A Survey on Sustainability in ICT
A Computing Perspective

Alexandru-Adrian Tantar, Emilia Tantar
Interdisciplinary Centre for Security, Reliability and Trust
University of Luxembourg
4 rue Alphonse Weicker
L-2721 Luxembourg, Luxembourg
{alexandru, emilia}.tantar@uni.lu

ABSTRACT

The rise of the data centers industry, together with the emergence of large cloud computing that require large quantities of resources to be maintained, brought the need of providing a sustainable development process. Through this paper we aim to provide an introductory insight on the status and tools available to tackle this perspective within the evolutionary and genetic algorithms community. Existing advancement are also emphasized and perspectives outlined.

Keywords

evolutionary computation; sustainable ICT; energy efficiency

1. INTRODUCTION

A constant increase in the ICT sector energy consumption and the rising concern for environmental impact led to what we now refer to as sustainable computing. And while the concept may seemingly sound straightforward, it is far from being obvious, as we will later see in this paper. The sustainable computing abstraction *per se* spans a wide semantic area. Among others, it connects to energy-aware management of resources, energy efficient computing or the impact such paradigms have in other fields. At the same time, most studies only offer and, what is more, help to create an incomplete picture, e.g. by referring to resource allocation strategies or power management alone. While still relevant, such studies tend however to offer an answer that needs to fit a problem more complex than we first imagined. End of life sustainability accounts for much more than just what energy management can offer. In the following we address sustainability by laying down a series of questions and problems which, when all connected, offer a more coherent picture of the impact ICT and computing have.

As a high level definition, sustainability is often described by relying on the notion of carbon footprint. Why exactly carbon? In ordinary terms, first due to the fact that all our actions, from simply being alive, can be quantified in terms of carbon emissions, i.e. offered an idea of how green a system is. And second due to the common connection we tend to make between carbon emissions and environment destruction. A natural way of approaching sustainability, in this setup, is to reduce the (equivalent) carbon footprint, even if this only means carbon offsets, i.e. investing in areas that result in carbon emission reduction [73]. As an alternative, one may refer to a decrease in energy consumption. What these 'symbols' stand for and what they conceal, i.e. with respect to carbon footprint and energy efficiency, is in most cases more than what is delivered. For a better understanding of the entire process, one needs to think of greenhouse gas emissions, e.g. carbon dioxide, methane or nitrous oxide, or to rely on information about the nature of the energy sources used like, for example, renewable vs nuclear power. This remains however difficult to achieve due to a lack of detail in the information we have available. The percentage of nuclear power in the input energy mix is a simple example of difficult to find information. The problem is even more opaque when considering a full technology lifespan. From production to overseas delivery, operation, maintenance and, finally, replacement and disposal, any ICT component implicitly results in emissions and indirect energy, e.g. processor cooling, being consumed. Accounting such quantities is however close to impossible. Furthermore, a device like, for example, a touchscreen, may need rare earth elements for its production. The long-term impact of out of use components disposal may be at least difficult to measure, e.g. resulting toxic waste and the effort needed in removing (or isolating) such byproducts from the environment. Operation is also a complex process. Dissipated heat is a main problem and solutions vary from placing data centers in cold regions to strategies like follow the Moon [16]. Then, a different view is to consider an ICT system as being sustainable via the impact it brings in other fields like smart buildings or the design of zero-emission buildings.

A different question, when moving from the overall ICT view to algorithms, is what kind of paradigms qualify the system for the 'sustainability' or 'green' label? Would a faster algorithm, e.g. Fast-Fourier-Transform [31], imply that a given computing architecture is improved in terms of sustainability? A faster execution time is in the end equivalent to a lower energy consumption and, in a classical data center, by direct correspondence, to less emissions. All computer science research would, by this approach, be a quest for energy-efficiency and emissions reduction. Faster algorithms

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.

http://dx.doi.org/10.1145/2598394.2605695.

unfortunately only equate to more data being processed in a same time unit interval and thus to no (actual) change in terms of carbon footprint or energy reduction. For an algorithm to be considered 'green', a most commonly accepted view is that it explicitly contains terms that address issues related to energy efficiency, e.g. processor stepping.

For the remainder of this paper, we focus on the computing side of sustainability. The following sections are organized as follows. Section 2 discusses sustainability and energy efficiency in IT at different levels, e.g. data center and networking. A series of metrics and benchmarks are presented in Section 3, followed by a brief overview of evolutionary computing as used in energy efficiency and sustainability in Section 4. Section 5 concludes the paper.

2. SUSTAINABILITY AND ENERGY EFFICIENCY IN IT: LEVELS OF GRANULARITY

All domains of our everyday life are concerned with and by reducing energy consumption. For areas like, for example, in sustainable buildings, external energy input may approach zero or, in some cases, buildings may sell energy back into the network. A few notable examples are NASA's Sustainability Base [27] or the Energy Plus building in Paris [29]. When it comes to computing, however, it may not be all that simple. Also, in order to understand why energy efficiency in computing is an important issue, we first need to understand where the limits are and how we are positioned with respect to these limits.

The ICT sector is comparable to the aviation industry in terms of carbon footprint, i.e. accounting for 2% of the worldwide emissions and estimated to reach 3% by 2020 [56]. This is in fact a rough view of where ICT stands: the direct impact it has is by no means negligible. A second aspect we need to consider is related to what is commonly referred to as Moore's law [52], i.e. a doubling of the number of transistors roughly every two years. At the same time, the number of operations per joule of dissipated energy doubled every one year and a half, cf. Koomey's law[1] [36]. A last element is Maxwell's daemon and Landauer's principle [38]. Landauer's principle, only recently verified via a scientific experiment [3, 17, 12], argued that destroying information leads to energy dissipation [38]. Erasing a bit of information would thus release an amount of heat equivalent to $k_B T \ln 2$, i.e. 3×10^{-21} J (at room temperature), where $k_B = 1.3806488(13) \times 10^{-23} \text{JK}^{-1}$ stands for Boltzmann's constant. And, while it may not seem the case, erasing information is a central part of how nowadays computers work. All classical computer architectures make use of what is called an irreversible gate, e.g. a NAND gate, which by concept discards information. Thus, even if we assume that all other dissipative processes are removed, a computing system that relies on irreversible gates would still release some amount of heat. This is in fact a lower and unavoidable limit. When correlated with Moore's and Koomey's laws, what follows is shattering: in a decade from now, a chip would release more heat per square centimeter than the surface of the Sun [53]. Reversible Turing machines, while not new as an idea [10], did not reach anywhere beyond the theory status for now. Thus, finding ways for dealing with heat dissipation and reducing the energy consumption in the process is essential. Heat dissipation, even if not yet at such an extreme level, is a main issue in nowadays data centers. As discussed below, a broad range of solutions are used.

[1]A large number of studies cite and refer to both Moore's and Koomey's observations as *laws.* A more correct description is that of a (linear) statistical model. Furthermore, industry also later mapped to follow this trend.

2.1 Data Centers and Cloud Computing

Energy efficiency in data centers and Cloud Computing is a topic that relates to all levels, from physical design and installation to operation. The need for finding an answer to reducing cooling related consumption led, among others, to solutions that take advantage of natural factors, e.g. data centers placed in cold geographic regions [47] or that make use of wind for cooling [33]. Other examples include water-based data center designs [15], compressed air [18] or smart cooling solutions [57]. An important conclusion of most studies, while referring to operation aspects only, is that sustainable computing is mainly about energy efficiency and renewable energy sources [24]. As such, the operation of a large scale computing facility needs to be optimized at all levels, i.e. design and physical arrangement down to dynamic CPU frequency and voltage scaling. Sustainability, in addition, also needs for renewable energy to be used. The main problem one is finally faced with is operation cost (read as amount to pay for consumption) which does not always equate with what sustainability stands for. Different operators may and will prefer nuclear or coal power, for example, over solar energy due to lower costs; for a worldwide view of energy use in data centers, please refer to [35]. Also, for a taxonomy of energy-efficient data centers and cloud computing systems, refer to [8]. A follow the Sun-like approach would imply moving the load where sufficient solar (renewable for the general case) energy is available. And while this scenario implies a high percentage of green energy being used, costs are likely to be above the ones of using conventional energy. A follow the Moon strategy, at the opposite end, translates into a lower energy consumption (less cooling is required) though not always in line with where renewable energy is available [16]. Last but not least, due to a law enforcement on data transfer, any of these strategies may be fully or partially limited [16].

When looking at large scale, energy efficient computing from an algorithmic and model oriented perspective, a wide range of studies can be found in literature on, among many others, heuristics [7, 26], machine learning, e.g. neural networks, auto-regressive stochastic state transition models [59], branching processes [63, 6], stochastic programming and optimization [45, 46, 28, 41], or nature inspired algorithms [14, 9]. Among the different criteria taken into account, one can mention, energy efficiency but also performance, availability, elasticity, bandwidth or Quality of Service (QoS). A large part of the existing studies rely on discrete models alone [45, 51, 26, 58], with very few exceptions bringing continuous attributes into discussion [63, 6]. A trend towards focusing more on high-level aspects in queuing, load balancing or scheduling can be observed, also with the additional support of technologies such as frequency scaling, e.g. SpeedStep [32], Cool'n'Quiet [2], or virtualization. With respect to Cloud computing, a very concise study on details that run from low-level to infrastructure related aspects is given in [11]. Among others it reviews recent advancements and research on methods or technologies for energy efficient cloud and data-center operation. The authors point at, as an example, wireless and wired networks power minimization, energy efficient hardware, energy-aware scheduling in multiprocessor and grid systems, state, e.g. slowing down the processor or powering off different parts of the chip or hardware, dynamic voltage and frequency scaling, virtualization, consolidation and protocols. Full, hosted and operating system layer virtualization are also mentioned. Last, an overview of energy efficient cloud computing and energy-aware data centers is given, ending with the impact communication oriented applications have. A series of surveys on mobile cloud computing, virtualization vulnerabilities, federation of clouds, data access and integrity or low resolution security issues can also be found in [22, 5, 66, 13, 75, 37]. A more detailed discussion can be found in [68].

A separate problem from computing alone is storage. As data needs to be readily available at all times in most cases, storage devices, e.g. hard drives, and all related communication and transfer infrastructure is typically kept alive even if not used. As opposed to computing however, idle storage is considered responsible for up to 80% of the energy usage, i.e. when compared to peak operation mode. Examples of possible ways of improvement with respect to storage can be found by referring to [65].

2.2 Sustainability in Networking

A first, although not the most energy consuming part of all IT equipment comes from the standby power consumption. According to the Standby Power project, supported by the Lawrence Berkeley National Laboratory, a computer desktop consumes in average 73.97 W/day when On (idle); 27.5 W/day when Off and 1.38W/day while in a Sleep mode. Although the measurements where performed in residential appliances, this easily translates to campus networks for example. Although this seems irrelevant at first, when considering the number of IT equipment of a small to medium network, the waisted energy amount increases quickly.

The first countermeasure one can think of in a network context, comes from enabling the centralized control of the switch-on/off of the network devices. Although multiple energy efficient applications are acknowledged in the networking area, they lack scalability and depend on the ability of the network administrator of integrating them and configuring the physical network accordingly. The newly emerged Software Defined Networking (SDN) paradigm [55], promoted by the Open Networking Foundation [1], enables a centralized controller layer, that, while provided with the proper applications, can intelligently manage the unused resources, being them equipment, services or subnetworks, for example.

SDN introduces a centralized and programmable way of designing networks, that completely changes the classical distributed networking approach. It not only introduces a separation between the data plane and the control plane, but also enables the programability of the network by means of external applications. Although the Southbound Interface acknowledges the OpenFlow protocol [50] as a *de facto* standard, promoted also by the Open Networking Foundation; the Northbound Interface that ensures the interaction between the controller and the application plane is still subject of research, with no unique solution identified.

On a base functioning grounds, the protocols used for broadcasting, routing for example dictate the energy efficiency of the network. On a higher level, SDN empowers the controller with the ability of having specific applications – interfacing with the network through the north bound – that oversee/control the network sustainable characteristics. A controller can be enabled also with traffic management applications, that reduce the network load, optimize paths and thus minimize the resources consumed for handling the network traffic.

3. QUANTIFYING THE IT IMPACT – METRICS AND BENCHMARKS

The rise of the data center industry and the emergency of cloud computing infrastructures implied the need of quantifying its impact through its multiple facets. A first tool are benchmarks offering the means for a solid analysis and comparison. Next, metrics allow following the over time evolution of a given data center with respect to specified operation targets. In the follow up, an overview of the metrics recommended as best practice by the community (through for example the Green Grid Global Taskforce outputs) is provided, with an emphasize on the sustainability oriented ones, with no intent of being exhaustive. Detailed metrics overviews are available in literature (e.g. [72]) and are beyond the scope of the current paper.

3.1 Metrics for Data Centers

The usance considers a data center management needs to bridge between requirements as performance, sustainability and QoS (Quality of Service). Therefore factors as energy efficiency or performance are seen as amongst the primary ones to be quantified. Besides it has been noted that from the total energy consumption, 50% is attributed to the cooling system [64], giving rise to the need of efficient **thermal metrics**.

Data center owners and managers have also the possibility of prioritizing amongst metrics in order to serve their specific needs [48].

The common assumption considered while measuring the energy efficiency is that the resources are thoroughly allocated. Nevertheless, unsustainable solutions include also idle equipment or otherwise stated maintaining servers, storage or in general a data center with low usage.

3.1.1 Power / Energy Efficiency Metrics

Even before designing a data center, the first aspect to consider is how to measure its energy efficiency, witnessed also by the interest of both operators and policy makers. One of the leading initiatives in this area if represented by the Green Grid Taskforce on *Global Harmonization of Data Center Efficiency Metrics.* The participants of the task force represent the major players in the area U.S. Department of Energy, Save Energy Now and Federal Energy Management Programs (March 2009 – October 2012); the U.S. Environmental Protection Agency ENERGY STAR Program; the European Commission Joint Research Centre Data Centres Code of Conduct; Japan's Ministry of Economy, Trade and Industry; Japan's Green IT Promotion Council; and The Green Grid Association.

One major output of the initiative consists in the reviewing of the existing metrics and the release of new energy efficiency metrics, with the intent of providing a common ground of understanding and improve the energy use of data centers.

PUE (Power Usage Effectiveness) - the above formulation follows the last proposed one by the Global Taskforce aforementioned [54]

$$\text{PUE} = \frac{\text{total data center source energy}}{\text{IT source energy}},$$

although it is worth noting that large companies are adapting the formulation to their particular conditions, see Google [25] or Dell [60]. If we take the example of the Dell case the main difference consist in the fact that the external power delivery systems, cooling systems,lighting and so on are ignored in the computation. This can also be explained by the novelty of the guide [54].

GEC (Green Energy Coefficient) [54] - the proportion of a facility's energy that comes from green sources

$$\text{GEC} = \frac{\text{green energy used by the data center}}{\text{total data center source energy}}$$

ERF (Energy Reuse Factor) [54] - the proportion of energy that is reused outside the data centre (e.g. in campus facilities)]

$$\text{ERF} = \frac{\text{reuse energy outside of the data center}}{\text{total data center source energy}}$$

CUE (Carbon Usage Effectiveness) [54] - measures the total carbon dioxide emission equivalents (CO2e) from the energy consumption of the data center divided by the total IT energy consumption, defined as follows:

$$\text{CUE} = \text{CEF} \times \text{PUE},$$

where CEF represents the Carbon Dioxide Emission Factor (kgCO2eq/kWh) of the Data Center. It can also be seen as the ratio between the Total $C0_2$ emissions caused by the Total Data Center Energy and the IT equipment Energy.

Besides the basic metrics mentioned, also variants of them coexist. The mostly used example comes from the reciprocal of the PUE, namely the **DCIE (Datacenter Infrastructure Efficiency)**. Furthermore, when needing more granularity for PUE, the definition can be revised as:

$$\text{DCiE} = \text{CLF} + \text{PLF} + 1.0,$$

where CLF represents the Cooling Load Factor and PLF stands for the Power Load Factor, with 1.0 the normalized IT Load [4].

Various variants of the PUE have been proposed by the Green Grid association over time, e.g. PUEx, where x can take a value between 0 and 3, and depends for example on where the measurements have been made.

Extensions based on PUEx also emerged as the **GPUE (Green Power Usage Effectiveness)**, defined as

$$\text{GPUE} = \text{G} \times \text{PUEx (for inline comparison of data centers)},$$

with G being the *weighed sum of energy sources and their lifecycle KG CO2/KWh.*

The abovementioned variants are by no means exhaustive and represent only useful pointers for a novice reader eager to tackle the area and searching for the proper starting point.

3.1.2 Cooling Efficiency

Besides the energy used directly for performing computations, a large percent of a system consumption comes from the cooling part of the system.

One of the first metrics emerging in the area concerned the HVAC (Heating, Ventilation, and Air conditioning) aspect. The metric was called the **HVAC system effectiveness** [48] and is measuring the system effectivness in terms of the ratio 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 lately concerns any other HVAC energy use like steam or chilled water. Studies detailing the various aspects, such as airflow efficiency or quantifying the usage of air/water economizers exist in the literature [49] and [61]. In a recent taxonomy of performance metrics in data centers, the following categories have been identified:

Humidity - can produce hardware failure and increase the cooling system usage.

Thermal metrics -relate with the control of thermal parameters in the storage space

- Cooling system efficiency - seen from a holistic perspective or divided in a individual components perspective. Various characteristics can be monitored and provide useful information: airflow efficiency, the sizing of the cooling system (hard to manipulate after acquisition, erroneous usage estimation increasing also the **Total Cost of ownership** metric); air/water economizer utilization.
- Data center temperature
- British Thermal Unit
- Airflow performance index
- Air/water economizer usage

3.1.3 Performance Metrics

Various performance goals need to be met in order to obtain a sustainable computing center, often contradictory or different by nature. Take for example energy efficiency and productivity in a data center. If the first one focuses on reducing the costs needed in order to achieve a certain performance or a specific work output, productivity focuses on providing besides the already specified performance, the maximum of useful work produced given a specific amount of resources. One metric in this direction and following a holistic perspective is the **DCeP (Data Center Energy Productivity)** that is used to measure the overall performance of a specific data center in terms of productivity.

This can be synthetically expressed as:

$$\text{DCeP} = \frac{\text{useful work produced}}{\text{total source energy for producing this work}}$$

As this metric measures data center specific productivity characteristics, a generic metric able to provide a common ground for comparison between the productivity of various data centers is needed. Data center productivity proxies represent one potential alternative, although no consensus has been reached in this regard yet. Many declinations aver possible as the *Data center energy productivity (DCeP)* measured by energy checker or the sample load. Other potential alternatives include the *weighted CPU utilization – SPECpower* or the *Computed Units per Second (CUPS).*

Name	Developer	Intended use	Architecture
EnergyBench	EEMBC [20]	industrial	processors, embedded systems
SpecPower	Spec [39]	industrial	client/server architecture (for web servers)
TPC-Energy	Transaction Processing Performance Council [74]	industry	servers, disk systems, other components
JouleSort	Rivoire *et al.* [62]	academic	client/server architecture, embedded systems
SWEEP	Du Bois *et al.* [19]	both	client/server architecture

Table 1: Synthetic view of available benchmarks for energy efficiency

3.2 Available Benchmarks

In order to support researchers by enabling energy efficiency testing on data centers or cloud computing facilities, various synthetic benchmarks have been developed. The panoply of benchmarks spans over different levels of granularity, from (1) *processors*to (2) *servers* or the (3) *holistic data center perspective*.

Covering the evaluation of a single processor's energy efficiency, the Embedded Microprocessor Benchmarking Consortium (EEMBC) provides a physical infrastructure for the embedded area [20], called the EEMBC EnergyBench. The metric used to evaluate processors' energy efficiency is the commonly employed, Performance Per Watt(PPW) metric [40]. Other related metrics, more tailored for performance criteria, as the energy-delay product coexist. The later being the equivalent of performance squared per Watt. At the servers level, the SPEC Power and Performance Committee has developed the SPECPower benchmark [39]. Together with the TPC-Energy [74] proposed by the Transaction Processing Performance Council, these benchmarks enable predicting the performance per Watt in e.g., a server farm. As an addition compared with the SPECPower, TPC-Energy enables measuring the energy consumption of other system components, besides servers, as disk systems or other items that consume power [74]. The class of workloads for energy-efficiency benchmarks acknowledges also more flexible (tunable) variants, as for example the Synthetic Workloads for Energy Efficiency and Performance evaluation (SWEEP) [19]. It represents a tunable synthetic workload generator, capable of creating instances that are compute-intensive, memory-intensive, I/O-intensive, or a mix of them.

Table 1 encloses a non-exhaustive overview of the reference benchmarks providing references, targeted segment and intended audience.

4. WHERE DOES EC PLAY A ROLE?

Sustainability (optimization) problems, due to their complexity, are beyond reach for most conventional approaches. Regardless of dealing with large scale computing systems, renewable energy plants or buildings, as an example, a series of hard limitations have to be considered. All these systems are subject to complex interactions, time dependent dynamics and external perturbations [70, 71]. A classical one-snapshot approach does not suffice. What is more, the number of objectives that typically need to be met in such problems is above what most optimization paradigms can deal with. Evolutionary Algorithms (EAs), as it will be discussed in the following, are a standard solution different studies turned to, mainly due to their sheer simplicity.

Even from the setup of testing environments, by means of synthetic benchmarks, genetic algorithms play a role. They are used for example, in the search and prune of workload attributes that maximize criteria as power or temperature [34].

Among the first areas EAs were used for, even before the entire green wave, e.g. energy-efficiency in computing, buildings and so forth, communications stands out, namely in wireless sensor networks and fixed networks [30]. Then, a large number of studies focused on results in energy-aware resource allocation and scheduling. For an overview see [7, 43, 67]. Evolutionary computing was also used from thermal design [44] to live application placement [42] or workload distribution [21]. Another approach explored in detail is Dynamic Voltage and Frequency Scaling (DVFS), where the power of individual computing nodes is modulated dynamically as to reflect momentary load. A discussion of DVFS and Dynamic Frequency Islands can be found in [23]. The main axes thus cover virtualization and live migration, resource allocation, workload mitigation and low-level tuning, e.g. via DVFS.

An interesting remark is that data centers are exceedingly complex and difficult to optimize even when using evolutionary algorithms. The large number of parameters one is given with, e.g. state of all nodes, processor stepping, etc., makes the problem intractable when the system is to be optimized as a single and indivisible entity. Instead, a more realistic approach is to break the system into smaller parts like, for example, computing nodes, where hierarchical agents control the state of a given part of the system via e.g. workload negotiation or transition among states [69]. Other examples of EC algorithms used in tuning consumption vs performance and energy-aware task scheduling can be found in [14, 9].

5. CONCLUSIONS

Energy efficiency and sustainability in ICT are major problems we have not solved yet. As new technologies replace the ones now in use, advancements in, for example, renewable energies or heat dissipation, allow us to go a step further. The need for even more extensive computations and the evolution of Cloud computing shadow however the small advantage technology offers. Thus, we need to rethink the way we manage resources, how computations are made or even how data is stored and later retrieved. The different metrics presented in this paper are an important tool for understanding how different solutions scale and how effective they are. There remains however a strong need for structured data analysis with a more proactive use of runtime information. On the algorithmic side, Evolutionary Computing and the heuristic paradigms in general, play an important role when solid practical solutions are needed. The large number of parameters one needs however to tackle and the number of objectives, now beyond what the state of the art can deal with, represent a major issue. The most important conclusion of this study is however that, if addressing any of the sustainability and energy efficiency issues in, for example, data centers is delayed, we will soon reach limits that force us to restrict the current evolution of how computing is made.

Acknowledgements

Alexandru-Adrian Tantar would like to acknowledge the support offered by the National Research Fund (Fonds National de la Recherche, FNR), Luxembourg, and the Centre National de la Recherche Scientifique (CNRS), France, in the framework of the Multi-Objective Metaheuristics for Energy-Aware Scheduling in Cloud Computing Systems INTER/CNRS/11/03 Green@Cloud project. Emilia Tantar would like to acknowledge that the publication is based in parts on work performed in the framework of the CoSDN project, INTER/POLLUX/12/4434480, funded by the Fonds National de la Recherche, Luxembourg.

6. REFERENCES

[1] Open Networking Foundation. http://www.opennetworking.org.

[2] AMD. Cool'n'Quiet Technology. http://www.amd.com/us/products/technologies/cool-n-quiet/Pages/cool-n-quiet.aspx, 2004.

[3] Janet Anders, Saroosh Shabbir, Stefanie Hilt, and Eric Lutz. Landauer's Principle in the Quantum Domain. In Cooper et al. [17], pages 13–18.

[4] John Pflueger Andy Rawson and Tahir Cader. The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE. http://www.thegreengrid.org/Global/Content/white-papers/The-Green-Grid-Data-Center-Power-Efficiency-Metrics-PUE-and-DCiE#sthash.tnUhZM7e.dpuf, Oct. 2007.

[5] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. A View of Cloud Computing. *Commun. ACM*, 53(4):50–58, April 2010.

[6] Marko Bajec and Johann Eder, editors. *Advanced Information Systems Engineering Workshops - CAiSE 2012 International Workshops, Gdask, Poland, June 25-26, 2012. Proceedings*, volume 112 of *Lecture Notes in Business Information Processing*. Springer, 2012.

[7] Anton Beloglazov, Jemal Abawajy, and Rajkumar Buyya. Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing. *Future Generation Computer Systems*, 28(5):755 – 768, 2012. Special Section: Energy efficiency in large-scale distributed systems.

[8] Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Y. Zomaya. A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems. *CoRR*, abs/1007.0066, 2010.

[9] S. Benedict, R.S. Rejitha, and C.B. Bright. Energy Consumption-Based Performance Tuning of Software and Applications Using Particle Swarm Optimization. In *Software Engineering (CONSEG), 2012 CSI Sixth International Conference on*, pages 1 –6, sept. 2012.

[10] C. H. Bennett. Logical Reversibility of Computation. *IBM J. Res. Dev.*, 17(6):525–532, November 1973.

[11] Andreas Berl, Erol Gelenbe, Marco Di Girolamo, Giovanni Giuliani, Hermann De Meer, Minh Quan Dang, and Kostas Pentikousis. Energy-Efficient Cloud Computing. *Comput. J.*, 53(7):1045–1051, September 2010.

[12] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, Mar 2012.

[13] Rajkumar Buyya, Rajiv Ranjan, and RodrigoN. Calheiros. InterCloud: Utility-Oriented Federation of Cloud Computing Environments for Scaling of Application Services. In Ching-Hsien Hsu, LaurenceT. Yang, JongHyuk Park, and Sang-Soo Yeo, editors, *Algorithms and Architectures for Parallel Processing*, volume 6081 of *Lecture Notes in Computer Science*, pages 13–31. Springer Berlin Heidelberg, 2010.

[14] Ying Chang-tian and Yu Jiong. Energy-Aware Genetic Algorithms for Task Scheduling in Cloud Computing. In *ChinaGrid Annual Conference (ChinaGrid), 2012 Seventh*, pages 43 –48, sept. 2012.

[15] J. Clidaras, D.W. Stiver, and W. Hamburgen. Water-based Data Center. http://www.google.com/patents/US7525207, April 28 2009. US Patent 7,525,207.

[16] Computer & Communications Industry Association. Cloud Computing. http://www.ccianet.org/wp-content/uploads/library/Cloud_Computing.pdf, 2009.

[17] S. Barry Cooper, Prakash Panangaden, and Elham Kashefi, editors. *Proceedings Sixth Workshop on Developments in Computational Models: Causality, Computation, and Physics, DCM 2010, Edinburgh, Scotland, 9-10th July 2010*, volume 26 of *EPTCS*, 2010.

[18] M.P. Czamara and O.P. Morales. Compressed Air Cooling System for Data Center. https://www.google.com/patents/EP2554032A1?cl=en, February 6 2013. EP Patent App. EP20,110,712,079.

[19] Kristof Du Bois, Tim Schaeps, Stijn Polfliet, Frederick Ryckbosch, and Lieven Eeckhout. SWEEP: Evaluating Computer System Energy Efficiency Using Synthetic Workloads. In *Proceedings of the 6th International Conference on High Performance and Embedded Architectures and Compilers*, HiPEAC '11, pages 159–166, New York, NY, USA, 2011. ACM.

[20] Embedded Microprocessor Benchmark Consortium (EEMBC). Energybench V1.0 Power/Energy Benchmarks. http://www.eembc.org/benchmark/power_sl.php.

[21] Eugen Feller, Louis Rilling, and Christine Morin. Energy-Aware Ant Colony Based Workload Placement in Clouds. In *Proceedings of the 2011 IEEE/ACM 12th International Conference on Grid Computing*, GRID '11, pages 26–33, Washington, DC, USA, 2011. IEEE Computer Society.

[22] Ian Foster, Yong Zhao, Ioan Raicu, and Shiyong Lu. Cloud Computing and Grid Computing 360-Degree Compared. *2008 Grid Computing Environments Workshop*, abs/0901.0(5):1–10, 2008.

[23] Siddharth Garg, Diana Marculescu, Radu Marculescu,

and Umit Ogras. Technology-driven Limits on DVFS Controllability of Multiple Voltage-frequency Island Designs: A System-level Perspective. In *Proceedings of the 46th Annual Design Automation Conference*, DAC '09, pages 818–821, New York, NY, USA, 2009. ACM.

[24] Greenpeace International Gary Cook. How Green is Your Cloud? Catalysing an Energy Revolution. `http://www.greenpeace.org/new-zealand/Global/international/publications/climate/2012/iCoal/HowCleanisYourCloud.pdf`, apr 2012.

[25] Google. Efficiency: How we do it. `http://www.google.com/about/datacenters/efficiency/internal/index.html#measuring-efficiency`.

[26] Hadi Goudarzi, Mohammad Ghasemazar, and Massoud Pedram. SLA-based optimization of power and migration cost in cloud computing. In *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012)*, CCGRID '12, pages 172–179, Washington, DC, USA, 2012. IEEE Computer Society.

[27] June A. Grant. Implementing Sustainable Practices @ NASA Ames, GreenGov Symposium, 2011 Washington, USA. `http://www.greengov2011.org/presentations/GreenFacilities/GreenGov-2011-GreenFacilities-S5-JuneGrant.pdf`, Oct 2011.

[28] Albert G. Greenberg and Kazem Sohraby, editors. *Proceedings of the IEEE INFOCOM 2012, Orlando, FL, USA, March 25-30, 2012*. IEEE, 2012.

[29] Katherine Griffiths. World's greenest office block set for Paris, The Telegraph. `http://www.telegraph.co.uk/finance/newsbysector/constructionandproperty/2781961/Worlds-greenest-office-block-set-for-Paris.html`, Jan 2008.

[30] M.N. Halgamuge, S.M. Guru, and A. Jennings. Energy Efficient Cluster Formation in Wireless Sensor Networks. In *Telecommunications, 2003. ICT 2003. 10th International Conference on*, volume 2, pages 1571–1576 vol.2, Feb 2003.

[31] Haitham Hassanieh, Piotr Indyk, Dina Katabi, and Eric Price. Nearly Optimal Sparse Fourier Transform. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 563–578, New York, NY, USA, 2012. ACM.

[32] Intel White Paper 30057701. Wireless Intel SpeedStep power manager: optimizing power consumption for the Intel PXA27x processor family. Technical report, Intel, 2004.

[33] Jeremy Kirk, IDG News Service, PCWorld. HP Opens First Ever Wind-cooled Data Center. `http://www.pcworld.com/article/188996/hp_opens_first_ever_wind_cooled_data_center.html`, 2010.

[34] Ajay M. Joshi, Lieven Eeckhout, Lizy Kurian John, and Ciji Isen. Automated Microprocessor Stressmark Generation. In *HPCA*, pages 229–239. IEEE Computer Society, 2008.

[35] Jonathan G Koomey. Worldwide Electricity Used in Data Centers. *Environmental Research Letters*, 3(3):034008, 2008.

[36] Jonathan G. Koomey, Stephen Berard, Marla Sanchez, and Henry Wong. Implications of Historical Trends in the Electrical Efficiency of Computing. *IEEE Annals of the History of Computing*, 33(3):46–54, 2011.

[37] K. Kumar and Yung-Hsiang Lu. Cloud Computing for Mobile Users: Can Offloading Computation Save Energy? *Computer*, 43(4):51–56, 2010.

[38] R. Landauer. Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.*, 5(3):183–191, Jul 1961.

[39] Klaus-Dieter Lange. Identifying Shades of Green: The SPECpower Benchmarks. *IEEE Computer*, 42(3):95–97, 2009.

[40] James Laudon. Performance/Watt: The New Server Focus. *SIGARCH Comput. Archit. News*, 33(4):5–13, November 2005.

[41] Wei-Ru Lee, Hung-Yi Teng, and Ren-Hung Hwang. Optimization of Cloud Resource Subscription Policy. In *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on*, pages 449 –455, dec. 2012.

[42] Bo Li, Jianxin Li, Jinpeng Huai, Tianyu Wo, Qin Li, and Liang Zhong. EnaCloud: An Energy-Saving Application Live Placement Approach for Cloud Computing Environments. In *Cloud Computing, 2009. CLOUD '09. IEEE International Conference on*, pages 17–24, Sept 2009.

[43] Dawei Li and Jie Wu. *Energy-aware Scheduling on Multiprocessor Platforms*. Springer Publishing Company, Incorporated, 2012.

[44] G. Li, M. Li, S. Azarm, J. Rambo, and Y. Joshi. Optimizing Thermal Design of Data Center Cabinets with a New Multi-objective Genetic Algorithm. *Distrib. Parallel Databases*, 21(2-3):167–192, June 2007.

[45] Qiang Li. Applying Stochastic Integer Programming to Optimization of Resource Scheduling in Cloud Computing. *JNW*, 7(7):1078–1084, 2012.

[46] Siva Theja Maguluri, R. Srikant, and Lei Ying. Stochastic Models of Load Balancing and Scheduling in Cloud Computing Clusters. In Greenberg and Sohraby [28], pages 702–710.

[47] Mark Gregory, BBC News. Inside Facebook's Green and Clean Arctic Data Centre. `http://www.bbc.com/news/business-22879160`, 2013.

[48] Paul A. Mathew, Srirupa Ganguly, Steve E. Greenberg, and Dale A. Sartor. Self-benchmarking Guide for Data Centers: Metrics, Benchmarks, Actions. Technical report, Lawrence Berkeley National Laboratory, Jul 2009.

[49] Sara McAllister, Van P. Carey, Amip Shah, Cullen Bash, and Chandrakant D. Patel. Strategies for Effective use of Energy-based Modeling of Data Center Thermal Management Systems. *Microelectronics Journal*, 39(7):1023–1029, 2008.

[50] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. OpenFlow: enabling innovation in campus networks. In *Proc. of the ACM SIGCOMM 2008 conference*, pages 38(2):69–74, 2008.

[51] M. Mezmaz, N. Melab, Y. Kessaci, Y. C. Lee, E. G. Talbi, A. Y. Zomaya, and D. Tuyttens. A Parallel

Bi-Objective Hybrid Metaheuristic for Energy-Aware Scheduling for Cloud Computing Systems. *J. Parallel Distrib. Comput.*, 71(11):1497–1508, November 2011.

[52] G. E. Moore. Cramming More Components onto Integrated Circuits. *Electronics*, 38(8):114–117, Apr 1965.

[53] Samuel K. Moore. Landauer Limit Demonstrated, IEEE Spectrum. `http://spectrum.ieee.org/computing/hardware/landauer-limit-demonstrated`, Mar 2012.

[54] The Green Grid Global Harmonization of Data Center Efficiency Metrics Task Force. Harmonizing Global Metrics for Data Center Energy Efficiency, March 2014.

[55] Open Networking Foundation. SDN architecture overview. ONF White paper, Dec. 2013.

[56] Parliament Office of Science and Technology. ICT and CO_2 Emissions. `http://www.parliament.uk/documents/post/postpn319.pdf`, Dec 2008.

[57] C.D. Patel, C.E. Bash, and A.H. Beitelmal. Smart Cooling of Data Centers. `https://www.google.com/patents/US6574104`, June 3 2003. US Patent 6,574,104.

[58] Dung H. Phan, Junichi Suzuki, Raymond Carroll, Sasitharan Balasubramaniam, William Donnelly, and Dmitri Botvich. Evolutionary Multiobjective Optimization for Green Clouds. In *Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference companion*, GECCO Companion '12, pages 19–26, New York, NY, USA, 2012. ACM.

[59] J.J. Prevost, K. Nagothu, B. Kelley, and M. Jamshidi. Prediction of Cloud Data Center Networks Loads using Stochastic and Neural Models. In *System of Systems Engineering (SoSE), 2011 6th International Conference on*, pages 276 –281, june 2011.

[60] Paul Rad, Max Thoene, and Tim Webb. Best Practices for Increasing Data Center Energy Efficiency, February 2008.

[61] Jeffrey Rambo and Yogendra Joshi. Modeling of Data Center Airflow and Heat Transfer: State of the Art and Future Trends. *Distributed and Parallel Databases*, 21(2-3):193–225, 2007.

[62] Suzanne Rivoire, Mehul A. Shah, Parthasarathy Ranganathan, and Christos Kozyrakis. JouleSort: A Balanced Energy-efficiency Benchmark. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, SIGMOD '07, pages 365–376, New York, NY, USA, 2007. ACM.

[63] Victor Romanov, Aleksandra Varfolomeeva, and Andrey Koryakovskiy. Branching Processes Theory Application for Cloud Computing Demand Modeling Based on Traffic Prediction. In Bajec and Eder [6], pages 502–510.

[64] Richard Sawyer. Calculating Total Power Requirements for Data Centers, 2011.

[65] Storage Networking Industry Association (SNIA). Storage Management Initiative Specification. `http://snia.org/tech_activities/publicreview`, 2013.

[66] S. Subashini and V. Kavitha. A Survey on Security Issues in Service Delivery Models of Cloud Computing. *Journal of Network and Computer Applications*, 34(1):1 – 11, 2011.

[67] Q. Tang, S. K S Gupta, and G. Varsamopoulos. Energy-Efficient Thermal-Aware Task Scheduling for Homogeneous High-Performance Computing Data Centers: A Cyber-Physical Approach. *Parallel and Distributed Systems, IEEE Transactions on*, 19(11):1458–1472, Nov 2008.

[68] A.-A. Tantar, Anh Quan Nguyen, P. Bouvry, B. Dorronsoro, and E.-G. Talbi. Computational Intelligence for Cloud Management Current Trends and Opportunities. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 1286–1293, June 2013.

[69] Alexandru-Adrian Tantar, Grégoire Danoy, Pascal Bouvry, and SameeU. Khan. Energy-Efficient Computing Using Agent-Based Multi-objective Dynamic Optimization. In JaeH. Kim and MyungJ. Lee, editors, *Green IT: Technologies and Applications*, pages 267–287. Springer Berlin Heidelberg, 2011.

[70] Alexandru-Adrian Tantar, Emilia Tantar, and Pascal Bouvry. A Classification of Dynamic Multi-objective Optimization Problems. In *Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation*, GECCO '11, pages 105–106, New York, NY, USA, 2011. ACM.

[71] E. Tantar, A. Tantar, and P. Bouvry. On Dynamic Multi-Objective Optimization, Classification and Performance Measures. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 2759–2766, June 2011.

[72] Lizhe Wang and SameeU. Khan. Review of Performance Metrics for Green Data Centers: A Taxonomy Study. *The Journal of Supercomputing*, 63(3):639–656, 2013.

[73] White Paper, Google. Google's Carbon Offsets: Collaboration and Due Diligence. `https://static.googleusercontent.com/media/www.google.com/en//green/pdfs/google-carbon-offsets.pdf`, 2011.

[74] Erik Young, Paul Cao, and Mike Nikolaiev. First TPC-Energy Benchmark: Lessons Learned in Practice. In Raghunath Nambiar and Meikel Poess, editors, *Performance Evaluation, Measurement and Characterization of Complex Systems*, volume 6417 of *Lecture Notes in Computer Science*, pages 136–152. Springer Berlin Heidelberg, 2011.

[75] Qi Zhang, Lu Cheng, and Raouf Boutaba. Cloud Computing: State-Of-The-Art and Research Challenges. *Journal of Internet Services and Applications*, 1:7–18, 2010.