International Journal of Physical Sciences Vol. 7(13), pp. 2052 - 2065, 23 March, 2012
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.1732
ISSN 1992 - 1950 ©2012 Academic Journals

Full Length Research Paper

Green Information Technology (IT) framework for energy efficient data centers using virtualization

Mueen Uddin[1]*, Muhammad Talha[2], Azizah Abdul Rahman[1], Asadullah Shah[3], Jameel Ahmed Khader[4] and Jamshed Memon[1]

[1]Department of Information Systems, Universiti Teknologi Malaysia, Malaysia.
[2]Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia.
[3]Department of Computer Science, Kulliyyah of Information and Communication Technology, International Islamic University, Malaysia.
[4]Arriyadh Community College, King Saud University, Riyadh, Saudi Arabia.

Accepted 31 January, 2012

The increasing demand for storage, networking and computation has driven the escalation of large complex data centers, the massive server farms that run many of today's Internet, financial, commercial and business applications. A data center can comprise many thousands of servers and can use as much energy as a small city. The massive amounts of computation power required to drive these server systems results in many challenges like energy consumption, emission of green house gases, backups and recovery issues, etc. The rising costs of oil and global warming are some of the biggest challenges of today's world. The research proposed in this paper discusses how virtualization can be used to improve the performance and energy efficiency of data centers. To prove this work, Green Information Technology (IT) based framework is developed to seamlessly and securely divide data center components into different resource pools depending on different parameters like energy consumption ratio, utilization ratio, workloads, etc. The framework highlights the importance of implementing green metrics like power usage effectiveness (PUE) and data center effectiveness, and carbon emission calculator to measure the efficiency of data center in terms of energy utilization and carbon dioxide (CO_2) emissions. The framework is based on virtualization and cloud computing to increase the utilization ratio of already installed servers from 10% to more than 50%.

Key words: Green Information Technology (IT), virtualization, server consolidation, energy efficient data centers, energy efficiency.

INTRODUCTION

The progress of Information and Communication Technology (ICT) based businesses and social practices have transformed many economies into e-economy and businesses into e-businesses. Technology has the potential to create sustainable businesses and societies with grim and green economics. ICT is increasingly playing critical role in transforming and generating economic opportunities. On the other hand, global warming and climate change coalescing with limited availability and rising cost of energy are posing serious challenges for the sustainability of the global digital economy (Molla et al., 2009). The worldwide agitation to achieve ecological, business and environmental sustainability is starting to redraw industrial landscape. The current status of global warming, ecological deterioration and the severity of its potential consequences explain the overwhelming popularity of environmental initiatives across the world. There has been an unprecedented increase on the level of concern regarding climate change and environmental sustainability. In recent years, business organizations have witnessed a global intention of employing environment friendly products and technologies to encounter global warming and achieve environment

*Corresponding author. E-mail: mueenmalik9516@gmail.com.

friendly and sustainable businesses. Environmental impact of Information Technology (IT) under the banner of “Green IT” has started being discussed by academia, media and government. IT professionals are expected to play significant roles in bringing Green IT to organizations, provided they are prepared, have developed or developing necessary capabilities to lead and support sustainability initiatives (Molla et al., 2009).

The increasing demand for storage, networking and computation has driven the demand for large data centers, the massive server farms that run most of today’s Internet, financial, commercial and business applications. They are found in nearly every sector of the economy, like financial services, media, high-tech, universities and government institutions to aid business processes, information management and communication functions (Daim et al., 2009). A data enter comprise of many thousands of servers and can use as much energy as a small city (EPA, 2009). All data centers are predominantly occupied by low cost underutilized volume servers also called x-86 servers as major components for performing different processes and services for the end users. With the recent development and growth in the size these server farms, the number of servers continuously increases as the demand for networking, storage, speed, backups and recovery and computation increases. These servers consume a lot of energy and power to perform processing, hence generate too much carbon dioxide (CO_2) and their utilization ratio is also low. In an average server environment, 30% of the servers are “dead” only consuming energy, without being properly utilized (Mueen and Azizah, 2010). The utilization ratio of all types of servers in a server based environment is only 5 to 10%. As new servers are being added continuously without considering the proper utilization of already installed servers, it causes superfluous and unavoidable increase in the energy consumption. This increased energy consumption causes an increase in the emission of green house gases (GHG) very hazardous for environmental health and global warming (Green Grid, 2009).

One of the areas of IT businesses where environmental sustainability is becoming imperative is data centers. They are found major culprits in consuming huge energy in their overall operations. Data centers are one of the largest energy consumers, accounting for approximately 2% of total global energy use (Koomey, 2007). As the demand for data centers continues to increase, these digital powerhouses are faced with several power, cooling, performance, and space constraints associate with environmental, technological, and economic sustainability issues. Improving the energy efficiency and environmental performance of data centers is therefore at the forefront of organization’s actions in ‘greening’ their IT (Koomey, 2007). In order to handle the sheer magnitude of today’s data, data centers have to use much more power as they become larger, denser, hotter, and significantly more costly to operate (Gartner, 2008). An EPA report to congress on servers and data center energy efficiency estimates that U.S data centers consume 1.5% of total U.S electricity consumption for a cost of $4.5 billion (EPA, 2007). From the year 2000 to 2006, data center electricity consumption has doubled in U.S and currently on a pace double again in 2011 to more than 100 billion kWh, equal to $7.4 billion in annual electricity costs (EPA, 2009). Data centers do not only consume huge energy but also major contributor towards company’s energy bill (McNamara et al., 2008). Gartner warns that, if data center managers remain unaware of energy problems, they will most probably run the risk of doubling their energy costs between 2005 and 2011 (Gartner, 2008). If energy costs continue to double every 5 years, they will substantially increase to 1,600% between 2005 and 2025 (Kumar, 208).

Figure 1 compares the purchasing dollars spent on new servers with the power and cooling cost since 1996 and projects those numbers until 2011 (Scaramella and Eastwood, 2008). According to IDC, the cost to power the servers will exceed the cost of servers by the next year.

Gartner emphasizes that ICT industry was responsible for about 2% of global CO_2 emissions almost equivalent to the aviation industry (Pettey, 2007). An EPA report presented to U.S congress in 2007 emphasized that, current energy consumption in data centers is leading to an annual increase in the emission of CO_2 (green house gases) from 42.8 million metric tons ($MMTCO_2$) in 2007 to 67.9 $MMTCO_2$ in 2011 (EPA, 2007). Emissions of GHG from aviations, shipping, transportations, telecommunications and manufacturing industry are rising fast, but the emissions from IT are mounting faster. Reductions achieved through the use of Green IT in key economic sectors would be five times greater than the growth in emissions from the IT sector itself. Continuous increase in these emissions from IT is projected to increase from 3% of total global emissions in 2009 to a whopping 6% by 2020 (“SMART 2020” Report in 2008 (on behalf of the Global e-Sustainability Initiative) (Table 1).

Ensuring a secure energy supply, preserving the environment and protecting the climate are central challenges facing today’s world. Environmentally friendly technologies are the key to sustainable economic activity (Koomey et al., 1996). In order to optimize the use of resources across the entire spectrum of global value chains, it is essential to tap the full potential of technology. Computer servers and the data center facilities in which they were located, are becoming more numerous and more electricity intensive. The first major efforts to understand server electricity use more deeply were spurred by a controversy over the total power used by IT equipment in which dubious claims were made about the contribution of IT equipment to total electricity use in the U.S (Huber et al., 1999). Later studies were built upon earlier work in order to create detailed measurements of data center power use in multiple

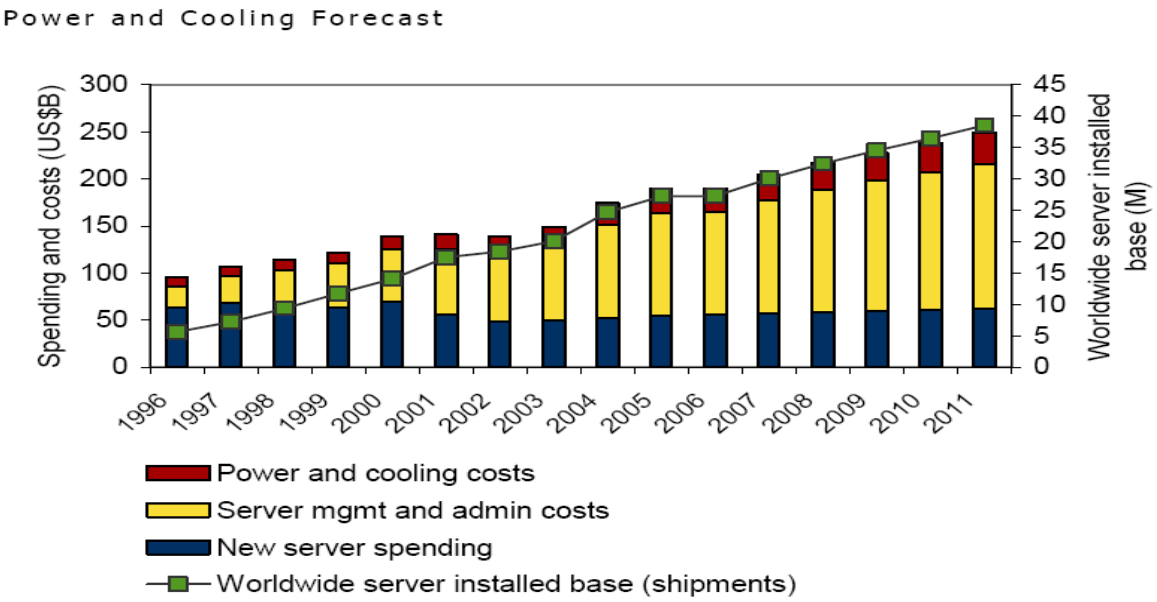

Figure 1. Worldwide cost to power and cool server installed base, 1996 to 2011.

Table 1. CO_2 emissions (carbon foot print) climate group and the global e sustainability initiative SMART 2020.

Sector	Emissions 2007 $MtCO_2e$	Percentage 2007 (%)	Emissions 2020 $MtCO_2e$	Percentage 2020 (%)
World	830	100	1430	100
Server farms/data centers	116	14	257	18
Telecom infrastructure and devices	307	37	358	25
PCs and peripherals	407	49	815	57

facilities (Greenberg et al., 2006; Tschudi et al., 2004; Tschudi et al., 2006; Tschudi et al., 2003).

Green computing is the study and practice of using computing resources efficiently (Hart, 1997). It is a way towards more environment friendly and cost-effective use of power and production technology. Green IT is defined as optimal use of ICT for managing the environmental sustainability of enterprise, its operations, products, services and resources (Mingay, 2007). It is a process that focuses on strategic deployment of operations and IT to dynamically, sustainably and responsibly align business oriented goals with green objectives for the entire duration of operations. It entails applications of environment friendly IT activities with business oriented organizational goals and creates functional bridge between corporate motivations and environment to create mutually satisfactory benefits (Hanuv et al., 2009). Green IT is used as an umbrella term for overlapping concepts like virtualization, recycling, telecommuting and power management, etc. The aspiration of Green IT is focused on achieving higher energy efficiency in the use of IT devices and to increase the utilization of already installed devices like servers in data centers. However, at the same time, organizations need to deliver new IT services in short time considering high reliability, performance and availability requirements as important issues in supporting the business processes (Byrd and Turner, 2000). To achieve these objectives, IT must have a clearly defined Green IT strategy or framework tuned with business strategy and goals to be able to realize and achieve both of these objectives at the same time. IT needs to quickly and transparently translate the IT strategy to IT operational processes to ensure effective business-IT alignment in most efficient way.

Latest emerging technology called virtualization can be used to achieve energy efficient data centers by providing different solutions like live migration, server consolidation, data deduplications, energy efficiency measurements and green metrics etc to overcome IT inefficiencies in data centers (Uddin and Rahman, 2011). Server consolidation increases the utilization ratio of different servers up to 50% or even more, saving huge energy and reducing CO_2 emissions by increasing the utilization ratio of already installed servers. It helps in implementing green

data centers to ensure that IT infrastructure contributes as little to the emission of green house gases, and helps to regain power and cooling capacity, recapture resilience and dramatically reducing energy costs and total cost of ownership (Uddin et al., 2011).

The research highlights the importance of Green IT for developing a framework using virtualization and green metrics to be implemented in large and complex data centers to overcome some of the problem and issues highlighted previously. The proposed Green IT framework provides an overall model to describe how data center managers can implement Green IT solutions in their existing or new data centers to make them more energy efficient and green. The proposed framework also contains a metrics based model to categorize data center components into measureable units depending on the workloads they execute and then defines certain green metrics to measure the performance of data center in terms of energy efficiency and CO_2 emissions.

PROBLEM STATEMENT

Data centers are plagued with thousands of servers performing the processing for businesses and end users to facilitate and accomplish large business goals. The problem with these servers is that almost 90% remain idle most of the time performing nothing but utilizing huge energy and generating enormous amount of CO_2 very hazardous for environmental sustainability and global warming. Developing countries like Pakistan, which are already facing huge energy predicament and deficits to meet the demands of domestic and business purposes; it becomes almost impossible to provide energy resources to these ever growing data centers resulting in overall degradation of economy of the country. There is a strong need to develop strategies, policies or frameworks for data centers so that they can cop up with ever growing demands from businesses.

PROPOSED WORK

Recent advances in hardware and software technologies including low-power processors, solid state drives and energy-efficient monitors have alleviated the energy consumption issue to a certain degree, a series of software approaches have significantly contributed to the improvement of energy efficiency. These two approaches (hardware and software) should be seen as complementary rather than competitive. User awareness is another non-negligible factor that should be taken into account when discussing Green IT. User awareness and behavior in general considerably affect computing workload and resource usage patterns; this in turn has a direct relationship with the energy consumption of not only core computing resources, but also auxiliary equipments. For example, a computer program developed without paying much attention to its energy efficiency may lead to excessive energy consumption and it may contribute to more heat emission, resulting in increases in the energy consumption for cooling.

The rise of the green movement has been on for a long time; perhaps the oil shortage and record gas prices mainstreamed the challenge for all business enterprises and government agencies. Regardless of how we finally reached this point in time, there is little argument that we are here. The environment and sustainable energy have become a hot topic of conversation everywhere from kitchen tables to political arenas. It is no secret that the environmental movement has reached mass acceptance as leading publications across the globe have put environmental issues on their covers in an effort to raise awareness. Whether driven by consumer demand, economics, politics, legislative mandate, or even corporate social responsibility objectives, the green movement has caught on and corporations across the globe are beginning to embrace it with open arms (McNamara et al., 2008).

It is very much obvious from the previous discussion that, IT business enterprises specifically data center managers must follow a virtualized Green IT framework, to save huge amount of energy and at the same time reduce the CO_2 emissions, which ultimately reduce the global warming effects. The proposed Green IT framework to improve the efficiency of data center consists of five steps to be followed to properly implement virtualization and then apply metrics to measure the efficiency of data center in terms of energy efficiency and CO_2 emissions. The proposed Green IT framework using virtualization technology and green metrics should be used and followed by data center managers to implement green initiatives in their data center to make it more energy efficient and green. The proposed framework consists of five steps.

Green IT framework using virtualization

The energy efficient and green data center is gathering momentum as organizations have started realizing its importance in energy conservation and sustainable development. It is applied to new technologies that can help in cutting down data center energy costs and in saving energy, which is synonymous to saving money. It has a big role to play in reducing power consumption in the data centre (Murugesan, 2007). The proposed framework does not only provide understanding of different aspects of Green IT, but can be used as a tool to evaluate the techniques needed to put Green IT into action, and to find out which departments out of strategy, IS/IT or operations would figure most heavily in data center's pursuit of 'greening' efforts. As a continuing valuation tool, the framework acts both as a strategic Green IT implementation bridge and as an interval reference point for policy-making and strategic goal progression.

The proposed Green IT framework also helps IT business enterprises specifically data center industry to follow a virtualized Green IT framework, to save huge amount of energy and at the same time reduce the CO_2 emissions, which ultimately reduces global warming effects. It consists of five steps to be followed to properly implement virtualization at different layers and levels and then helps to identify and employ green metrics to measure the efficiency of data center in terms of energy efficiency and CO_2 emissions. The proposed Green IT framework using virtualization and green metrics should be used and followed by data center managers to implement green initiatives in their data center to make it more energy efficient and green. It consists of the following five phases:

1) Develop plan for green data centers.
2) Categorize data center into measureable components.
3) Identify green metrics and set benchmarks.
4) Identify and implement virtualization type.
5) Measure the performance in terms of energy efficiency and CO_2 emissions.

Develop plan for green data centers

The green data center has moved from the theoretical to the realistic, with IT leaders being challenged to construct new data centers or retrofit existing ones with energy saving features, sustainable materials and other environmental efficiencies. The

green data center is an energy-efficient, dense computing ecosystem where:

1) Software technologies control data growth and shrink capacity demands.
2) Managers use service level agreements to manage energy usage.
3) Energy efficient computing infrastructure optimizes performance and utilization levels.
4) Physical plant is engineered for maximum energy efficiency.

Benefits of green data center: These include

1) Lower server/storage temperature.
2) Better system reliability.
3) Better uptime.
4) Extends life of current data center.
5) Increased reliability of server/storage.
6) Lower total cost of ownership.
7) Lower energy use.
8) Maximize server/storage density.
9) Lower carbon footprints.
10) Maximizes hardware utilization.
11) Maximizes software utilization.
12) Increases environmental sustainability.
13) Increases business sustainability.
14) Lower heating and cooling costs.
15) Environment friendly and sustainable business.
16) Reduces the effects of global warming (Green Grid, 2009).

The first step in developing Green IT framework is to identify current awareness of Green IT initiatives in tier level data centers which are the largest data centers consuming more energy and generate more CO_2 emissions. Then identify best green IT practices like proper hardware utilization, live migration, server consolidation, installing energy efficient equipments etc that can be implemented in these data centers considering cost and other issues to achieve energy efficient data centers as shown in Figure 2. This phase also consists of establishing efficiency teams and then assigning goals to be achieved. This phase is subdivided into the following categories:

1) Perform survey to get the awareness of green data centers.
2) Identify best practices for green data centers.
3) Research potential green data center advantages.
4) Establish efficiency teams and goals.

Categorize data center components into measureable units

Data centers are huge entities which consist of many different components and devices performing different tasks to meet the end user needs. These components should be categorized into measureable resource pools depending on the workloads they execute, so that green metrics can be applied to measure their performance and efficiency individually and then to find out the overall efficiency of data center, because it is difficult to manage and measure the efficiency of whole data center collectively. Some of the major data center components are as follows:

1) Servers
2) Storage devices
3) Uninterruptible power supplies (UPS)
4) Switch gear
5) Chillers
6) Computer room air conditioners
7) Direct expansion (DX) units
8) Pumps
9) Cooling tower
10) Generators
11) Distribution losses external to the racks
12) Power distribution units (PDUs)
13) Batteries
14) Lighting etc.

Currently, there is no proposer mechanism available to categorize data center components into measureable units so that green metrics can be applied to them separately to measure their performance. This paper proposes a metrics based energy efficiency model to categorize data center into measureable units and then applied green metrics available so that efficiency can be measured and benchmarking can be set.

Metrics based energy efficiency model: The major barrier in improving energy efficiency in data center is the lack of appropriate metrics. This paper proposes what metrics could a company or an organization use to measure the effectiveness of their green initiatives for energy savings in the data center. It also provides IT managers with two most important industry acceptable metrics power usage effectiveness (PUE) and its reciprocal, Data center efficiency (DCE) and data center productivity (DCP). These metrics enables data center operators to quickly estimate the energy efficiency of their data centers, compare the results against other data centers and determine if any energy efficiency improvements need to be made.

To mitigate the problems associated with a multitude of metrics and thereby inconsistent measurements of data center performance, the US department of energy DOE and EPA conducted a workshop in July, 2008 regarding the collaboration between the government and IT industry to improve energy efficiency in data centers (National data center energy efficiency strategy workshop, 2008). The areas of improvement discussed include:

1) Defining energy efficient data centers, creating better transparency in the energy use in data centers and IT equipment through metrics, standards and best practices.
2) Advancing energy efficient data centers: focusing on adoption of energy efficient technologies and practices in data centers through knowledge creation and management.
3) Rewarding energy efficient data centers to help organizations better quantify and understand internal rewards from energy efficiency.

To achieve some of the objectives highlighted, Interviews were conducted from top managers including directors from IT data centers operations, environmental initiatives and enterprise infrastructures. The survey results from the interviews are then normalized and it seemed surprising that performance and cost were the most important attributes. From these surveys, it was found that three industry metrics server consolidation, power management and data center cooling were considered important. From these results, it was also found that managing capital costs and operating expenses are vital to a company's viability. But due to the lack of appropriate metrics, it becomes difficult for IT managers to measure the cost and performance of data centers to make them energy efficient and green.

This paper proposes a model which categorizes data center components into layers and then applies green metrics discussed earlier to achieve the desired results. A good metric would be something that measures the efficiency, the sustainability and the cost of a green initiative. The biggest problem any metrics faces when applied for calculating energy efficiency is the lack of standardized system of categorizing different resources of data center. The proposed model identifies different hardware

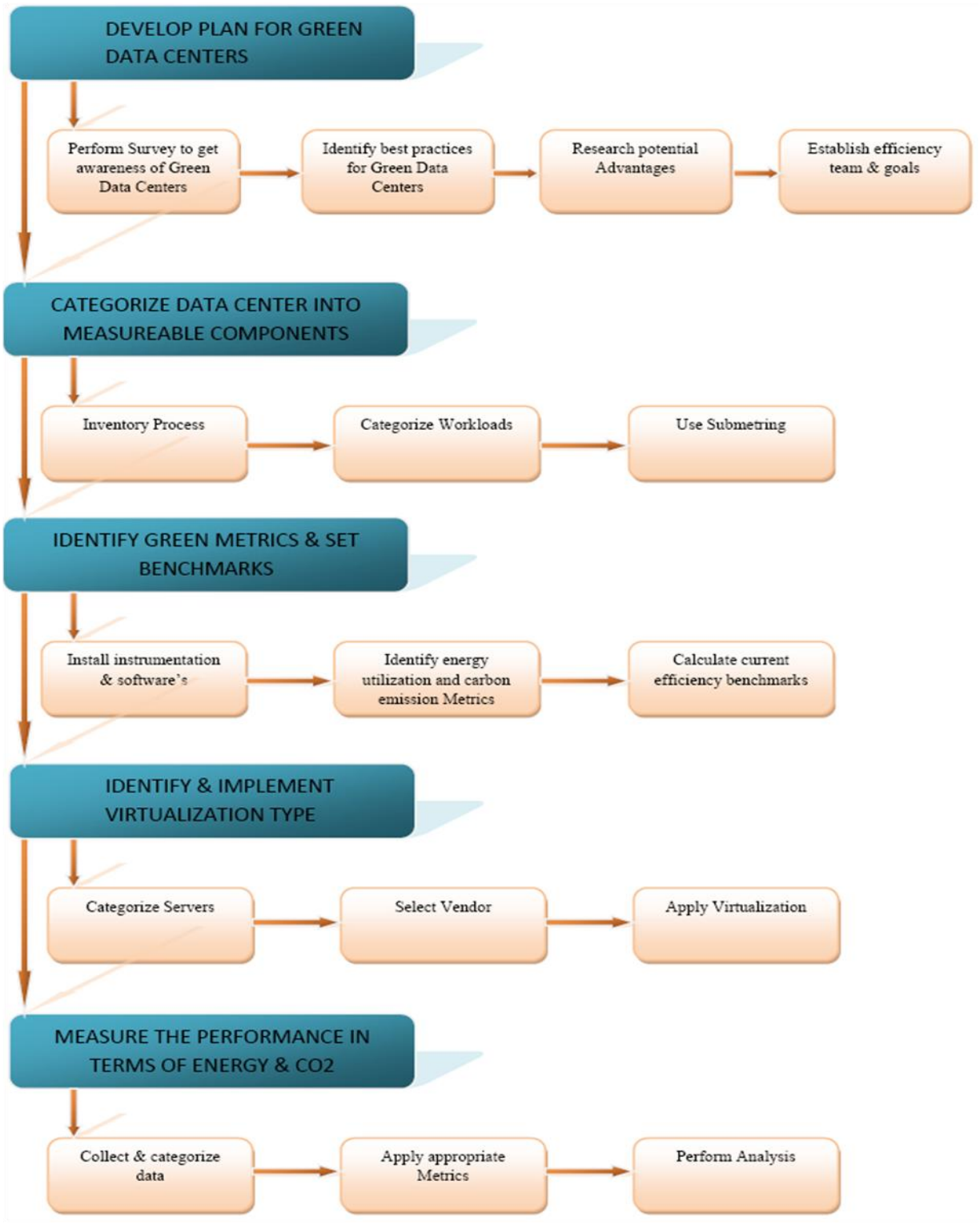

Figure 2. Green IT framework using virtualization.

components that need to be measured for energy efficiency and then creates a relationship between these hardware components and creates a relational layer between them. The first step in developing the model is to identify the components and then group them in different categorizes according to their measurement efficiency and workloads. As this research mainly focuses on calculating energy efficiency and carbon footprints that is, CO_2 emissions, so devices related to power energy will be used for

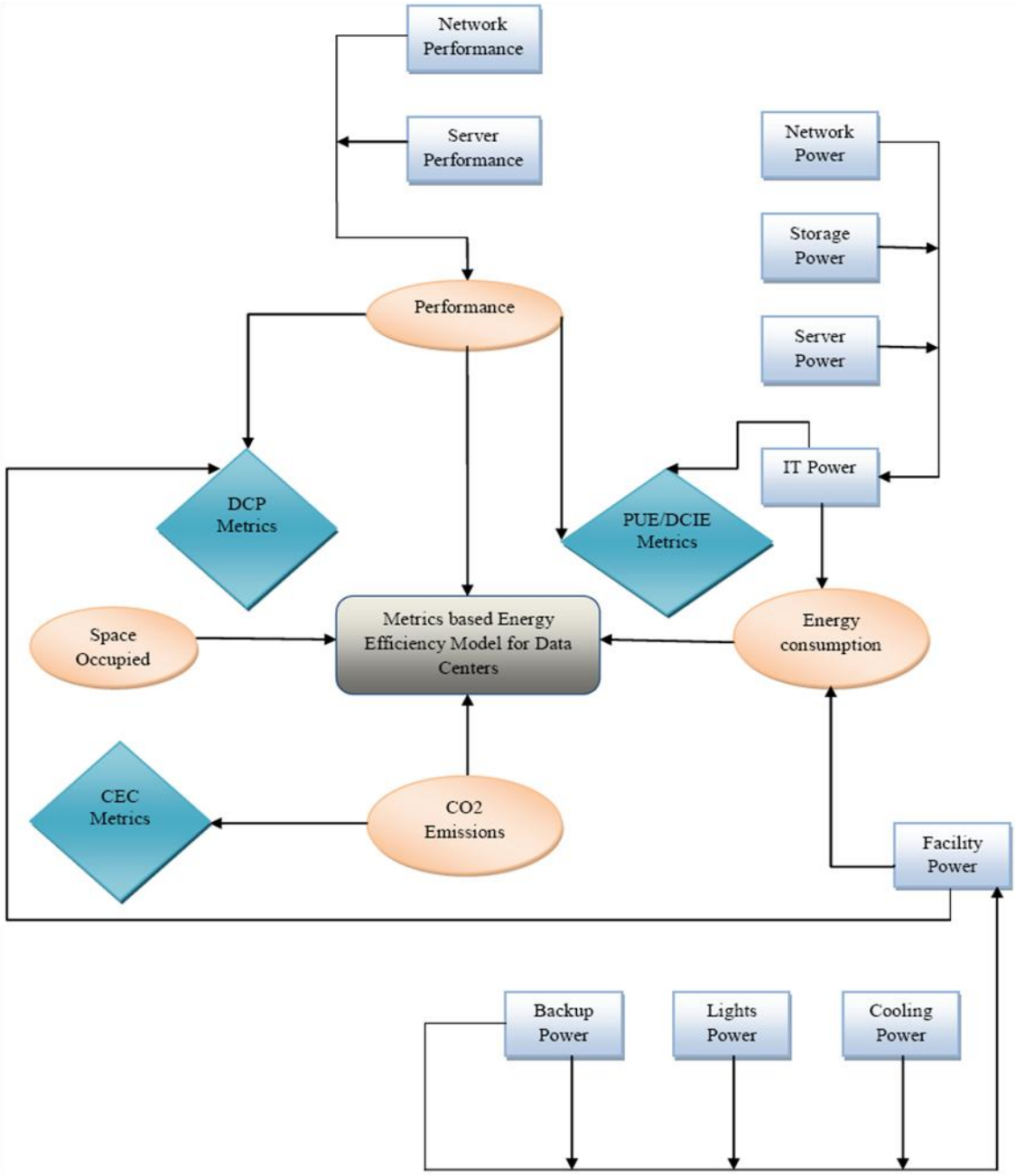

Figure 3. Metrics based energy efficiency model.

calculating energy efficiency.

The proposed model divides data center into four key areas so as to measure them separately. The areas considered are then grouped and these are: energy consumption, performance of the system, space used within the facility and CO_2 emissions. All these elements are arranged in a layer called services layer. After that, a component layer will be developed beneath the services layer which consists of small components from top layer elements. For example, under energy consumption, there are subcategories like IT power and facilities power and under these are further sub components like server, network and storage power. Each of these components has an associated set of metrics that deal solely with that component. Each element in the component layer is identified in the model with a blue rectangle. The metrics found relate the major modules of the system which include PUE, DCP and CEC (Belady, 2008). These metrics sum up the relationships between major constituents of the data center. This is the relational layer and is associated with the blue boxes in the model as depicted in Figure 3. This model can be useful in communicating the management of energy efficiency outside of the IT organization. The metrics used in the relational layer (blue boxes) are the ones that would typically be communicated to the corporate ranks on a continuing basis. The proposed model uses the following phases to categorize data center into measureable units. These phases are:

1) Inventory process (Identify and classify DC equipment according to some criteria).
2) Categorize workloads.

Inventory process: The inventory process identifies and classifies all the components of a data center and classifies them according to different parameters like energy use, carbon emission, utility ratio, type of equipment, life time etc. The data center inventory process classifies all equipment into two major categories:

A) Total facility power
B) IT equipment power

***A) Total facility power*:** It is defined as the power measured at the utility meter. This power is dedicated solely to the data center. It includes everything that supports the IT equipment components such as:

1) Power delivery components:

a) Uninterruptable power supply
b) Switch gear
c) Generators
d) Power distribution units
e) Batteries
f) Distribution losses external to the IT equipment.

2) Cooling system components:

a) Chillers
b) Computer room air conditioning units (CRACs)
c) Direct expansion air handler (DX) units
d) Pumps
e) Cooling towers
f) Other miscellaneous components such as data center lighting etc.

***B) IT equipment power*:** The IT equipment power is defined as the equipment that is used to manage, process, store or route data within the raised floor space in the data center. It includes the components associated with all of the IT equipments, such as:

1) Computing equipments:

a) Server machines
b) Client machines
c) Network equipment
d) Storage equipment
e) The decreased efficiency of UPS equipment when run at low loads.

2) Supplementary equipments:

a) KVM switches
b) Monitors
c) Workstations/ laptops used to monitor or otherwise control the data center

Categorize workloads: After inventory process, workloads are categorized according to their usage and resource requirements. Servers are the major components performing most of the processing; so they should be categorized according to the services they provide and resources they require in order to process the requests.

In a traditional data center, each server is devoted to a specific function. For instance, an e-mail server deals only with e-mail and a payroll server handles only payroll. But this traditional way is inefficient because the e-mail server could only run at 65% capacity, for instance, during business hours to accommodate spikes in demand. And the same server would use significantly less energy during non business hours. The payroll server, on the other hand, might run at only 5% capacity during business hours as a few changes and queries are processed by personnel, holding the remainder of its capacity in reserve for the larger job of payroll processing after hours. Using virtualization, the e-mail server and payroll server could share the same machine, with e-mail processing using the bulk of the capacity during business hours and payroll processing using the bulk during off hours (Brill, 2007). Using this method, governments purchase and maintain less equipment. They also save on the cost of housing, powering and cooling huge server farms that only use a fraction of their processing power. Bringing all of government's processing needs together under one roof can bring immense efficiency and cost benefits to the government organization. A single data center also gives organizations more flexibility to align their processes with their goals especially when services-based approaches, such as SOA, and an increased use of industry standards begin bridging and breaking down information silos.

A) Categorize server resources: After creating server inventory information, the next step is to categorize the servers and their associated resources and workloads into resource pools. This process is performed to avoid any technical political, security, privacy and regulatory concern between servers, which prevent them from sharing resources. Once analysis is performed, we can categorize each server roles into groups. Server roles are categorized into the following service types (Mueen and Azizah, 2010):

1) Network infrastructure servers
2) Identity management servers
3) Terminal servers
4) File and print servers
5) Application servers
6) Dedicated web servers
7) Collaboration servers
8) Web servers
9) Database servers.

B) Categorizing application resources: After categorizing servers into different resource pools, applications will also be categorized as:

1) Commercial versus in-house
2) Custom applications
3) Legacy versus updated applications
4) Infrastructure applications
5) Support to business applications
6) Line of business applications
7) Mission critical applications.

C) Allocation of resources: After creating the workloads, the next process is to allocate computing resources required by these different workloads and then arranging them in normalized form; but for normalization, the processor utilization should be at least 50%. It is very important to normalize workloads so as to achieve maximum efficiency in terms of energy, cost and utilization. The formula proposed in this paper for normalization is to multiply utilization ratio of each server by total processor capacity that is, (maximum processor efficiency * number of processors * number of cores). Servers are the major consumers of energy as they are in huge quantity and perform most of the processing. If we apply different efficiency calculators to calculate different efficiencies like energy efficiency and CO_2 emissions then, there must be a mechanism to categorize these different devices into different categories so that they can easily be evaluated and measured. But unfortunately,

there is no standard mechanism available to categorize these devices according to some criteria (Koomey, 2008).

D) Server categorization: Virtualization allows organizations to get more value from their existing data center hardware by running multiple "virtual machines" on one physical server (Anderson et al., 2001). It enables running two or more operating systems simultaneously on a single machine. It uses virtual machine monitor (VMM) or hypervisor, a system software inserted between the guest operating systems and the underlying hardware to provide a platform that hosts multiple operating systems running concurrently and sharing different resources among each other to provide services to the end users depending on the service levels defined before the processes are being performed (Talaber et al., 2009). Hypervisor is responsible for allowing multiple operating system images and their running applications to share the resources of a single hardware. Each operating system assumes that it has all the resources under its control; instead hypervisor is controlling all resources and transparently ensures that resources are properly and securely partitioned between different operating system images and their applications. The hypervisor manages all hardware structures, such as the memory management unit (MMU), I/O devices and direct memory access (DMA) controllers, and presents a virtualized abstraction of those resources to each guest operating system (Crosby and Brown, 2006).

A technique called "server consolidation" has been proposed which increases the utilization ratio up to 50% or even higher and saves energy and cost (Mueen and Azizah, 2010). The proposed technique divides the underutilized volume servers into three categories on the basis of workloads; these servers perform and applications they execute.

These classifications may vary from business to business with different infrastructures and requirements, as servers are generally used to create, maintain and execute solutions on behalf of businesses, architectures, processes and infrastructures. After categorizing, server consolidation is applied on all categories of servers depending on their utilization ratio in the data center. This process reduces the number of servers by consolidating the load of multiple servers on one server. It also increases the utilization ratio of servers. These categories are:

1) Innovation servers
2) Production servers
3) Mission critical servers

Identify green metrics and set benchmarking

After categorization, the next step is to identify the suitable metrics for benchmarking. But unfortunately, due to the lack of both standard metrics and benchmarking formulas, there is no industry metrics available to calculate both efficiencies that is, (energy and carbon footprints) at the same time using the same tool. Hence, this research proposes the following metrics that are considered as industry standard in measuring the energy efficiency and CO_2 emissions:

1) PUE (Green Grid, 2009)
2) Data center effectiveness (Green Grid, 2009)
3) Carbon emission calculator.

It is important to define what a metric is and the characteristics that make a metric useful. Metrics are instruments to measure and serve as an indicator of progress (Malone and Belady, 2007). A good metric would be something that measures the efficiency, the sustainability and the cost of a green initiative. The importance of rational, measurable metrics therefore, becomes imperative to measure and manage data center (Rivoire, 2007). The metrics must consider some properties defined as attributes for the selection to be an effective metrics. These attributes are:

1) Definition of metrics
2) Measurement capability
3) Usage of metrics

Green Grid (2007) defined following the characteristics that data center metrics possess to be an efficient metric:

1) The metric name should be clear
2) Be intuitive
3) The metric should be capable of scaling according to the purpose for which it was initially created and should factor technological, economic, environmental changes etc
4) Scientifically accurate and used precisely
5) Granular enough to analyze individual aspects
6) Should be capable of providing data driven decisions.

Identify and implement virtualization type

It is an architecture that separates an operating system from the underlying platform resources. Virtualization technology is now becoming an important advancement in IT especially for business organizations and has become a top to bottom overhaul of the computing industry. It combines or divides the computing resources of a server based environment to provide different operating environments using different methodologies and techniques like hardware and software partitioning or aggregation, partial or complete machine simulation, emulation and time sharing (Bugnion et al., 1997). Virtualization is a technique for abstracting (or hiding) the physical characteristics of computing resources from the way in which other systems, applications or end users interact with those resources.

It is very important for an organization to know in advance the total content of its infrastructure before implementing virtualization. This is the most important step in any virtualization project. There are many tools available from different vendors for performing initial analysis of an organization (Mueen and Azizah, 2010). Virtualization can be applied to operating systems, desktops, applications, storage and networks, depending on the end user business requirements. Virtualization technology is provided by many vendors including VMware, Cisco, Citrix, EMC, Hewlett-Packard, IBM, Microsoft, NetApp, Novell, Oracle, Quest, Sun, Symantec and Virtual Iron.

Server virtualization has become popular in data centers since it provides an easy mechanism to cleanly partition physical resources, allowing multiple applications to run in isolation on a single server. It categorizes volume servers into different resource pools depending on the workloads they perform, and then server consolidation is applied. This technique decouples softwares from hardware and splits multi processor servers into more independent virtual hosts for better utilization of the hardware resources, allowing services to be distributed one per processor. In server consolidation, many small physical servers are replaced by one large physical server to increase the utilization of expensive hardware resources, reducing the consumption of energy and emission of CO_2 (Mueen and Azizah, 2010).

Microsoft baseline security analyzer (MBSA) tool provides different information like IP addressing, operating system, installed applications and most importantly vulnerabilities of every scanned system. After analyzing, all generated values are linked to MS Visio, which generates a complete inventory diagram of all components and also provides details about each component being analyzed. Microsoft assessment and planning toolkit (MAP) is another tool for the assessment of network resources. It works with windows management instrumentation (WMI), the remote registry service or

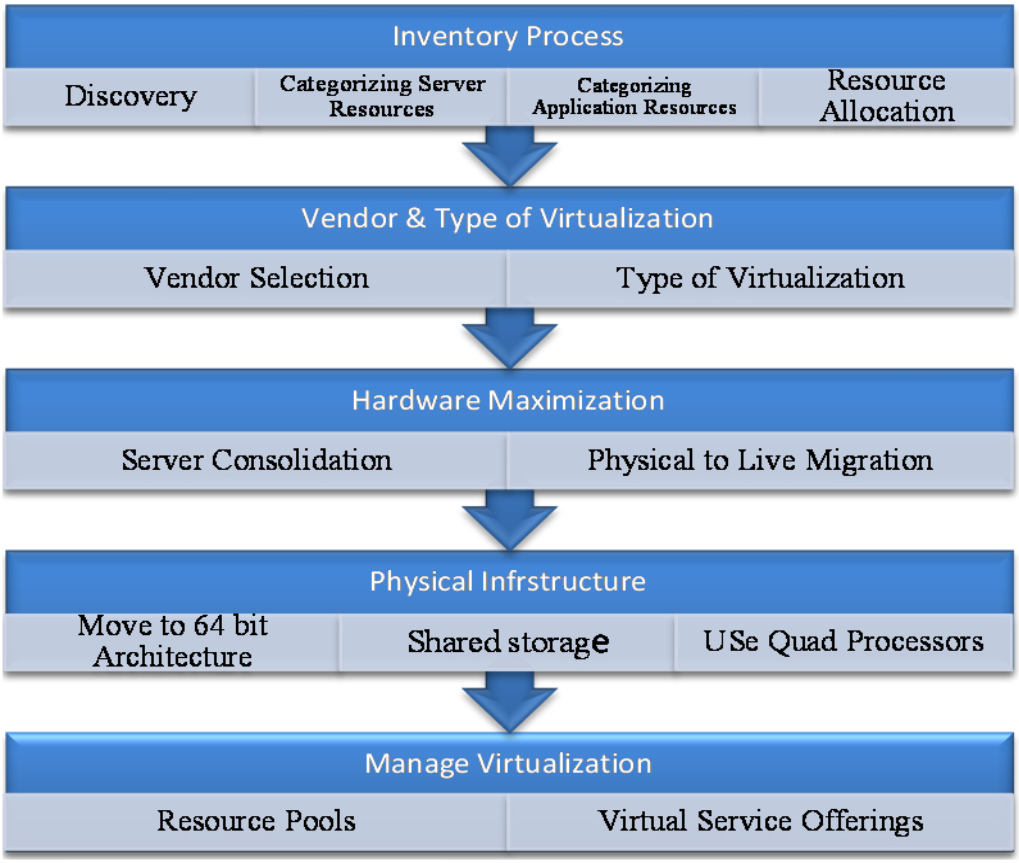

Figure 4. Components of the proposed model.

with simple network management protocol to identify systems on network. VMware, the founder of X-86 virtualization, also offers different tools for the assessment of servers that could be transformed into virtual machines. VMware guided consolidation (VGC), a powerful tool assesses network with fewer than 100 physical servers. Since VGC is an agent less tool, it does not add any overhead over production server's workload.

To properly implement virtualization, I presented a model called layered model consisting of five layers. Each layer defines more detailed processes. These components provide a detailed treatment of state of the art and emerging challenges faced by data centers managers to implement and manage virtualization properly in their data centers to achieve desired objectives as given in Figure 4. The proposed model defines that, the process of virtualization should be structured and designed in such a way that it must fulfill the necessary requirements and should be within the scope and infrastructure domain already installed in the data center. It is therefore, much more than simply loading a virtualization technology on different servers and transforming one or two workloads into virtual machines. Rather, it is a complex and rigorous process that need to be implemented and monitored properly (Mueen and Azizah, 2011). The proposed model defines five key steps that need to be followed at different stages in a structural way to achieve the efficiency required in the data center. The components of the proposed model (Figure 4) are as follows:

1) Inventory process
2) Type and nature of virtualization
3) Hardware maximization
4) Architecture
5) Manage virtualization

The reasons to consolidate and virtualize IT infrastructure that is, server are wide-ranging. Mostly, IT managers focus on just the cost of the physical IT infrastructure, but there are many additional benefits that can be derived from server consolidation like:

1) Proper utilization of servers
2) Proper energy utilization
3) Reduction in consumption of energy
4) Reduction in emission of green house gases
5) Reduction in global warming effects
6) Achieving sustainable businesses
7) Simplified management
8) Improved data protection
9) Improved resource utilization
10) Easier revision control and flexibility
11) Easier data protection and security
12) Reduced server and software costs etc

Servers are the major consumers of energy and they need to be consolidated using virtualization. This process comprises of the following sub phases:

1) Categorize server resources (Mueen and Azizah, 2010).
2) Select vendor (VMware, Citrix, Microsoft etc).
3) Apply server consolidation (Mueen and Azizah, 2010).

Measure performance in terms of energy and CO_2

This phase deals with measuring the performance of data center regularly from time to time using the metrics selected; this is done by performing the following two steps:

1) Collect and categorize data (Facilities data, Infrastructure data, IT equipment data)
2) Analysis using energy and CO_2 efficiency calculators.

In this phase, data center managers need to collect data regularly related to energy usage, carbon emissions, utilization ratio, life span, etc. of all categories of equipment from data center. After collecting data, analysis will be performed by using different tools like SPECpwer_ssj (SPECpower, 2008). This tool will generate outputs on the basis of comparison between performance and power. This process should be repeated until maximum efficiency in terms of power to performance is achieved.

VALIDATION OF PROPOSED GREEN IT FRAMEWORK FROM DIFFERENT DATA CENTERS

The proposed framework will be applied to case study to test its validity and legitimacy. It will be tested according to the following criteria:

1) The overall structure is a sensible and suitable to be implemented
2) It is feasible, easily understood and can be implemented with ease
3) It is a comprehensive approach and covers all the major aspects of the benchmarking Implementation for data centers
4) It provides a straight forward guide, which could simplify the process even to someone who is new to the greening and benchmarking concepts
5) It gives a clear and effective way of presenting the general Green IT implementation concept
6) It is a practical, realistic and uncomplicated framework, which can easily be used in real working environment
7) It provides a simple approach for incorporating measuring, benchmarking and recycling concepts of greening data centers
8) It adds continuous improvement and equipment utilization in the key performance measures section by providing latest energy efficiency techniques
9) Human resource development and training should be focused on educating the employees to improve positive work cultures practices
10) Feasible for overall data center performance
11) Feasible for overall data center management to implement it properly
12) Meets data center industry standards
13) Reduces overall cost of ownership
14) Achieves objectives like measuring the energy efficiency and CO_2 reduction and reducing overall cost of ownership of data center.

Criteria for verification and validation of proposed Green IT framework

The framework was tested for validity, credibility, reliability and legitimacy at five large tier level data centers in Pakistan. The case study focuses to test whether the proposed framework is feasible to be implemented with ease for greening data centers or not. It was tested whether the proposed framework fulfils the desired requirements of business sustainability and provides solutions to overcome energy issues and making the environment CO_2 free. The framework was also tested to find out whether it provides effective solutions in making the data center industry green and environment friendly. The results were based on group discussion, face to face semi structured interviews and survey questionnaire from different data center managers responsible for performing technical and operational functionalities at their respective data centers.

RESULTS AND DISCUSSION

Having developed the proposed energy efficiency and low carbon enabler Green IT framework for implementing energy efficient, green and environment friendly data centers considering green metrics as shown in Figure 2, it was evaluated and validated in this case study being performed on multiple tier level data centers in Pakistan. These data centers were chosen as case study companies because they were little bit easy to reach and their top management that is, managers provide enough support and commitment towards achieving the desired goals highlighted in this case study.

The researcher discusses about the general and specific questions, comments, criticisms and suggestions made by case study data center managers involved concerning the framework's strengths and weaknesses about its validation and verification while implementing it in those data centers. The author begins the discussion by conducting a cross-case analysis across all the five case study data centers about the validation and verification of implementation of proposed Green IT framework in tier level data centers in Pakistan.

As shown in Table 2, all five case study data center managers gave very positive comments on the proposed Green IT framework. The results highlighted clearly show that Green IT framework is acceptable to data center industry which measures energy efficiency and CO_2 emissions and is easy to be implemented without overhead. It provides cost effective means to be implemented with ease. Most of the data center managers commented that framework is feasible, easily understood and can be implemented with ease. It provides a comprehensive approach and covers all the major aspects of benchmarking and implementing energy efficient and green data centers. It also provides a straightforward guide, which could simplify the implementation process by dividing the implementation process into five categories with each category further subdivided into sub-components providing enough details, techniques, policies and metrics and so on for data center managers and even for someone who is new to this field. In short, the framework could be used as a base for implementing green and energy efficient data centers.

Apart from that, most of the data center managers agreed that proposed Green IT framework provides a sagacious and sensible approach towards highlighting the importance of measuring the performance of data center at regular time intervals so that benchmarking

Table 2. Validation of proposed Green IT framework from different data centers.

S/No.	Questions for validating the proposed Green IT framework for data centers	Data center 1 PTCL	Data center 2 HBL	Data center 3 NBP	Data center 4 PTCL	Data center 5 UBL
1	It helps data center managers to implement green IT techniques to achieve energy efficient and environment friendly data centers	✓	✓	✓	✓	✓
2	Its overall structure is simple, sensible and suitable for data center managers	✓		✓	✓	✓
3	It is feasible, easily understood and can be implemented with ease	✓	✓	✓		✓
4	It has comprehensive approach and covers all major aspects of green data centers	✓	✓	✓	✓	✓
5	It provide a straight forward guide, which could simplify the process even to someone who is new to this field		✓	✓	✓	✓
6	It is a practical, realistic and uncomplicated framework, which can easily be used in real working environment	✓	✓	✓	✓	✓
7	It adds continuous improvement and equipment utilization in key performance measures section by providing latest energy efficiency techniques	✓	✓	✓	✓	✓
8	It's a complete Green IT framework for implementing green data centers	✓	✓	✓	✓	
9	it highlights the importance of measuring the efficiency of data centers in terms of energy consumption and CO_2 emissions	✓		✓	✓	✓
10	It highlights the importance green metrics for measuring the efficiency of data centers in terms of energy consumption and CO_2 emissions	✓	✓	✓	✓	✓
11	It highlights the importance of proper recycling and procurement policy for implementing green data centers	✓		✓		✓
12	It follows a comprehensive approach and covers all the major aspects of benchmarking data center values for achieving efficiency	✓	✓	✓	✓	
13	It provides a simple approach for incorporating measuring, benchmarking and recycling concepts of greening data centers.				✓ ✓	✓
14	It focuses on human resource development and training for data center employees on educating employees to improve positive work cultures practices	✓	✓		✓	✓
15	It meets data center industry Standards and service level agreements (SLAs)				✓ ✓ ✓	
16	It achieves objectives like measuring the energy efficiency and CO_2 reduction and reducing overall cost of ownership of data center	✓	✓	✓	✓	✓

Table 2. Contd.

17	It identifies and categorizes key data center components into different resource pools (workloads) using green metrics	✓	✓	✓	✓	✓
18	It helps data center managers to implement new Green IT techniques like virtualization, cloud computing, thin provisioning, consolidation, physical to live migration and green metrics	✓	✓	✓	✓	✓

values can be set and changed at times to achieve desired objectives, as measuring the performance will help to quantify which parts of data center are up to the standard and which needs improvement particularly that which involves manufacturing, recycling and purchasing sectors. Most of the data center managers highlighted top management's roles and responsibilities in the key areas of the framework and asked that they should be in-place first before embarking on the actual implementation effort in achieving the vision towards business competitiveness. For example, developing implementation strategies, policies, vision and mission for energy efficiency should form an integral part of framework.

Conclusion

In recent years, energy efficiency has emerged as one of the most important design requirements for modern computing systems, such as data centers, as they continue to consume enormous amounts of electrical power. Apart from high operating costs incurred by computing resources, this leads to significant emissions of CO_2 into the environment. For example, currently IT infrastructures contribute about 2% of total CO_2 footprints. Unless energy-efficient techniques and algorithms to manage computing resources are developed, IT's contribution in the world's energy consumption and CO_2 emissions is expected to rapidly grow. This is obviously unacceptable in the age of climate change and global warming. This paper proposes a Green IT framework using virtualization technology to achieve power and energy efficiency in data centers.

The power management problem becomes more complicated when considered from the data center level. In this case, the system is represented by a set of interconnected computing nodes that need to be managed as a single resource in order to minimize the energy consumption. Live and offline migrations of VMs offered by the virtualization technology have enabled the technique of dynamic consolidation of VMs according to current performance requirements. However, VM migration leads to time delays and performance overhead, requiring careful analysis and intelligent techniques to eliminate non-productive migrations that can occur due to the workload variation. Virtualization provides the following advantages:

1) Virtualization of the infrastructure to support hardware and software heterogeneity and simplify the resource provisioning.
2) Application of VM migration to continuously adapt the allocation and quickly respond to changes in the workload.
3) Ability to handle multiple applications with different service level agreement (SLA) owned by multiple users.
4) Guaranteed meeting of the QoS requirements for each application.
5) Support for different kind of applications, mixed workloads.
6) Decentralization and high performance of the optimization algorithm to provide scalability and fault tolerance.

This paper proposed a new Green IT framework that used virtualization technology as basics for implementing energy efficient data centers. The framework provided an imminent solution to the data center managers to improve the performance of their existing data center by implementing the proposed Green IT framework. It also helped them to reduce the emission of GHG so that global warming effects can be eliminated or reduced. It also provided with a set of green metrics to be implemented in data centers to measure the performance in terms on energy efficiency and CO_2 emissions.

ACKNOWLEDGEMENT

The authors are thankful to the Deanship of Scientific Research, King Saud University Riyadh for funding the work through the research Group project No RGP-VPP-076.

REFERENCES

Anderson D, Thakar P, Vahdat A, Doyle R (2001). Managing Energy and server resources in hosting centers. In Proceedings of 18th ACM Symposium on Operating Systems Principles, (SOSP), pp. 103–116.
Belady C (2008). Green Grid Data Center Power Efficiency Metrics: PUE and DCIE. The Green Grid, San Francisco, CA.
Brill KG (2007). The Invisible Crisis in the Data Center: The Economic Meltdown of Moore's Law. Uptime Institute White Paper, Santa Fe, NM.
Bugnion E, Devine S, Rosenblum M (1997). DISCO: Running Commodity Operating Systems on Scalable Multiprocessors. In SOSP, pp. 143–156.
Byrd TA, Turner DE (2000). Measuring the Flexibility of Information Technology Infrastructure: Exploratory Analysis of a Construct. J. Manage. Inform. Syst., 17(1): 167-208.

Crosby S, Brown D (2006). The Virtualization Reality, http://queue.acm.org/detail.cfm?id=1189289.

Daim T, Justice J, Krampits M, Letts M, Subramanian G, Thirumalai M (2009). Data center metrics: An energy efficiency model for Information Technology managers, Manage. Environ. Qual. Int. J., 20(6): 712 – 731.

Environmental Protection Agency (EPA) (2007). Report to Congress on Server and Data Center Energy Efficiency – Public Law 109-431, Environmental Protection Agency, Washington, DC.

Environmental Protection Agency (EPA) (2009). EPA ENERGY STAR Program Requirements for Computer Systems - Draft 4, Environmental Protection Agency, Washington, DC.

Gartner (2008). Sustainable IT, A Gartner Briefing, Gartner, Dublin.

Gartner (2009). Cost Optimization and Beyond: Enabling Business Change and the Path to Growth. A Gartner Briefing, Gartner, London.

Greenberg S, Mills E, Tschudi W, Rumsey P, Myatt B (2006). Best Practices for Data Centers: Results from Benchmarking 22 Data Centers. Proceedings of the 2003 ACEEE Summer Study on Energy Efficiency in Buildings. Asilomar, CA: American Council for Energy Efficient Economy, in Washington, DC. August.

Green Grid, Using virtualization to improve data center efficiency. Available at: www.thegreengrid.org/ (accessed February 2009)

Green Grid Metrics (2007). Describing the data center power efficiency, http://www.thegreengrid.org/~/media/WhitePapers/Green_Grid_Metrics_WP.ashx?lang=en.

Hanuv M, Grant G, Mann IJS (2009). Green IT: An Implementation Framework. AMCIS 2009 Proceedings.

Hart SL (1997). Beyond Greening: Strategies for a Sustainable World. Harvard Bus. Rev., 75(1): 66-76.

Huber P, Mark PM (1999). Dig more coal-the PCs are coming. Forbes, pp. 70-72.

IDC Predictions (2008). An Economic Pressure Cooker Will Accelerate the IT Industry Transformation, http://www.idc.com/getdoc.jsp?containerId=215519 (last accessed: 2009-05-12).

Koomey J, Piette MA, Cramer M, Eto J (1996). Efficiency Improvements in U.S. Office Equipment: Expected Policy Impacts and Uncertainties. Energy Pol., 24(12): 1101-1110.

Koomey JG (2007). Estimating Total Power Consumption by Servers in the US and the World. Lawrence Berkeley National Laboratory, Berkeley, CA.

Kumar R (2008). Media Relations. Gartner, Available at: www.gartner.com/it/page.jsp?id¼781012.

Koomey J (2008). Worldwide electricity used in data centers." Environ. Res. Lett., 3(034008) (8 pp), doi:10.1088/1748-9326/3/3/034008.

McNamara W, Seimetz G, Vales KA (2008). Best Practices for Creating the Green Data Center. http://www.it-solutions.siemens.com/b2b/it/en/us/Documents/publications/best-practices-green-it-white-paper_PDF_e.pdf.

Malone C, Belady C (2007). Metrics and an infrastructure model to evaluate data center efficiency. In Proceedings of the Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition (IPACK).

Mingay S (2007). Green IT: A New Industry Shockwave, Gartner.

Molla A, Cooper VA, Pittayachawan S (2009). IT and Eco-sustainability: Developing and Validating a Green IT Readiness Model. Proceedings of International Conference of Information Systems, December 15-18, Phoenix, AZ.

Mueen U, Azizah AR (2010). Pre Requisites for implementing energy efficient and cost effective data centers using virtualization. J Comput., 2(11): 95-101.

Mueen U, Azizah AR (2010). Server consolidation: An Approach to make Data Centers Energy Efficient and Green. Int. J. Sci. Eng. Res. (IJSER), 1(1): 1-7.

Mueen U, Azizah AR (2011). Virtualization Implementation Model for Cost Effective and Efficient Data Centers. Int. J. Adv. Comput. Sci. Appl. (IJACSA), 2(1): 59-68.

Murugesan S (2007). Going Green with IT: Your Responsibility toward Environmental Sustainability. Cutter Business-IT Strategies Executive Report, 10(8): 1-25.

National Data Center Energy Efficiency Strategy Workshop (2008). Energy efficiency in data centers: Recommendations for government-industry coordination, presented at National Data Center Energy Efficiency Strategy Workshop, Redmond, WA.

Pettey C (2007). Gartner estimates ICT industry accounts for 2 percent of global CO_2 emissions, http://www.gartner.com/it/page.jsp?id=503867.

Rivoire S, Shah MA, Ranganathan P (2007). A balanced energy efficiency benchmark." Proceedings of the ACM SIGMOD International Conference on Management of Data.

Scaramella J, Eastwood M (2008). Delivering Improved Energy Efficiency for the Data center. http://www.nec.com/global/prod/express/library/pdf/idc_ecocenter.pdf

SMART 2020 (2008). Smart 2020: enabling the low carbon economy in the information age', http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf .

SPEC SSJ Design (2008). SPEC.org, available at:www.spec.org/power_ssj2008/docs/SPECpower_ssj2008-Design_ssj.pdf.

Talaber R, Brey T, Lamers L, (2009). Using virtualization to improve data center efficiency.

Tschudi B, Tengfang Xu, Dale S, Bruce N, Jon K, Osman S (2004). Energy Efficient Data Centers. Berkeley, CA: LBNL, for the California Energy Commission. LBNL-54163. M.

Tschudi W, Evan M, Steve G, Peter R (2006). Measuring and Managing Data Center Energy Use: Findings–and resulting best practices– from a study of energy use in 22 data centers. Heating, Piping, Air conditioning (HPAC) Eng., pp. 45-51.

Tschudi W, Priya S, Tengfang Xu, David C, Paul R (2003). Data Centers and Energy Use – Let's Look at the Data. Proceedings of the 2003 ACEEE Summer Study on Energy Efficiency in Industry, Rye Brook, NY: American Council for Energy Efficient Economy, in Washington, DC.

Uddin M, Rahman AA, Memon J (2011). Carbon sustainability framework to reduce CO2 emissions in data centres. Int. J. Green. Econ., 5(4): 353–369.

Uddin M, Rahman AA (2011). Techniques to implement in green data centres to achieve energy efficiency and reduce global warming effects. Int. J. Global Warm., 3(4): 372–389.