Journal of Theoretical and Applied Information Technology
20th July 2013. Vol. 53 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPLEMENTATION OF VIRTUALIZATION IN DATA CENTERS TO INCREASE PROFICIENCY AND PERFORMANCE

[1]MUEEN UDDIN, [2]ASADULLAH SHAH, [3]RAED ALSAQOUR

[1,2]Kulliah of Information Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia
[3]School of Computer Science, Faculty of Information Science and Technology, University Kebangsaan Malaysia, Bangi, 43600, Selangor, Malaysia

ABSTRACT

A traditional data center is used to host computer systems and associated components. Electricity usage has become an increasing fraction of the total cost of ownership for data centers. It is difficult to reduce the electrical consumption of tier level data centers through appropriate design of the network-critical physical infrastructure and through the design of the IT architecture. This paper explains how virtualization can be implemented in data centers to increase proficiency and productivity. The paper is a guideline for data center managers to properly design and implement virtualization at different levels in data centers. It helps to achieve sustainable energy efficient data centers to ensure that electrical consumption contributes exceedingly little to the overall emission of greenhouse gases and implementing sustainable businesses.

Keywords: *Data Centers; Data Center Performance; Energy Efficiency; Server Consolidation; Virtualization Implementation.*

1. INTRODUCTION

The progress of Information and Communication Technology (ICT) based businesses and social practices has transformed many economies into e-economy and businesses into e-business. Continuous increase in Internet based businesses has led the way for large server based network facilities called data centers [1]. These gigantic data centers consume an enormous amount of electricity to process business requests leading to massive demands for power and energy. They are identified as crucial concerning areas for implementing environmental sustainability [2]. They are business amenities serve a wide diversity of facilities and services offered via Internet including e-commerce, web-hosting, social networking, backups, recovery. Other services like Internet as a service (IAAS), software as a service (SAAS), platform as a service (PAAS), infrastructure as a service (IAAS), virtualization and cloud computing [3].

Data centers are used and operated in different organizations to help business processes, information management, storage, backups and recovery functions [4]. High performance computing data centers are rapidly increasing in both size and magnitude, due to wide acceptance of hosting and performing extensive services and applications by large scale heterogeneous, multi-tier server clusters called blades and resource management in numerous businesses, internet and e-commerce applications.

Thousands of servers are densely packed in machine rooms to provide services to scamper businesses efficiently and cost effectively [5]. The increased load on these facilities is pushing the demand for installing millions of servers to match ever-increasing business needs [6].

Data center IT equipments consume an enormous amount of power and energy resulting in increased energy costs along with growing demands for the purchase of additional cooling and other functioning equipments [7]. These data centers are populated with racks of blade servers, these racks use almost all power required by IT equipments to operate i.e. 25 to 30 kWh of power. The problem with these racks is that they are underutilized but still consume power and energy. The cooling equipments required to cool servers and other equipments also consume an additional 15 to 20 kWh of power, equivalent to peak electricity demands for more than 30 typical homes in California [8]. The data center equipments contribute one half to one watt for every watt of power consumed by IT equipments [9]. Research is

Journal of Theoretical and Applied Information Technology
20th July 2013. Vol. 53 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

going on to create techniques and opportunities to help data centers make better decisions in ensuring secure energy supply, protecting the environment by reducing the effects of global warming, reducing server footprints and creating sustainable businesses.

Energy efficient computing is now being recognized as the most important aspect in computing systems ranging from embedded devices to large scale data centers. Large installation of server machines and other additional equipments known as data centers are increasingly become popular as they are attached with almost all organizations to provide critical services for business enterprises [10]. There is significant potential for energy efficiency improvements in data center industry. Many technologies are either commercially available or will soon be available that could improve the energy efficiency of microprocessors, servers, storage devices, network equipment and other infrastructure systems [11].

This paper focuses on the use of virtualization to overcome energy problems in data centers. In this paper, we proposed a five-step process to implement virtualization in a data center to save energy and at the same time increases the productivity of servers with little or no additional energy consumption. Also, this paper explores the importance of virtualization technology and helps data center managers to implement different types of virtualizations at different levels in a tier level data center. Virtualization has become popular in data centers since it provides an easy way to partition physical resources, allowing multiple applications to run in isolation on a single machine. The creation and management of virtual machines is called virtualization [12]. It decouples software from hardware and splits multi-processor servers into more independent virtual hosts for better utilization of hardware resources, allowing services to be distributed one per processor (Figure 1). It promises to dramatically change how data centers operate by breaking the bond between physical servers and the resource shares granted to customers [13]. It helps to consolidate the load of multiple underutilized volume servers onto few servers thereby reducing costs, energy consumption and cooling requirements. Enterprise Management Association research shows that over 90% of all enterprises are very keen to implement virtualization, and the market is growing at over 20% per year [14].

Virtualization can be applied to operating systems, desktops, applications, storage, and networks, depending on the end user and business requirements. There are many vendors to provide virtualization like VMware, Cisco, Citrix, EMC, Hewlett-Packard, IBM, Microsoft, NetApp, Novell, Oracle, Quest, Sun, Symantec, Virtual Iron, and many more. This paper explains and discusses different requirements to be fulfilled before implementing virtualization in any data center.

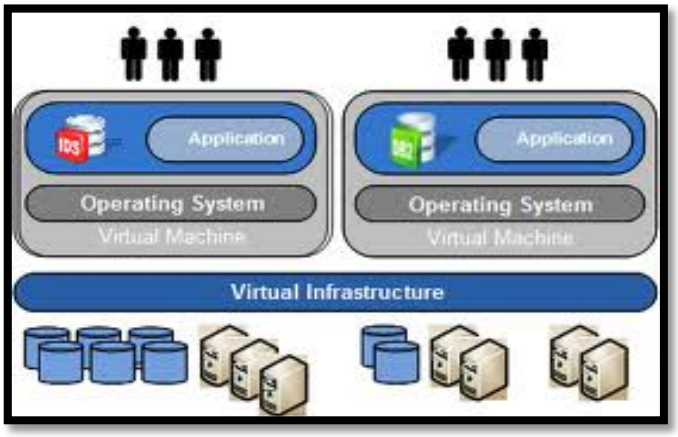

Figure 1. Process of Virtualization (VMWare)

2. PROPOSED WORK

Virtualization is performed on a given hardware platform by introducing a software layer that creates a simulated computer environment, a virtual machine, for its guest software [14]. The guest software runs just as if it were installed on a stand-alone hardware platform. The guest software itself is in many cases a complete operating system, and can also be different from the operating system hosted by the physical machine. Usually, many virtual machines run on a single physical machine; their number is limited by the host hardware capability, such as core number, CPU power, and RAM resources. The guest system is able to access specific peripheral devices, exploiting interfaces to those devices, e.g. Hard disk drive, network interface card, graphic and audio card. There are different types of virtualization depending on what is being virtualized and up to what extent, for instance platform virtualization, hardware virtualization, software virtualization, application virtualization, server virtualization and storage virtualization [15].

Some of the major benefits of virtualization technology are:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1. Easier, faster disaster recovery/business continuity planning.
2. Fast, agile, and responsive self-service provisioning.
3. Effective management of IT services to meet business needs.
4. Rapid scale-up (and scale-down) of compute resources to satisfy user load.
5. Faster, cheaper test and development for in-house applications.
6. Simple, responsive, utility-style 'cloud computing' infrastructure.
7. Hardware and software maximization
8. Workload configuration and utilization
9. Performance and energy measurement techniques.
10. CO_2 emission reductions
11. Reducing the effects of global warming
12. Saving the total cost of ownership
13. It reduces the total amount of hardware used in your environment
14. Idle virtual servers can be powered off
15. The virtualized server will have much less idle time and waste less
16. The total volume of space, air, and rent will be reduced. Data centers can use up to 100 times.

Virtualization is a technology that abstracts physical hardware from IT services. It is a disruptive innovation that is taking IT by storm. It improves server utilization, reduces hardware costs, increases uptime, makes IT faster to react, and drives significant return on investment (ROI). It also provides multiple execution environments for different applications and workloads to execute freely with all resources available. Debugging, performance, sandboxing, resource sharing and testing are the major characteristics of virtualization when implemented properly in a networking environment [16].

2.1 Energy saving techniques in data centers

The development and implementation of energy efficient resource management strategies in data centers is now becoming a fundamental prerequisite to implement energy efficient green and environment friendly data centers. The actual energy consumed by data center does not affect the cost of infrastructure, but is reflected in the electricity cost consumed by the system during the period of operations.

Considering power consumption in data centers, the major problem is the minimization of peak power required to feed completely utilized system. To achieve power efficiency, system level power management solutions should be developed through careful integration of hardware solutions such as asymmetric multi-core microprocessor design, efficient packaging techniques, and energy proportional hardware designs that focus on memory and disk subsystems which can reduce the power consumption of central processing units [17]. Table 1 shows six different energy savings techniques with their advantages and disadvantages.

Table 1. Energy Saving Techniques in Data Centers

Technique	Description	Advantages	Disadvantages
Virtualization	Multiple operating systems are setup on one or more servers	Users do not notice a difference	Setup cost
Consolidation	Minimize servers that sit idle using fewer servers actively	Fewer systems to maintain	Does not allow extra machines to accommodate future capacity
Improved Cooling	Techniques to cool IT equipment with less energy	No need to downsize capacity	Hard to measure the impact without sufficient modeling
Green Metrics	Measure efficiency in terms of energy utilization and CO_2 emissions	Improves the efficiency of data center	No standardize metrics available to measure efficiency
Green Acquisition	Through programs such as electronic product environmental assessment tool (EPEAT) and energy star, buyers buy products which are easily recyclable and made of less toxic materials	Encourage manufacturers to build more green products	Limits what you can buy
Green Building	Using green building practices such as leadership energy and environmental design (LEED) certification to build more energy efficient buildings	Energy savings; LEED prestige great public relations	Can be implemented when only building new data center

2.2 Plan for virtualization

The purpose of the planning phase is to plan all project processes and activities required to ensure successful implementation of proposed virtualization technique and to create a comprehensive set of plans. It is a process of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

defining an orderly arrangement of activities and resources needed to implement the strategy.

The reasons to consolidate and virtualize IT infrastructure (servers) are extensive. Most IT managers focus on just the cost of the physical IT infrastructure, but there are many additional benefits that can be derived from server virtualization include proper utilization of servers, proper energy utilization, reduction in consumption of energy, achieving sustainable businesses, simplified management, improved data protection, improved resource utilization, easier revision control and flexibility, easier data protection and security, reduced server and software costs, reduction in emission of greenhouse gases and reduction in global warming effects.

Like other IT projects, implementation of virtualization also needs to be structured and designed in such a way that they must fulfill the necessary requirements and should be within the infrastructure domain already installed. It is much more than simply loading a virtualization technology on a server and transforming one or two workloads into virtual machines. It involves five key steps to be followed while implementing virtualization. These steps are important because they provide a detailed description about requirements to be fulfilled before implementation.

a) Inventory process.
b) Virtualization type.
c) Hardware maximization.
d) Virtualization architecture.
e) Virtualization management.

a) Inventory Process

Data centers are huge entities consisting of many different components and devices performing different tasks to meet end user and business needs. These components should be categorized into measureable resource pools depending on the workloads they execute. The inventory process identifies and then classifies all the components according to different parameters like energy use, carbon emission, utility ratio, type of equipment, life time. This process helps to identify:

- Make and model of the processor.
- Types of processors (socket, core, threads, cache).
- Memory size and speed.
- Network type (number of ports, speed of each port).
- Local storage (number of disk drives, capacity, redundant array of independent disks).
- Operating system and their patch levels (service levels).
- Applications installed.
- Running services.

There are many tools available from different vendors for performing initial analysis of an organization. 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 Microsoft visio, which generates a complete inventory diagram of all components and also provides details about each component being analyzed. Microsoft assessment and planning (MAP) toolkit is another tool for the assessment of network resources. It works with windows management instrumentation (WMI), the remote registry service or with simple network management protocol to identify systems on a 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 agentless tool it does not add any overhead over production server's workload.

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 following service types:

- Network infrastructure servers.
- Identity Management servers.
- Terminal servers.
- File and print servers.
- Application servers.
- Dedicated web servers.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Collaboration servers.
- Web servers.
- Database servers.

Categorizing Application Resources

Data centers are the backbone of today's businesses. They execute different applications from different users. These applications are mostly performed by servers responsible for executing different tasks. The categorization phase also categorizes these applications into different categories depending on their execution requirements and type. These categories are commercial versus in-house, custom applications, legacy versus updated applications, infrastructure applications, support to business applications, line of business applications and mission critical applications.

Allocation of Computing Resources

After creating the workloads, the next phase is to allocate computing resources required by these different workloads and arrange them in normalized form, but for normalization the processor utilization should be at least 50%. It is very important to normalize workloads to achieve maximum efficiency in terms of energy, cost and utilization. The formula proposed 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).

b) Virtualization Type

After analyzing and categorizing servers, applications and associated workloads, the next phase defines virtualization in detail, its advantages, its types, layers and most importantly vendor identification. Virtualization is the faithful reproduction of entire architecture in software, which provides the illusion of a real machine to all software running above it. It is a framework or methodology of dividing the resources of a computer into multiple execution environments, by applying hardware and software partitioning, time-sharing, partial or complete machine simulation, emulation, quality of service, and many others. It is implemented at either software level or hardware level or even at both levels depending on the requirements and infrastructure.

It is very significant for an organization to know in advance the total content of its infrastructure before implementing virtualization technology. 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. These tools help data center managers to have an idea of physical infrastructure and requirements fulfilled before actually implementing virtualization at any layer.

c) Hardware Maximization

This is the most important step of virtualization process. Since servers are now going running multiple virtual workloads, it is important to consider hardware issues because already available hardware is not enough and suitable for providing high availability of virtual workloads. A change is required to install new hardware that supports and delivers the best price and performance. This process ensures high availability of virtual workloads and also provides leaner and meaner resource pool of resources for these virtual workloads.

d) Virtualization Architecture

The architecture of a machine consists of set of different instructions that allow inspecting or modifying machine state trapped when executed in any or most probably the privileged mode. To support proper hardware utilization, it is important to update and revise whole datacenter architecture. To protect virtual workloads, x-64 systems should be linked to shared storage and arranged into some form of high availability clusters so as to minimize the single point of failure.

e) Virtualization Management

It is a continuous process to provide a better understanding and relationship between the components of virtualization and the process as a whole to achieve desired objectives. The process of managing data centers is implemented in two phases. In the first phase, all the activities and requirements needed for proper implementation of virtualization should be monitored and managed properly from time to time. In the second phase, the process of management should continue even after the implementation of virtualization to properly manage the applications, their workloads, processes and the resources available to meet those objectives.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.3 Benefits of Virtualization

Virtualization promises to radically transform computing for the better utilization of resources available in the data center reducing overall costs and increasing agility. It reduces operational complexity, maintains flexibility in selecting software and hardware platforms and product vendors. It also increases agility in managing heterogeneous virtual environments. Some of the benefits of virtualization are:

2.3.1 Server Consolidation

Server consolidation is a logical and compelling approach for reducing operational and purchasing costs and improving the reliability of data center [18]. This technique is getting popularity acceptance to manage and utilize underutilized systems. It describes a variety of ways of reducing capital and operating expenses associated with running servers as shown in Figure 2. It also provides an opportunity to address management issues such as resilience and scalability, storage consolidation and disaster recovery (which may be impossible if servers are dispersed throughout the organization). It provides cost savings in different ways to reduce the overall total cost of ownership of data center. These savings can be categorized as:

Staff Costs: Supporting staff is recruited on the basis of how many servers an organization has (one member of staff per twenty servers is often used as a benchmark). Most server consolidation projects aim to reduce costs by freeing staff from mundane server maintenance tasks. Gartner (2008) suggests that more than 70% of potential savings from a typical project will come from reduced staff requirements [19].

Hardware Costs: Consolidation can reduce costs through better server utilization, by reducing the total requirement for storage and by enabling the use of more cost effective back-up/restore mechanisms. Centralized purchasing may also enable better discounts to be negotiated with hardware suppliers [18].

Software Costs: Consolidation may also reduce the total number of licenses needed while standardizing fewer applications, that may allow better deals to be negotiated with suppliers. With many (but not all) applications, the incremental cost of software licenses decreases as the number of user's increases [20].

Facilities Costs: Server consolidation reduces the amount of floor space needed for data centers. This is a particular benefit if one or more existing locations are full [20].

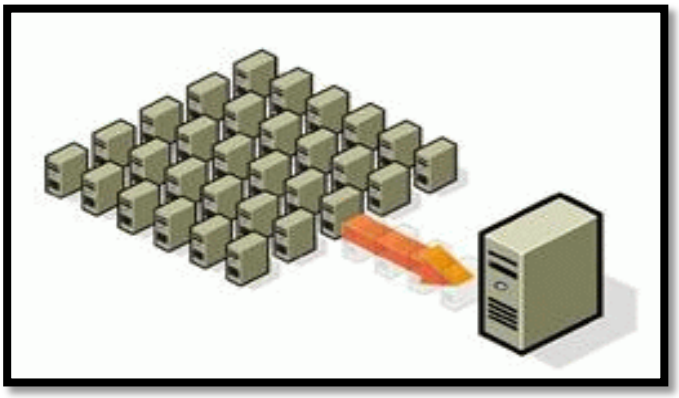

Figure 2. Server Consolidation (VMWare)

2.3.2 Simplified Management

A virtual machine can be more easily managed, configured and controlled from outside than a physical one. While it is possible to switch off a physical host remotely, it is impossible to switch on a turned off physical machine. In the virtual environment, all the actions on the virtual machine can be managed from a remote location, including the power on and power off. By implementing virtualization, the number of physical machines reduces automatically hence management becomes simpler and cost effective [21].

2.3.3 Sandboxing

Virtualization provides secure, quarantined sandboxes for running entrusted applications. Examples include address obfuscation. They also provide fault and error containment by isolating applications and services they execute, ultimately providing better behavior of these different faults.

2.3.4 Multiple Execution Environments

Virtual machines are used to create operating systems and execution environments to guarantee resource management and using resource management schedulers with resource limitations. Virtual machines provide the illusion of hardware configuration such as SCSI devices. Virtualization can also be used to simulate networks of independent computers. It enables to run multiple operating systems simultaneously having different versions, or even different vendors sharing and executing the workloads o different applications being processes.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.3.5 Debugging and Performance

Virtual machines allow powerful debugging and performance monitoring tools that can be installed in the virtual machine monitor to debug operating systems without losing productivity.

2.3.6 Physical to Live Virtual (P2V) Live Migration

A virtual machine can easily be moved from one physical machine and relocated to another, if needed [22]. Typical examples of virtual machine migration are: virtual machine cloning (copy more instances of the same virtual machine), virtual machine relocation (to free the original machine hosting the virtual machine). This scenario helps to increase the mobility of server machines in data center resulting in high availability and disaster recovery solutions.

3. CONCLUSION

This paper highlights the importance of virtualization technology being implemented in data centers to save the cost and maximize the efficiency of different resources available. We proposed a five-step model to properly implement virtualization. It starts by categorizing servers and their associated applications and resources into different resource pools. It is important to consider that virtualization not only needs to characterize the workloads that are planned to be virtualized, but also target the environments into which the workloads are to be applied. It is important to determine the type of servers, their current status whether idle or busy, how much it will cost to implement server virtualization, the type of technology needed to achieve the service levels required and finally meet the security/privacy objectives. It is also important for the data center to check whether it has the necessary infrastructure to handle the increased power and cooling densities arise due to the implementation of virtualization.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of this work by the Centre for Research and Instrumentation Management (CRIM), University Kebangsaan Malaysia (UKM), Malaysia. Grant numbers: UKM-GGPM-ICT-035-2011 and UKM-GUP-2012-089.

REFERENCES

[1] C. Lefurgy, K. Rajamani, F. Rawson, W. Felter, M. Kistler and T. W. Keller, "Energy management for commercial servers," *Computer*, vol. 36, pp. 39-48, 2003.

[2] G. Schulz, *The green and virtual data center*: Auerbach Publications, 2010.

[3] K. Kant, "Data center evolution: A tutorial on state of the art, issues, and challenges," *Computer Networks*, vol. 53, pp. 2939-2965, 2009.

[4] T. Daim, J. Justice, M. Krampits, M. Letts, G. Subramanian and M. Thirumalai, "Data center metrics: An energy efficiency model for information technology managers," *Management of Environmental Quality: An International Journal*, vol. 20, pp. 712-731, 2009.

[5] J. G. Koomey, "Worldwide electricity used in data centers," *Environmental Research Letters*, vol. 3, p. 034008, 2008.

[6] J.-M. Pierson, "Green task allocation: taking into account the ecological impact of task allocation in clusters and clouds," *Journal of Green Engineering*, vol. 1, pp. 129-144, 2011.

[7] M. Uddin and A. A. Rahman, "Energy efficiency and low carbon enabler green IT framework for data centers considering green metrics," *Renewable and Sustainable Energy Reviews*, vol. 16, pp. 4078-4094, 2012.

[8] R. Bianchini and R. Rajamony, "Power and energy management for server systems," *Computer*, vol. 37, pp. 68-76, 2004.

[9] M. Uddin and A. A. Rahman, "Techniques to implement in green data centres to achieve energy efficiency and reduce global warming effects," *International Journal of Global Warming*, vol. 3, pp. 372-389, 2011.

[10] S. Zeadally, S. U. Khan and N. Chilamkurti, "Energy-efficient networking: past, present, and future," *The Journal of Supercomputing*, vol. 62, pp. 1093-1118, 2012.

[11] S. Sharavanan, D. Poongodi and A. Kumar, "Towards Green Computing: energy efficient data centers," *Journal of Mathematics and Technology*, pp. 131-135, 2010.

[12] C. Liu, C.-L. Sia and K.-K. Wei, "Adopting organizational virtualization in B2B firms: An empirical study in Singapore," *Information & management*, vol. 45, pp. 429-437, 2008.

[13] X.-J. Chen and J. Zhang, "Overview of virtualization technology and its application in manufacturing information," *Jisuanji Gongcheng yu Yingyong(Computer*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Engineering and Applications), vol. 46, pp. 25-30, 2010.

[14] A. E. Outlook, "Energy Information Administration," in *Department of Energy*, ed, 2010.

[15] G. Goth, "Virtualization: Old technology offers huge new potential," *Distributed Systems Online, IEEE*, vol. 8, pp. 3-3, 2007.

[16] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy and G. Jiang, "Power and performance management of virtualized computing environments via lookahead control," *Cluster computing*, vol. 12, pp. 1-15, 2009.

[17] L. Benini, A. Bogliolo, G. A. Paleologo and G. De Micheli, "Policy optimization for dynamic power management," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 18, pp. 813-833, 1999.

[18] W. Min, L. Jing, F. Zhong-Lei and X. Lu, "A service model for virtual resource management and its implementation," *Chinese Journal of Computers*, vol. 5, pp. 31-34, 2005.

[19] A. Jones and S. Mingay, "Executive Summary: Going Green: The CIO's Role in Enterprisewide Environmental Sustainability," *Gartner Exp Premier, Accessed June*, vol. 10, p. 2008, 2008.

[20] J. Luo, Y.-y. Xu and R.-f. Li, "Dynamical resource allocation algorithm research in network virtualization," *Journal of China Institute of Communications*, vol. 32, pp. 64-70, 2011.

[21] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt and A. Warfield, "Xen and the art of virtualization," *ACM SIGOPS Operating Systems Review*, vol. 37, pp. 164-177, 2003.

[22] T. Garfinkel, B. Pfaff, J. Chow, M. Rosenblum and D. Boneh, "Terra: A virtual machine-based platform for trusted computing," in *ACM SIGOPS Operating Systems Review*, 2003, pp. 193-206.