Nurun Nahar,2021, 9:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

A Systematic Literature Review of Emerging Energy Efficient Approach for Cloud Data Centers

Nurun Nahar, Ahmed Afif Monrat, MD Masuduzzaman, MD Shakib Mahamud
South China University of Technology
Guangzhou, China
nurunnaharakhey@gmail.com, mdshakibmahamud@gmail.com, masuduzzaman811@gmail.com
Lulea˚ University of Technology Skelleftea˚, Sweden
ahmed.monrat@ltu.se

Abstract- On a pay-per-use model, cloud computing has emerged as a popular paradigm for providing IT infrastructure, tools, and services. The broader adoption of cloud and virtual- ization technology in recent years has led to the establishment of massive data centers with tremendous energy usage, affecting not only high operating costs, but also adding to the environ-ment's carbon footprint. Energy efficiency is, therefore, becoming increasingly important for cloud data centers. This paper pro-vides a detailed overview of power management techniques to optimize energy consumption in datacenters. We discussed some elementary issues regarding cloud computing and data centers, for example, data center energy savings technologies, resource allocation and scheduling algorithms, provisioning policies and power saving metrics and discussing on recent state-of-the-art models that prove to be efficient for cloud data center.

Keywords:- virtualzation, environment's, pro-vides

I. INTRODUCTION

The massive transformation of the entire computing industry in the 21st century projected this vision of computing utilities based on a service provisioning model, whereby computing resources will be readily accessible on-demand, as are other utility services available today [1]. Similarly, when using computer resources,users (cons umers) must pay providers. At present, customers no longer have to spend heavily or face challenges in building and maintaining complex IT infrastructure [2]. These utilities have continued to improve the accessi bility of leading users to computing services more int egrated than ever and more sophisticated based on their demands,regardless of where the hosts are loca ted.This paradigm was historically referred to as utili tybased computing,and later as cloud computing [3]. The former definition describes the infrastructure as a"cloud" (IaaS) from which businesses and customers can access services from anywhere in the world as applications on demand. Therefore, cloud computing is considered a modern paradigm for the perfor mance optimization of state-of-the-art computing re sources provided by data centers that commonly use virtual machine (VM) technologies for consolidation and environmental exclusion objectives while main tain their industrial business model.

Global companies suchas Google,Amazon, Microsoft, and IBM2 are massively building up data centers in various locations with different time zones to satisfy millions of users worldwide accessing services provid ed by cloud storage technologies [4].

From the viewpoint of the consumer, the reason for this growth is cost-effectiveness and quality of experi ence. It delivers infrastructure, networks, and apps (applications) as services that are accessible under the pay-as-you-go model to consumers as subscri ption-based services. Such services are referred to in the industry as Infrastructure as a Service (IaaS), Application as a Service (PaaS), and Software as a Service (SaaS).Cloud Computing, that long-held dream of computing as a utility, has the potential to transform a large part of the IT industry,making software even more attractive as a service. A recent Berkeley research states,"Cloud Computing is intend ed to provide advanced cloud storage as it seeks to virtually run in an energy-efficient direction where data centers can function efficiently" [5].

However,in a recent online report by (IEA, 2020) sugg ests that the global internet trends as shown in Fig1. [6]

Since 2010, the number of internet users worldwide has increased, while global internet traffic has grown 12-fold.Nevertheless, significant improvements in per formance have helped the energy crisis to be redu ced through data centers and data transmission net works accounting for about 1% of global electricity consumption in 2019 [7].

Traffic increased by almost 40 percent from February to mid-April 2020, driven by growth in video stream ing, video conferencing, online gaming, and social networking [8].

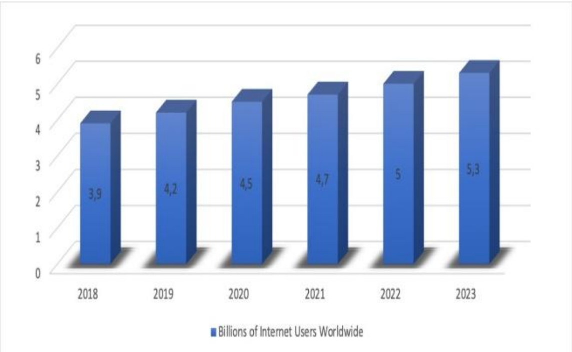

Fig 1. Global Internet Users (Billions).

The findings such as 205 terawatt-hours (TWh) were likely consumed by global data centers in 2018 [9], or 1percent of global electricity consumption, is in stark contrast to previous extrapolation- based figures that have shown increasingly growing data center energy use over the past decade.

This paper conducts a systematic literature survey of power management techniques to optimize energy consumption in data centers.Several important aspec ts of data centers such as, data center energy savings technologies,resource allocation and scheduling algo rithms,provisioning policies and power saving metr ics have been discussed. The rest of this paper is organized as follows. Section II introduces power consumption challenges in data centers. Section III discusses energy efficient solutions for data centers. Section IV introduces metrics for determining energy consumption. Section V concludes the paper.

II. POWER CONSUMPTION CHALLENGES IN DATA CENTERS

This paper aims to evaluate Data centers' power cons umption,as discussed in the introduction section that energy consumption has increased in recent years due to growth in size and number of data centers, which comprise cooling and power provision systems storage drives,andservers.Ina2016,a report suggested that in United States alone,"In the near future, energy consumption is projected to continue to rise margin ally, rising by 4 percent from 2014-2020, at the same pace as in the past five years" [10].

U.S. data centers are expected to consume approxi mately 73billion kWh in 2020, based on current trend projections." as indicate in Fig. 2.

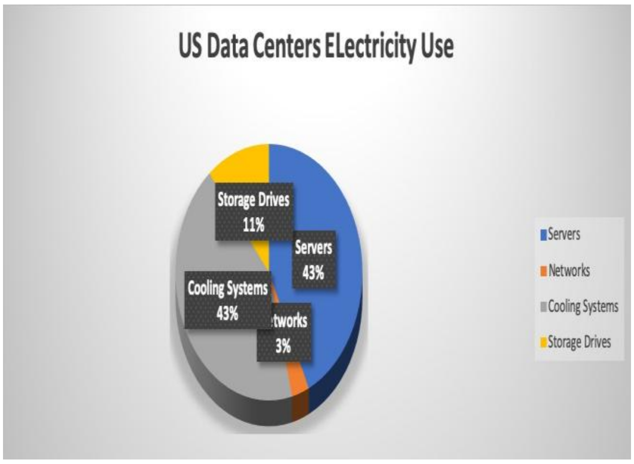

Fig 2. US Data Center electricity use in 2016.

The growing energy demand is because a typical datacenter that delivers cloud services contains tens of thousands of servers. This growth is in reaction to the massive increase in digital services over the last decade: since 2010,the number of internet users worldwide has doubled, while internet traffic world wide has risen by 12 times. Strong government and industry efforts on energy conservation, procurement of renewable sources are required to minimize grow th in energy demand and emissions over the next decade.However,despite the rapid increase in dema nd for information services over the past decade, between 2010 and 2018, global energy usage in data centers is likely to increase by just 6 percent. These new results are focused on the incorporation of different recent data-sets that better characterize installed stocks, operating characteristics and energy usage of data center IT equipment compared to previous studies, as well as on structural changes in the data center industry.

As data centers take advantage of virtualization tech nology to host multiple virtual machines (VMs) on a single physical server,e Electricity cost for powering servers forms a significant portion of the operational cost of data centers [11]. It is expected that the electricity demand for data centers to rise more than 66% over the period 2011–2035 [12]. It is estimated that energy costs may contribute even more than the cost of IT soon. Cloud service providers need to implement energy efficient management of data center resources to meet the increasing demand for cloud computing services and ensure low costs. Hence, there is a growing interest in saving energy consumption of cloud data centers. Due to massive power consumption levels of data centers, energy-saving techniques have become essential to maintain both energy and cost- efficiency [13].

III. ENERGY EFFICIENT SOLUTIONS FOR DATA CENTERS

Since the huge demand of consumer services are risi ng,more and more new users are being connected to IT resources,so does the growth of energy to operate massive infrastructures.Statistically, according to data center knowledge. com, in 2018 alone, 205 terawatt hours of electricity was consumed by world data centers,or about 1 percent of all electricity consumed worldwide that year as illustrated in Fig. 3.

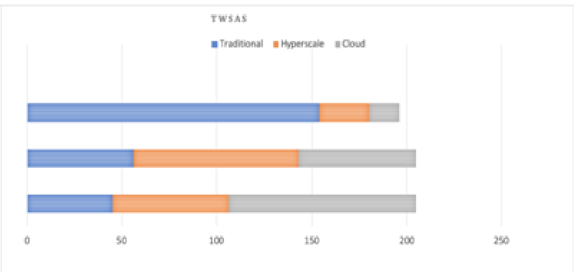

Fig 3. Energy consumption indicator among traditional, hyperscale and cloud.

In a wider view,climate change is inevitable.Some might sugg est that we are living in the carbon age, facing huge amounts of greenhouse gas and carbon dioxide CO2 which is being released into the planet's natural eco-system,impacting some of the fragile and indigenous areas of the world and keeping them at potential risk of disaster. Which is why green energy solutions being necessary. The objective of this study is to discuss such state-of-the-art methods which help optimize energy.

1. Energy Efficient GREEN Solutions:
Many multinational companies are taking steps to mitigate the detrimental effects of their decisions on the environment. The United Nations Framework Con venation on Climate Change(UNFCC) is an international environmental treaty that seeks to stabilize green house gas emissions from the atmosphere toa degree that would discourage detrimental anthropogenic ecosystem interference.

Sustainable development means the development, without adverse needs, of future generations. That is to accomplish human growth objectives while protec ting the natural resources and ecosystems on which society relies. Energy saving, renewable energy and eco-friendly technical advances are now becoming an integral part of typical IaaS, PaaS and SaaS Applications in their respective architectures. This study is about new and optimal solutions in data centers that help save resources.

Since data centers emit a huge amount of energy usage into the atmosphere, scientists around the world may take a broad approach by using technolo gies that pledge complex resource alloca tion methods and power management techniques to curb this phenomenon to the best of their capacity, while supporting providers of data centers and providing customers with the best services.

Green computing is gaining considerable significance in terms of raising awareness of the environmental effects of computing. With rising global warming, energy use and e-waste, the concept of green comp uting is commonly taken into consideration as a com mitment to moral practices for environmental impro vement by both governments and companies [15].

Since1992,withthelaunch of the Energy Star initiative, the green computing or greenIT term has been used, offering a voluntary mark awarded to computing goods that provide maximum performance while consuming less energy.

Refrigerators, televisions, monitors, air conditioners, and other household currencies were given the rati ng. Soon after that, computer-related products such as USBs, printers, displays, networking systems, servers, and network systems were protected by the word Green Computing [16].

In modern times, a few researchers have suggested a green computing architecture for the allocation of energy-efficient cloud storage services as shown in Fig. 4.

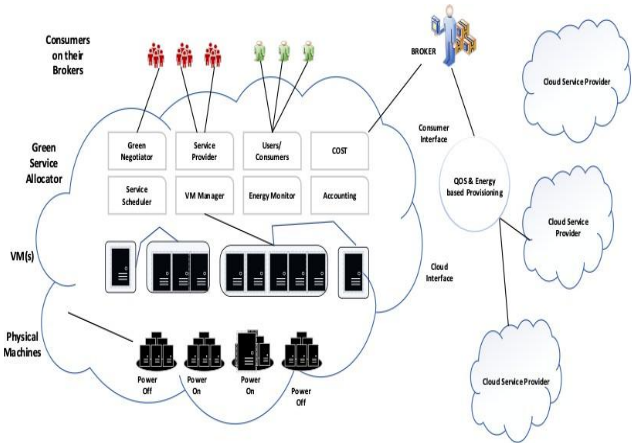

Fig 4. Energy Efficient Cloud Infrastructure Model.

Four key components are defined by this model such as User broker, Green resource allocator, Virtual mac hines and Physical machines.

1.1 User broker: Cloud customers or their brokers send service requests to the Cloud from anywhere in the world. It is important to remember that there may be adistinction between consumers of deployed services and cloud customers.

1.2 Green Resource Allocator: serves as the gateway be- tween users and the cloud infrastructure.

1.3 Virtual Machines: satisfy accepted demands, multi pleVMs can be dynamically started and stopped on a single physical machine, thus offering full flexibility to configure various resource partitions on the same physical machine to different specific service request requirements.

1.4 Physical Machines: In order to satisfy service dem ands, the underlying physical computing servers provide hard- ware infrastructure to build virtualized services.

2. Energy Efficient Resource Allocation:

Resource allocation or scheduling is one of the most important tasks in cloud computing. It consists in identifying and assigning resources to each incoming user request insuch a way that the user requirements are fulfilled, and specific goals of the cloud provider are satisfied.

These goals could be optimizing energy consumption or cost optimizing, etc. Based on the resource information like resource usage and monitoring, the requests information and the Cloud provider goal, the resource allocator or scheduler finds out resource
allocation solutions[17].

Schedul ers could just ensure the initial and static resource allocation after request arrival or ensure both static and dynamic resource allocation to manage reso urces in a continuous way and to further optimize and readjust the old requests [18].

3. Virtualization:

Virtualization involves making a virtual version, rather than a real version of something. Virtualization in cloud computing is a technique of digitally replic ating a version of something real. It helps multiple clients and organisations to share a single physical resource or an application. Virtualization is defined as software that enables a single hardware calcula ting device to be automatically partitioned into single/multi-assumed devices.

Virtualization is defined as software that enables a single hardware calculating devicetobe automatically partitioned into single/multi-assumed devices.

It is possible to use and monitor any of these virtual devices easily, thus minimizing costs by increasing the use of infrastructure as well as providing the agility required to speed up IT processes [19].

This technology allows multiple VMs to run simulta neously on a single physical host machine(HM) using a hardware resource capacity partition, where each VM hosts its applications, operating system, and middle-ware(memory ,CPU power, network bandwidth, and store capability) [20].

Users need resources, requiring a lot of investment in physical infrastructure to respond to user require ments; thus, cloud infrastructure providers fix and deal with this situation by delivering higher quality and low cost VM services based on user requests. There are different forms of virtualization, including disk,server,device,desktop, and network virtualization [21].

In an online study (Spiceworks.com, 2020) Server virtualization,which is used by 92 % of organizations, is estimated to be ubiquitous. Other forms of virtual ization, however, have some similar things to do. Storage virtualization (also called software-defined storage) with a 40 percent widespread adoption is the most common among emerging virtualization technologies, followed by application virtualization at 39 percent and virtual desktop infrastructure (VDI) technology at 32%.

Besides, network virtualization (also defined as software-defined networking) and data virtualization both have an acceptance rate of 30 percent. Looking ahead, research from Spice works reveals that by 2021,more than half of organizations intend to use storage virtualization and application virtualization.

4. The Static Threshold VM Consolidation:
Besides virtualization technologies there are other technologies such as: The static threshold process, the upper and lower limits for the workload, and the allocation and relocation of virtual machines are based on the specified threshold. Based on variables, such as minimum migration time, maximum correlation, and minimum consumption, this virtual machine is selected. By leveraging the trade-offs between service quality and service level agreement, the quantity of power that is being used in the data center can be controlled [22].

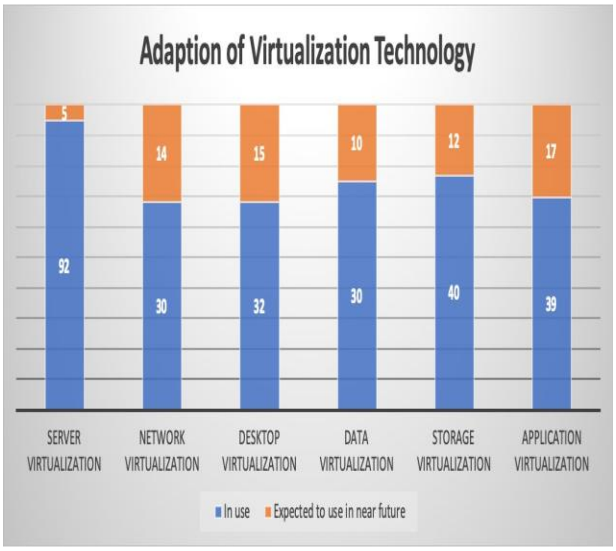

Fig 5. Virtualization Adoption.

IV. METRICS FOR DETERMINING ENERGY CONSUMPTION IN DATA CENTERS

Around the world, effective and new technologies are currently taking place. It is so quick to change the IT world that the cloud paradigm is becoming an important and integral part of the technological age. Cloud Storage firms therefore face problems in terms of energy costs about the economic impact. Energy utilization depends on many variables, e.g., service level agreement, strategies for choosing virtual mach ines, optimization policies, types of workload, etc. Many organizations search for the appealing cloud functionality of cloud technology. This includes nume rous sets of facilities, including high-perform ance storage processing output. This includes various entities that carry different data to servers, such as servers, networks, cooling systems, wires. Although cloud storage is mostly used for data centers, by adding heavy energy consumption, heat emission, toxic gaseous release, it contributes to an environ ental hazard. Table I depicts a measurement of data centers efficiency is depicted, taken from a thorough study of multiple sources [28].

Table 1. A Measurement Of Data Centers Efficiency Metrics.

Metrics for Energy Efficiency	Explanation	Formula
PUE (Power usage) [23]	The ratio of the total energy consumed by data centers or servers to the total energy consumed by computing devices is determined by PUE.	PUE = TotalPowerof DataCentersorServers/Total Powerof DataCentersorServers
CUE (Carbon usage Effectiveness) [24]	CUE is the fraction of the overall ambient discharge of CO_2 from the data center to the total energy used by the IT units.	CUE = TotalEmissionof Co2TEIT (TECO)/ Total Energy consumed by IT devices(TEIT)
ERE (Energy Reuse Effectiveness) [25]	ERE is a fraction of Data Center Reuse Energy Profit by total energy consumed by IT devices.	ERE = Total Energy(TE) – Reuse Energy (RE)/ Total energy consumed by IT devices(TEIT)

DCiE (Data center Infrastructure Efficiency) [26]	Data Centre Infrastructure Efficiency is a method of performance enhancement to define a data center's energy efficiency. It measures the effect of just reversing the PUE by calculating the power consumed by dividing IT equipment by the power consumed by the data center.	DCiE =Total Finished Jobs/ Total Resources
DCP (Data Centre Productivity) [27]	The DCP calculates the overall valuable work doneby the Data Centre by the total resources needed to complete the mission.	DCP =1/ PUE

V. CONCLUSION

This paper conducts a systematic literature survey of power management techniques to optimize energy consumption in data centers. Several important aspe cts of data centers suchas,data center energy savings technologies,resource allocation and scheduling algo rithms, provisioning policies and power saving met rics have been discussed.

We would like to extend this research by conducting some quantitative experi ments on cloud platforms, for instance, Azure, AWS, IBM etc, by pushing some workloads to determine the performance and scala bility along with the energy efficiency and quality of experience.

REFERENCES

[1] B.Ahmad, Z.Maroof, S.McClean, D.Charles, and G. Parr,"Economic impact of energy saving techni ques in cloud server," Cluster Computing, pp. 1–11, 2019.
[2] L. Kleinrock, "A vision for the internet," ST Journal of Research, vol. 2, no. 1, pp. 4–5, 2005.
[3] R.Buyya,S.K.Garg, and R.N.Calheiros,"Sla-oriented resource provi- sioning for cloud computing: Challenges, architecture, and solutions," in 2011 international conference on cloud and service computing. IEEE, 2011, pp. 1–10.
[4] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud compu ting," Future generation computer systems, vol. 28, no.5, pp. 755–768, 2012.
[5] KKonstantinos, M Persefoni, FEvangelia, M.Christos, and N. Mara," Cloud computing and economic growth,"in Proceedings of the 19th Panhellenic Conference on Informatics, 2015, pp. 209–214.
[6] G. Peters, R. Andrew, J. Canadell, P. Friedlingstein, R. Jackson, J.Ko- rsbakken, C.Le Que´re´, and A. Peregon, "Carbon dioxide emissions continue to grow amidst slowly emerging climate policies," Nature Climate Change,vol.10,no.1, pp.3–6, 2020.
[7] R.Hintemann and S.Hinterholzer,"Energy consum ption of data centers worldwide," in The 6th Inter national Conference on ICT for Sustainabil- ity (ICT4S). Lappeenranta, 2019.
[8] E.Masanet, A.Shehabi, N.Lei, S.Smith, and J. Koo mey,"Recalibrating global data center energy-use estimates,"Science, vol.367, no.6481, pp.984–986, 2020.
[9] E.Masanet and N.Lei,"Aspen global change instit ute energy project."
[10] A.Shehabi, S.Smith, D.Sartor, R. Brown, M. Herrlin, J.Koomey,E.Masanet, N.Horner,I.Azevedo,and W. Lintner,"United states data center energy usage report," 2016.
[11] Z.Li and Y.Lin,"Energy-saving study of green data center based on the natural cold source," in 2013 6thInternational Conference on Information Mana gement, Innovation Management and Industrial Engineering, vol. 3. IEEE, 2013, pp. 355–358.
[12] J.M.Kaplan,W.Forrest,andN.Kindler,"Revolutionizi ng data center efficiency," in Uptime Institute Sy mposium, 2008.
[13] D.Kliazovich,P.Bouvry,and S.U.Khan, "Greencloud: a packet-level simulator of energy-aware cloud computing data centers,"The Journal of Super computing, vol. 62, no. 3, pp. 1263–1283, 2012.
[14] A.Kusiak,Z.Zhang,and A.Verma,"Prediction, oper ations,and condition monitoring in wind energy," Energy, vol. 60, pp. 1–12, 2013.
[15] M. A. Mohammed, D.A. Muhammed, and J. M. Abdullah,"Green computing beyond the tradit ional ways,"Int. J. of Multidisciplinary and Current research, vol. 3, 2015.
[16] L.Liu, H. Wang, X. Liu, X. Jin, W. B. He, Q. B. Wang, and Y.Chen,"Green cloud: a new architecture for green data center,"in Proceedings of the6th inter national conference industry session on

Autono mic computing and communications industry
session, 2009, pp. 29–38.

[17] N.Akhter and M.Othman, "Energy aware resource allocation of cloud data center: review and open issues,"Cluster computing, vol.19, no.3, pp. 1163–1182, 2016.

[18] M.I.Alam, M.Pandey,and S.S.Rautaray,"A proposal of resource allocation management for cloud computing,"International Journal of Cloud Com puting and Services Science,vol.3,no.2,p.79, 2014.

[19] G.Warkozek,E.Drayer,V.Debusschere,andS. Bacha, "A new approach to model energy consumption of servers in data centers,"in 2012IEEE Inter national Conference on Industrial Technology. IEEE, 2012, pp. 211–216.

[20] J.V. Wang, K.-Y. Fok, C.-T. Cheng, and K. T. Chi, "A stable matching-based virtual machine allocation mechanism for cloud data centers," in 2016 IEEE World Congress on Services (SERVICES). IEEE, 2016, pp. 103–106.

[21] M. J. Usman, A. Samad, H. Chizari, A. Aliyu et al., "Energy-efficient virtual machine allocation techn ique using interior search algorithm for cloud datacenter," in 2017 6th ICT International Stud ent Project Conference (ICT-ISPC). IEEE, 2017, pp. 1–4.

[22] T.Deepika and A.N.Rao,"Active resource provision in cloud comput- ing through virtualization," in 2014IEEE International Conference on Computa tional Intelligence and Computing Research. IEEE, 2014, pp. 1–4.

[23] G. A. Brady, N. Kapur, J. L. Summers, and H. M. Thompson, "A case study and critical assessment in calculating power usage effectiveness for a data centre," Energy Conversion and Manage ment, vol. 76, pp. 155–161, 2013.

[24] D.Azevedo,M.Patterson, J. Pouchet, and R. Tipley, "Carbon usage effectiveness (cue): a green grid data center sustainability metric," White paper, vol. 32, 2010.

[25] B.Tschudi, O. Vangeet, J. Cooley, and D. Azevedo, "Ere: A metric for measuring the benefit of reuse energy from a data center," White Paper, vol. 29, pp. 82–83, 2010.

[26] B.Radibratovic,K.Meagher,and B.Needle, "Method for predicting power usage effectiveness and data center infrastructure efficiency within a real-time monitoring system," Mar.19 2013, uS Patent 8,401,833.

[27] D.Mukherjee,S.Chakraborty,I.Sarkar,A.Ghosh, and S.Roy,"A detailed study on data centre energy efficiency and efficient cooling techniques," Inter national Journal, vol. 9, no. 5, 2020.

[28] A.Varasteh and M.Goudarzi,"Server consolidation techniques in virtualized data centers: A survey," IEEE Systems Journal, vol. 11, no. 2, pp. 772–783, 2015.