A SURVEY ON ENERGY EFFICIENCY IN INFORMATION SYSTEMS

MONICA VITALI

Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano,
Piazza Leonardo Da Vinci 32, Milan, 20133, Italy
vitali@elet.polimi.it

BARBARA PERNICI

Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano,
Piazza Leonardo Da Vinci 32, Milan, 20133, Italy
barbara.pernici@polimi.it

Concerns about energy and sustainability are growing everyday involving a wide range of
fields. Even Information Systems are being influenced by the issue of reducing pollution
and energy consumption and new fields are rising dealing with this topic. One of these
fields is Green IT, which deals with energy efficiency with a focus on IT. Researchers have
faced this problem according to several points of view. The purpose of this paper is to
understand the trends and the future development of Green IT by analyzing the state of
the art and classifying existing approaches to understand which are the components that
have an impact on energy efficiency in Information Systems and how this impact can be
reduced. At first, we explore some guidelines that can help to understand the efficiency
level of an organization and of an Information System. Then we discuss measurement
and estimation of energy efficiency and identify which are the components that mainly
contribute to energy waste and how it is possible to improve energy efficiency, both at
the hardware and at the software level.

Keywords: Energy Efficiency; Green IS; Green IT; Information Systems; Assessment;
Adaptivity.

1. Introduction

The general interest about environmental issues is growing every day in a wide range
of fields. In the last years, the problem of being green and of saving energy has been
faced also in Information Systems, giving birth to a new research branch known
as Green IT. Examining any IT system, it is evident that any single component of
the system has a cost in terms of energy consumption or pollution. At the same
time, IT is also an effective source of solutions for monitoring and improving energy
efficiency. This twofold role of IT for energy efficiency is analyzed in this paper.

The purpose of this paper is to analyze the maturity of the approaches available

in the state of the art about assessing and improving the sustainability in Information Systems, from a computer science and engineering point of view. A similar effort has been recently published in Ref. 1, where the authors faced the problem of energy efficiency from several perspectives: hardware, operating system, virtualization, and data center levels. The focus is on the hardware component contribution to energy consumption and on tools available to improve their efficiency.

On the other hand, the focus of this work is most of all on IT, with a special attention to the contribution of the business processes and of the setting of the organization to both energy waste and saving. A thorough analysis of the literature allows us to identify three main areas. The first area is related to the definition of general instruments and guidelines to assess the maturity level of sustainability in an organization. The second one involves measurement and estimation of energy consumption of an entire Information System or of individual components, that can be both hardware and software. Available metrics to assess and monitor greenness are also discussed. The last area focuses on the definition of a set of techniques that can be used to reduce the waste of energy, both at the hardware and at the software level.

The research questions that we want to inspect in this work are: *What are Green IT and Green IS?*; *How it is possible to state when an Information System is green?*; *How can we measure energy efficiency of an Information System?*; and *What can be done to improve energy efficiency in an Information System?*.

As energy efficiency in Information Systems is a recent scientific field, the authors have performed a systematic review of journals and conference proceedings. Only representative research approaches were selected based on the quality of their contribution and their significance and relevance to the scope of this article by also considering the respective publication year and the classification of the conference or journal in which the approaches were published. Different criteria were exploited to assess the quality of a document, such as the document expressiveness, richness, formality, and innovation. Research approaches of conferences or journals with low classification were not considered, unless giving an interesting contribution or perspective not detected in other papers with a higher classification. We examined the proceedings of the main conferences related to Information Systems and, at the same time, the first special sessions and special issues dedicated to the energy efficiency topic. We also examined the work of the best known researchers in the area, selecting relevant citations from their work in order to have a comprehensive view of each topic. The main keywords in the research of documents were: *Energy Efficiency*, *Green IT*, *Green IS*, *Energy Estimation*, *Energy Efficiency Improvement*, *Energy-Aware Information Systems*, *Energy Efficiency Assessment*, *Energy Metrics*, and *Energy-Aware Adaptation*.

The rest of the paper is divided into four sections. First of all, an overview about efficiency in Information Systems and Information Technology is given with some definitions (Sec. 2). After that, three main areas are identified for dealing with greenness in Information Systems. This division is represented through a chart in Fig. 1.
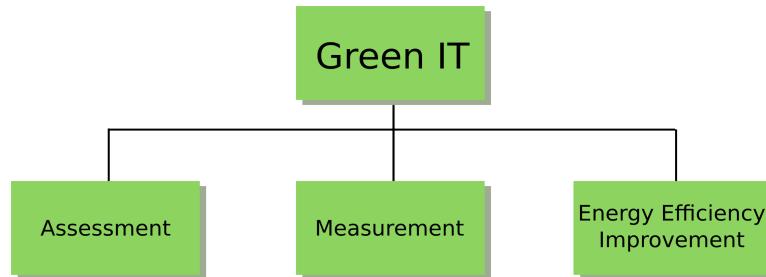
Fig. 1. Different approaches to Green IT

Each of the areas represented is analyzed in detail: Sec. 3 introduces the problem of assessing energy efficiency of an organization, Sec. 4 discusses approaches to measure energy in an Information System, and Sec. 5 shows several techniques adopted to improve energy efficiency. Finally, Sec. 6 summarizes the concepts emerged from the analysis and it discusses future trends in Green IT.

2. An Overview of Green Perspectives for Information Systems

In this section we give some definitions and perspectives on energy efficiency both at the IT and IS level, trying to answer the research question "*What are Green IT and Green IS?*".

The most efficient driver for energy efficiency can be identified in governmental regulations. A big number of regulations exist and cannot be described in details in this work. To understand the amount of different existing regulations, a survey published in 2009 by the Organization for Economic Co-operation and Development (OECD) analyzed 90 of them.[2] Just to cite some relevant regulations, in the USA we can find Energy Star, developed by the US Environmental Protection Agency (EPA), for producing energy efficient devices.[3] In Europe, the Code of Conduct for data centers has been developed to give directions about how to improve their greenness.[4]

Focusing on energy efficiency at the Information Technology level, the field of Green IT has been analyzed in the recent years. A definition for Green IT can be found in Ref. 5:

> "*Green IT is a systematic application of ecological-sustainability criteria (such as pollution prevention, product stewardship, use of clean technologies) to the creation, sourcing, use and disposal of the IT technical infrastructure as well as within the IT human and managerial practices.*"

As it can be inferred from this definition, this field is really wide and can be faced at different levels. A detailed analysis of this problem is described in Ref. 6, where the authors examine green issues related to an Information System at each

level, looking at IT both as a source of pollution and as a solution to it.

According to Ref. 6, reasons of pollution due to IT are multiple. The computer pollutes because it uses energy, because it is made with toxic materials and because nowadays disposal of computers is so fast that a great amount of old machines needs to be demolished. So, in order to understand the impact of a machine on the environment, its whole life cycle has to be considered.

Therefore, the approach to Green IT should cover four areas:[6]

- *Usage* of IT: energy efficient machines are needed and good practices should be enacted. At the level of personal computers, some simple practices can dramatically reduce energy consumption. For instance, energy saving features can be enabled, or the computer could be turned off when it is not in use. Even the use of screen savers can reduce energy consumption and in an organization with a large amount of employees this saving can be considerable. It is worth noticing that a computer in power saving mode consumes about 25% of the normal mode power. At the level of data centers an efficient cooling system or the migration from old mainframes to more recent servers could enable energy efficiency improvement. Another good practice can be virtualization, which optimize resources usage;
- *Design* of energy efficient components and equipments: some techniques to improve energy efficiency could be used. For instance, switching from single to multiple core architectures improves energy efficiency. The cache could be divided into sub-segments that could be powered only when needed. About storage, using less devices with a greater capacity can contribute to reduce energy usage by the 50%;
- *Manufacturing* components with the minimal possible environmental impact: the materials chosen to build servers or personal computers have a big impact on Green IT;
- *Disposal* of IT: reusing old machines and appropriately disposing of the ones that are useless. At the end of the life cycle, reusing and recycling of old machines should be mandatory.

In addition, IT can also be used as a tool to improve energy efficiency in other fields. In fact, IT gives the ability to build models for predicting the behavior of a given system, to run simulations for predicting its energy consumption under given circumstances, and to design decision support tools that can be used in a wide range of areas for reducing energy consumption. This concept has been synthesized in Ref. 7 with the name of *energy informatics*. Energy informatics is concerned with the analysis, design and implementation of Information Systems able to improve the energy efficiency of a system. The core idea is that enriching the system with monitored information the energy can be reduced. Actors involved in a typical system can be divided into three categories: suppliers who provide the service consuming energy, clients who ask for the services, and governments who intervene enacting some regulations to lead suppliers toward a more efficient management of the re-

sources. It is important to find an agreement between service providers and clients. Usually a Service Level Agreement (SLA) is ratified between them that expresses constraints about the quality that the service has to respect to satisfy the client. When energy efficiency becomes part of the equation, a trade off between energy efficiency and the quality of the service might be needed and SLA should to be renegotiated.

This twofold aspect of Green IT for sustainability is studied by Vom Brocke.[8] He moves the attention from IT to Information Systems, touching the area of Green IS:[9]

> *"Green Information Systems initiatives generally focus on designing, building and operating sustainable IS in order to improve the sustainability of organizations."*

In his analysis of sustainability, he focuses on business processes, as discussed in a panel section on sustainability of Information Systems.[10] Vom Brocke claims that an improvement of energy efficiency in an organization is possible only if there are enablers for the adoption of sustainability practices. These enablers are classified in four major categories: definition of a clear *strategy for sustainability*, *organizational support* from the top management, *motivation* for applying strategies, and *traceability* in order to be able to see the results of the adopted practices. This approach shifts the focus from a technological perspective to an organizational one. A proximate area to Green IS is the topic of Green Business Process Management (Green BPM), defined and analyzed in details in Ref. 9:

> *"Green Business Process Management concerns the understanding, documenting, modeling, analyzing, simulating, executing, and continuously changing of business process with dedicated consideration paid to the environmental consequences of these business processes."*

In Ref. 11, Green BPM is seen as the area of intersection between Green IS and BPM.

Given these premises, in the following sections we analyze approaches to manage energy efficiency in IT, with an Information Systems perspective.

3. Assessing the Energetic Maturity of an Information System

As discussed before, IT systems have a relevant impact on the total amount of CO_2 emissions and more and more companies are now starting to deal with this problem. As reported by Greenpeace in Ref. 12, in 2008 ICT was responsible for the 2% of the global greenhouse gas (GHG) emissions. Moreover, considering the electricity demand of data centers and clouds as the one of a country, and comparing it with the one of the other countries, the former would be in the fifth position. This demand is also growing: in an era where recession made electricity consumption stable, the

2007 electricity consumption
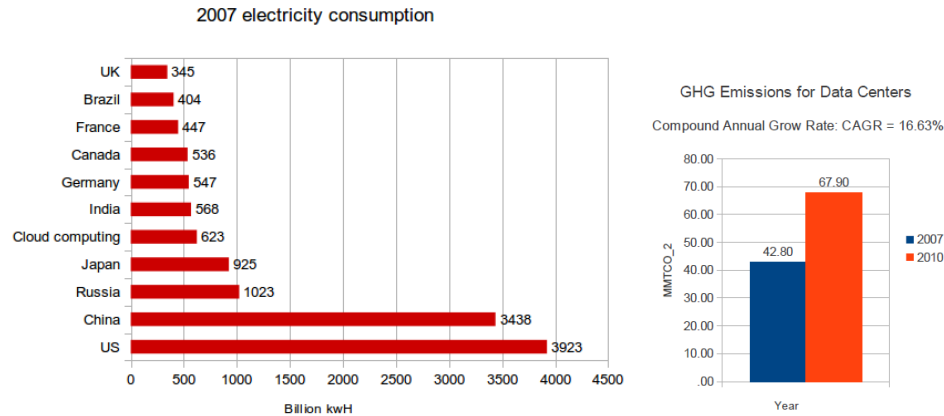
(a) Electricity demand of ICT compared with countries (source Greenpeace)

(b) ICT GHG emissions growth

Fig. 2. GHG emissions and energy demand of ICT

demand from data centers and clouds has been growing of the 56% from 2005 to 2010. In Ref. 1, authors report data from the EPA, stating that GHG have increased from 42.8 million metric tons of CO_2 (MMTCO$_2$) in 2007 to 67.9 MMTCO$_2$ in 2011 (see Fig. 2).

The issue of CO_2 emissions is related but separated from energy efficiency. In fact, improving energy efficiency results in a reduction of CO_2, but this is not the only aspect to be considered. According to Greenpeace, also the source of energy has to be considered, and big companies have the responsibility of pushing towards the adoption of policies for renewable energy. In the report, Greenpeace analyzes the trends in energy consumption of well known companies in the IT sector (e.g. Google, Apple, Akamai, and Amazon), reporting their clean energy index, compared to the percentage of energy originated from non renewable sources. The levels of greenness of these organizations are significantly different, due to different investments that they have reserved to the green perspective.

The effort in greening an organization can be motivated by different goals. Some companies are constrained by new regulations about environmental impact and they need to modify their behavior to respect these regulations. Other companies are trying to become greener for reasons related to public image. Nowadays, customers are getting sensitive to environmental issues and are starting to prefer companies which demonstrate an effort in this direction. As an example, the big company Facebook, was moved to adopt a green direction for its data center energy provisions by campaigns conducted by Greenpeace and several user associations[a].

[a]http://www.greenpeace.org/international/en/news/features/Victory-Facebook-friends-renewable-energy/
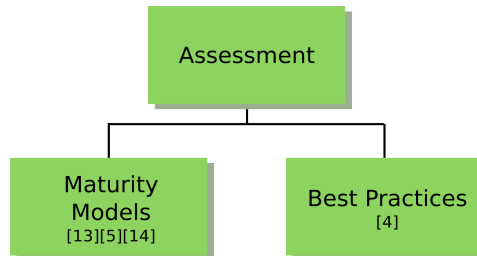
Fig. 3. Assessment of energy efficiency

In this section we try to answer the research question *"How it is possible to state when an Information System is green?"*. The problem for companies that set off towards a greener asset is twofold. First, they need to be able to understand their current maturity level in sustainability. For maturity level we mean the level at which sustainability is addressed inside the organization; this level should be measurable using some comparison parameters. Second, they need some guidelines to decide what to do to improve their status. These two trends, discussed in this section, are shown in Fig. 3. This effort requires the participation of a large number of subjects that are involved in the process. Different approaches exist looking at different aspects of an organization: while some scholars analyze the whole organization, others focus their attention only on IT. The general approaches analyze the production process, people involved in the organization and managerial factors. Approaches focused on IT usually analyze different levels of the infrastructure such as the facilities (power, cooling, etc.) or the resources (disks, network, etc.). In the rest of the paper we see that dealing with only one of these aspects can be not enough to improve the efficiency of the organization, so both IT and organizational factors, such as the business processes, need to be considered.

Some efforts made for understanding the sustainability maturity level of an organization have generated some capability maturity models related to sustainability. A capability maturity model can be seen as a methodology used to develop and refine an organization's aspect, in this case its sustainability level.

The Green Grid consortium[b] proposed a capability maturity model addressed to data centers. The consortium is composed of companies, government agencies, and educational global institutions that are dedicating their research to improve energy efficiency in data centers. The model, presented in Ref. 13, captures some relevant aspects related to data center energy efficiency. The maturity level of an organization can be one of six levels, from 0 to 5. Level 0 represents a company which is not enacting any effort for sustainability, between 1 and 2 characteristics of a data center actuating sustainable practices are described, with level 2 being the state of the art best practice. Levels from 3 to 5 express future sustainable data

[b]http://www.thegreengrid.org/

center and level 5 is expected to be reached in 2016. Moreover, under each level, a distinction between different aspects of the data center is made. Two main areas are analyzed: facility and IT. The facility area stresses the attention on cooling and power, considering also the management of the data center. The IT area stresses the importance of compute, storage, and networks components including also a more general category which describes the level of utilization and the workload distribution in the servers. At each level, each of these areas describes which should be the state of the data center in order to belong to that level of maturity. As an example, considering the general IT category, a state of the art data center should use virtualization and consolidation to handle its workload and should have a CPU average usage rate of at least 20%. CPU average utilization should incrementally increase to 60% to reach level 5. Mapping an organization in the model makes it simpler to understand which efforts are needed to move from one level to another. The model can be an important tool for the organization to understand which are the parts of the data center that are excelling and which parts present the largest opportunity for improvement.

Another approach to sustainability in Green IT is presented in Ref. 5. Here, the authors define the G-readiness concept as *"an organizations capability in Greening IT (that is applying environmental criteria to its IT technical infrastructure as well as within its IT human and management practices across the key areas of IT creation, sourcing, operations, and disposal) in order to reduce IT, business process, and supply chain related emissions, waste and water use; improve energy efficiency; and generate green economic rent"*. So, greening IT should involve both technology and related people. Also here five categories are distinguished which are *attitude*, *policy*, *practice*, *technology*, and *governance*. The authors also divide the readiness into 7 levels and represent the G-readiness of an organization using a graph in which the seven levels and the five categories are depicted. Through this graph, the organization is able to understand which are the fields that need more attention.

Starting from this model, a new maturity model has been proposed by the Innovation Value Institute (IVI) as complementary of pre-existing approaches. The model is described in Ref. 14. A framework is presented for increasing the sustainability in ICT executing four key actions: (i) defining the scope and goal of a sustainable ICT for the company; (ii) understanding the current capability maturity level; (iii) developing and managing sustainability in a systematic way using some capability building blocks; (iv) assessing and managing progress over time. The scope depends on the strategy of the organization that can range from satisfying customers requests about greenness and sustainability, to gaining a role of leadership in the field. Even in this case, the organization needs to understand its own maturity level that is distinguished into five levels: *initial* (where sustainable ICT is not considered as a business asset), *basic* (partially considered), *intermediate* (a strategy exists with plans and priorities), *advanced* (sustainable ICT is a core component of the business strategy), and *optimizing* (the adopted strategies are considered as an example for other organizations). As in the Green Grid Maturity

Model, each level is decomposed into different categories that cover different aspects of the system. In this case, the categories highlighted by the model are *strategy and planning*, *governance*, *people and culture*, and *process management*. Each of them is decomposed into a small number of subcategories to focus on more detailed aspects. Each IT process consists of building blocks against which maturity levels are defined. The framework can be used to assess the maturity of the organization and systematically improve it in a measurable way to meet the sustainability objectives.The importance level of each category can also be considered by the organization and the comparison with its maturity level can lead the organization to focus its efforts on a category rather than an another.

Considering approaches to assessment based on best practices, the Codes of Conduct, proposed since 2007 by some researchers of the European Commission, and revised in 2013, proposes an action with the aim of improving the energy efficiency of information and communication technologies. Some specific documents give some indications related to data centers. In Ref. 4 a set of best practices is described and classified. Their role is to identify and implement measures to improve the energy efficiency of data centers. All organizations willing to obtain a *participant* status need to implement the non optional best practices. A score is associated to each best practice, indicating its impact and consequently its priority. The score is between 1 and 5. For all the practices, participants have to specify if they are already implemented or endorsed to outsource partners, or if there is a plan to implement them. Otherwise, the participant has to specify the reason why the practice is not implemented or is not applicable to its data center. Practices are grouped in categories which are identified in the document as: *Data Center Utilization, Management and Planning*, *IT Equipment and Services*, *Deployment of New IT Services*, *Cooling*, *Data Center Power Equipment*, *Other Data Center Equipment*, *Data Center Building* and *Monitoring*. The authors also define a scope for the suggested practices that can be requested only for part of the data center. Five scopes are identified: entire data center, new software, new IT equipment, new build or retrofit, and optional tasks. This last category covers all the practices which are suggested, but not required.

A comparison between the presented models is summarized in Tab. 1, which reports the main features of the analyzed models. The level of operation can be of two kinds: Data Center when the model is used to assess the maturity at the Data Center level; Organizational in case of a more general approach including several aspects of the whole organization. All approaches divide the assessment of the sustainability levels in different categories and subcategories, but the Green Grid maturity model seems to give more detailed and technical information to help the organization to place their system in one level or another, while the other approaches describe categories and levels in a more qualitative way.

Another issue for a company should be how much to invest in Green IT. This decision should depend from the asset of the company. This problem is discussed in Ref. 15. The paper analyzes the factors that can influence the decision of the

Table 1. Comparison Between Assessment Models for Green IT

Maturity Model/ Best Practices	Level of operation	Number of Levels	Number of Categories	Level of Detail
Green Grid[13]	Data Centers - Facility & IT	6	2 main categories + several subcategories	high, each subcategory contains detailed information
G-Readiness Framework[5]	Organization - Business Processes & IT	7	5 and subcategories	low, levels are seen as a sort of score for each sub category
IVI Maturity Model[14]	Organization - Business Processes & IT	5	4 and subcategories	low, levels are seen as a sort of score for each sub category
EU Code of Conduct[4]	Data Centers - Facility & IT	N.A.	8	high, each practice highlights a needed action to be implemented

kind of governance for Green IT of an organization. Three levels of governance are identified: (i) *centralized*: no coordinator for Green IT but only an extension of other jobs, low importance; (ii) *federal*: a coordinator dedicated to the Green IT management, medium importance; (iii) *decentralized*: a committee as coordinator for Green IT, high importance, Green IT can lead the organization strategy. The contingency factors individualized are: competitive strategy, firm size, organization structure, performance strategy, environmental impact, environmental strategy, IT infusion, and IT diffusion. Supported by the model, companies can select the most promising Green IT governance form or compare their existent governance against the model, which can confirm or provide strategic directions for improving it.

4. Measuring Energy and Energy Efficiency

To assess the maturity level of an organization, the metrics proposed in the previous section need to be computed. Precise indications in the maturity models and best practices on how to compute these metrics are still work in progress. To this purpose, another branch of Green IT focuses its attention on measurements. In order to improve energy efficiency or to understand the state of the system, we need to be aware of its energy consumption. This branch can be divided into two parts, as shown in Fig. 4. In Sec. 4.1 we deal with energy consumption estimation: in some systems, measuring the amount of energy used by IT components is a non trivial problem. In Sec. 4.2 we see that the state of an IT system can be measured looking at metrics other than energy itself. This section addresses the research question *"How can we measure energy efficiency of an Information System?"*.

4.1. *Energy usage measurement and estimation in IT systems*

The first step that has to be taken in order to improve energy efficiency in an IT system consists in measuring the energy consumption of the system itself and

```
                    ┌──────────────┐
                    │ Measurement  │
                    └──────┬───────┘
              ┌────────────┴────────────┐
     ┌────────┴────────┐       ┌────────┴────────┐
     │    Energy       │       │   Indicators    │
     │   Estimation    │       │   Definition    │
     │    [21]-[25]    │       │  [31]-[37][12]  │
     └─────────────────┘       └─────────────────┘
```
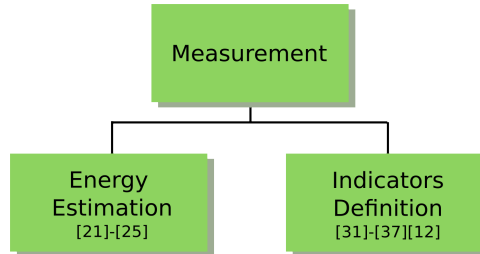
Fig. 4. Measuring energy and energy efficiency

eventually of the individual components of the system. Measuring the real energy consumption of a system under a specific workload is usually very hard and, most of the time, the precision of the results is very poor. In fact, energy consumption of devices changes in a non linear way depending on the load. That means that complex models need to be learned to estimate the power consumption of a device given its load. Some benchmarks have been defined to compare the behaviors of different configurations stressing the importance of both performance and power.[16] As an example, the Standard Performance Evaluation Corporation (SPEC) proposes a power and performance benchmark, SPECpower.[17] It provides a means to measure power in conjunction with a performance metric, helping IT managers to consider power characteristics along with other selection criteria to increase efficiency of data centers. SPEC includes power measurements even in other benchmarks, such as SPECweb, used for evaluating web server performance, and including also a power workload and a performance/power metric. The same approach can be found in SPECvirt, addressing performance evaluation of data center servers considering consolidation in virtualized environments. Another benchmark dealing with energy has been provided by the Transaction Processing Performance Council with the name of TPCEnergy.[18] This benchmark enriches former ones with information about energy with the purpose of helping IT managers in choosing the best configuration taking this aspect into account. A recent benchmark has been proposed by the GreenIT-BB network for showing potential energy and cost saving.[19] The benchmark is designed for energy and resource efficiency using environmentally sound performance parameters both at the hardware and software level. Data centers are classified in peer groups based on the number of servers and CPUs, the use area, and the operational purpose. In the end, thirty comparison values are given and a traffic light representation for each of them suggests dimensions that have to be improved to conform the average of the data centers in the same peer group.

Most of the benchmarks with attention on energy are still under development and some scholars prefer to use more consolidated benchmarks as a starting point to build a model for estimating power consumption. In Ref. 20, authors refers to TPC-C benchmark because they wanted to study the behavior of transactional systems. Starting from the system configuration explicit in the TPC-C tests, the authors

build a model for estimating power consumption taking into account the main components of the system: CPU, memory, disks, server chassis, and disk enclosures. While the energy consumption of CPU, memory and disks is estimated considering the peak power consumption reported in the manufacturer's specifications, the consumption of the server chassis has been estimated as the 30% of its component consumption plus a fixed overhead of 100 watts. The model was compared with real measured values and showed an error between 14% and 17% (see Ref. 20 for more details).

The principal utilization of benchmarks is in the comparison of several systems with the intent of choosing the best option. This decision is usually based on data describing the behavior of the system when running the benchmark. In other cases, the goal can be to improve an already existing system. In this scenario, metrics to evaluate a system need to be defined.

Measuring energy in case of hardware components means, for instance, to measure the energy consumption of a single processor or hard disk. If it is not possible to measure the energy consumption directly on the component, it can be estimated starting from the usage of the component itself and from the data sheet released by the manufacturer. Data about usage are provided by the operating system.

In addition, it is also useful to measure energy consumption of a software or middleware component. In some cases, it can be necessary to split the energy consumption of servers in the energy consumption of the virtual machines deployed on them or, at a higher granularity, to estimate the energy used by running applications. A considerable amount of work has been done in this direction, but the problem is still open. In Ref. 21, the authors try to estimate the energy consumption of a single application based on the resources used and the total amount of time, also considering the interaction among system resources. The resources considered are CPU, network, and disk. The estimation of the energy consumption is performed starting from the battery usage of a laptop. A similar approach is described in Ref. 22. Here a distinction is made about the consumption of each component in activity and in idle state.

Other approaches focus on the Virtual Machine level trying to produce models that approximate the energy consumption of each running VM. An example of this approach is described in Ref. 23, where the account of energy consumption is based on performance monitoring counters (PMC) that enable CPU and memory energy consumption estimation. PMCs are a hardware facility which collects data about the occurred events. The approach is based on the production of a power model collecting data for different loads of the components, which is then used to train the model. The model is created using linear regression to compute the weight in power consumption for each component. The per-VM accounting can be performed because of the abstraction provided by the operating system, able to gather per process PMCs. Another approach builds a parametric model to compute energy based on resources usage.[24] In the model only CPU, memory, and disk are considered. The contribution of each component is modeled as a linear function of

its usage. The accounting of each VM can be made knowing the amount of each physical resource used by it. This information is available from the hypervisor. The parameters of the linear functions need to be learned using linear regression.

In Ref. 25, authors try to estimate energy consumption in a cloud environment in order to compare it with conventional computing. They consider three approaches using cloud: (i) storage as a service, in which only storage is in the cloud; (ii) process as a service, in which cloud is used only for large computational tasks; (iii) software as a service, in which the cloud hosts both software and storage devices. Different models for energy consumption are described to represent the transport network, the data center, and even terminals and equipments owned by customers. A distinction is made between private and public clouds, which have a different cost for accessing to the network due to the devices and distances involved. For user devices and data centers, energy consumption is estimated from hardware specifications and from the assumption that the average power use in a data center is 66% of the maximum power. The energy consumption of the network is estimated considering all the devices involved, knowing their power consumption and their capacities expressed in bits per second.

An overview of the approaches for estimating energy efficiency discussed in this section can be seen in Tab. 2, highlighting the components that are considered in each of the presented approach and the level at which energy is estimated (Virtual Machine, Application, or Cloud). Estimation of energy is still an open issue and all the presented techniques use approximation in defining the energy consumption model.

Beyond IT resources energy consumption evaluation, other factors should also be considered. It is worth noticing that even software has a relevant impact on energy consumption, as emphasized in Ref. 26, due to a different resource management. Here, an empirical measurement of energy consumption has been made to compare Management Information Systems (MIS) running similar applications from the efficiency point of view. Three kinds of benchmarks have been used: 2 ERPs, 2 CRMs, 4 DBMSs, and significant differences in energy consumption for similar applications emerged. Even if the relation is non-linear, the fastest application proved to be also the most energy efficient. The Operating System itself has an impact, and Linux proved to be 50% more efficient than Windows, on average. This demonstrates the importance of being aware of the energy consumption of an application and of being able to measuring it. In Ref. 27, authors claims that a bad software can waste efficiency obtain with optimal hardware. They also identifies factors that influence software performance as computational efficiency (e.g. efficient algorithms and data structures, multithreading), data efficiency (e.g. data movement minimization, efficient use of cache memories), and context awareness (application that can adapt their behavior to changes in the environment).

Another aspect that is usually omitted when measuring energy and the impact of a system over the environment is related to the source of energy. Different sources of energy can not be considered as equal because some of them can be greener than

others. This concept is well described in Ref. 28 where authors make a distinction between green and non green energy, arguing that the true impact in the production of energy has to be considered together with the amount of energy consumed. This difference should be taken into account when allocating tasks to servers in the cloud. A similar approach is also faced by Ref. 29 where authors highlight the importance of using green energy to reach sustainability in IT systems and propose a load balancing algorithm that takes into account if the source of energy is green or not. This issue is also faced by the Eco2Clouds European project, which deals with energy efficiency and CO2 emissions in cloud environment, where emissions are computed starting from the knowledge of the energy mix of the source of energy for the data center.[30]

4.2. *Monitoring and assessment of energy efficiency*

In complex systems, measuring or estimating energy is not sufficient to enact an effective strategy for sustainability. Other features of the system can be monitored and evaluated with the intent of understanding its energy efficiency. For this reason, a wide set of indicators has been defined in the state of the art.

The most common way to measure efficiency of a system is the use of the PUE (Power Usage Effectiveness) metric. It measures the total energy of the data center divided by the amount of energy consumed by the IT. The Data Centre Metrics Coordination Taskforce was set up in 2010 to ensure clarity and consistency in data

Table 2. Comparison Between Energy Estimation Approaches for Applications and Virtual Machines

ID	Level	Resources	Description
Do2009[21]	Application	CPU, Network interface, Disk	Energy consumption is inferred monitoring the battery usage (Watt), considering the amount of resources used and time
Kansal2008[22]	Application	CPU, Disk	Energy is estimated distinguishing between idle and active mode of components and assuming to know the power consumption in the two modes
Bertran2010[23]	VM	CPU, Memory	PMC based power model with different weight for each component contribution. Parameters learned through linear regression
Kansal2010[24]	VM	CPU, Memory, Disk	Energy estimation using a model based on a linear function of usage of resources. Parameters learned through linear regression
Baliga2011[25]	Cloud	Network, CPU, Disk, Memory, GPU	Three kinds of clouds services are compared with conventional computing. Network, data center, and customer equipment are considered in energy computation

center metrics[c]. It published a white paper giving details of how to compute PUE in an effective way.[31] Other metrics are also defined to measure renewable energy technology and reuse of energy such as GEC (Green Energy Coefficient) measuring the percentage of green energy, ERF (Energy Reuse Factor) measuring the percentage of energy reused outside the data center, and CUE (Carbon Usage Effectiveness) measuring the amount of GHG emissions relative to IT energy consumption.

Even if PUE is very extensively used, it is most of the time a misleading metric in stating the energy efficiency of a data center. As discussed in Ref. 12, if the data center manager reduces energy consumption of the IT by managing resources in a better way (e.g. turning off a unused server), PUE will increase though the performance of the data center is improved. For this reason, PUE can not be considered as a standalone metric for stating greenness of a data center.

Another category of metrics has been defined for measuring efficiency at the application and service level. A well known set of indicators, related to the performance of the system, is composed by the Key Performance Indicators (KPIs). This set of indicators includes the most relevant metrics that are able to express the ability to respond to the requests in terms of quality and success of the whole organization, of a process or of a single activity. In order to meet the green requests, other indicators have been recently defined to express the energy efficiency and not only the performance of a system or organization. These indicators can be referred in the state of the art with the names of Green Performance Indicators (GPIs)[32] or Key Ecological Indicators (KEIs)[33] with the purpose of measuring the environmental impact of a monitored system, process or activity. Authors of Ref. 32 cluster indicators in four categories:

- *Application Lifecycle KPIs*: include all the indicators defined for monitoring the performance of a system, an application or an activity. They are only indirectly related to energy;
- *Organisational GPIs*: the organizational factors that help to define guidelines for managing the organization according to laws and regulations related to energy;
- *Energy Impact GPIs*: this cluster describes the impact of a system or application on the environment considering energy related factors;
- *IT resource usage GPIs*: this cluster is strictly related to IT and it is composed by metrics expressing the efficiency of resources in terms of their usage level.

Tab. 3 summarizes these four clusters of indicators making some examples. GPIs are usually associated with a threshold indicating the minimum and/or

---

[c]Members of the task-force are: Green Grid, the US Department of Energy's Advanced Manufacturing Office and Federal Energy Management Programs, the US EPA Energy Star Program, the European Data Centre Code of Conduct, Japan's Ministry of Economy, Trade and Industry's (METI) and Japan's Green IT Promotion Council (GIPC)

Table 3. Green Performance Indicators Clusters

Cluster Name	Definition	Examples
Application Lifecycle KPIs	Indicators defined for monitoring the performance of a system, an application or an activity	Response Time, Throughput, Availability, Recoverability, Reliability, Process Engineering
Organizational GPIs	Indicators describing organizational factors that helps to define guidelines for managing the organization according to laws and regulations related to energy	Human Resource, Compliance, Infrastructural Cost, Carbon Credit
Energy Impact GPIs	Indicators describing the impact of a system or application on the environment considering energy related factors	Application Performance, DCIE, PUE, CO_2 Emission, IOPS/Watt, Capacity/Watt
IT resource usage GPIs	Indicators expressing the efficiency of IT resources in terms of their usage level	CPU, Memory, I/O and Storage Usage, CADE, DH-UR, DH-UE

maximum desired value for that metric. Three levels of satisfaction are defined for GPIs: green if they are fully satisfied, yellow for partial satisfaction, and red for violation.[34]

The concept of KEIs proposed in Ref. 33 refers to a set of indicators used to measure the environmental impact of a process or of part of it, considering features such as energy consumption, water consumption, CO_2 emission, recycling, or regulatory compliance. KEIs are defined as a tuple consisting of a metric and a target value representing the goal that has to be achieved. This goal is defined by the system manager.

Ref. 35 summarizes metrics used in the assessment of the energy efficiency of data centers and cloud resources, and to assess energy-costs of application and workload execution for various data center granularity levels and sites. The work compares the approaches used by two EU projects: CoolEmAll and Eco2Clouds.

In Ref. 36, a new perspective about which kind of metrics should be used is presented. The authors find out that consumption-based metrics reflect energy savings better than considering only resource efficiency once the performance level (expressed as the number of operations to be executed) is fixed. In fact, while resource efficiency reflects the level of performance given the amount of energy used to do so, consumption based metrics show the energy needed to obtain a fixed performance. Using this kind of metrics, resources are easily comparable in relation to their energy efficiency. As an intuitive example, the authors use the comparison between the miles per gallon (MPG) metric (a resource efficiency metric) and the gallons per 100 miles (GPM) metric (a consumption-based metric) for comparing vehicles. They argue that the MPG metric can be misleading because it obscures the fact that small improvements in efficiency lead to important reductions in fuel consumption, highlighted by the GPM metric. The authors propose to apply this reasoning to decide whether it is recommendable to change a hardware device or not. A similar argument is defended in Ref. 37, in which "usage metrics" are introduced. Authors make a distinction between two perspectives for measuring indicators: the *manufacturer perspective* and the *usage perspective*. The former focuses on the potential

energy efficiency of the single component of the system, while the latter considers the efficient usage of these components. Authors stress the usage perspective suggesting that how a resource is used in a data center configuration is as important as the potential efficiency of the resource itself. So, the optimization of the actual usage of the components of a data center is necessary for improving energy efficiency.

In Ref. 12, Greenpeace proposes a methodology for assessing the greenness of big companies using the Clean Energy Index. Companies are scored considering the following aspects:

- *Coal and nuclear intensity*: the total percentage of coal and nuclear generated electricity used to power the data center;
- *Energy transparency*: the level of detail made publicly available related to the energy consumption of the IT infrastructure;
- *Infrastructure siting*: the siting criteria and investments for allowing the usage of cleaner energy;
- *Energy efficiency and GHG mitigation*: the strength of the strategies to mitigate the demand for not renewable energy generated by the IT infrastructure;
- *Renewable energy investment*: the commitment to renewable energy investments and to wide-scale renewable energy generation and use.

All the indicators described in this section can be used to evaluate the state of the system and to understand its efficiency level. The energy assessment can lead to some repair or adaptation actions if the efficiency is not satisfying. Techniques to improve the energy efficiency are discussed in Sec. 5.

5. Methods and Techniques to Improve Energy Efficiency

Metrics introduced in Sec. 4 can be used to detect situations of inefficiency in the monitored system. Once an inefficiency is detected, a set of techniques can be used to deal with it.

This section addresses the last research question: "*What can be done to improve*
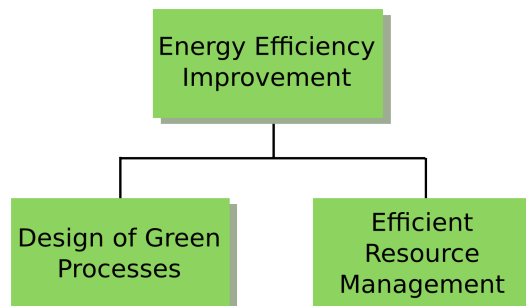

Fig. 5. Improving energy efficiency

energy efficiency in an Information System?", classifying some techniques and dividing them in two categories (Fig. 5). Most of the work at the state of the art about energy efficiency focuses on physical resources management, however this is not the only way to deal with energy saving. Some studies face the problem with a focus at a higher level, taking into account the process itself and how it can be modified in order to obtain a greener solution. So, in Sec. 5.1 the problem is considered from the point of view of designing processes, while in Sec. 5.2 the techniques of resource management for improving energy efficiency are analyzed. Given the wide body of literature on some specific aspects, this section does not want to be comprehensive of all the possible techniques adopted to deal with energy efficiency improvement, but only to give an overview of which parts of the system can be improved and available strategies to do that.

5.1. *Designing green processes*

Physical resources management is not the only way to deal with energy saving. Some studies face the problem with a focus at a higher level, taking into account the process itself and how it can be modified in order to obtain a greener solution.

According to Ref. 38, the role of business processes to create green IS and environmental sustainable organizations has to be explored with attention. Without a design and implementation of business processes with attention to green features, it is not possible to reach sustainability. So, according to the authors, the greening of an organization has to include a process re-engineering. To this purpose, four relevant areas are identified: (1) Process Design, (2) Process Measuring, (3) Process Improvement and Change, and (4) Process Implementation, covering the life cycle of a business process. Business process design and implementation should take into consideration environmental constraints and it should be possible to measure sustainability in order to reach transparency and awareness. Through awareness, humans involved in the process are more cooperative in working for sustainability. Moreover, awareness allows the management to change some aspects of the process to improve its efficiency.
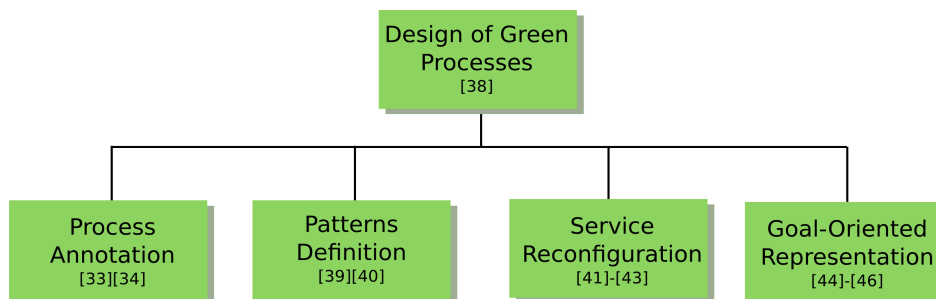

Fig. 6. Business Process design

The four areas identified in Ref. 38 are explored with some variations in other research papers. In fact, several approaches can be used to reach the goal. In this section we classify the approaches in the following categories: process annotation, patterns definition, service reconfiguration, and goal-oriented representation (Fig. 6).

5.1.1. *Business process annotation*

A research branch proposes the enrichment of the business process description with additional information that can be used to understand its behavior with respect to energy consumption.

In Ref. 33, the authors analyze the problem proposing a methodology to improve energy efficiency. It consists in re-engineering the business process considering its composing activities, enriched with information about their own energy efficiency. The re-engineering process is divided in four phases: (1) Strategy; (2) Sensing and Monitoring; (3) Analysis and Management and (4) Adaptation. In the first phase, a set of performance (KPI) and ecological (KEI) indicators are defined. These indicators are used to reflect the ecological objectives and the quality of service requirements of the process. In the sensing and monitoring phase, data to compute the indicators are collected and linked to the activities that use the monitored resources. The result of this phase is a process annotated with indicators. During the Analysis and Management phase, indicators are analyzed to identify the activities with the higher negative impact on the environment. From the set of activities only the ones that are modifiable and that violate the thresholds of the indicators are considered, and a weight is associated to them. The weight depends on the amount of the violation. The last phase, the adaptation, consists in applying some actions to improve energy efficiency that can involve a restructuring of the process or the modification of the allocated resources. The impact of an adaptation strategy is evaluated computing the value that KPIs and KEIs would assume after its execution. This evaluation enables the choice of the best strategy. However, there is no automatic decision maker, but the choice is performed by an expert knowing in advance all the needed information about the outcome of an adaptation action (examples of possible adaptation actions are described below in the Pattern-based approach).

Another adaptive approach to energy efficiency improvement in business processes is described in Ref. 34. Here, the business process is enriched with information related to past observations of the workflow, such as the probability of execution of a branch and the set of indicators (KPIs and GPIs) used to monitor the whole process or each single activity. A set of adaptation actions is also defined and linked to the violation of indicators by the definition of "rules". A rule is composed by an antecedent, which is the violation of an indicator, and a consequent which is the action that can be addressed to solve the violation. Some examples of actions are the reconfiguration of the resources allocated to the business process or to its specific activities or the reconfiguration of the business process itself by including

or excluding optional activities in the workflow. Some parameters are associated to each rule to describe the reliability (*confidence*) and the impact (*importance*) of executing that action in the current context. These parameters are computed considering the past history of the system. An algorithm is proposed to decide the best strategy to be enacted taking into account a hierarchical organization of the indicators. In this case all the violations are considered in the same way, without considering the entity of the violation. Through the described algorithm, the system is able to support the system administrator in the decision of appropriate actions to achieve goals based on KPIs and GPIs, suggesting all the available strategies that fit the actual situation and ordering them considering the confidence and importance parameters. Using this approach, the strategy can also be automatically selected, requiring a human intervention only when no improvement occurs.

5.1.2. *Patterns to design business processes*

One qualitative approach to improve energy efficiency of a business process is described in Ref. 39, where the authors propose some patterns to design green business processes. In this work the concept of green Business Process (BP) modeling as an extension of the classical BP modeling approach is introduced. While BP modeling considers four dimensions (cost, time, flexibility and quality), green BP modeling adds as a fifth dimension, the ecological one. Improving the classical four dimensions usually has an improvement over energy efficiency as a side effect. A set of patterns can be identified to lead the organization to reduce its environmental impact. Authors in Ref. 39 identify nine different patterns, which explore different ways to modify the business process toward a greener dimension. These are Green Compensation, Green Variant, Resource Change, Green Feature, Common Process Improvement, Process Automatization, Human Process Performance, Outsourcing and Insourcing, summarized in Tab. 4. A decision tree helps deciding which pattern/s can be activated proposing a set of questions about the process and the organization.

The approach has been further developed in Ref. 40, where the infrastructure level has been considered through the definition of a set of cloud patterns. The paper proposes a pattern-driven adaptation divided into three main phases: preparation (annotate patterns, identify ecological problems, identify green patterns), action (migration decision, migration), and follow-up analysis (monitoring and analysis). In the cloud context, business process components can be migrated inside the cloud infrastructure to reach a more efficient configuration. The application of the green business process patterns is supported by a set of cloud patterns, describing how to create and manage applications and services in a way supporting the ecological adaptation of the application. The cloud patterns considered in the approach can be grouped into three categories: *cloud environment* (private, public, hybrid), *cloud elasticity* (component removal, start and stop on demand), and resource sharing through *multi-tenancy*. Each of these patterns can be related to the business process

patterns described in Tab. 4, as described in Tab. 5[d].

5.1.3. *Service reconfiguration*

Some scholars faced the problem of sustainable business processes focusing their attention on business process reconfiguration. The general approach consists in decomposing the process in sub-tasks that can be substituted with similar but greener ones. This is the case of Ref. 41 where the authors propose a technique to re-design business processes basing the decision on qualitative and quantitative metrics, such as CO_2 emissions, air quality, and damages to fauna and flora. These measures are

---

[d]Other cloud patterns are described at `http://cloudpatterns.org/` and `http://en.clouddesignpattern.org/index.php/Main_Page`

Table 4. Patterns to reduce the environmental impact of a business process

Pattern Name	Pattern Description
Green Compensation	If the business process cannot be modified, a compensation activity can be started to compensate its environmental impact
Green Variant	An alternative green process is used achieving the same result in a greener way
Resource Change	Change the resources used by some of the activities with greener ones
Green Feature	Replace all the features for which a green alternative exists
Common Process Improvement	Some process or activities are improved using their KPI
Process Automatization	Some of the activities within the BP are automatized
Human Process Performance	Replace activities performed by machines that pollute with human activities
Outsourcing	Outsource certain activities from the owner to another partner providing an equivalent but greener service
Insourcing	Integrate services from third parties into its own organization

Table 5. Cloud and BP patterns relations

BP Pattern	Cloud Patterns
Green Compensation	Cloud environment
Green Variant	Cloud environment, cloud elasticity, multi-tenancy
Resource Change	Cloud environment, cloud elasticity
Green Feature	Cloud environment, multi-tenancy
Common Process Improvement	Cloud environment
Process Automation	Cloud environment, cloud elasticity, multi-tenancy
Human Process Performance	none
Outsourcing	Cloud environment, cloud elasticity, multi-tenancy
Insourcing	Cloud environment, cloud elasticity, multi-tenancy

annotated to the business process and are propagated to obtain a single cumulative value for each possible path. The business process is considered as composed by fragments. Each fragment can be substituted using a library of similar fragments, in order to find the greener solution to the problem. In the exploration of the possible solutions, a requirement of *design proximity* is used. It means that the system tries to obtain the best solution maintaining the status quo as much as possible.

Another example is discussed in Ref. 42. Here, the authors take into consideration both quality of service and energy efficiency metrics proposing a technique to aggregate and combine them in a non linear way. The process is represented in an abstract way and it is decomposed in sub-blocks. For each of the abstract sub-blocks several implementations exist with different costs in terms of energy efficiency and quality. The problem consists in finding the best configuration using service composition techniques and optimizing a general function for the whole problem. In order to do that, authors propose to use the CSOP (Constraint Satisfaction Optimization Problem) technique, under the assumption that the probability of execution of branches is known. Solutions that do not respect the constraints for unlikely paths are considered with a penalty, so that a wider range of solutions can be found. However, this approach can present some computability issues, since the resolution algorithm is NP-Hard. For this reason, it is effective only for relatively stable compositions.

A similar approach is presented in Ref. 43. The architecture in this case is divided into three levels. The higher one is the abstract business process composed by activities implemented through services that compose the middle level. Services are deployed over physical machines that constitute the lower level. Also in this case the problem of service composition is instantiated as a CSOP. The decision is based over a QoS value for each service, comprehensive of performance and green indicators. The main difference consists in the proposal of a model to compute energy consumption at the service level. Energy is estimated considering the percentage of the resources used by each service weighted with parameters computed observing the components behavior under different workloads. When deciding whether to modify the actual implementation of the abstract process, the cost of this modification is considered, in order to change it only if the gain is real.

5.1.4. *Goal-oriented representation*

In Ref. 44, authors propose a process representation using Goal Requirement Language to improve energy efficiency. In this model, energy efficiency goals are modeled as non functional requirements. That means that even if they are not mandatory, they are taken into account when selecting the best strategy. The model includes several elements, such as goals, soft-goals, tasks, resources, and beliefs. A goal is usually the objective that the organization wants to achieve and can be either a business or a system goal. Soft-goals are quality attributes representing the non-functional requirements. Tasks are different strategies that can be used to achieve

a goal, and they can be decomposed in subtasks. Tasks also use resources that can be available or not. All the elements are related to each other with different kinds of links, expressing the kind and the strength of the relation. Authors suggest an analysis strategy divided in four steps to help the decision process of a strategy to satisfy the goals. A five steps process is proposed: (i) *modeling IT system design objectives*, where the goals are identified; (ii) *generating IT system design alternatives* to satisfy the detected goals, they are represented as tasks connected to the related goal; (iii) *building environmental analysis models for the IT system design*, where the most important aspects of the IT systems environmental impact are identified together with their relation to the operational task that influence them; (iv) *evaluating the environmental impact of the design alternatives*, based on the positive and negative impact that each possible path has over detected environmental goals; (v) *improving the design alternatives and making design decisions* considering the results obtained at the previous step to compare the different solutions. This approach gives also the possibility to deal with performance and energy issues at the same time, giving more importance to performance.

Another goal-oriented approach is described in Ref. 45. Authors use a goal-oriented model for taking decisions about how to improve energy efficiency while maintaining the required QoS. Three layers are identified: goals, events and treatments. The contribution focuses principally on the Event Layer. Goals are represented by a set of metrics related to both energy efficiency and QoS, for which the value has to be maintained inside a given interval. When a violation occurs, an event is raised. The adaptation process is composed of two phases: the event creator and the adaptation strategy selector. In the first phase, the state of the system is analyzed in order to identify the goal violations that can be considered significant. In fact, not every violation can be categorized as an event. Temporary violations can be ignored since they can automatically recover without enacting any strategy. So, in order to raise an event, the violation should last in time. According to this approach, events are described by: a duration window, a direction (increasing/decreasing), a priority and a severity. When one or more events are raised, actions can be enacted to lead the system towards a normal state. An action is selected based on the events and past history. The same model is also considered in Ref. 46 where a prediction mechanism is proposed in order to anticipate goals violations in a service-based process. This mechanism is based on historical data and is addressed using an Auto-Regressive Moving Average (ARMA) model. Due to prediction, the system provides both pro-active and reactive adaptation behaviors. Once an actual or predicted violation occurs, an action ranker prescribe to the system administrator the best possible repair action to apply, considering both positive and negative outcomes of this action.

The four categories discussed in this section show the several possibilities available in the literature to improve energy efficiency acting over the process. A summary is

shown in Tab. 6, where the main features of each of them are highlighted. The enrichment of the process with annotations imposes constraints over energy consumption and highlights the features of the process itself that can be used to take green decisions about it. Examples of the available decisions can be represented through patterns that describe general solutions available. Service reconfiguration enables the modification of the process choosing the best solution in a pool of services, while goal-oriented representation allows the process designer to include greenness as one of the goals of the organization. All these approaches are not exclusive, and can be combined as a composed green solution.

5.2. *Efficient resource management*

The most common and intuitive approach to improve energy efficiency is related to IT resource management. Many techniques, born to improve the performance of IT systems, have a great impact also on energy. This set of techniques are widely applicable and flexible. In fact, resources are easy to be monitored and managed and techniques about resource management can be applied to each kind of system at every moment, without the need of an a-priori design. In this section resource management techniques are discussed considering three categories: virtualization and consolidation, cloud computing and networks, and storage energy efficiency

Table 6. Approaches to Green Process Design

Approach ID	Category	Short Description
Nowak2011[33]	Process Annotation	Activities annotated with information about their greenness and weighted accordingly
Cappiello2011[34]	Process Annotation	BP and activities annotated with information about their greenness and probability of execution
Leymann2011[39]	Patterns Definition	Definition of a set of general patterns to increase real or perceived greenness of a BP
Leymann2012[40]	Patterns Definition	Definition of a set of general patterns to increase real or perceived greenness of a BP in a cloud environment
Ghose2009[41]	Service Reconfiguration	Decomposition of the BP in interchangeable fragments associated with an energy efficiency parameter
Ferreira2009[42]	Service Reconfiguration	Decomposition of the BP in blocks and composition implemented as an optimization problem
Oliveira2010[43]	Service Reconfiguration	Decomposition of the BP in blocks and composition implemented as an optimization problem. Energy estimated at the service level
Zhang2010[44]	Goal-Oriented Representation	Use of the GRL language to represent the process and reason about the alternative solution and their energetic impact
Ferreira2013[45]	Goal-Oriented Representation	Modeling of a goal based model to represents the effect of indicators and to prescribe adaptation strategies focused on events detection
Plebani2013[46]	Goal-Oriented Representation	Modeling of a goal based model to represents the effect of indicators and to prescribe adaptation strategies enriched with violations prediction

improvement (Fig. 7).

5.2.1. *Virtualization and consolidation*

Every system uses a set of resources that need to be allocated to it. If the allocation is oversized, unused resources are wasted, and the result is a waste of energy too. In the last years, a lot of studies have demonstrated the important contribution of efficient resource allocation in energy saving, enabled by common techniques such as virtualization and cloud computing.

Virtualization consists in allocating virtual resources instead of physical ones to a machine, allowing the system to run more than one machine on a single server and to distribute physical resources in a more efficient way. The use of Virtual Machines allows an efficient and dynamic management of the available resources that would be impossible using only physical machines. In this way, resources that are underutilized can be reallocated. At the same time, an application requesting more resources than expected can be empowered to satisfy the Service Level Agreement stipulated with the user. A lot of efforts have been made to find a trade off between efficient resource allocation and performance, creating a field known as server consolidation. The aim of server consolidation consists of allocating virtual machines to physical devices in the most efficient way. Trying to concentrate the allocated resources over less servers, servers that are not used can be turned off or put in stand by with a consequent energy saving. The problem and the factors that influence consolidation and its performance are analyzed in Ref. 47. The three factors considered are: the amount of incoming workload, burstiness and SLA constraints. Two time scales are considered: slow (the time needed from the Data Center to change its configuration), fast (the variability of request arrivals). Results have shown that burstiness negatively affects the cost saving of dynamic consolidation, since more resources are needed if the workload is not constant in a relatively long time. Strict constraints in the SLA have a negative effect only if the arrival rate is high. The peak-to-mean ratio (workload incoming rate) has also an important role, the higher the better for cost saving opportunities. According to this, consolidation is effective only where
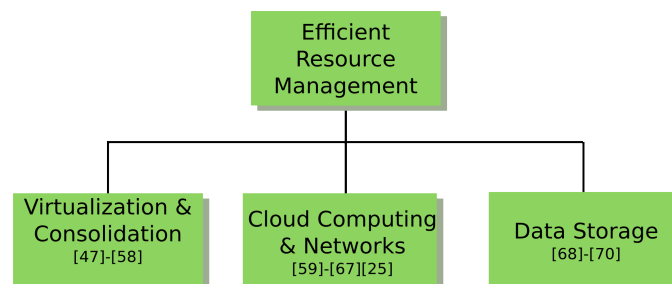

Fig. 7. Efficient resource management

burstiness is limited and principally when the workload changes significantly over time. Given these premises, several researchers analyzed possible techniques to allow dynamic consolidation in a virtualized environment. Simple solutions can be used in case of static workload considering only performance.[48] Here the reconfiguration is lead by a cost and utility function, considering only CPU and memory as devices. The algorithm is able to refine its parameters during the execution, but it is not designed to deal with dynamic loads. The gain in performance is around 25%. A similar context is described in Ref. 49. It considers the problem of VM allocation in case of static workload, where the only factor to consider is the initial allocation of a new VM. Several algorithms are compared and three different objectives are considered: job performance maximization under power consumption constraints, power consumption minimization under job performance constraints, and multi-objective optimization of both power consumption and job performance. A heuristic is needed since all the algorithms are NP-Hard to solve. The multi-objective approach seems to perform better, even if it requires a longer elaboration time, that can be a problem in case of big data centers. The gain in energy terms for the best algorithm is around 10%.

As stated before, virtualization techniques were initially used to improve performance and not energy. Only few authors focus their attention also on energy aspects. An example is in Ref. 50, where a comprehensive approach is presented to efficiently allocate resources maximizing SLAs while reducing costs. A set of possible interventions to improve energy efficiency are coupled with virtualization. These interventions are divided into two categories that we adopt in the analysis in this section: *short term* and *long term*. As "short term" the authors classify all the actions that can be actuated in a small time range such as load balancing, capacity allocation, and frequency scaling. As "long term" they identify more complex actions such as server switching to power off and resource allocation to applications in servers and reconfiguration. Short term interventions can be enacted with a higher frequency than long term ones. An autonomic computing approach is presented to allocate resources. The service center is modeled as a queue of requests that can be satisfied by one or more applications. The problem consists in deciding the best solution about which application to use and which resource to allocate. A cost model is presented consisting of utility functions in which penalties are related to the performance level and to the energy cost associated with the use of servers. A first solution is approximated using a heuristic approach.

Simpler systems concentrate their efforts in allocating virtual machines to available processing units trying to keep the frequency of the processor as low as possible.[51] This is because a big amount of energy consumed by a server is due to the processor and the lower is the frequency, the lower is the energy consumption. Optimization techniques are used to distribute the load to the different virtual machines in an efficient way. Before being allocated, virtual machines are modeled in terms of required processor frequency and execution time. The allocation algorithm processes VM allocation one by one by allocating first VMs with a higher

processing frequency requirement.

Other approaches tend to optimize resource allocation taking advantage of some features of the virtualization technology, such as min, max and share parameters, to set the amount of each resource requested.[52] Thorough these parameters it is possible to specify the minimum amount of a resource needed to run the virtual machine, the maximum amount that can be allocated, and how much it is possible to share the allocated resources with other VMs. Guided by application utilities, high utility applications get most of the resources. The improvement in the overall utility of the data center is around 47%.

In Ref. 53, authors propose a two level control architecture to guarantee response time while minimizing energy. At the first level, load balancing is used to guarantee that each VM respects response time constraints. The second level controls CPU frequency to increase energy efficiency. In Ref. 54 a control loop is designed to dynamically reconfigure VMs, monitoring and evaluating the current state, and planning and executing modifications. System load and response time are used to assess the level of performance, and reconfiguration operations are defined, such as switching off servers and frequency scaling. A similar approach is proposed in Ref. 55. The main difference consists in the fact that authors consider a multi-tier environment composed of three tiers: web server, application server, and data base server. A methodology named PowerTier is proposed for improving energy saving by reducing the number of active servers and setting the optimal processor frequency. The Lagrangian optimization method is used to optimize power consumption while satisfying response time constraints. The algorithm succeeds in determining the number of servers needed and the optimal frequency both at design and at run time.

Artificial intelligence techniques are often used to deal with uncertain information and to predict the future behavior of the system. The approach proposed by Ref. 56 deals with resource allocation and switching off servers. The approach, focused on CPU intensive workloads, uses machine learning to predict the SLA level in a new configuration helping in deciding the goodness of the solution. The system tries to consolidate tasks into as few machines as possible, without affecting performance. Models to predict power consumption and client satisfaction of the future configuration are learned from collected data. The algorithm reduces power consumption by 10%. In Ref. 57 a simpler system is presented. Here, a reinforcement learning approach is used to apply efficient system reconfiguration to deallocate unused resources and to increase insufficient resources. The resources considered are only memory and CPU. A Markov decision process is used to choose the best policy. The policy consists in the best reconfiguration action given a state and can be learned by enacting random actions at the initial steps and collecting information about performance of each VM in the new configuration. This approach is dynamic, in the sense that a new policy is automatically learned when the external environment changes. Two different algorithms are compared in Ref. 58 to control the CPU frequency in order to save energy. This has been done working on P-states, which

are a pre-defined set of working states of a CPU. The number of P-states depends on the specific processor, switching from a P-state to another allows clock rate reduction. The authors propose a human immune system inspired approach and a fuzzy logic approach. The former is divided into two steps. In the initialization stage, data are collected and associations between workloads and P-states of the CPU are acquired, forming "detectors". In this phase, also "effectors" are acquired, associating actions to a given state. The second step is the self-optimization stage, following the MAPE (monitoring, analyzing, planning, executing) paradigm. Here, collected data are compared with optimal data about maximum and minimum workloads for each P-state, and a strategy is actuated if thresholds are not satisfied, which consists in changing the P-state. The second approach uses a fuzzy technique, defining three sets for each P-state: HIGH, LOW and MEDIUM. A fuzzy controller, depending from this value, is evaluated regularly and a P-state modification occurs when its value reaches 0 or 1. From tests, the immune inspired techniques seems to be more successful because it is more able to capture dynamic modification in the workload.

An overview of different approaches to an efficient allocation of the resources can be seen in Tab. 7, highlighting the main features of each approach and the long term and short term strategies adopted. As it can be seen in the table, the approaches differ for the combination of long and short term techniques adopted and for the decision algorithm used to decide when and how to change the configuration.

5.2.2. *Cloud computing and networks*

Similar approaches can be used to deal with cloud computing. Cloud computing consists in the dynamic allocation of shared resources to a virtual machine. Differently from the previous approach, in this case the user of the system is not aware of the physical location of the resources allocated to its service. As before, the resource allocation problem consists in the search of the best physical machine where to place the virtual machine. The choice is not limited to a single data center as in the previous class of approaches, but it is usually extended to a set of servers positioned in different data centers, geographically distributed. As a consequence, issues related to network transportation costs are often considered by the allocation algorithm. An example can be found in Ref. 59. Here the problem of resource allocation is considered as a bin packing problem and an energy-aware heuristic algorithm is proposed to find a solution. The algorithm also deals with the dynamic resource demand of applications by saving around 10% of energy. An other approach is proposed in Ref. 60, where authors propose a technique for improving energy efficiency in a cloud environment using three basic actions: VM reconfiguration, VM migration, and physical server powering off/on. In the model proposed by the authors, also the cost of each of these actions is considered, in the attempt to keep the number of SLA violations as low as possible. The reconfiguration action has the aim to optimize the resource usage level, avoiding under- and over-utilization. Several algorithms are proposed for migration, giving penalties for overloaded hosts and

Table 7. Resource Allocation Improvement Strategies

ID	General Approach	Short Term Strategies	Long Term Strategies
Soror2010[48]	Static load balancing using utility and cost functions	none	resource allocation, only CPU and memory
Pierson2012[49]	multi-objective optimization of energy and performance	none	resource allocation, only CPU and memory
Ardagna2010[50]	Autonomic computing approach using an optimization model, a heuristic is used to find the first solution	load balancing, capacity allocation, frequency scaling	server switch off, resource allocation
Laszevski2009[51]	Ordered allocation of processing unit keeping frequency as low as possible	frequency scheduling	resource allocation
Cardosa2009[52]	Resource allocation using hypervisor features (min, max and share parameters)	none	resource allocation
Wang2008[53]	Response time guarantee through a two layer architecture	load balancing, frequency scaling, CPU allocation	none
Petrucci2009[54]	Loop procedure for reconfiguring VM dynamically	frequency scaling	server switch off, resource allocation
Wang2013[55]	Langrangian optimization for multi-tier configuration	frequency scaling	server switch off
Berral2010[56]	Machine Learning approach to consolidate VM keeping SLA satisfaction, energy and SLA prediction in future configurations	none	server switch off, resource allocation
Rao2009[57]	Dynamic resource allocation using reinforcement learning with reward based on performance	none	resource allocation
Salomie2011[58]	Immune-inspired and fuzzy-based techniques	frequency scaling	none

for number of migrations, while giving a reward for every empty host. Maximum energy saving using only reconfiguration is estimated to be around 61.6%; adding migration it can increase to 37%. Similarly, in Ref. 61 the resource allocation takes into consideration several factors, related both to energy and QoS. In the specific, three factors are considered: total energy consumption, number of SLA violations (expressed as MIPS not allocated to a VM when needed), and number of migrations. Resource allocation is divided into two phases: new VMs placement and old VMs allocation optimization. The first problem is solved using a bin packing algorithm, while trying to keep CPU utilization of the server between a maximum and a minimum threshold. The second part consists in deciding when a VM should be migrated. The decision is again based on the thresholds over CPU utilization: when the server has low utilization, its VMs are migrated to put it in sleep mode; when

it has high utilization some VMs has to be moved to ensure SLA satisfaction. The possible energy saving declared by the authors is between 53% and 77% if compared to a non energy-aware algorithm, with a performance degradation of 5.4%.

In Ref. 62, authors propose a framework for energy-efficiency awareness and adaptation. Starting from real power measurements at the server level, they propose models for assessing and predicting efficiency at several levels: server, virtual machine, infrastructure, and service. Optimization is achieved through an algorithm considering actions as consolidation and migration of virtual machines, and deployment and undeployment of services. The objective function is the maximization of the total resulting efficiency of all nodes. Authors declare an improvement in energy efficiency of 14%.

When dealing with a cloud environment, an other factor can have a relevant impact on the allocation of tasks to resources: the source of energy. This issue has been discussed in Sec. 3 and at the end of Sec. 4.1. Recent research tends to consider the aspect of an efficient usage of renewable energy and of the avoiding of peak demand of energy from the electricity providers. An example is discussed in Ref. 63 proposing to use Geographical Load Balancing (GLB) to shift workloads and avoid peak power demands. The incoming workload and peak demand need to be predicted and this is achieved using the SARIMA prediction method and the RHC technique. Since accuracy decreases with the length of the prediction window, prediction error should be considered. GLB is modeled as a network flow optimization, considering workload split, power state of servers, renewable energy sources and energy storage devices. Workload and renewable energy production have been modeled using real traces. GLB results more effective than performance balancing in maximizing renewable utilization. However, prediction has to be improved to improve the management of the peak power. The same issue is faced in Ref. 64. Authors propose a combination of two techniques for avoiding peak loads demands on the energy network: workload shifting and local power generation usage. Savings in energy usage can be up to 40% than using either techniques alone and emission can be reduced of 10-15%. The algorithms are defined for data centers participating in Coincident Peak Pricing (CPP) programs, charging more when the whole network is in peak of requests. These peaks are impossible to be predicted. Peak pricing is composed of: (i) connection charge, (ii) usage charge, (iii) peak demand charge (an incentive to a uniform electricity demand), (iv) coincident peak charge. A cost model is proposed considering the power supply (local backup power and renewable supply) and demands (IT and cooling) of a data center. The algorithm optimizes local generation and workload shifting for reducing energy cost.

As stated before, network transportation is a relevant component to be considered when dealing with clouds. Some recent studies focused their attention on that, stating that network transportation of data needs energy and under some circumstances using a cloud system can be more expensive than using a conventional one.[25] In fact, energy consumption in transport and switching can be significant compared to the total energy consumption in cloud computing. In particular, cloud

computing could not be the best solution in services in which there is an high rate of access to storage devices that are placed in the cloud. That means that, when deciding whether to use cloud computing, considerations related to the nature of the workload have to be made.

Network efficiency is a very important variable in distributed systems. As discussed in Ref. 65, in order to obtain a green ICT system, a lot of data about monitoring and configuration have to be transmitted over the network. But this transmission has a cost that shouldn't overcome the benefits. Some scholars have been studying techniques to improve energy efficiency in networks working on routing control while keeping QoS level high. As an example, in Ref. 66, the routing optimization is modeled as an optimization problem, using as goal a function depending on network power consumption and delay of packets transmission at each node. The algorithm, solved using gradient descendant, can help in load balancing, distributing the traffic over the network and taking into consideration both QoS and energy.

An overview about reaching energy efficiency in a cloud environment is given in Ref. 67, where authors analyze the problem of software development life cycle in a cloud environment with a green perspective. Authors also propose a green cloud computing framework in which they identify opportunities for energy efficiency in efficient hardware and software selection, network optimization, VM scheduling and management, clean energy source selection, and efficient data center design.

5.2.3. *Improvement of energy efficiency for data storage*

A branch about resource management is related to storage energy efficiency. Storage devices impact on energy is really important and a good management of these devices can reduce energy waste significantly. Two main approaches can be detected in this field. The first approach reduces energy waste by working with disk speed. Disk can work with different speed modes, wasting more or less energy. An example of this approach is discussed in Ref. 68. Here the authors provide an algorithm able to determine the optimal disk speed setting to optimize energy efficiency. In the optimization function, also the cost of changing the disk mode is considered, and a modification can be made only after a minimum time period if there is no need for performance issues, avoiding changing mode too frequently. The energy performance improvement of the approach is estimated around 29%. The other approach deals with data migration in order to consolidate data in a smaller amount of devices, enabling turning off of unused devices or putting them in stand by mode. An example is discussed in Ref. 69 where authors use a conservation technique called Popular Data Concentration (PDC) that migrates frequently accessed data to a subset of the disks. In this way, the rest of the devices can be turned to lower-power modes. Data are organized hierarchically, based on their access rate. It is worth noticing that data in the cache will be placed in the lower level of the hierarchy. The approach can conserve up to 30-40% of the disk energy. There are

Table 8. Comparison Between Assessment Models for Green IT

	Why Important	Sub Topics	What Directions	Challenges
Efficiency Assessment	It helps understanding the maturity level of the system and identifying future steps for improvement	Maturity Models, Best Practices	Regulations and constraints can raise the attention about energy efficiency and greenness, pushing towards a faster improvement in the field	Definition of a general set of goals for a heterogeneous set of organizations, definition of significant goals that can orient development towards efficiency
Efficiency Measurement	It allows assessing the level of maturity and measuring the effects of the enacted strategies	Energy Estimation, Indicators Definition	Definition of models for measuring energy at hardware and software level, definition of new indicators for measuring energy efficiency	Measuring energy consumption with a high level of precision and at different levels of the organization, definition of a significant set of metrics to be used in evaluating the system
Efficiency Improvement	It allows the reduction of the amount of energy consumed, it allows an adaptive behavior as a reaction to non-optimal situations	Design of Green Processes, Efficiency Resource Management	Using efficiency improvement techniques can significantly improve energy efficiency only with a coordination of techniques at different levels, considering the whole life-cycle of the system	The definition of a comprehensive approach for energy efficiency is still an open issue; efficiency is sometimes in contrast with quality of service and it requires a redefinition of SLAs

also approaches that combine the two techniques already described, as in Ref. 70. Here, the authors control disk dynamically, taking into account the actual usage of data and disk characteristics. The action performed are orchestrated by two controllers: the former controls disk modes and the latter manages data migration. The behavior of the controllers is lead by a set of policies expressed as fuzzy rules and clustered in three categories: policies for new data allocation (which is the best disk to store new data depending from the expected access rate), policies for disk modes (deciding when it is the case to switch a disk in quiet mode and to switch it back in normal mode) and policies for data migration (positioning of data in disks depending on their access frequency).

6. Concluding Remarks and Expected Evolution

In this paper we have analyzed the state of the art related to the recent Green IT topic, considering it mostly from an Information Systems perspective. From this analysis three main areas emerged, linked to different research questions. The first area, answering the question *"How it is possible to state when an Information System is green?"*, is related to the definition of models and suggestion of practices to assess and improve the sustainability level of an organization. The second area,
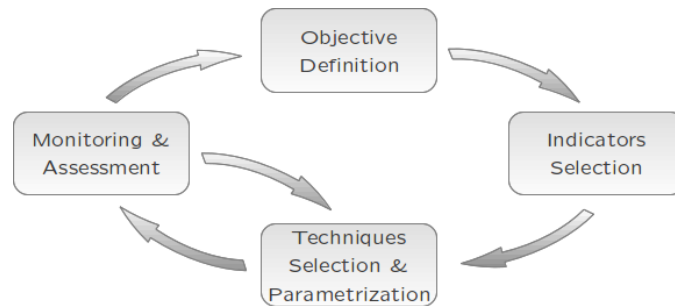
Fig. 8. Design process of a Green IT system

addressing the question "*How can we measure energy efficiency of an Information System?*", faces the problem of measuring energy and of defining useful metrics for monitoring and controlling the greenness of the system. In the last area, answering the question "*What can be done to improve energy efficiency in an Information System?*", techniques for improving energy efficiency are analyzed. These techniques can be distinguished into two main categories, that are high level techniques which affect the business process of the organization, and low level techniques which deal with efficient resource management. A summary of the techniques discussed and of the trends and challenges related to them can be seen in Tab. 8.

The three detected areas and their relative sub-categories cannot be considered as independent, but some relations can be found. For instance, the assessment area needs metrics and requests the adoption of efficient techniques. Conversely, the monitoring of the system state requests the definition of thresholds for the measured metrics in order to detect non-green situation in the system. These thresholds could be defined explicitly or implicitly in the assessment models. Moreover, adaptation techniques require metrics selection for verifying their efficacy or for driving decisions. However, these relations are not always exploited and the three areas seem to be completely independent in most of the research work analyzed above. This can be seen as the main drawback of the approaches we discussed here. An approach taking into considerations all the identified areas is still missing.

We can imagine that future trends in research about Green IT for Information Systems will try to combine the three areas in order to obtain a comprehensive approach able to increase energy efficiency. In our opinion, the design of a green IT system should be structured as in Fig. 8.

The first activity for the design of a Green IT system should be the assessment of the actual system taking as an example one or more maturity models described in Sec. 3. The assessment is obtained by comparing the desired state with the real one, that can be observed from the data collected through a monitoring system. Once the maturity level has been established, a set of objectives has to be defined. The objectives should depend either on the desired maturity level to reach or on a set of best practices that the organization wants to enact. The next step consists in selecting

the indicators that can help the organization to assess the accomplishment of the objectives selected in the previous phase. A monitoring system is needed to support the collection of data from which the indicators values can be computed. Finally, techniques to improve energy efficiency can be selected, both related to business processes and to resources allocation. These techniques should be parametrized according to the objectives. This can be done setting thresholds for the indicators or deciding the lower and upper bounds for resources usage. After an observation time depending on the type of techniques selected, two directions can be taken. If the objectives are still not reached, additional techniques can be applied or parameters can be tuned. Otherwise, if the goal is reached, the organization can decide to move to the next maturity level, setting new objectives and improving energy efficiency further on. The evolution and maintenance of the system are intrinsically supported by this model, since the model is independent from the specific process, and any changes in the objective definition automatically leads to the reconsideration of the other states, providing a dynamic and context-aware approach.

In any case, the current interest toward greenness and energy efficiency makes us believe that this field will continue its development and that soon all organizations will consider this aspect in their business plans.

Acknowledgements

References

1. Beloglazov, A., Buyya, R., Lee, Y.C., Zomaya, A.: A taxonomy and survey of energy-efficient data centers and cloud computing systems. Advances in Computers **82**(2) (2011) 47–111
2. OECD: Towards Green ICT Strategies: Assessing Policies and Programmes on ICT and the Environment. Technical report (June 2009)
3. EPA: ENERGY STAR® Program Requirements for Computer Servers. Technical report (2007)
4. Newcombe, L., Acton, M., Booth, J., Flucker, S., Latham, P., Strutt, S., Tozer, R.: 2012 Best Practices for the EU Code of Conduct on Data Centres. Technical report, European Commission (2011)
5. Molla, A., Cooper, V., Pittayachawan, S.: The Green IT Readiness (G-Readiness) of Organizations: An Exploratory Analysis of a Construct and Instrument. Communications of the Association for Information Systems **29**(1) (2011) 66–96

6. Murugesan, S.: Harnessing Green IT: Principles and Practices. IEEE IT Professional **10**(1) (2008) 24–33

7. Watson, R.T., Boudreau, M.C., Chen, A.J.: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. Management Information Systems Quarterly **34**(1) (2010) 23–38

8. Seidel, S., Recker, J.C., Pimmer, C., Vom Brocke, J.: Enablers and barriers to the organizational adoption of sustainable business practices. In: Proceedings of the 16th Americas Conference on Information Systems: Sustainable IT Collaboration around the Globe. (2010)

9. Vom Brocke, J., Seidel, S.: Green Business Process Management: Towards the Sustainable Enterprise. Springer (2012)

10. Pernici, B., Aiello, M., Vom Brocke, J., Donnellan, B., Gelenbe, E., Kretsis, M.: What IS can do for environmental sustainability: a report from CAiSE'11 panel on green and sustainable IS. Communications of the Association for Information Systems **30**(1) (2012)

11. Houy, C., Reiter, M., Fettke, P., Loos, P., Hoesch-Klohe, K., Ghose, A.: Advancing Business Process Technology for Humanity: Opportunities and Challenges of Green BPM for Sustainable Business Activities. In: Green Business Process Management. Springer (2012) 75–92

12. Cook, G.: How Clean is Your Cloud? Technical report, Greenpeace International (April 2012)

13. The Green Grid Consortium: Data Center Maturity Model. White Paper (2011)

14. Donnellan, B., Sheridan, C., Curry, E.: A Capability Maturity Framework for Sustainable Information and Communication Technology. IEEE IT Professional **13**(1) (2011) 33–40

15. Schmidt, N.H., Kolbe, L.: Towards a contingency model for green IT Governance. ECIS 2011 Proceedings (2011)

16. Rivoire, S., Shah, M.A., Ranganatban, P., Kozyrakis, C., Meza, J.: Models and metrics to enable energy-efficiency optimizations. Computer **40**(12) (2007) 39–48

17. SPEC. http://www.spec.org/benchmarks.html

18. TPC: TPC-Energy Specification - Standard Specification Version 1.2.0. Technical report (2010)

19. Israel, Y., Leitert, T.: The GreenIT DC-benchmarking tool: from scientific theory to real life. In: Energy Efficient Data Centers. Springer (2012) 47–53

20. Poess, M., Nambiar, R.O.: Energy cost, the key challenge of today's data centers: a power consumption analysis of TPC-C results. Proceedings of the VLDB Endowment **1**(2) (2008) 1229–1240

21. Do, T., Rawshdeh, S., Shi, W.: pTop: A Process-level Power Profiling Tool. In: Workshop on Power Aware Computing and Systems (HotPower'09). (2009)

22. Kansal, A., Zhao, F.: Fine-grained energy profiling for power-aware application design. ACM SIGMETRICS Performance Evaluation Review **36**(2) (2008) 26–31

23. Bertran, R., Becerra, Y., Carrera, D., Beltran, V., Gonzalez, M., Martorell, X., Torres, J., Ayguade, E.: Accurate Energy Accounting for Shared Virtualized Environments using PMC-based Power Modeling Techniques. In: International Conference on Grid Computing. (2010)

24. Kansal, A., Zhao, F., Liu, J., Kothari, N., Bhattacharya, A.A.: Virtual machine power metering and provisioning. In: Proceedings of the 1st ACM symposium on Cloud computing. (2010) 39–50

25. Baliga, J., Ayre, R.W.A., Hinton, K., Tucker, R.S.: Green cloud computing: Balancing energy in processing, storage, and transport. Proceedings of the IEEE **99**(1) (2011)

149–167

26. Capra, E., Francalanci, C., Slaughter, S.: Measuring Application Software Energy Efficiency. IEEE IT Professional **14**(2) (2012) 54–61

27. Sabharwal, M., Agrawal, A., Metri, G.: Enabling Green IT through Energy-Aware Software. IT Professional **15**(1) (2013) 19–27

28. Pierson, J.M.: Green Task Allocation: Taking into Account the Ecological Impact of Task Allocation in Clusters and Clouds. Journal of Green Engineering **1**(02) (2011) 11761

29. Liu, Z., Lin, M., Wierman, A., Low, S.H., Andrew, L.L.H.: Geographical load balancing with renewables. In Proc. of Sigmetrics 2011 (2011)

30. Wajid, U., Pernici, B., Francis, G.: Energy Efficient and CO2 Aware Cloud Computing: Requirements and Case Study. In: Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on. (2013) 121–126

31. Data Centre Metrics Coordination Taskforce: Harmonizing Global Metrics For Data Center Energy Efficiency. Whitepaper

32. Kipp, A., Jiang, T., Fugini, M., Salomie, I.: Layered Green Performance Indicators. Future Generation Computer Systems **28**(2) (2012) 478–489

33. Nowak, A., Leymann, F., Schumm, D., Wetzstein, B.: An Architecture and Methodology for a Four-Phased Approach to Green Business Process Reengineering. In: Proceedings of the 1st International Conference on ICT as Key Technology for the Fight against Global Warming, ICT-GLOW 2011, August 29 - September 2, 2011, Toulouse, France. Volume 6868 of Lecture Notes in Computer Science (LNCS)., Springer-Verlag (2011) 150–164

34. Cappiello, C., Fugini, M.G., Ferreira, A.M., Plebani, P., Vitali, M.: Business Process Co-Design for Energy-Aware Adaptation. In: Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on. (2011) 463–470

35. Volk, E., Tenschert, A., Gienger, M., Oleksiak, A., Sisó, L.: Improving Energy Efficiency in Data Centers and federated Cloud Environments. In: International workshop on European actions towards eco-friendly data centers, Karlshrue. (2013)

36. Larrick, R.P., Cameron, K.W.: Consumption-Based Metrics: From Autos to IT. Computer (2011) 97–99

37. Chen, D., Henis, E., Cappiello, C., et al.: Usage Centric Green Performance Indicators. In: Proceedings of the Green Metrics 2011 Workshop (in conjunction with ACM SIGMETRICS 2011). (2011)

38. Seidel, S., Vom Brocke, J., Recker, J.C.: Call for Action: Investigating the Role of Business Process Management in Green IS. Sprouts: Working Papers on Information Systems **11**(4) (2011)

39. Nowak, A., Leymann, F., Schleicher, D., Schumm, D., Wagner, S.: Green Business Process Patterns. In: Proceedings of the 18th Conference on Pattern Languages of Programs, ACM (2011)

40. Nowak, A., Binz, T., Fehling, C., Kopp, O., Leymann, F., Wagner, S.: Pattern-driven green adaptation of process-based applications and their runtime infrastructure. Computing **94(6)** (2012) 463–487

41. Hoesch-Klohe, K., Ghose, A.: Carbon-Aware Business Process Design in Abnoba. Service-Oriented Computing (2010) 551–556

42. Mello Ferreira, A., Kritikos, K., Pernici, B.: Energy-Aware Design of Service-Based Applications. In: International Conference on Service-Oriented Computing, Springer (2009) 99–114

43. De Oliveira, A., Frederico, G., Ledoux, T.: Self-Optimisation of the Energy Footprint in Service-Oriented Architectures. Proceedings of the 1st Workshop on Green

Computing (GCM'10) (2010)

44. Zhang, H., Liu, L., Li, T.: Designing IT systems according to environmental settings: A strategic analysis framework. The Journal of Strategic Information Systems (2011)

45. Ferreira, A.M., Pernici, B.: Using Intelligent Agents to Discover Energy Saving Opportunities within Data Centers. In: Second International Workshop on Requirements Engineering for Sustainable Systems (RE4SuSy). (2013)

46. Miracca, A., Plebani, P.: Increasing energy saving with service-based process adaptation. In: 2nd International Conference on Smart Greeds and Green IT Systems. (2013)

47. Wang, K., Lin, M., Ciucu, F., Wierman, A., Lin, C.: Characterizing the impact of the workload on the value of dynamic resizing in data centers. In: IEEE Infocom. Volume 40. (2013)

48. Soror, A.A., Minhas, U.F., Aboulnaga, A., Salem, K., Kokosielis, P., Kamath, S.: Automatic virtual machine configuration for database workloads. ACM Transactions on Database Systems (TODS) **35**(1) (2010)

49. Borgetto, D., Casanova, H., Da Costa, G., Pierson, J.M.: Energy-aware service allocation. Future Generation Computer Systems **28**(5) (2012) 769–779

50. Ardagna, D., Panicucci, B., Trubian, M., Zhang, L.: Energy-Aware Autonomic Resource Allocation in Multitier Virtualized Environments. Services Computing, IEEE Transactions on **5**(1) (2012) 2–19

51. von Laszewski, G., Wang, L., Younge, A.J., He, X.: Power-aware scheduling of virtual machines in DVFS-enabled clusters. In: Cluster Computing and Workshops, 2009. CLUSTER'09. IEEE International Conference on. (2009) 1–10

52. Cardosa, M., Korupolu, M.R., Singh, A.: Shares and utilities based power consolidation in virtualized server environments. In: Integrated Network Management, 2009. IM'09. IFIP/IEEE International Symposium on. (2009) 327–334

53. Wang, Y., Wang, X., Chen, M., Zhu, X.: Power-efficient response time guarantees for virtualized enterprise servers. In: Real-Time Systems Symposium, 2008. (2008) 303–312

54. Petrucci, V., Loques, O., Niteroi, B., Mossé, D.: Dynamic configuration support for power-aware virtualized server clusters. In: WiP Session of the 21th Euromicro Conference on Real-Time Systems. Dublin, Ireland. (2009)

55. Wang, S., Munawar, W., Liu, X., Chen, J.J.: Power-Saving Design in Server Farms for Multi-Tier Applications under Response Time Constraint. In: 2nd International Conference on Smart Greeds and Green IT Systems. (2013)

56. Berral, J.L., Goiri, Í., Nou, R., Julià, F., Guitart, J., Gavaldà, R., Torres, J.: Towards energy-aware scheduling in data centers using machine learning. In: Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking. (2010) 215–224

57. Rao, J., Bu, X., Xu, C.Z., Wang, L., Yin, G.: VCONF: a reinforcement learning approach to virtual machines auto-configuration. In: Proceedings of the 6th international conference on Autonomic computing. (2009) 137–146

58. Anghel, I., Cioara, T., Salomie, I., Copil, G., Moldovan, D., Pop, C.: Dynamic frequency scaling algorithms for improving the CPU's energy efficiency. In: Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on. (2011) 485–491

59. Li, B., Li, J., Huai, J., Wo, T., Li, Q., Zhong, L.: EnaCloud: an energy-saving application live placement approach for cloud computing environments. In: Cloud Computing, 2009. CLOUD'09. IEEE International Conference on. (2009) 17–24

60. Borgetto, D., Maurer, M., Da-Costa, G., Pierson, J.M., Brandic, I.: Energy-efficient

and SLA-aware management of IaaS clouds. In: Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet. (2012)

61. Beloglazov, A., Abawajy, J., Buyya, R.: Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Generation Computer Systems **28**(5) (2012) 755–768

62. Katsaros, G., Subirats, J., Oriol Fitó, J., Guitart, J., Gilet, P., Espling, D.: A service framework for energy-aware monitoring and VM management in Clouds. Future Generation Computer Systems (2012)

63. Abbasi, Z., Pore, M., Gupta, S.: Impact of workload and renewable prediction on the value of geographical workload management. In: Second International Workshop on Energy Efficient Data Centers (E2DC). (2013)

64. Liu, Z., Wierman, A., Chen, Y., Razon, B., Chen, N.: Data center demand response: Avoiding the coincident peak via workload shifting and local generation. Performance Evaluation **70**(10) (2013) 770–791

65. Andersson, G., Ilic, M.D., Madani, V., Novosel, D.: Network Systems Engineering for Meeting the Energy and Environmental Dream. Proceedings of the IEEE **99**(1) (2011) 7–14

66. Gelenbe, E., Morfopoulou, C.: A Framework for Energy-Aware Routing in Packet Networks. The Computer Journal **54**(6) (2011) 850–859

67. Chauhan, N.S., Saxena, A.: A Green Software Development Life Cycle for Cloud Computing. IT Professional **15**(1) (2013) 28–34

68. Zhu, Q., Chen, Z., Tan, L., Zhou, Y., Keeton, K., Wilkes, J.: Hibernator: helping disk arrays sleep through the winter. In: ACM SIGOPS Operating Systems Review. Volume 39. (2005) 177–190

69. Pinheiro, E., Bianchini, R.: Energy conservation techniques for disk array-based servers. In: Proceedings of the 18th annual international conference on Supercomputing, June. (2004)

70. Cappiello, C., Hinostroza, A., Pernici, B., Sami, M., Henis, E., Kat, R., Meth, K., Mura, M.: ADSC: Application-Driven Storage Control for Energy Efficiency. In: Information and Communication on Technology for the Fight against Global Warming. Springer (2011) 165–179