iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** II **Month of publication:** February 2024

DOI: https://doi.org/10.22214/ijraset.2024.58271

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

Estimating Energy Consumption of Cloud, Fog and Edge Computing Infrastructures

Hitesh Parihar[1], Suman M[2]
B.Tech in Computer Science & Engineering, Dept of CSE, Presidency University

Abstract: ***To address the imperative need for enhanced locality considerations, innovative Cloud-related architectures, such as Edge Computing, have emerged. Despite the burgeoning popularity of these architectures, their energy consumption remains a relatively unexplored domain. To delve into this critical question, we present a comprehensive approach. Initially, we establish a taxonomy delineating diverse Cloud-related architectures. Subsequently, we introduce an intricate energy model designed to assess their consumption patterns.***

What sets our model apart is its inclusivity, encapsulating not only the energy usage of computing facilities but also factoring in cooling systems and the energy consumption of network devices connecting end users to Cloud resources. In contrast to prior proposals, our model provides a holistic view of energy consumption within the Cloud ecosystem.

The instantiation of our model on various Cloud-related architectures, ranging from fully centralized to entirely distributed ones, facilitates a nuanced comparison of their energy consumption. Our findings reveal that a fully distributed architecture, by eschewing intra-data center networks and large-scale cooling systems, exhibits a notable energy efficiency advantage, consuming between 14% and 25% less energy than fully centralized and partially distributed architectures, respectively.

Notably, our work represents a pioneering effort in introducing a model that empowers researchers to analyze and compare the energy consumption of diverse Cloud-related architectures.

Keywords: Cloud computing, energy consumption, dis tributed Clouds, Fog computing, Edge computing, P2P.

I. INTRODUCTION

The energy consumption of Cloud Computing solutions has emerged as a critical challenge in contemporary computing facilities, with data centers (DCs) being significant contributors to this issue. A 2016 report from Lawrence Berkeley National Laboratory (LBNL) revealed that DCs consumed an estimated 70 billion kWh, approximately 1.8% of total U.S. electricity consumption, and projected an increase to around 73 billion kWh by 2020 in the U.S.

Simultaneously, the rise of Internet of Things (IoT) applications has spurred the development of more distributed cloud-related architectures, leveraging resources deployed across the network and at its edge. These architectures, often referred to as Fog, Edge, and occasionally P2P computing infrastructures, aim to meet the low latency and high bandwidth requirements of IoT applications. Despite the certainty of their deployment, the energy footprint of these emerging architectures has not been thoroughly explored.

This paper addresses this gap by presenting a generic energy model designed to precisely assess and compare the energy consumption of these novel Cloud-related architectures. Recognizing the evolving frontiers of cloud-related architectures, we initially outline existing definitions and key features of each architecture. We then advocate for identifying each as a distinct and well-defined architecture, ranging from centralized to fully distributed structures.

To achieve this classification, we introduce a taxonomy that categorizes available Cloud-related architectures based on their characteristics and classifies the hardware elements within each infrastructure. Building upon this taxonomy, we propose an energy model that accounts for the nuanced differences among architectures. This model serves a dual purpose: accurately estimating the energy consumption of a specific architecture and facilitating meaningful comparisons across diverse architectures. The holistic nature of our approach positions it as a valuable tool for researchers seeking to address the energy challenges associated with modern Cloud-related infrastructures.

Our primary contributions can be summarized as follows:

1) Development of a comprehensive taxonomy for Cloud-related architectures and components, offering an extensive coverage in terms of technical features, services, and user satisfaction. The primary goal of our proposed taxonomy is to delve into the distinctive features of diverse architectures and components within similar paradigms, providing a foundation for categorizing both present and future architectures.

©IJRASET: All Rights are Reserved

2) Introduction of a generic and precise model designed to estimate the energy consumption of existing Cloud-related infrastructures. This model not only forms the basis for a thorough analysis and clear comprehension of the current Cloud landscape but also offers insights into the underlying technologies currently deployed in the domain.

The subsequent sections of this paper are organized as follows: Section II delves into related work, providing context for our contributions. Section III employs our proposed taxonomy to categorize Cloud-related architectures. In Section IV, we introduce the energy model we have developed. The analysis of energy consumption across different architectures is presented in Section V. Finally, Section VI concludes our work, summarizing our key findings and contributions.

II. OBJECTIVES

The objectives of estimating the energy consumption of Cloud, Fog, and Edge Computing Infrastructures are multifaceted and encompass various aspects related to efficiency, sustainability, and informed decision-making. Some key objectives include:

1) *Resource Optimization:* By estimating energy consumption, one can identify areas of inefficiency and optimize resource utilization within Cloud, Fog, and Edge Computing infrastructures. This optimization contributes to cost-effectiveness and reduces the overall environmental impact.
2) *Environmental Impact Assessment:* Understanding the energy footprint of these infrastructures allows for a comprehensive assessment of their environmental impact. This information is crucial for adopting environmentally sustainable practices and meeting regulatory requirements.
3) *Infrastructure Design and Planning:* Accurate energy consumption estimates aid in the design and planning of Cloud, Fog, and Edge Computing infrastructures. This involves making informed decisions about hardware, cooling systems, and network architecture to enhance overall energy efficiency.
4) *Performance Evaluation:* Energy consumption is often correlated with system performance. Estimating energy usage enables the evaluation of performance trade-offs, helping to strike a balance between optimal system operation and energy efficiency.
5) *Cost Analysis:* Energy costs are a significant component of operating Cloud, Fog, and Edge Computing infrastructures. Estimating energy consumption allows for more accurate cost analysis, aiding in budgeting, pricing strategies, and overall financial management.
6) *Policy Development:* Energy consumption data can inform the development of policies and guidelines for sustainable computing practices. This is especially important as energy efficiency becomes a focal point in global efforts to mitigate climate change.
7) *Technology Innovation:* Insights into energy consumption can drive innovation in technology development. This includes the creation of more energy-efficient hardware, cooling solutions, and network technologies for Cloud, Fog, and Edge Computing.
8) *Comparative Analysis:* Estimating energy consumption across different architectures facilitates comparative analysis. This helps stakeholders, including researchers, businesses, and policymakers, in making informed decisions about the most suitable architectural choices based on energy efficiency considerations.
9) *Predictive Modeling:* Energy consumption estimates can be used to develop predictive models, enabling proactive measures to be taken to address potential energy-related issues before they occur.
10) *Societal and Economic Impact:* Understanding the energy implications of Cloud, Fog, and Edge Computing infrastructures contributes to broader societal and economic discussions. This knowledge is crucial for fostering responsible and sustainable development in the digital age.

III. LIMITATIONS

Estimating the energy consumption of Cloud, Fog, and Edge Computing Infrastructures faces several inherent limitations. The complexity and dynamic nature of these infrastructures, characterized by diverse hardware and software components, pose challenges in creating accurate and comprehensive models. Workload variability, rapid technological advancements, and the lack of standardized metrics hinder precise estimations. Additionally, the heterogeneity of data center locations and environmental conditions, coupled with privacy and security concerns, further complicate the assessment. Limited transparency into infrastructure details, especially in public Cloud environments, restricts the depth of analysis. The absence of real-world validation and vendor-specific challenges, where different providers may have proprietary architectures and reporting practices, contribute to the overall complexity and uncertainty in estimating the energy consumption of these evolving and multifaceted computing environments. Addressing these limitations necessitates concerted efforts in research, standardization, and collaboration among stakeholders.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. LITERATURE SURVEY

The literature survey [1] conducted by Shehabi et al. (2016) in the "United States Data Center Energy Usage Report," published by Lawrence Berkeley National Laboratory, offers a comprehensive examination of energy usage in data centers within the United States. The study investigates the critical issue of energy consumption in data centers, providing valuable insights into the environmental impact of these facilities. Through extensive data analysis, the authors reveal that data centers consumed an estimated 70 billion kWh in 2014, accounting for approximately 1.8% of the total U.S. electricity consumption. The report not only quantifies the energy usage but also projects a further increase to around 73 billion kWh by 2020. The survey emphasizes the significance of understanding and addressing the energy challenges associated with data centers, making it a pivotal resource for policymakers, researchers, and industry professionals engaged in enhancing the sustainability and efficiency of data center operations.

The literature survey [2] conducted by Mahmud et al. (2018) in their work "Fog Computing: A Taxonomy, Survey and Future Directions," published in the book "Internet of Everything, Internet of Things," provides a comprehensive examination of Fog Computing. The survey encompasses a taxonomy that systematically categorizes the key elements of Fog Computing, offering a structured understanding of this paradigm. The authors conduct a thorough survey of existing literature, summarizing the state-of-the-art in Fog Computing, its applications, and challenges. Furthermore, the paper outlines future research directions, guiding the trajectory of this evolving field. This survey serves as a valuable resource for researchers, practitioners, and policymakers interested in gaining a holistic perspective on Fog Computing, fostering its development, and addressing the emerging challenges in the Internet of Things landscape.

In [3] their comprehensive literature survey published in the IEEE Communications Surveys and Tutorials in 2018, Mouradian et al. present a thorough examination of Fog Computing, encompassing the state-of-the-art developments and highlighting prevalent research challenges. The survey meticulously reviews existing literature, offering a detailed overview of Fog Computing's architecture, applications, and deployment scenarios. By synthesizing a wealth of information, the authors contribute to a deeper understanding of the current landscape, emphasizing the diverse applications and potential benefits of Fog Computing. Additionally, the survey addresses critical research challenges, providing insights that are instrumental for researchers, practitioners, and industry professionals seeking to navigate and contribute to the evolving field of Fog Computing.

[4] The ETSI white paper titled "Mobile-Edge Computing - Introductory Technical White Paper," published in 2014, serves as a foundational literature survey providing a comprehensive overview of Mobile-Edge Computing (MEC). In this technical document, ETSI explores the fundamental concepts, architecture, and key principles underlying MEC. The white paper delves into the potential applications and benefits of pushing computing resources closer to the network edge, enhancing the efficiency and responsiveness of mobile networks. By offering insights into the technical aspects of MEC, this publication acts as a valuable resource for researchers, industry practitioners, and stakeholders interested in understanding the foundational elements of Mobile-Edge Computing and its implications for the evolution of mobile networks.

[5] The literature survey conducted by Taleb et al. in their article "On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration," published in IEEE Communications Surveys & Tutorials in 2017, offers a comprehensive overview of Multi-Access Edge Computing (MEC) within the context of the emerging 5G network architecture. The authors systematically review the key components, architectural principles, and orchestration mechanisms associated with MEC, providing a thorough understanding of the evolving landscape. The survey emphasizes the integration of edge cloud capabilities in 5G networks, exploring the potential applications, challenges, and deployment scenarios. With an extensive examination of the state-of-the-art developments, this survey serves as a valuable resource for researchers, industry practitioners, and stakeholders seeking insights into the intricacies of MEC and its pivotal role in the 5G ecosystem. The literature survey [6] conducted by Mao et al. in their article titled "A Survey on Mobile Edge Computing: The Communication Perspective," published in IEEE Communications Surveys & Tutorials in 2017, provides a comprehensive examination of Mobile Edge Computing (MEC) with a focus on communication aspects. The authors systematically review the key communication challenges, architectural components, and enabling technologies associated with MEC. The survey delves into the communication protocols, resource allocation schemes, and networking paradigms relevant to MEC, offering valuable insights into the communication-centric considerations of this emerging paradigm. By addressing various aspects, including latency reduction, energy efficiency, and spectral efficiency, the survey contributes to a holistic understanding of MEC's role in enhancing communication performance. This work serves as a pivotal resource for researchers, practitioners, and stakeholders interested in comprehending the communication dynamics of MEC and its implications in the broader context of mobile computing. The literature survey [7] conducted by Shi et al. in their article "Edge Computing: Vision and Challenges," published in the IEEE Internet of Things Journal in 2016, presents a comprehensive exploration of the vision and challenges associated with Edge Computing.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The authors systematically survey the key concepts, fundamental principles, and potential applications of Edge Computing, providing valuable insights into the evolving landscape of distributed computing. Focusing on the Internet of Things (IoT) context, the survey addresses the unique challenges posed by Edge Computing, such as latency reduction, bandwidth optimization, and enhanced privacy. The authors contribute to a deeper understanding of the visionary aspects of Edge Computing while delineating the inherent challenges that must be addressed for successful implementation. This work stands as an essential resource for researchers, practitioners, and stakeholders seeking a comprehensive overview of Edge Computing and its pivotal role in shaping the future of IoT and distributed computing systems. The literature survey [8] conducted by Bonomi et al. in their paper "Fog Computing and its Role in the Internet of Things," presented at the Mobile Cloud Computing workshop in 2012, offers an early exploration of Fog Computing and its significance in the context of the Internet of Things (IoT). The authors systematically examine the key principles, architectural components, and potential applications of Fog Computing, emphasizing its role in addressing the challenges posed by IoT deployments. This survey lays the groundwork for understanding how Fog Computing, situated at the edge of the network, can enhance the efficiency, responsiveness, and scalability of IoT systems by offloading computation and storage tasks from centralized cloud servers. The work provides valuable insights into the conceptualization of Fog Computing and its potential impact on the evolving landscape of IoT, serving as a foundational resource for subsequent research and developments in this burgeoning field. The literature survey [9] conducted by Hu et al. in their paper "Survey on Fog Computing: Architecture, Key Technologies, Applications, and Open Issues," published in the Journal of Network and Computer Applications in 2017, offers a comprehensive exploration of Fog Computing. The authors meticulously survey the architectural frameworks, essential technologies, applications, and unresolved challenges associated with Fog Computing. The survey spans various aspects, including communication protocols, resource management, security considerations, and the diverse applications of Fog Computing in different domains. By providing a holistic overview, the authors contribute to a deeper understanding of the state-of-the-art developments in Fog Computing and highlight key research directions. This survey stands as a pivotal resource for researchers, practitioners, and stakeholders seeking an in-depth understanding of the architecture and dynamics of Fog Computing, guiding further advancements in this emerging paradigm.

The literature survey [10] by Ahmed and Saumya Sabyasachi in their paper "Cloud Computing Simulators: A Detailed Survey and Future Direction," presented at the IEEE International Advance Computing Conference in 2014, provides an extensive examination of cloud computing simulators. The authors systematically review the landscape of cloud computing simulation tools, offering insights into their capabilities, features, and applications. The survey encompasses various simulators, including CloudSim, GreenCloud, and CloudAnalyst, among others, and critically evaluates their strengths and limitations. Additionally, the paper identifies emerging trends and outlines future directions for the development of cloud computing simulators. This survey serves as a valuable resource for researchers, practitioners, and educators seeking a comprehensive understanding of the simulation tools available for studying and analyzing the complexities of cloud computing environments, guiding the direction of future research in this domain.

V. RELATED WORK

A. Virtualized Architectures

Cloud computing has undergone a shift towards more distributed architectures in response to the increasing demands for mobility, high data volume processing, and low latency from users. In 2014, the European Telecommunications Standards Institute (ETSI) introduced a novel Cloud architecture termed Mobile Edge Computing (MEC), initially defined as providing IT and cloud-computing capabilities within the Radio Access Network (RAN) in close proximity to mobile subscribers. The MEC group within ETSI evolved in 2016, expanding its reach to heterogeneous networks and being renamed Multi-access Edge Computing (MEC). These infrastructures can be managed hierarchically or through distributed control based on specific applications and may incorporate mobile devices to offer computation for smaller devices. In parallel, driven by the surge in Internet of Things (IoT), Cisco introduced the concept of Fog Computing in 2012. Fog Computing, as defined by the Open Fog Consortium, constitutes a system-level horizontal architecture that distributes computing, storage, control, and networking resources and services across the continuum from Cloud to Things. Specifically designed to handle time-sensitive and voluminous data generated at the network edge by IoT devices, Fog Computing extends cloud capabilities to devices proximate to the Things, providing computing, storage, and network connectivity. This architecture exhibits a hierarchical structure comprising IoT devices, Fog nodes, and Cloud data centers, with fog nodes managed either in a decentralized or distributed manner depending on the specific applications. Clarifying distinctions between these models, particularly from an industrial perspective, remains a challenge, prompting initiatives like the Open Glossary of Edge Computing to provide precise definitions in this domain.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Energy Models and Simulation Tools

Recent studies [12], [13] have identified over 20 cloud and data center (DC) simulators proposed in the last decade, with CloudSim [14] standing out as one of the most popular choices. While CloudSim boasts advantages such as modeling large-size DCs, supporting federation models, and accommodating a large number of virtual machines, it has limitations, including restricted communication and network energy models. Various efforts, such as those by Li et al. [17] with DartCSIM, aimed at enhancing CloudSim by simulating energy and network concurrently. Alternative cloud simulators like iCanCloud [18] exist but lack support for distributed DC architectures. CloudNetSim++ [19] by Malik et al. offers a GUI-based framework for modeling DCs in OMNeT++ but neglects the impact of packet length on energy consumption. The SimGrid toolbox accurately models virtual machines and server energy consumption but lacks consideration for network component energy. Despite these existing tools, a straightforward model or simulator accommodating diverse cloud-related architectures, from fully centralized to completely distributed, without delving into intricate programming details, remains absent. Addressing this gap, we introduce a generic energy model for comparing different cloud-related architectures. Our work represents the first attempt to provide a model enabling researchers to analyze and compare the energy consumption of various cloud-related architectures comprehensively. Existing studies by Jalali et al. [22], Li et al. [23], and C. Fiandrino et al. [24] comparing energy consumption in centralized data centers, fog computing, and IoT applications, respectively, lack accuracy in their proposed models for assessing the energy footprint of diverse cloud-related architectures. These models often overlook critical factors, such as the size and intra-network topology of centralized data centers, limiting their ability to analyze intra-DC network components accurately. Our generic energy model aims to overcome these limitations and provide a comprehensive framework for assessing and comparing energy consumption in different cloud-related architectures.

VI. PROPOSED TAXONOMY

A. Cloud-related Architectures

A Cloud-related architecture encompasses multiple data centers (DCs) and a network connecting end users through telecommunication networks. End users function as request senders and service receivers, while DCs provide diverse cloud services and consist of Compute Nodes (CNs) hosting Virtual Machines (VMs), along with Service Nodes (SNs) serving as infrastructure management or storage nodes. The primary distinctions among Cloud-related architectures revolve around the location, number, size, and roles (e.g., controller, storage, compute) of their DCs. The proposed taxonomy, depicted in Figure 1, categorizes architectures into four main groups based on DC characteristics: Fully Centralized (FC), Partly Distributed (PD), Fully Distributed-Centralized Controller (FD-CC), and Fully Distributed (FD). In the Fully Centralized (FC) architecture, all end users connect to a central large-size DC to receive services. The Partly Distributed (PD) architecture involves several DCs distributed across different locations, collaborating to serve end users, with a telecommunication network linking the Cloud DCs and connecting them to end users. PD architectures position DCs closer to end users compared to FC architectures, due to a higher number of DCs, although these DCs are not situated on end users' premises.

1) Fully Distributed with Centralized Controller (FD-CC)

Fully Distributed with Centralized Controller (FD-CC) architecture, multiple data centers (DCs) are situated on end users' premises, typically catering to one or a few end users each. However, controllers (designated as Service Nodes or SNs) are strategically positioned in one or more DCs within the network core. The DCs on end users' premises are generally sized equivalent to a Physical Machine (PM). Importantly, the remaining DCs located in the core of the network solely serve management functions and are not engaged in computing or storage services. This architectural model is characterized by a distributed deployment of computing resources at end users' locations, while centralizing control functions in dedicated DCs at the network core for efficient management.

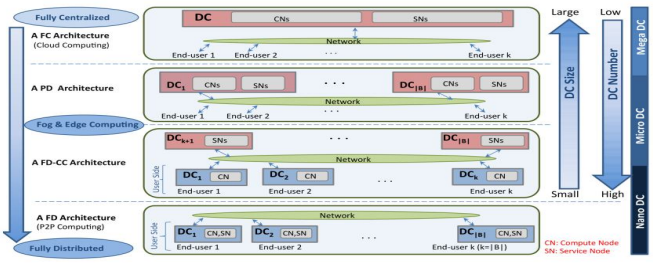


Fig. 1. Considered architectures; ranging from fully centralized (FC) to fully distributed (FD).

2) Fully Distributed (FD)

In an FD architecture, unlike the FD-CC model, the management system is also distributed across end users' locations, eliminating the presence of a data center (DC) in the network core responsible for resource control—meaning no independent Service Node (SN) is designated for this purpose. The employed abbreviations are summarized in Table I. The FC architecture represents centralized Cloud infrastructures, while examples of PD architectures are evident in decentralized Fog infrastructures [10]. The FD-CC architecture encompasses hierarchical Fog and Edge infrastructures with distributed nodes and decentralized management nodes, potentially hosting Virtual Network Functions' (VNFs) orchestration [5], [10]. Lastly, the FD architecture characterizes fully distributed edge infrastructures [6]. It's important to note that, in the context of this work, we do not consider the wide variety of wireless network technologies [26], and incorporating architectures with mobile fog nodes is acknowledged as potential future work.

B. Architecture's Components

Figure 2 provides a comprehensive overview of the Information and Communication Technology (ICT) equipment for all Cloud-related architectures. Each architecture consists of two main components: Data Center(s) (DCs) and the telecommunication network, which are further detailed below. Additionally, DCs are categorized based on their size: mega (with more than 10,000 hosted Physical Machines or PMs), micro (with less than 500 PMs), and nano (comprising only a few PMs). In the extreme case of a nano DC, it consists of a single PM without any additional ICT equipment like intra-DC network elements (e.g., switches and links). It's important to note that all DCs, except for nano ones, also feature non-ICT equipment such as independent cooling systems. Reflecting the energy consumption of this non-ICT equipment in DCs (ranging from micro to mega), the Power Usage Effectiveness (PUE) serves as a well-known indicator of DC energy efficiency. PUE represents the ratio of the total facility energy consumption to the IT equipment energy consumption [27]. In essence, the overall energy consumption of a DC can be estimated by multiplying the energy consumption of its ICT equipment by its PUE value.

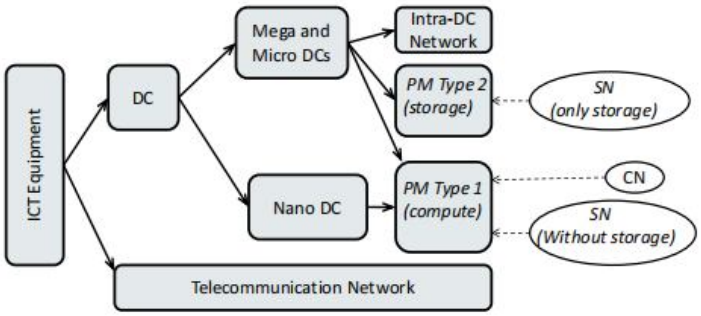


Fig. 2. ICT equipment composing the different architectures

As depicted in Figure 2, Physical Machines (PMs) can be categorized into two primary types: Type 1 PMs are specialized for computing purposes (e.g., Computing and Cloud management nodes), while Type 2 PMs are designated for storing large datasets due to their specific storage backends (i.e., storage nodes). Each type of PM exhibits a distinct energy profile owing to the additional storage devices. In data center (DC) terminology, Compute Nodes (CNs) correspond to Type 1, while Service Nodes (SN) are Type 1 if they manage the infrastructure or Type 2 if employed for storage purposes. In the extreme scenario of an FD architecture with a single CN-based PM, SN functions are fulfilled by a Virtual Machine (VM) directly hosted on the CN, utilizing local disks as the default storage backend.

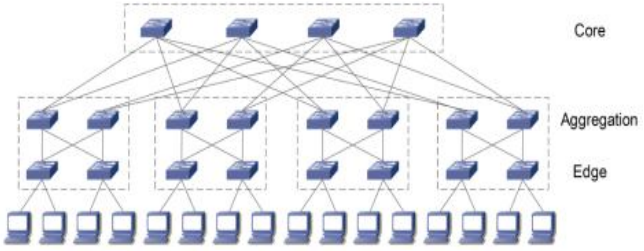


Fig. 3. 4-ary fat tree topology from [30]

Beyond PMs, the ICT equipment of a mega or micro DC encompasses network elements such as routers, switches, and links to interconnect PMs. The number and type of network elements required to link PMs within a DC depend on its network topology. Various network topologies have been proposed for DCs since 2008, with more than thirty options available. In this study, we adopt a fat tree switch-centric architecture, specifically the n-ary fat tree topology, reflecting the network topology of numerous contemporary DCs. The fat tree topology offers advantages such as the utilization of small-size and identical switches, facilitating cost-effectiveness and simplifying system models and analyses compared to architectures with non-identical switches. This flexible and scalable topology can adapt to different DC sizes, and in this context, we consider an n-ary fat tree topology with three layers.

The telecommunication network serves to interconnect all Data Centers (DCs) and end users, typically comprising four main logical layers [31]: core, backbone, metro, and feeders (or end users). The metro layer facilitates end users' access to the network, the backbone layer ensures policy-based connectivity, and the core layer provides high-speed transport to meet the requirements of the backbone layer devices. In the Fully Centralized (FC) architecture, we assume the single DC is directly connected to the core layer of the telecommunication network. For the Partly Distributed (PD) architecture, DCs are connected to either backbone or metro routers. In the Fully Distributed with Centralized Controller (FD-CC) architecture, DCs with Compute Nodes (CNs) are linked to either backbone or metro routers, while DCs with Service Nodes (SNs) connect to the core layer. In the Fully Distributed (FD) architecture, DCs are connected to the feeder layer. It is assumed that the routing policy consistently employs the shortest path.

VII. THE ENERGY MODEL

A. General Overview

As illustrated in Figure 4, our model breaks down the energy consumption of Information and Communication Technology (ICT) equipment into static and dynamic components [32]. The static energy consumption represents the energy used when resources are idle, irrespective of any workload. In contrast, the dynamic component is computed based on the current utilization of Cloud resources by active Virtual Machines (VMs).

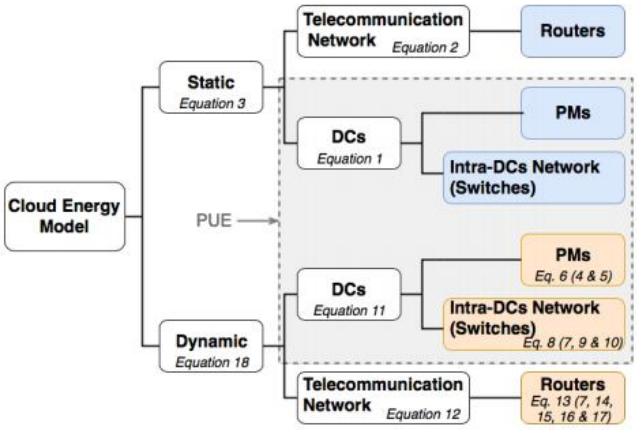


Fig. 4. General view of the proposed energy model and the equations used in the following to express the different parts.

Before delving into the static and dynamic aspects of the energy model, we introduce technical assumptions that streamline the model's complexity without sacrificing generality:

1) Virtual Machines (VMs) are assumed to be homogeneous in size concerning Central Processing Unit (CPU) and Random Access Memory (RAM). While modeling heterogeneous VMs is possible, it is achieved by presuming that a large-sized VM encompasses a set of smaller-sized VMs, serving as the basic unit of resources.
2) Throughout the entire time period T, all V VMs are consistently operational.
3) VMs for each architecture are evenly distributed across all active Compute Node (CN)-based Physical Machines (PMs). This simplifying assumption eases the design and utilization of the model, although it can be relaxed by introducing a scheduling algorithm that specifies the number of VMs on each CN and the traffic exchanged with other VMs.
4) The intra-DC network adheres to a fat tree topology, while the telecommunication network follows a hierarchical topology, as previously discussed.
5) Switch interfaces are considered bi-directional, and the capacities of uplink and downlink are assumed to be similar. In cases where there is no traffic in one direction, the entire interface remains active.

B. Static Energy Part of the Model

As depicted in Figure 4, the static energy component of a Cloud-related architecture encompasses the idle power consumption of Physical Machines (PMs) and switches. For a PM, the idle power consumption refers to the power consumed by the server when powered on but not executing any tasks [32]. Similarly, the energy consumption of the load-independent portion of a switch is regarded as its idle power consumption, which includes the power consumption of elements such as the chassis power supply and fans [33].

The static energy consumption of each device, whether PM or switch, corresponds to its idle power consumption over the specified time period T, as outlined in Table II. Given that idle power consumption remains constant over time, the total energy consumption over the time period T is obtained by multiplying it by T.

For the telecommunication network segment, the static energy consumption pertains to the idle power consumption of routers used to connect the Data Centers (DCs) and end users. The overall static energy consumption is then calculated as the sum of the static energy consumption of telecommunication network devices and the static energy consumption of each DC, including its cooling-related energy consumption (expressed by the Power Usage Effectiveness or PUE). The subsequent discussion outlines the dynamic energy consumption attributed to PMs, intra-DC switches, and telecommunication routers.

VIII. MODEL EXPLOITATION

A. Model Instantiation

To ensure a fair comparison among Cloud architectures, we maintain an identical number of Compute Nodes (CNs) across all scenarios, amounting to 1,000 CNs, each operating for one hour. It is essential to note that in Fully Distributed (FD) and Fully Distributed with Centralized Controller (FD-CC) architectures, CNs play a dual role by also supporting Service Nodes (SN) and/or storage services, rendering them more powerful. Specifically, the Fully Centralized (FC) architecture features one large-sized Data Center (DC) housing the 1,000 CNs. The Partly Distributed (PD) infrastructure consists of eight medium-sized homogeneous DCs, each accommodating 125 CNs.

In the FD-CC architecture, 1,000 nano DCs, each with a single CN, coexist with eight central small-sized controller DCs incorporating the SNs. Similarly, the FD architecture mirrors FD-CC, comprising 1,000 nano DCs, each hosting one CN. The count of SN-based Physical Machines (PMs) is determined by the number of CNs and the unique characteristics of each architecture. For mega/micro DCs in FC or PD architectures, we assume three global SNs (ensuring high availability) and an additional SN for every 100 CNs.

In the context of FD-CC, we consider three SNs within each of the eight small-sized DCs located in the core of the network. As for the FD architecture, a single multi-purpose PM is present per nano DC.

To derive realistic energy profiles for Physical Machines (PMs), experiments were conducted on the French experimental testbed Grid'5000, utilizing Taurus servers (12 cores, 32 GB memory, and 598 GB/SCSI storage) for FC and PD cloud PMs and Parasilo servers (16 cores, 128 GB memory, and 5*600 GB HDD + 200 GB SSD/SAS storage) for FD and FD-CC cloud PMs. The stress benchmark was employed to obtain consumption values per loaded core (Px,j cores). Power measurements based on external wattmeters on each Taurus server were obtained from the Grid'5000 testbed, providing one value per second with a 0.125 Watts accuracy.

For the FC architecture with a large-sized DC (1,000 CNs and 13 SNs), a 16-ary Fat Tree switch topology (320 switches) was considered. In PD architecture, with a DC comprising 125 CNs and 3 SNs, an 8-ary Fat Tree topology was used. In the FD-CC category, a central control DC with 3 SN-based PMs required one switch per DC. The FD architecture, being fully distributed, did not necessitate an intra-DC network. The telecommunication network topology, inspired by real ISP configurations, included 8 core routers, 52 backbone routers, 52 metro routers for aggregation, and 260 feeder routers in the access layer (372 routers connected with 718 links).

For VMs, it was assumed that each VM reserved one CPU core, utilizing an average of 15% of its reserved CPU. The energy profile of routers was based on measurements, with idle power consumption values of 11,000, 4,000, 2,000, and 2,000 Watts for Core, Backbone, Metro, and Feeder routers respectively. The switches inside DCs, following a 16-ary fat tree topology, consumed approximately 280 Watts. The traffic distribution was assumed with 70% remaining inside a mega or micro DC. Constant values for energy-related parameters were assumed, such as Er_pkt = 1,300 nJ/pkt for routers and switches, and Er_SF = 14 nJ/Byte for all network elements. A Power Usage Effectiveness (PUE) value of 1.2 was assigned to all mega and micro DCs in FC and PD architectures. All PMs and network elements were assumed to be continuously active to eliminate the influence of varying energy management techniques on cloud energy consumption evaluation.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Simulation Results

The In a general scenario where 10,000 VMs are requested by a set of end users and are continuously running for the considered time period (1 hour), hosting an average of 10 VMs per CN, the proposed model is used to compute the energy consumption of each architecture. The variables include the average traffic generated by one PM (λ: 250, 500, and 1000 Mbps) and the average packet length (L: 100, 576, and 1000 bytes). The packet length varies for a given amount of traffic, influencing the number of packets rather than the total data volume.

Figure 8 illustrates the total energy consumption of the four Cloud-related architectures. In all cases, the FD architecture outperforms the others in terms of energy consumption. Specifically, for the highest traffic rate (1000 Mbps), it consumes between 14% and 25% less energy than FC and PD architectures, respectively. FD-CC demonstrates performance close to FD with the considered traffic rates (up to 1000 Mbps per PM).

The breakdown of static energy consumption reveals insights into architecture-specific patterns. The FD architecture generally exhibits lower total static energy consumption than FC and PD architectures. Due to the variety in PM sizes and energy consumption, the total static energy consumption of CNs in FD and FD-CC architectures is slightly higher than in FC and PD architectures. The static energy consumption of SNs is influenced by their higher number in PD and FD-CC clouds, resulting in higher energy consumption than the FC architecture in this component. However, the total static energy consumption of SNs in the FD-CC architecture is less than in the FC Cloud due to a lower PUE. In the intra-DC network, the PD architecture consumes more static energy due to its increased use of switches. The FD architecture notably saves energy in this component by eliminating the need for any switches. For telecommunication network routers, Core routers consume the most individually, but the overall static energy consumption is dominated by the Feeder routers due to their higher number.

In the static energy consumption part, the number of intra-DC switches plays a crucial role in the energy consumption of a cloud-related architecture. Other factors influencing static energy efficiency include the number of SNs, the energy efficiency of PMs, and the Power Usage Effectiveness (PUE).

Figure 9 illustrates the total dynamic energy consumption of the four architectures. While the packet length has a minimal impact on dynamic energy consumption at low traffic rates (250 Mbps), its effect becomes more noticeable at higher rates (500 and 1000 Mbps). The FD architecture consistently consumes less dynamic energy than others, primarily due to the absence of an intra-DC network. As FD-CC and FD architectures employ more energy-consuming CNs, their dynamic energy consumption due to PMs is higher than the other architectures. Additionally, the FD-CC architecture utilizes the highest number of SN nodes, resulting in higher energy consumption than the others. The dynamic energy consumption of switches is null for the FD infrastructure. FD-CC cloud consumes less energy than FC and PD architectures since FD-CC only requires a few switches to connect its SNs. While there is generally a negligible difference between the results of FC and PD architectures, PD slightly outperforms FC when increasing traffic due to longer paths between PMs in larger-sized data centers. Regarding the dynamic energy consumption of routers, the location of DCs and the impact of traffic on energy consumption are directly related. Although FD exhibits slightly lower dynamic energy consumption for routers than FD-CC, a more significant difference is expected for applications with heavy traffic exchange between controllers (SNs) and CNs.

IX. DISCUSSION

Numerous In summary, several key findings emerge from the study:

1) *Telecommunication Network:* The energy consumption of the telecommunication network, connecting DCs and end users, significantly dominates the overall consumption. This aligns with existing literature emphasizing that telecommunication networks constitute a substantial portion of ICT consumption, reaching 37% in 2014, including end-user devices. Future distributed Cloud architectures could potentially mitigate this by keeping traffic more localized, reducing the reliance on network routers.

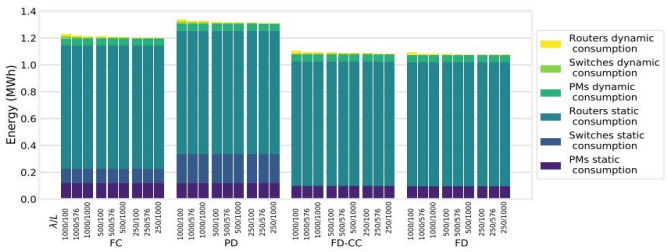


Fig. 8. Total energy consumption of different architectures for 1 hour

2) *Packet Length:* The study reveals that in scenarios with heavy traffic rates, the packet length has a notable impact on the dynamic energy consumption of all architectures. Understanding and optimizing packet length can be crucial for enhancing energy efficiency.
3) *Power Usage Effectiveness (PUE):* PUE plays a crucial role in the energy consumption of architectures with medium and large-sized DCs. High PUE values can offset gains achieved through improvements in the energy efficiency of ICT devices. Therefore, efforts to minimize PUE are essential for enhancing overall energy efficiency.

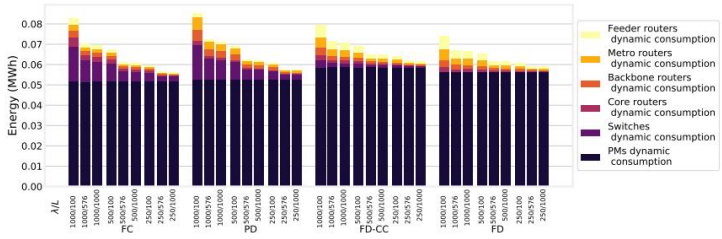


Fig. 9. Dynamic energy consumption of different architectures for 1 hour

4) *Energy Consumption of Intra-DC Network:* In FC and PD architectures, intra-DC switches can consume over 50% of total DC energy and up to 20% of the total infrastructure energy. Careful selection of network architecture and switch size emerges as a critical factor in significantly improving energy efficiency within an infrastructure.
5) *Overall Energy Consumption:* The FD architecture, devoid of intra-DC switches and DCs, results in a PUE of 1 (eliminating the need for air conditioning equipment). This makes FD and even FD-CC architectures more energy-efficient compared to others. It's important to note that the study assumes similar performance for CNs across all architectures, and future Fog and edge computing infrastructures may introduce heterogeneity in end-user devices, impacting these findings.

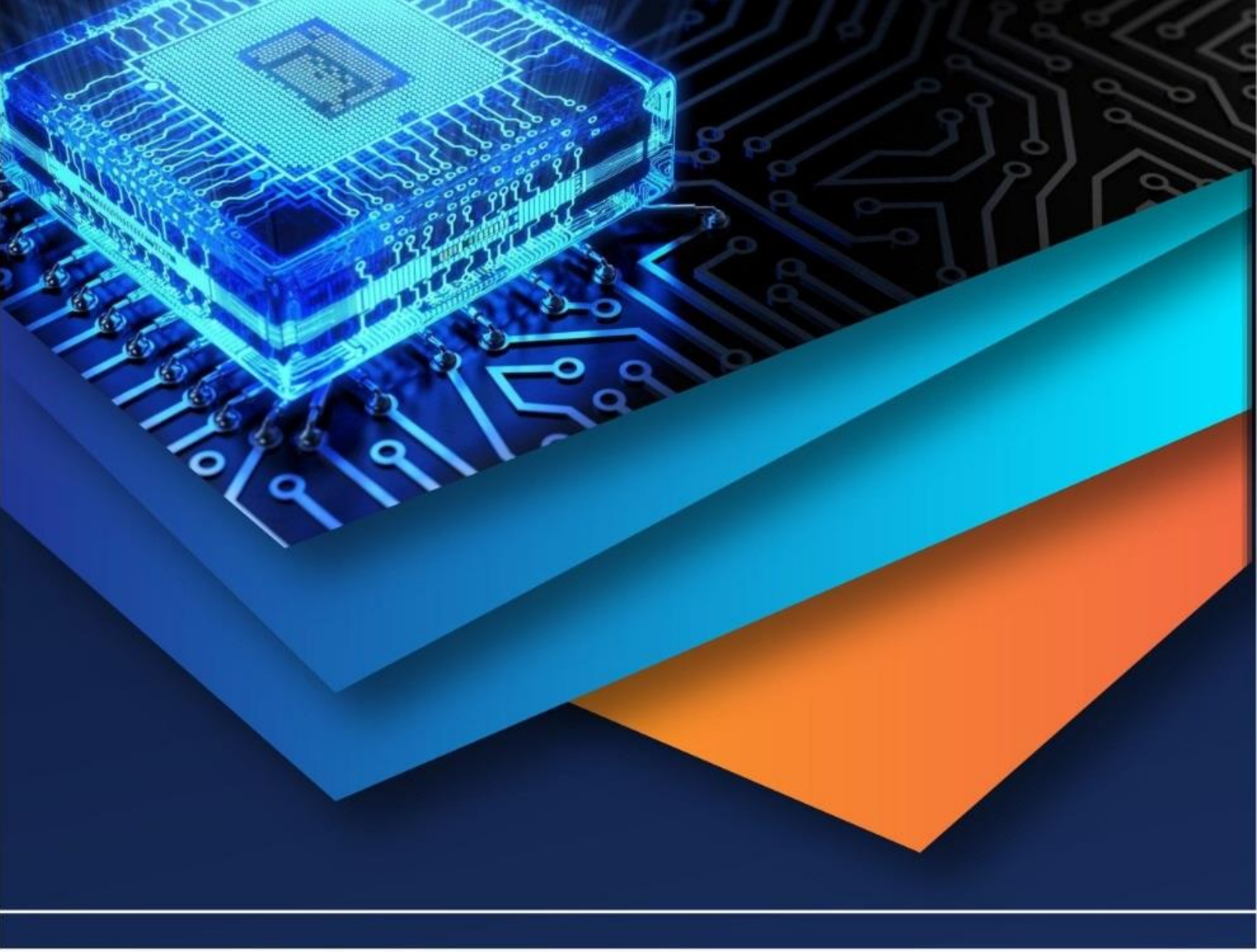
X. CONCLUSION

By examining a range of Cloud-related architectures, from fully centralized to completely distributed, this paper addresses the crucial question of energy efficiency, a key concern for Cloud researchers and providers. The proposed taxonomy provides a structured classification of these architectures, elucidating their main topological distinctions. Building upon this taxonomy, the paper introduces an energy model that is comprehensive, easy-to-use, and scalable. This model effectively highlights architectural distinctions in terms of energy consumption. The analysis of four distinct Cloud-related architectures (FC, PD, FD-CC, and FD) using this energy model reveals that a fully distributed architecture (FD) consumes between 14% and 25% less energy compared to fully centralized (FC) and partly distributed (PD) architectures, respectively. The incorporation of the Power Usage Effectiveness (PUE) metric within the model captures the influence of cooling systems, an aspect currently of non-negligible importance in energy consumption. The findings underscore the significance of focusing on improving telecommunication network utilization, enhancing the efficiency of computing infrastructures (as indicated by PUE), and leveraging heterogeneous infrastructures to align more closely with user needs in future efforts to green emerging cloud architectures. While the model addresses many crucial technical considerations, potential improvements for future work include accounting for uneven distribution of end-users in real-world scenarios and extending support for other technologies such as containers and Virtual Network Functions (VNFs), as well as heterogeneous telecommunication networks and architectures.

REFERENCES

[1] Shehabi, S.J. Smith, N. Horner, I. Azevedo, R. Brown, J. Koomey, E. Masanet, D. Sartor, M. Herrlin, W. Lintner, “United States Data Center Energy Usage Report”, Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-1005775, 2016.
[2] R. Mahmud, R. Kotagiri, R. Buyya, “Fog Computing: A Taxonomy, Survey and Future Directions”, In B. Di Martino, KC. Li, L. Yang, A. Esposito (eds) Internet of Everything, Internet of Things (Technology, Communications and Computing). Springer, 2018
[3] C. Mouradian, D. Naboulsi, S. Yangui, R. Glitho, M. J. Morrow, P. A.Polakos, “A Comprehensive Survey on Fog Computing: State-of-the-artand Research Challenges”, IEEE Communications Surveys and Tutorials Vol. 20(1), 2018.
[4] ETSI, “Mobile-Edge Computing - Introductory Technical White Paper”, ETSI White paper, 2014.
[5] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta and D. Sabella, “On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration”, in IEEE Communications Surveys & Tutorials, 19(3), pp. 1657-1681, 2017.
[6] Y. Mao, C. You, J. Zhang, K. Huang and K. B. Letaief, “A Survey on Mobile Edge Computing: The Communication Perspective”, in IEEE Communications Surveys & Tutorials, 19(4), pp. 2322-2358, 2017.
[7] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, “Edge Computing: Vision and Challenges,” in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, 2016.
[8] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, “Fog computing and its role in the internet of things”, MCC workshop on Mobile cloud computing, pages 13-16, 2012.

[9] The OpenFog Consortium, https://www.openfogconsortium.org, accessed December 2018.
[10] P. Hu, S. Dhelim, H. Ning and T. Qiu, "Survey on fog computing: architecture, key technologies, applications and open issues", Journal of Network and Computer Applications, Volume 98, pages 27-42, 2017.
[11] The Linux Foundation, "Open Glossary of Edge Computing", https: //github.com/State-of-the-Edge/glossary, 2018.
[12] A. Ahmed and A. Saumya Sabyasachi, "Cloud Computing Simulators: A Detailed Survey and Future Direction", IEEE International Advance Computing Conference, 2014.
[13] F. Fakhfakh, H. Hadj Kacem, and A. Hadj Kacem, "Simulation Tools for Cloud Computing: A Survey and Comparative Study", IEEE International Conference on Computer and Information Science, 2017.
[14] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Al gorithms", Software: Practice and Experience, vol. 41, pp. 23-50, 2011.
[15] M. Bux and U. Leser, "DynamicCloudSim: Simulating heterogeneity in computational clouds", Future Generation Computer Systems, vol. 46, pp. 85-99, 2015.
[16] A. Zhou, S. Wang, C. Yang, L. Sun, Q. Sun, and F. Yang, "FTCloudSim: support for cloud service reliability enhancement simulation", Journal of Web and Grid Services, vol. 11, no. 4, pp. 347-361, 2015.
[17] X. Li, X. Jiang, K. Ye and P. Huang, "DartCSim+: Enhanced CloudSim with the Power and Network Models Integrated", IEEE International Conference on Cloud Computing (CLOUD), 2013.
[18] A. Nez, J. Vzquez-Poletti, A. Caminero, G. Casta, J. Carretero, and I. Llorente, "iCanCloud: A Flexible and Scalable Cloud Infrastructure Simulator", Journal of Grid Computing, vol. 10, no. 1, pp. 185-209, 2012.
[19] A. W. Malik, K. Bilal, S. U. Malik, Z. Anwar, K. Aziz, D. Kliazovich, N. Ghani, S. U. Khan, R. Buyya, "CloudNetSim++: A GUI Based Framework for Modeling and Simulation of Data Centers in OMNeT++", IEEE Transactions on Services Computing, vol. 10, pp. 506–519, 2017.
[20] T. Hirofuchi, A. Lebre and L. Pouilloux, "SimGrid VM: Virtual Machine Support for a Simulation Framework of Distributed Systems", IEEE Transactions on Cloud Computing, pp. 1-14, 2018.
[21] C. Heinrich, T. Cornebize, A. Degomme, A. Legrand, A. Carpen Amarie, S. Hunold, A.-C. Orgerie M. and Quinson, "Predicting the Energy Consumption of MPI Applications at Scale Using a Single Node", IEEE Cluster, pp. 92-102, 2017.
[22] F. Jalali, K. Hinton, R. Ayre, T. Alpcan, and R. S. Tucker, "Fog Computing May Help to Save Energy in Cloud Computing", IEEE J. on Selected Areas in Communications, 34(5), pp. 1728-1739, 2016.
[23] Y. Li, A.-C. Orgerie, I. Rodero, B. L. Amersho, M. Parashar and J.-M. Menaud, "End-to-end Energy Models for Edge Cloud-based IoT Plat forms: Application to Data Stream Analysis in IoT", Future Generation Computer Systems, pp.1-12, 2017.
[24] C. Fiandrino, D. Kliazovich, P. Bouvry, and A. Y. Zomaya, "Performance and Energy Efficiency Metrics for Communication Systems of Cloud Computing Data Centers", IEEE Transactions on Cloud Computing, vol. 5, no. 4, 2017.
[25] T. Baker, M. Asim, H. Tawfik, B. Aldawsari and R. Buyya, "An energy-aware service composition algorithm for multiple cloud-based IoT applications", Journal of Network and Computer Applications, Volume 89, pages, 96-108, 2017.
[26] Y. Li, A.-C. Orgerie, I. Rodero, B. Lemma Amersho, M. Parashar, and J.-M. Menaud, "End-to-end energy models for Edge Cloud-based IoT platforms: Application to data stream analysis in IoT", Future Generation Computer Systems, vol. 87, pages 667-678, 2018.
[27] ISO/IEC 30134-2:2016, "Information technology – Data centres – Key performance indicators – Part 2: Power usage effectiveness (PUE)", 2016.
[28] T. Chen, X. Gao and G. Chen, "The features, hardware, and architectures of data center networks: A survey", Journal of Parallel and Distributed Computing, Elsevier, Volume 96, pp. 45–74, 2016.
[29] Cisco, "Cisco's Massively Scalable Data Center: Network Fabric for Warehouse Scale Computer", Cisco Systems, Inc., 2010.
[30] T. Wang, Z. Su, Y. Xia and M. Hamdi, "Rethinking the Data Center Networking: Architecture, Network Protocols, and Resource Sharing", IEEE Access, vol. 2, pp. 1481-1496, 2014.
[31] L. Chiaraviglio, M. Mellia and F. Neri, "Energy-aware Backbone Net works: a Case Study", International Workshop on Green Communications (GreenComm), 2009.
[32] M. Kurpicz, A.-C. Orgerie and A. Sobe, "How much does a VM cost? Energy-proportional Accounting in VM-based Environments" Europicro PDP, pp. 651-658, 2016.
[33] A. Vishwanath, K. Hinton, R.W.A. Ayre and R.S. Tucker, "Modeling Energy Consumption in high-capacity routers and switches", IEEE Jour nal on selected areas in communication, 32(8), pages 1524–1532, 2014.
[34] Y. Li, A.-C. Orgerie and J.-M. Menaud, "Opportunistic Scheduling in Clouds Partially Powered by Green Energy", GreenCom: IEEE Int. Conf. on Green Computing and Communications, pp. 448-455, 2015.
[35] OpenStack Configuration Reference, [online] https://docs.openstack.org/ mitaka/config-reference/compute/scheduler.html, accessed May 2018.
[36] V. Sivaraman, A. Vishwanath, Z. Zhao and C. Russell, "Profiling per packet and per-byte energy consumption in the NetFPGA Gigabit router", INFOCOM Workshops, pages 331-336, 2011.
[37] A. C. Orgerie, B. L. Amersho, T. Haudebourg, M. Quinson, M. Rifai, D. L. Pacheco and L. Lef`evre, "Simulation toolbox for studying energy consumption in wired networks", International Conference on Network and Service Management, 2017.
[38] F. Yao, J. Wu, G. Venkataramani and S. Subramaniam, "A compara tive analysis of data center network architectures", IEEE International Conference on Communications (ICC), 2014, pp. 3106-3111.
[39] D. Balouek, et al.,"Adding Virtualization Capabilities to the Grid'5000 Testbed", In I. Ivanov, M. Sinderen, F. Leymann, and T. Shan, editors, Cloud Computing and Services Science, (367) of Communications in Computer and Information Science, pp. 3-20. Springer, 2013.
[40] F. Jalali, "Energy Consumption of Cloud Computing and Fog Computing Applications", PhD thesis, University of Melbourne, Australia, 2015.
[41] Cisco's power consumption calculator. [Online]. Available: http://tools.cisco.com/cpc
[42] Cisco, "Cisco Global Cloud Index: Forecast and Methodology, 2015– 2020", Cisco publications, 2016.
[43] Google Data Centers Efficiency, accessed in May 2018, [online] https: //www.google.com/about/datacenters/efficiency/internal/
[44] W. Van Heddeghem, S. Lambert, B. Lanoo, D. Colle, M. Pickavet and P. Demeester, "Trends in worldwide ICT electricity consumption from 2007 to 2012", Computer Communications 50, 64-76, 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)