxonomy for green data center performance metrics. The taxonomy includes basic performance metrics, for example, greenhouse gas emission, humidity, power & thermal metrics, and extended metrics. The key driving features to develop above metrics include:

– Economic features: For most industry data centers, to save money is the most concern to make their data centers “green”.
– Environmental concern: a green data center can save energy, thus making itself environment respect, for example by reducing CO_2 emission.

Currently, there are numerous performance metrics available and some prospective metrics will be proposed in the near future. However, there still are some requirements to be addressed:

– Metrics for balancing purchasing cost and operation cost.
  In this paper we mostly discuss the performance metrics for operating a data center. It is known that a green data center may be expensive in the initial phases of building the center and purchasing the resources for the center. Therefore, metrics are needed to show how the operation cost of a data center responses to its initial purse cost.

– Metrics for balancing compute performance and green performance.
  A green data center must balance the requirements of achieving compute performance and green performance. A performance metric still required is the "scale-down" factor, which reflects how computing performance behaves when green performance changes.

References

1. 42U (2009) Website http://www.42u.com/power/data-center-temperature.htm
2. 42U (2011) DCeP: Data Center energy Productivity. Web Page http://www.42u.com/measurement/dcep.htm
3. American Society of Heating, Refrigerating and Air-Conditioning Engineers (2011) Web Page http://www.ashrae.org/
4. ASHRAE (2008) ASHRAE environmental guidelines for datacom equipment—expanding the recommended environmental envelope. Technical report, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc
5. Belady C (2007) The green grid data center efficiency metrics: PUE and DCIE. Technical report, The Green Grid, Feb. 2007
6. Belady C (2007) The green grid productivity indicator. Technical report, The Green Grid, Feb. 2007
7. Chen Y, Das A, Qin W, Sivasubramaniam A, Wang Q, Gautam N (2005) Managing server energy and operational costs in hosting centers. In: SIGMETRICS '05: proceedings of the 2005 ACM SIGMETRICS international conference on measurement and modeling of computer systems, New York, NY, USA, 2005. ACM Press, New York, pp 303–314
8. Das R, Kephart JO, Lefurgy C, Tesauro G, Levine DW, Chan H (2008) Autonomic multi-agent management of power and performance in data centers. In: AAMAS '08: proceedings of the 7th international joint conference on autonomous agents and multiagent systems, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems, New York, pp 107–114
9. Data center humidity levels source of debate (2009) Website, June 2009
10. Dongarra J, Bunch J, Moler C, Stewart G (2011) Web Page http://www.netlib.org/linpack/
11. Feng WC, Cameron KW (2007) The Green500 list: encouraging sustainable supercomputing. Computer 40(12):50–55
12. Freeh VW, Lowenthal DK (2005) Using multiple energy gears in mpi programs on a power-scalable cluster. In: Pingali K, Yelick KA, Grimshaw AS (eds) PPOPP. ACM Press, New York, pp 164–173
13. Green Computing Performance Index (2011) Web Page http://www.sicortex.com/green_index
14. Greenberg S, Khanna A, Tschudi W (2009) High performance computing with high efficiency. Am Soc Heat Refrig Air-cond Eng (ASHRAE) Trans 179
15. Greenberg S, Mills E, Tschudi B, Rumsey P, Myatt B (2006) Best practices for data centers: lessons learned from benchmarking 22 data centers. Technical report, Lawrence Berkeley National Laboratory
16. Haas J et al (2008) A framework for data center energy productivity. Technical report, The Green Grid, Feb 2008
17. Herrlin MK (2005) Rack cooling effectiveness in data centers and telecom central offices: the Rack Cooling Index (RCI). Am Soc Heat Refrig Air-cond Eng (ASHRAE) Trans 111
18. Herrlin MK (2007) Improved data center energy efficient and therman performance by advanced airflow analysis. In: Digital power forum, San Francisco, CA, USA, September 2007
19. Hoelzle U, Barroso LA (2009) The datacenter as a computer: an introduction to the design of warehouse-scale machines. Morgan and Claypool, San Rafael
20. Jiang N, Parashar M (2009) Enabling autonomic power-aware management of instrumented data centers. In: IPDPS '09: proceedings of the 2009 IEEE international symposium on parallel & distributed processing, Washington, DC, USA. IEEE Computer Society, Los Alamitos, pp 1–8
21. Jung G, Joshi KR, Hiltunen MA, Schlichting RD, Pu C (2009) A cost-sensitive adaptation engine for server consolidation of multitier applications. In: ACM/IFIP/USENIX 10th international middleware conference, pp 163–183
22. Kiehl JT, Trenberth KE (1997) Earths annual global mean energy budget. Bull Am Meteorol Soc 78:197–208

23. Lange K-D (2009) Identifying shades of green: the specpower benchmarks. Computer 42:95–97
24. Mathew P, Ganguly S, Greenberg S, Sartor D (2009) Self-benchmarking guide for data centers: metrics, benchmarks, actions. Technical report, Lawrence Berkeley National Laboratory, Berkeley, California, July 2009
25. McAllister S, Carey VP, Shah A, Bash C, Patel C (2008) Strategies for effective use of exergy-based modeling of data center thermal management systems. Microelectron J 39(7):1023–1029
26. Mehta S, Neogi A (2008) Recon: a tool to recommend dynamic server consolidation in multi-cluster data centers. In: IEEE/IFIP network operations and management symposium: pervasive management for ubioquitous networks and services, pp 363–370
27. Patterson MK, Costello DG, Grimm PF, Loeffler M (2007) Data center TCO; a comparison of high-density and low-density spaces. Technical report, Intel Corp, Jan 2007
28. Patel CD, Sharma R, Bash CE, Beitelmal A (2002) Thermal considerations in cooling large scale high compute density data centers. In: Thermal and thermomechanical phenomena in electronic systems, 2002, ITHERM 2002. The eighth intersociety conference on, pp 767–776
29. Rambo J, Joshi Y (2007) Modeling of data center airflow and heat transfer: state of the art and future trends. Distrib Parallel Databases 21(2–3):193–225
30. Recommended Data Center Temperature & Humidity (2009) Website www.avtech.com/About/Articles/AVT/NA/All/-/DD-NN-AN-TN/Recommended_Computer_Room_Temperature_Humidity.htm, June 2009
31. Report to Congress on Server and Data Center Energy Efficiency (2011) Available at www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf
32. Rivoire S, Shah MA, Ranganathan P, Kozyrakis C (2007) Joulesort: a balanced energy-efficiency benchmark. In: SIGMOD '07: proceedings of the ACM SIGMOD international conference on management of data, New York, NY, USA, 2007. ACM Press, New York, pp 365–376
33. Rivoire S, Shah MA, Ranganathan P, Kozyrakis C, Meza J (2007) Models and metrics to enable energy-efficiency optimizations. Computer 40(12):39–48
34. Sawyer R (2004) Calculating total power requirements for data centers. Technical report, American Power Conversion
35. Sharma RK, Bash CE, Patel RD (2002) Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers. In: 8th ASME/AIAA joint thermophysics and heat transfer conf, 2002
36. SiCortex (2011) Web Page http://www.sicortex.com/
37. Singh A, Korupolu MR, Mohapatra D (2008) Server-storage virtualization: integration and load balancing in data centers. In: Proceedings of the ACM/IEEE conference on high performance computing, p 53
38. SWaP (Space, Watts and Performance) Metric (2011) Web Page http://www.sun.com/servers/coolthreads/swap/
39. Tang Q, Gupta SKS, Varsamopoulos G (2008) Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: a cyber-physical approach. IEEE Trans Parallel Distrib Syst 19(11):1458–1472
40. The green grids opportunity: decreasing datacenter and other IT energy usage patterns (2007) Technical report, the green grid, Feb. 2007
41. van den Berghe S (2009) Metrics and an API for green computing on the cloud. Web Page, May 2009
42. Vanderster DC, Baniasadi A, Dimopoulos NJ (2007) Exploiting task temperature profiling in temperature-aware task scheduling for computational clusters. In: Asia–Pacific computer systems architecture conference, pp 175–185
43. VanGilder JW, Shrivastava SK (2006) Real-time prediction of rack-cooling performance. Am Soc Heat Refrig Air-cond Eng (ASHRAE) Trans 112:151–162
44. VanGilder JW, Shrivastava SK (2007) Capture index: an airflow-based rack cooling performance me. Am Soc Heat Refrig Air-cond Eng (ASHRAE) Trans
45. Verdun G (2007) The Green Grid metrics: data center infrastructure efficiency (DCiE) detailed analysis. Technical report, The Green Grid, Feb 2007. Available at http://www.thegreengrid.org/Global/Content/white-papers/DCiE-Detailed-Analysis
46. Wang L (2011) IEEE TCSC technical area of green computing. Web Page http://sites.google.com/site/greencomputingproject/
47. Wang L, Khan S, Dayal J (2011) Thermal aware workload placement with task-temperature profiles in a data center. J Supercomput. doi:10.1007/s11227-011-0635-z

48. Wang L, von Laszewski G, Huang F, Dayal J, Frulani T, Fox G (2011) Task scheduling with ann-based temperature prediction in a data center: a simulation-based study. Eng Comput. doi:10.1007/s00366-011-0211-4
49. Weaver T (2011) Cooling your datacenter/Server Room Temperature Control. http://www.bestpricecomputers.ltd.uk/servers/datacenter-cooling.htm