A TAXONOMY AND SURVEY OF ENERGY-EFFICIENT DATA CENTERS AND CLOUD COMPUTING SYSTEMS

Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya

Prepared by: Dr. Faramarz Safi
Islamic Azad University, Najafabad Branch,
Esfahan, Iran.

Introduction

- The primary focus of designers of computing systems and the industry has been on the improvement of the system performance.
- According to this objective, the performance has been steadily growing by more efficient system design and increasing density of the components described by Moore's law [1].
- Although the performance per watt ratio has been constantly rising, the total power draw by computing systems is hardly decreasing.
- Oppositely, it has been increasing every year that can be illustrated by the estimated average power use across three classes of servers presented in Table 1 [2].
- If this trend continues, the cost of the energy consumed by a server during its lifetime will exceed the hardware cost [3].

Table 1. Estimated average power consumption per server class (W/Unit) from 2000 to 2006 [2].

Server class	2000	2001	2002	2003	2004	2005	2006
Volume	186	193	200	207	213	219	225
Mid-range	424	457	491	524	574	625	675
High-end	5,534	5,832	6,130	6,428	6,973	7,651	8,163

Introduction

- The problem is even worse for large-scale computing infrastructures, such as clusters and data centers.
- It was estimated that in 2006 IT infrastructures in the US consumed about 61 billion kWh for the total electricity cost about 4.5 billion dollars [4].
- The estimated energy consumption is more than double from what was consumed by IT in 2000. Moreover, under current efficiency trends the energy consumption tends to double again by 2011, resulting in 7.4 billion dollars annually.

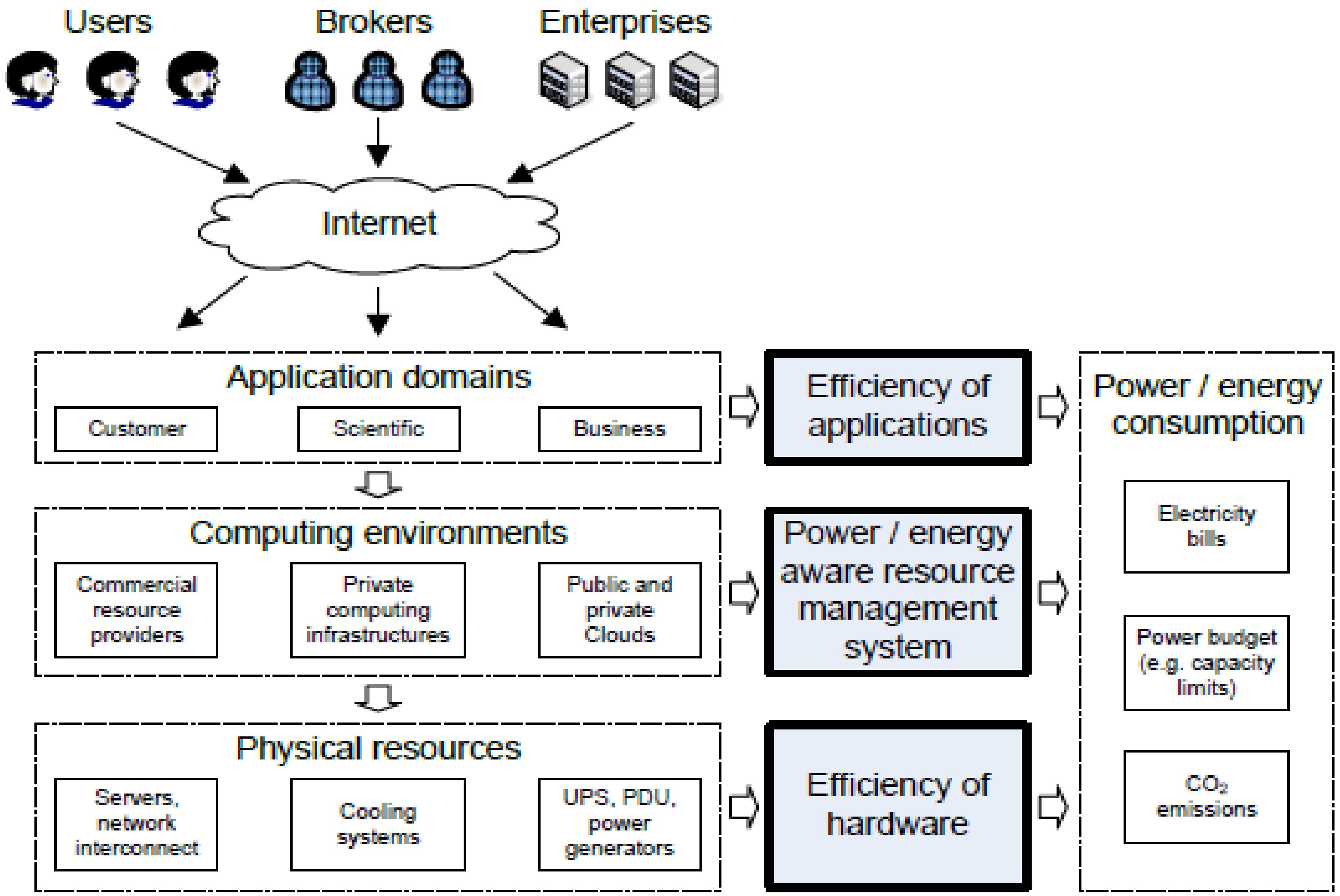
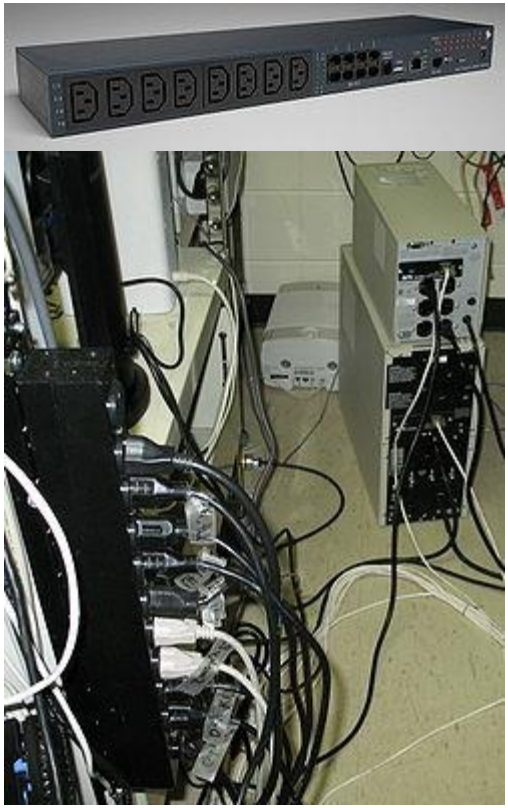


Figure 1. Energy consumption at different levels in computing systems.

Introduction

- The energy consumption is not only determined by the efficiency of the physical resources, but it is also dependent on the resource management systems deployed in the infrastructure and efficiency of applications running in the system. This interconnection of the energy consumption and different levels of computing systems can be seen from Figure-1.
- Energy efficiency impacts end users in terms of:
  - Resource usage costs, which are typically determined by the Total Cost of Ownership (TCO) incurred by a resource provider.
  - Higher power consumption results not only in boosted electricity bills, but also in additional requirements to a cooling system and power delivery infrastructure, i.e. Uninterruptible Power Supplies (UPS), Power Distribution Units (PDU), etc.



A 10-outlet rack-mount PDU without built-in surge protection, connected to a UPS (bottom unit on right)

Introduction

- Apart from the overwhelming operating costs and the Total Cost of Acquisition (TCA), another rising concern is the environmental impact in terms of carbon dioxide (CO_2) emissions caused by high energy consumption.
- Therefore, the reduction of power and energy consumption has become a first-order objective in the design of modern computing systems.
- The roots of energy-efficient computing, or Green IT, practices can be traced back to 1992, when the U.S. environmental protection Agency launched Energy Star, a voluntary labeling program which is designed to identify and promote energy-efficient products in order to reduce the greenhouse gas emissions. Computers and monitors were the first labeled products.
- This has led to the widespread adoption of the sleep mode in electronic devices. At that time the term "green computing" was introduced to refer to energy-efficient personal computers [7].
- Energy-efficient resource management has been first introduced in the context of battery feed mobile devices, where energy consumption has to be reduced in order to improve the battery lifetime.
- Although techniques developed for mobile devices can be applied or adapted for servers and data centers, this kind of systems requires specific methods.

Power and Energy Models

- To understand power and energy management mechanisms it is essential to clearly distinguish the background terms.
- Electric current is the flow of electric charge measured in Amperes (Amps). Amperes define the amount of electric charge transferred by a circuit per second.
- Power and energy can be defined in terms of work that a system performs. Power is the rate at which the system performs the work, while energy is the total amount of work performed over a period of time.
- Power and energy are measured in watts (W) and watt-hour (Wh) respectively. **Work** is done at the rate of **one watt** when **one Ampere** is transferred through a potential difference of **one volt**.
- A kilowatt-hour (kWh) is the amount of energy equivalent to a power of 1 kilowatt (1000 watts) running for 1 hour.

Power and Energy Models

- Formally, power and energy can be defined as follows:

$$P = \frac{W}{T},$$

$$E = P \cdot T,$$

- P is power, T is a period of time, W is the total work performed in that period of time, and E is energy.
- The difference between power and energy is very important, because reduction of the power consumption does not always reduce the consumed energy.
- For example, the power consumption can be decreased by lowering the CPU performance. However, in this case a program may require longer time to complete its execution consuming the same amount of energy.
- On one hand, reduction of the peak power consumption will result in decreased costs of the infrastructure provisioning, such as costs associated with capacities of UPS, PDU, power generators, cooling system, and power distribution equipment. On the other hand, decreased energy consumption will lead to reduction of the electricity bills.
- The energy consumption can be reduced temporarily using **Dynamic Power Management (DPM)** techniques or permanently applying **Static Power Management (SPM)**.
- **SPM** utilizes knowledge of the real-time resource usage and application workloads to optimize the energy consumption. However, it does not necessarily decrease the peak power consumption.
- In contrast, **DPM** includes the usage of highly efficient hardware equipments, such as CPUs, Disk Storage, Network devices, UPS and Power Supplies. These structural changes usually reduce both the energy and peak power consumption.

Static and Dynamic Power Consumption

- The main power consumption in Complementary Metal-Oxide-Semiconductor (CMOS) circuits comprises static and dynamic power.
- The static power consumption, or leakage power, is caused by leakage currents that are present in any active circuit, independently of clock rates and usage scenarios.
- This static power is mainly determined by the type of transistors and process technology. Reduction of the static power requires improvement of the low-level system design; therefore, it is not in the focus of this chapter.
- Dynamic power consumption is created by circuit activity (i.e. transistor switches, changes of values in registers, etc.) and depends mainly on a specific usage scenario, clock rates, and I/O activity.
- The sources of the dynamic power consumption are short-circuit current and switched capacitance. Short-circuit current causes only 10-15% of the total power consumption and so far no way has been found to reduce this value without compromising the performance.

Static and Dynamic Power Consumption

- Switched capacitance is the primary source of the dynamic power consumption; therefore, the dynamic power consumption can be defined as follows, where a is the switching activity, C is the physical capacitance, V is the supply voltage, and f is the clock frequency.

$$P_{dynamic} = a \cdot C \cdot V^2 \cdot f$$

- The values of **switching activity** and **capacitance** are determined **by the low-level system design**. Whereas combined reduction of the **supply voltage** and **clock frequency** lies in the roots of the widely adopted DPM technique called **Dynamic Voltage and Frequency Scaling (DVFS)**.

Sources of Power Consumption

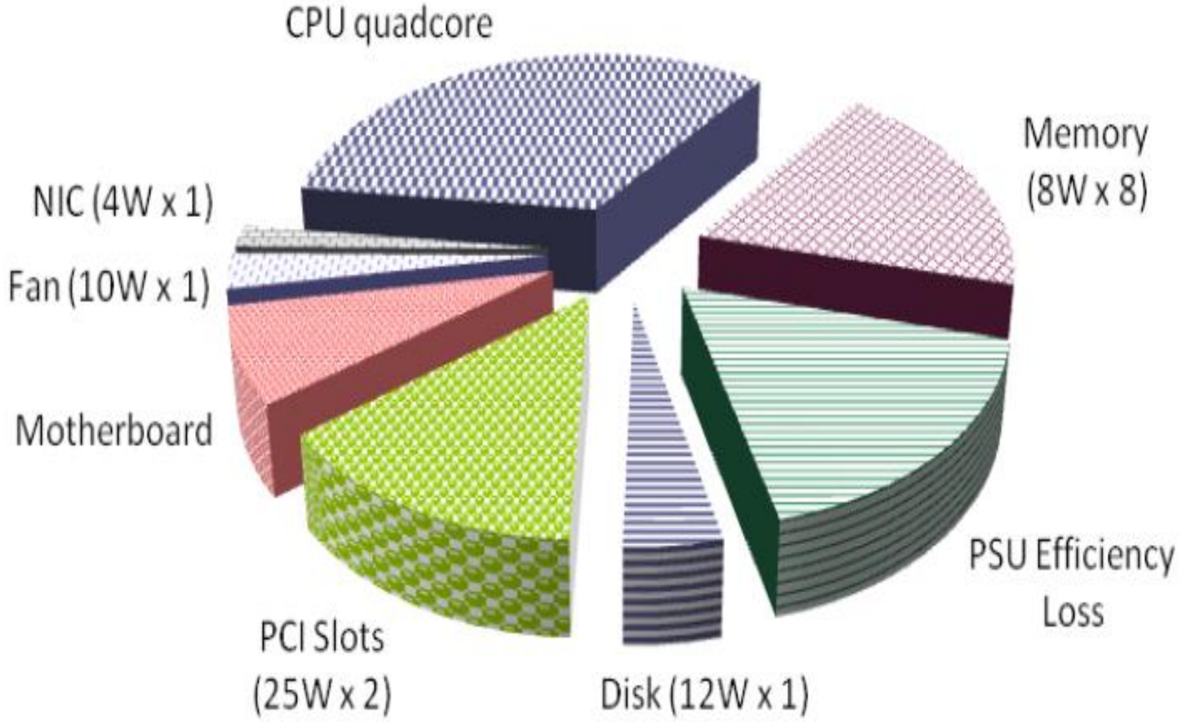


Figure 2. Power consumption by server's components [5].

Modeling Power Consumption

- To develop new policies for DPM and understand their impact, it is necessary to create a model of dynamic power consumption.
- Such a model has to be able to predict the actual value of the power consumption based on some run-time system characteristics.
- One of the ways to accomplish this is to utilize power monitoring capabilities that are built-in modern computer servers.
- This instrument provides the ability to monitor power usage of a server in real time and collect accurate statistics about the power usage.
- Based on this data it is possible to derive a power consumption model for a particular system. However, this approach is complex and requires collection of the statistical data for each target system.

Modeling Power Consumption

Fan et al. [10] have found a strong relationship between the CPU utilization and total power consumption by a server. The idea behind the proposed model is that the power consumption by a server grows linearly with the growth of CPU utilization from the value of power consumption in the idle state up to the power consumed when the server is fully utilized. This relationship can be expressed as:

$$P(u) = P_{idle} + (P_{busy} - P_{idle}) * u$$

where P is the estimated power consumption, P_{idle} is the power consumption by an idle server, P_{busy} is the power consumed by the server when it is fully utilized, and u is current CPU utilization.

The authors have also proposed an empirical non-linear model as follows, where r is a calibration parameter that minimizes the square error and has to be obtained experimentally. For each class of machines of interest a set of calibration experiments must be performed to fine tune the model.

$$P(u) = P_{idle} + (P_{busy} - P_{idle}) \cdot (2u - u^r)$$

Modeling Power Consumption

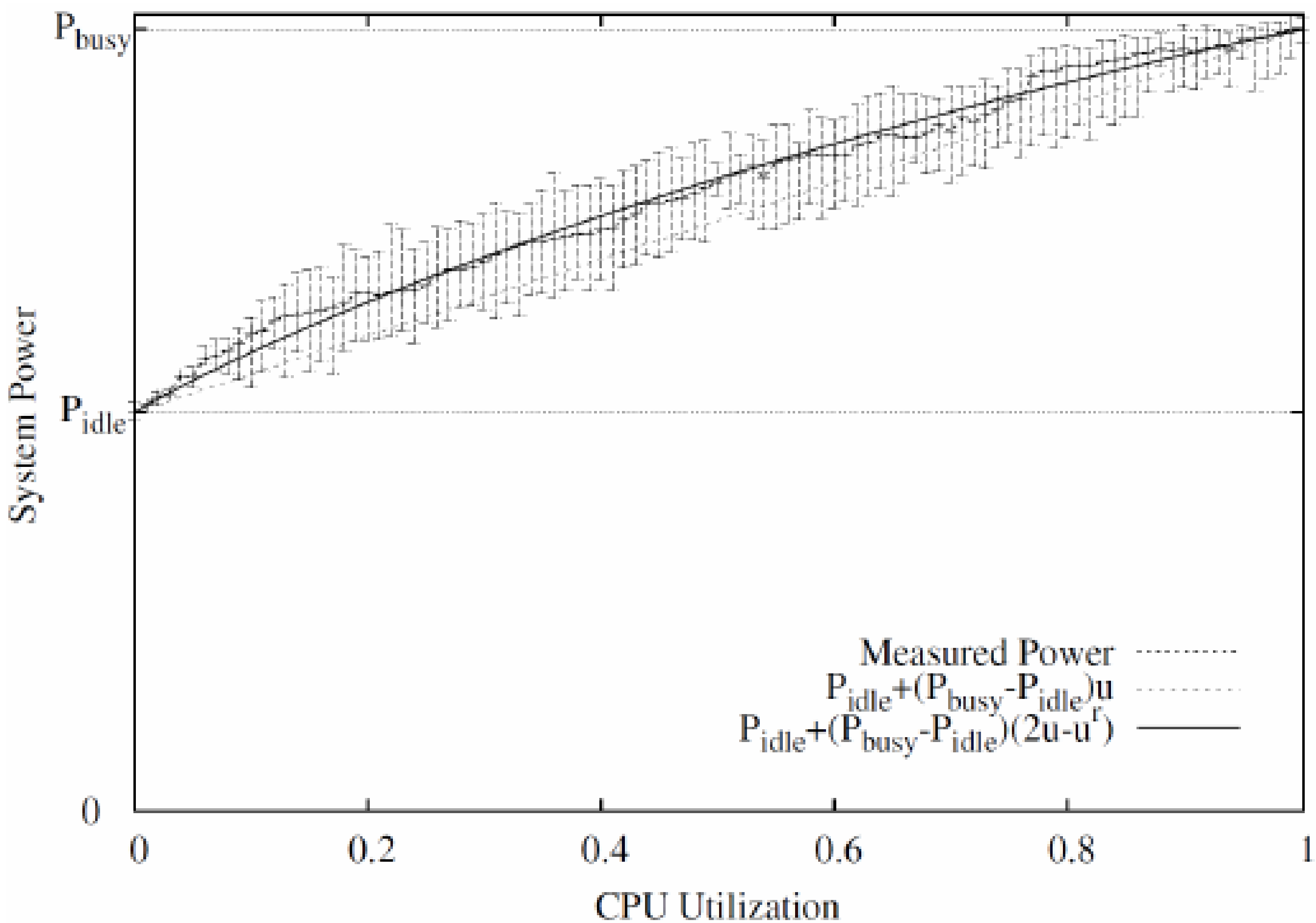


Figure 3. Power consumption to CPU utilization relationship [10].

Modeling Power Consumption

Dhiman et al. [13] have found that although regression models based on just CPU utilization are able to provide reasonable prediction accuracy for CPU-intensive workloads, they tend to be considerably inaccurate for prediction of power consumption caused by I/O- and memory-intensive applications.

The authors have proposed a **power modeling methodology** based on **Gaussian Mixture Models** that **predicts power consumption by a physical machine running multiple VM instances.**

To perform predictions, in addition to CPU utilization, the model relies on **run-time workload characteristics, such as the number of Instructions Per Cycle (IPC) and the number of Memory accesses Per Cycle (MPC)**.

The proposed approach requires a training phase to perceive the relationship between the metrics of the workload and the power consumption. The authors have evaluated the proposed model via experimental studies involving different types of the workload.

The obtained experimental results have shown that the model predicts the power consumption with high accuracy **(<10% prediction error)**, which is consistent over all the tested workloads.

Problems of High Power and Energy Consumption
High Power Consumption

- Considering the power consumption, the main problem is the minimization of the peak power required to feed a completely utilized system.
- The main reason of the power inefficiency in data centers is low average utilization of the resources.
- Data provided as a part of the CoMon project, a monitoring infrastructure for PlanetLab.
- The data of the CPU utilization by more than a thousand servers located at more than 500 places around the world.
- The data have been collected each five minutes during the period from the 10th to 19th of May 2010.
- The distribution of the data over the mentioned 10 days along with the characteristics of the distribution.

Problems of High Power and Energy Consumption
High Power Consumption

- The data confirm the observation made by Barroso and Holzle [9]: the average CPU utilization is 36.44%.
- Background tasks (e.g. incremental backups), or distributed databases or file systems.
- Despite the fact that the resources have to be provisioned to handle theoretical peak loads, it is very unlikely that all the servers of large-scale data centers will be fully utilized simultaneously.
- Another problem of high power consumption and increasing density of server's components is the heat dissipation. Much of the electrical power consumed by the computing resources gets turned into heat. The amount of heat produced by an integrated circuit depends on how efficient the component's design is, and the voltage and frequency at which the component operates.

Problems of High Power and Energy Consumption

High Energy Consumption

The **energy consumption** is defined by the **average power consumption over a period of time**. The actual energy consumption by a data center does not affect the cost of the infrastructure. On the other hand, it is reflected in the electricity cost consumed by the system **during the period of operation**, which is the main component of a data center's operating costs. Problems caused by growing energy consumption:

- High operating costs: in most data centers 50% of consumed energy never reaches the computing resources: it is consumed by the cooling facilities or dissipated in conversions within the UPS and PDU systems.
- High carbon dioxide (CO2) emissions, which contribute to the global warming. According to Gartner [16] in 2007 the Information and Communications Technology (ICT) industry was responsible for about 2% of global CO2 emissions that is equivalent to the aviation.

Problems of High Power and Energy Consumption
High Energy Consumption

- Overheating of the components can lead to decreased lifetime and high error-proneness.
- For each watt of power consumed by computing resources an additional 0.5 to 1 W is required for the cooling system [6].
  - For example, to dissipate 1 W consumed by a High-Performance Computing (HPC) system at the Lawrence Livermore National Laboratory (LLNL), 0.7 W of additional power is needed for the cooling system [15].
  - This fact justifies the significant concern about efficiency and real-time adaptation of cooling system operation.
  - Moreover, modern high density servers, such as 1U and blade servers, further complicate cooling because of the lack of space for airflow within the packages.

Taxonomy of Power / Energy Management in Computing Systems

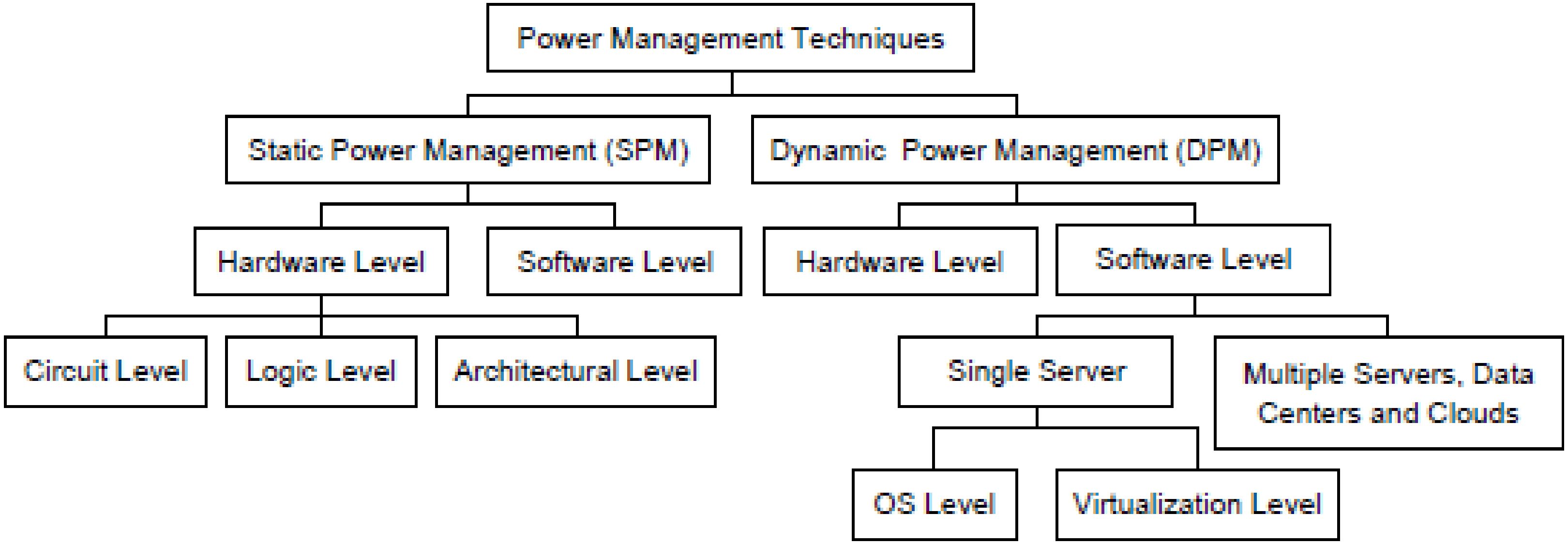


Figure 5. High level taxonomy of power and energy management.

- Large volume of research work has been done in the area of power and energy-efficient resource management in computing systems.
- As power and energy management techniques are closely connected, from this point we will refer to them as power management.
- As shown in Figure 5, from the high level power management techniques can be divided into static and dynamic.

Taxonomy of Power / Energy Management in Computing Systems
Hardware Point of View

From the hardware point of view:

- Static Power Management (SPM) contains all the optimization methods that are applied at the design time at the circuit, logic, architectural and system levels [17]. Circuit level optimizations are focused on the reduction of switching activity power of individual logic-gates and transistor level combinational circuits by the application of a complex gate design and transistor sizing.
- Optimizations at the logic level are aimed at the switching activity power of logic-level combinational and sequential circuits. Architecture level methods include the analysis of the system design and subsequent incorporation of power optimization techniques in it.
- In other words, this kind of optimization refers to the process of efficient mapping of a high-level problem specification onto a register-transfer level design.

Taxonomy of Power / Energy Management in Computing Systems
Software Point of View

From the software point of view:

- Apart from the optimization of the hardware-level system design, it is extremely important carefully consider the implementation of programs that are supposed to run in the system. Even with perfectly designed hardware, poor software design can lead to dramatic performance and power losses.
- However, it is impractical or impossible to analyze power consumption caused by large programs at the operator level, as not only the process of compilation or code-generation but also the order of instructions can have an impact on power consumption.
- Therefore, indirect estimation methods can be applied. For example, it has been shown that faster code almost always implies lower energy consumption [18]. Nevertheless, methods for guaranteed synthesizing of optimal algorithms are not available, and this is a very difficult problem.

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM)

- DPM techniques include methods and strategies for run-time adaptation of a system's behavior according to current resource requirements or any other dynamic characteristic of the system's state.
- The **major assumption** enabling DPM is that systems experience variable workloads during the operation time allowing the dynamic adjustment of power states according to current performance requirements.
- The **second assumption** is that the workload can be predicted to a certain degree.
- As shown on Figure 5, DPM techniques can be distinguished by the level at which they are applied: hardware or software:
    - Hardware DPM varies for different hardware components, but usually can be classified as Dynamic Performance Scaling (DPS), such as DVFS, and partial or complete Dynamic Component Deactivation (DCD) during periods of inactivity.
    - Software DPM techniques utilize interface to the system's power management and according to their policies apply hardware DPM.

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM)

- The introduction of the Advanced Power Management (APM) and its successor, the Advanced Configuration and Power Interface (ACPI), have drastically simplified the software power management and resulted in broad research studies in this area.
- The problem of power efficient resource management has been investigated in different contexts of
  - Device specific management,
  - OS level management of virtualized and non-virtualized servers, followed by multiple-node system, such as homogeneous and heterogeneous clusters, data centers and Clouds.

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM) - DVFS

- DVFS creates a broad dynamic power range for the CPU enabling extremely low-power active modes. This flexibility has lead to the wide adoption of this technique and many policies that scale CPU performance according to current requirements, while trying to minimize performance degradation [19].
- Subsequently, these techniques have been extrapolated on multiple-server systems providing coordinated performance scaling across them [20].
- To narrow overall dynamic power range of servers in a data center, it has been found beneficial to consolidate workload to a limited number of servers and switch off or put to sleep / hibernate state idle nodes [21].

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM) - Virtualization

- Another technology that can improve the utilization of resources, and thus reduce the power consumption is virtualization of computer resources.
- Virtualization technology allows one to create several Virtual Machines (VMs) on a physical server and, therefore, reduce the amount of hardware in use and improve the utilization of resources.
- Several commercial companies and open-source projects now offer software packages to enable a transition to virtual computing. Intel Corporation and AMD have also built proprietary virtualization enhancements to the x86 instruction set into each of their CPU product lines, in order to facilitate virtualized computing.
- Among the benefits of virtualization are improved fault and performance isolation between applications sharing the same computer node (a VM is viewed as a dedicated resource to the customer).
- The ability to relatively easily move VMs from one physical host to another using live or off-line migration; and support for hardware and software heterogeneity.
- The ability to reallocate VMs in run-time enables dynamic consolidation of the workload, as VMs can be moved to a minimal number of physical nodes, while idle nodes can be switched to power saving modes.

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM) - Terminal Servers

- Terminal servers have also been used in Green IT practices. When using terminal servers, users connect to a central server; all of the computing is done at the server level but the end user experiences a dedicated computing resource.
- It is usually combined with thin clients, which use up to 1/8 the amount of energy of a normal workstation, resulting in a decrease of the energy consumption and costs. There has been an increase in the usage of terminal services with thin clients to create virtual laboratories.
- Examples of terminal server software include Terminal Services for Windows, the Aqua Connect Terminal Server for Mac, and the Linux Terminal Server Project (LTSP) for the Linux operating system.
- Thin clients possibly are going to gain a new wave of popularity with the adoption of the Software as a Service (SaaS) model, which is one of the kinds of Cloud computing [22], or Virtual Desktop Infrastructures (VDI) heavily promoted by virtualization software vendors.
  - VMware View (VMware VDI) Enterprise Virtual Desktop Management. http://www.vmware.com/products/view/
  - Citrix XenDesktop Desktop Virtualization. http://www.citrix.com/virtualization/desktop/xendesktop.html
  - Sun Virtual Desktop Infrastructure Software. http://www.sun.com/software/vdi/

Taxonomy of Power / Energy Management in Computing Systems
Dynamic Power Management (DPM)

- Cloud computing naturally leads to power-efficiency by providing the following characteristics:
  - Economy of scale due to elimination of redundancies.
  - Improved utilization of the resources.
  - Location independence – VMs can be moved to a place where energy is cheaper.
  - Scaling up and down – resource usage can be adjusted to current requirements.
  - Efficient resource management by the Cloud provider.

Operating System Level

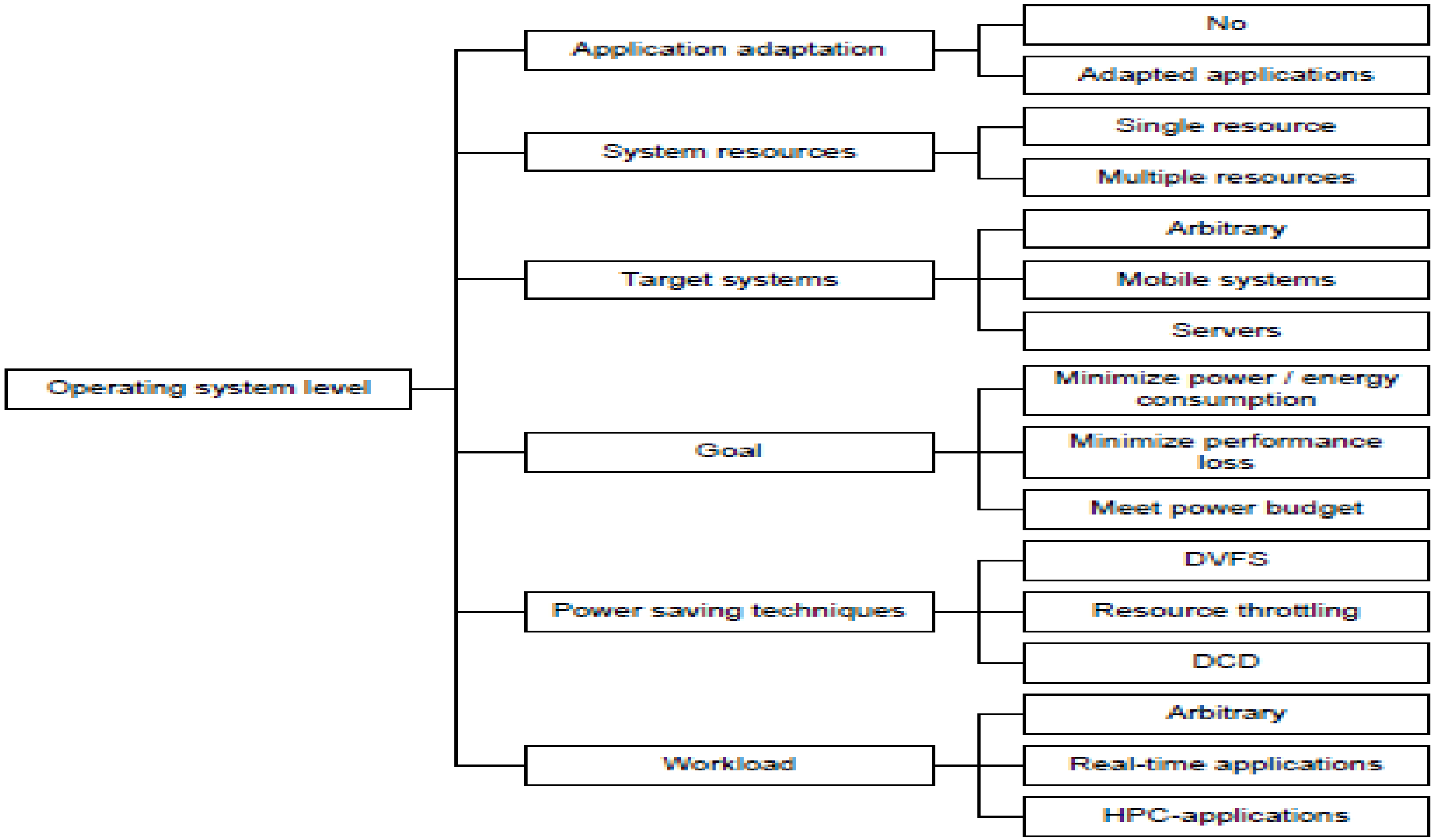


Figure 6. Operating system level taxonomy