# Integrated Resource Allocation Model for Cloud and Fog Computing: Toward Energy-Efficient Infrastructure as a Service (IaaS)

**Chapter** · July 2019

DOI: 10.1007/978-3-319-69889-2_7



**7 authors**, including:

Mohammed Joda Usman
Bauchi State University
36 PUBLICATIONS 480 CITATIONS



Abdul Samad Ismail
Universiti Teknologi Malaysia
111 PUBLICATIONS 1,278 CITATIONS



Hassan Chizari
University of Gloucestershire
54 PUBLICATIONS 615 CITATIONS



Haruna Chiroma
University of Hafr Al Batin
186 PUBLICATIONS 2,631 CITATIONS

# Integrated Resource Allocation Model for Cloud and Fog Computing: Toward Energy-Efficient Infrastructure as a Service (IaaS)

**Mohammed Joda Usman, Abdul Samad Ismail, Hassan Chizari, Abdulsalam Ya'u Gital, Haruna Chiroma, Mohammed Abdullahi and Ahmed Aliyu**

**Abstract** Cloud is transmigrating to network edge where they are seen as virtualized resources called "Fog Computing" that expand the idea of Cloud Computing perspective to the network edge. This chapter proposes an integrated resource allocation model for energy-efficient Infrastructure as a Service (IaaS) that extends from the network edge of the Fog to the Cloud datacenter. We first developed a new architecture and introduced a policy on the Fog end where a decision will be made to either process the user request on the Fog or it will be moved to the Cloud datacenter. We developed a decision model on top of the architecture. The decision model takes into consideration of the resource constraints of CPU, Memory, and Storage. Using this will improve resource utilization as well as the reduction in energy consumption by

M. J. Usman (✉) · A. Aliyu
Department of Mathematical Sciences, Bauchi State University Gadau, 75110 Itas-Gadau, Nigeria
e-mail: umjoda@gmail.com

A. Aliyu
e-mail: ahmedaliyu8513@gmail.com

A. S. Ismail
School of Computing, Universiti Teknologi Malaysia, Skudai Johor 81310, Malaysia
e-mail: absamad@utm.my

H. Chizari
School of Computing and Technology, University of Gloucestershire, Cheltenham Park Campus GL50 4SH, UK
e-mail: h.chizari@imperial.ac.uk; chizari@utm.my

A. Y. Gital
Department of Maths and Computer, Abubakar Tafawa Balewa University Bauchi, Bauchi 740272, Nigeria

H. Chiroma
Department of Maths and Computer Science, Federal College of Education Technical Gombe, Gombe 072158, Gombe, Nigeria
e-mail: freedonchi@yahoo.com

M. Abdullahi
Department of Computer Science, Ahmadu Bello University Zaria, Zaria 810211, Kaduna, Nigeria
e-mail: abdullahilwafu@abu.edu.ng



T. Herawan et al. (eds.), *Advances on Computational Intelligence in Energy*,
Green Energy and Technology, https://doi.org/10.1007/978-3-319-69889-2_7

a datacenter. Finally, we addressed future research direction considering the model components and its performance.

**Keywords** Virtual machine · Cloud datacenter · Resource allocation · Energy efficiency · Fog Computing

## 1 Introduction

Cloud Computing is a universal term used for the delivery of hosted services and resources connected with the Internet that are pre-packaged using virtualization technology. The technology allows the creation of duplicate servers or computer system known as a virtual machine (VM). The concept of utilizing Cloud Computing allows companies, industries, and organization to have their Computing resources as a utility that is on a pay-per-use basis instead of building and keeping computing infrastructures in your premises. More so, Cloud Computing has rapidly become known as a successful paradigm for providing ICT infrastructure support. There is an enormous growth in this area for the past few years. The broader acceptance of Cloud Computing domain contributes to the development of large-scale datacenters. The datacenters provide services through their respective deployment and service models presented in various means and at different stages of the datacenters as shown in Fig. 1. The services are IaaS, Software as a Service (SaaS), and Platform as a Service (PaaS) [1, 2]. There are a lot of benefits using Cloud which includes pay-per-use, instant on-demand self-service provisioning, rapid elasticity, and resource pooling [3].

Cloud datacenters are the strength of today's demanding ICT infrastructure. Therefore, energy-efficient resource utilization is growing gradually and becoming important [4, 5]. Furthermore, the Cloud is transmigrating to the network edge where they are seen as virtualized resources called "Fog Computing" that expand the idea of Cloud Computing perspective to the network edge. The Fog and Cloud used similar resources that are identical in characteristics such as networking and computing devices with storage. They also share the same techniques associated with energy-efficient resource allocation and utilization. Although Fog Computing relies on routers and edge devices that are usually away from the hub datacenter. On the other hand, there are applications such as mobility-based services, health-monitoring devices, sensor networks with actuators, and emergency response services that require real-time response. If Cloud Computing alone will be used, it will require many resources from the sources to the edge and round it back to the Cloud datacenter. This leads to inefficient usage of the Cloud datacenter infrastructure, poor scheduling policies, and resource wastage are the reason for high-energy consumption in datacenters, not their size. Therefore, energy efficiency remains a serious problem for datacenter service providers.

Among the means to address this problem is to reduce the potentialities of the Cloud datacenters due to the used of virtualization technology [2]. The virtualization

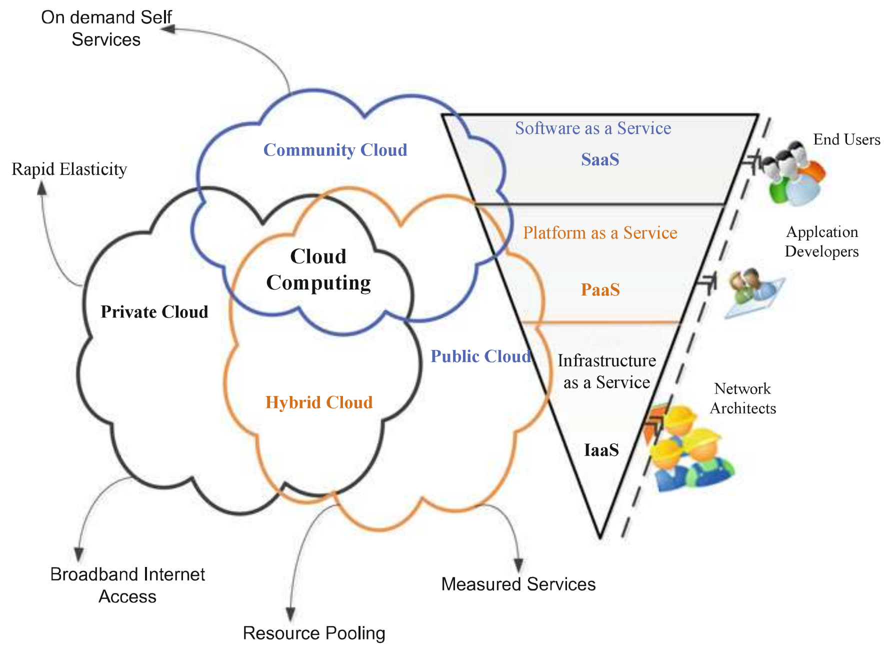


**Fig. 1** Classification and models of Cloud Computing

allows Cloud administrators to create many VM instances on a physical machine (PM) or server. Thereby demanding few physical resources unlike in the past, thus the resource utilization is improved. Alternatively, reducing energy consumption will be realized by turning off or switching PMs that are in the idle state to low-power mode state (i.e., sleep, hibernation). This approach also reduces the idle power consumption in the datacenters. This method enhances the condition of power consumption in the datacenter, but it is often not sufficient, due to the dynamic and heterogeneous nature of the Cloud environment. Yi et al. [6] in their survey they have find out unsolved Cloud Computing challenges because of their underlying problems such as resource utilization, central access, inadequate mobility support, and location awareness. Therefore, Fog Computing can address these challenges if properly apply in the Cloud domain. Combining the Fog with Cloud Computing is one of the most logical steps for building an adaptable and scalable platform that strengthens the IoT [7]. This will continue to support and increase the inescapable interplay of the Fog and the Cloud domain in years to come. Therefore, there is a need for integrating the Cloud and Fog Computing architecture for realizing energy efficiency in Cloud datacenter IaaS.

The focus of this research work is how to realize energy efficiency using resource allocation strategies and policies that can be applied to virtualized Cloud datacenter. Unlike existing studies [8–13], our approach will integrate Cloud IaaS with Fog Computing resources on a single datacenter. The integration will bring a new form of applications and services within the Cloud datacenter environment that reduces resource under-utilization and energy consumption. We also develop a decision model that takes into consideration of the resource constraints and their capacity at the Fog edge. Furthermore, the model determines whether the user's request will be executed on the computational infrastructure of Fog or Cloud datacenter. Also, we recommend future research direction considering the operational components, execution environment, and performance of the integrated environment.

The rest of the book chapter is organized as follows. Section 2 discusses the related work. Section 3 presents the Cloud and Fog Computing overview. Section 4 describes the needs for Cloud and Fog Computing with the proposed model. Section 5 outlines the future issues and direction of the research and Section 6 concludes the book chapter.

## 2 Related Works

Deng et al. [14] and Al Faruque and Vatanparvar [15] presented their work with resource provisioning of Fog and its related services. They focus mainly on the energy consumption of Fog–Cloud interaction of devices installed at smart houses or cities. Furthermore, they consider the different stages of energy consumption as well as the latency trade-off between the Cloud–Fog. Deng et al. [14] and Tang et al. [16] introduce a distributed architecture of Fog Computing that has big data analysis capabilities used in smart cities. They designed a four-layered architecture that supports immediate response and intelligence report for future smart cities. The computing nodes at each layer of the architecture execute applications that are latency-sensitive and support quick control loop to protect the safeness of some vital infrastructure resource components. They focus on the integration of Cloud IaaS with next-generation smart cities. It provides high computing performance that supports future intelligence smart cities. However, the authors consider response time only. Thus, leaving abundant resources under-utilized. Stojmenovic [17] expands the concept of Fog and Cloud Computing to the decentralized smart controller in survey paper. They recognize cloudlets often known as task as defined in Fog domain. The authors describe the cloudlets using scenarios in software-defined networks (SDN). They make a recommendation on the existing approach, which needs modification to adopt two layers Fog–Cloud Computing. Smart Cargo concept has been introduced by Costa et al. [18]. The authors integrate traditional technologies and frameworks ascertain from the cargo environment. Besides, the concept combined the capabilities of a Fog Computing centralization and scalability offered by Cloud paradigm applied to the freight transport domain. They explore home energy management (HEM) and a micro-grid-level energy management over a Fog Computing platform with a customized controller. This method shows a paradigm shift that takes away computing application and data services control from Cloud to Fog Computing. Therefore, the approach makes multimodal freight transport attained real-time awareness with future environmental prediction, but they did not consider the limitation of the Fog devices and/or nodes.

Fog Computing has been used as a new platform for energy management by Vatanparvar et al. [19] and Al Faruque and Vatanparvar [15]. The platform allows the interaction of different devices' interoperability and operational flexibility. Furthermore, they implemented two prototypes of HEM and micro-grid-level energy management that separates control and data communication layers. The platform is adaptive and scalable due to its open architecture. However, they focus on energy management of residential houses, not datacenters. Masip-Bruin et al. [20] proposed a layered architecture that shows the need for the coordinated management of Fog–Cloud systems. The authors use medical emergency scenario to evaluate the performance of the architecture. The F2C (Fog 2 Cloud) architecture optimizes and improves the execution time of applications compared with traditional Cloud services.

Madsen et al. [21] present the challenges and reliability posed by current Computing paradigms for reliable Fog platforms. The approach incorporates the networks of smart devices communicating with the Cloud. While it may be true, the authors conclude that the design and development of Fog Computing projects are demanding. Therefore, this strategy did not offer a new approach that is reliable for the network of smart devices in Fog Computing standard. Thus, the acceptance of Fog-based access call for the implementation of algorithms that deals with network reliability of smart devices. The authors focus on Fog devices that are operating under specific circumstances that inquire for fault tolerant technique, not resource utilization or energy management. Similarly, Bonomi et al. [22] analyze key distinct features of Fog and how it complements Cloud Computing. Main goal of this research work is to examine the used instances that motivate the need for Fog. This approach emphasizes the relevancy of Fog to IoT and Big Data Space (BDS) paradigm. Finally, the authors provide a high-level characterization of Fog's software architecture, giving prominence to the different components and the technological requirement to achieve the Fog vision.

A placement resource migration technique for Cloud and Fog is presented by Ottenwälder et al. [23]. The authors demonstrate how application understanding of complex event processing (CEP) system is applied to make live VM migration better. The needed bandwidth for the technique to perform the VM migration has been reduced substantially. The resources that are considered to be network-intensive are directed on distributed Fog nodes. On the other hand, the computationally intensive resources were hosted in the datacenter. This method reduces the migration cost due to the direct selection of the destination of resources to be migrated which assures sufficient network utilization over a period of time. The network utilization and mobility pattern have been improved. Thus, other resources and workload mobility have not been optimized due to the computationally intensive task by the Fog devices. The immense demand for geo-distribution, mobility support, low latency, and location-based application is becoming high in both Fog and Cloud domains. Therefore, to solve the above-mentioned problem, a resource-provisioning mechanism for VM-based tactical cloudlets has been proposed by Lewis et al. [24]. It is a strategy that can be hosted on a vehicle or other platforms for providing IaaS that endorses computation offloading and data forwarding at the edges of the Fog devices. The aim is to increase position awareness that will improve the decision making of moving the resources used in Cloud close to the Fog. This will encourage energy efficiency in mobile devices, lower latency and change provisioning policy in tactical environment. Although the cloudlets has been successful in supporting the mobile devices, the approach did not consider the limitation of the mobile devices in terms of running applications.

Aazam and Huh [25] proposed a solution using smart gateway-based communication (SMBC) with Fog Computing to realize smart communication and resource management. It alleviates the core network communication overhead of the Cloud datacenter. Using this method makes it easier to create better services for the Cloud in an efficient manner. However, resource utilization and energy saving are not their focus. Similarly, Aazam and Huh [26] presented another service-oriented resource management model (SORMM) for Fog Computing. The model performs adequate resource management for IoT deployments. It also uses a framework for resource management to predict user's resource usage and to also pre-allocate resources based on the probabilistic behavior that determines the future usage by the user. The used of this prediction strategy by the consumed resources bring good standard and efficiency. However, they did not present any evaluation or simulation result that supports their model. Nishio et al. [27] presented a mathematical framework for service-oriented utility in heterogeneous resource environment. In mobile Cloud concept, the heterogeneous resources are measured in standards, such as power, bandwidth, and latency. Neighboring nodes within a local area network (LAN) form a cluster group termed local Cloud. The nodes shared the resources within the LAN and other nodes from the clusters in the same local Cloud. Then from there, the Fog device is selected from the nodes in each local Cloud by local resource coordinator. However, the model did not take into consideration of resource utilization and energy consumption.

Dsouza et al. [28] proposed a relevant schema with a preliminary policy management framework followed by a detail description of the policy. The authors identified Fog Computing novel features. These key features have been identified with challenges that are inherent in managing the policies. It becomes vital to endorse ensure resource sharing and collaboration to reuse data within heterogeneous Fog environment. The model is effective in resource management for enterprise used but not Cloud datacenter. Recent development of Fog Computing is presented by Dastjerdi et al. [29], and the authors design a reference architecture for Fog and discussed their challenges. They pointed out issues that need urgent attention including security, resource management, and energy usage minimization are still in needs of the solution.

## 3 Overview of Cloud–Fog Computing

The Cloud Computing is the Cloud datacenter structural system, which is a building or room full of computing servers on premises or remote location. It also includes other related equipment and resources. Whereas the Fog Computing is being considered as a term that is originally coined by Cisco, that is, in many ways the same or nearly the same with edge Computing or nano-Cloud. The main distinct feature of the Fog when compare with the Cloud is that the platform and resources used by Fog are always at the network edge of an enterprise or institution with limited computational capabilities. The attribute of the Fog and other associated components and platforms are discussed in the following sections.

## 3.1 The Cloud Computing Architecture

The Cloud Computing architecture is the structure and organization of the system, which comprise on premises and computing resources (physical or logical) such as available services, middleware, software components, storages, geo-location, the externally visible properties of resources and their relationships to provide services to users through the network [30]. Recently, Cloud Computing reference architecture recognizes the major actors and classifies them based on types, models, as well as their behavior. More so, Cloud architecture has been considered as a generic high level that is proposed to help the understanding of the user demands, characteristics, and principles used in this domain.

This architecture comprises two parts: the front end and the back end. The front cannot operate independently without the support of the back and vice versa. All requests coming into the Cloud mostly come from the front end. This connects usually with the back end through a network communication system. What differentiates the front with the back end is the source of the request. The request always comes from the users who are usually known as client or consumers of services from the front end. The back end is where the requested services by the clients are hosted which is the Cloud side of the system. The users do not care where the request or information is stored or processes rather they are after service delivery and reliability from the Cloud side. Therefore, the front end is composed of user's resources such as computing and network resources with the required application to access the Cloud Computing resources. The Cloud Computing resources usually include physical machines (PM) also known as servers, storages, network communications, and different applications and operating systems running concurrently to service user's request. Requests of application in the Cloud are usually run on virtual machines (VM) hosted on the PM. These resources are managed by the administrator who manages and monitors the resource as well as the traffic within the Cloud environment. The management of the resources in the Cloud has to follow a set of rules known as a protocol with the special help of middleware. The middleware will allow the computers to communicate with each other over the network [30].

### 3.2 The Fog Computing Architecture

The Fog Computing system-level architecture has been modified to achieve maximum efficiency for the scattered Computing nodes used to perform the computational task with storages, processes, and networking capabilities according to various criteria within an Internet of Everything (IoE) domain [31]. The Fog Computing encompasses a tiered organization of Fog nodes all over the network amid sensors and the Cloud at the central core of the network system [32]. The Fog architecture constitutes a change from conventional access systems and trust on Cloud models, to the latest computation model that moves computational request close to the network edge often referred to edge Computing [33]. The nodes are not completely fixed to the edge but should be seen as a variable within the system of connectivity.

The Fog architecture provides the correct support and attempts to bring into the balance between resource content of the three most basic capabilities at the edge of the network level where they are optimally located. The Fog layers as shown in the figure above are located between different network levels of the Cloud and the Fog resources known as nodes. As a result, the Fog makes a connected graph, with the corresponding Fog nodes located at different levels, which forms an interconnected topology between core smart objects and layers Cloud components [32–35]. Storage and network bandwidth are the essential attributes of Fog architecture found at different levels.

## 4 The Need for Integrating Cloud and Fog Computing

There is always a need to utilize the Computing power available at the edge of the network which is the Fog. This can only be realized by taking Cloud concept which includes flexibility, agility, and distributed nature as well as network storages and moving them toward the devices that generate the data. Integrating Fog with Cloud Computing is more convenient than migration. The Cloud has been standardized in terms of information exchange and emerging use of services such as web, machine-to-machine communication (M2M) and used open sources. As the Fog devices are becoming smarter and intelligence that serve as the extension of the Cloud where user's requests are no longer centralized. Therefore, bringing in new concept of decentralization due to the massive computation at the edges of the network. The integrated architecture has been designed based on the Cloud and Fog characteristics as shown in Fig. 2.

This clearly shown the limitation of each of the architecture in terms of latency, hardware used, location awareness, and response time. Moreover, this has also lead to the need for their complimentary services toward energy efficiency in the Cloud datacenter. The integration of the Fog–Cloud architecture take into consideration the limitation of both Cloud and Fog Computing, thereby providing a new solution that improves energy efficiency, resource utilization, and data optimization.

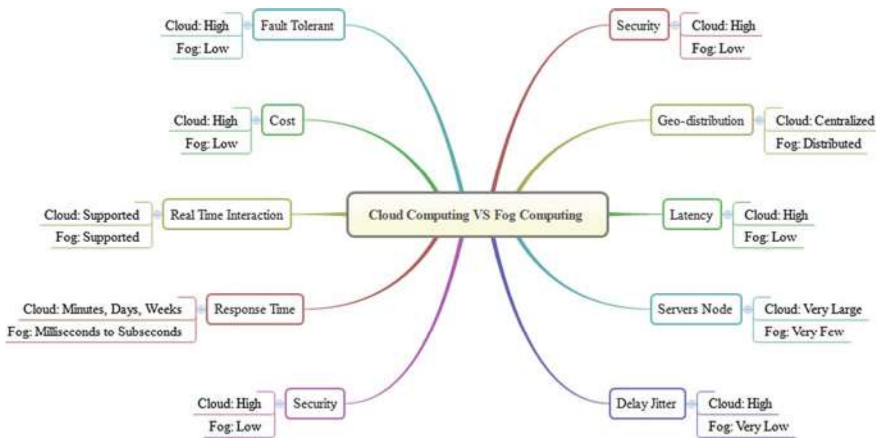


**Fig. 2** Challenges of Cloud and Fog Computing architecture

### *4.1 Operational Framework of Cloud–Fog Computing*

Figure 3 shows the framework for the integrated model of Cloud–Fog Computing. The framework is divided into four distinct parts with the objective of achieving energy efficiency in the Cloud datacenter. The users are the consumers of the services and always make application request. These requests are usually processed at the datacenter which consumed high energy due to inefficient resource utilization. Therefore, we have introduced a new concept whereby a user's request is process on the Fog using workload allocation model. This model has three decision criteria due to the inherent limitation of the Fog resources in terms of processing, storage, and memory else the request will be sent to the Cloud. The Cloud will now look at the request and allocate appropriate resource for it using resource allocation linear model. We also show how the Fog interacts with the Cloud through the gateway.

### *4.2 The Proposed System Architecture of Cloud–Fog Computing*

Figure 4 shows the overall system component of the proposed architecture with a brief explanation of their role while Fig. 5 shows the integrated architecture of Fog/Cloud.

The proposed system architecture shows the integration of Cloud and Fog architecture. The datacenter resources such as server processor and network which enable on-demand and dynamic allocation of resources while minimizing the datacentre energy consumption. The approach is model-based using resource allocation linear model of the IaaS that is handled by the Cloud Management System (CMS) which include brokers, service monitors, and energy optimizers. While for the Fog is handle by the integrated resource management system (IRMS) which includes the sensor as well as resource constraints. The technique will overcome the limitations of both the Cloud and the Fog.

The Cloud Computing players include users, application request, service providers, and CMS and the other side of the architecture is connected with the communication gateway through the IRMS. The Fog physical resources are organized based on their application areas and execution environment. In Cloud datacenter setting, users are regarded as one of the important entities since they are the end users

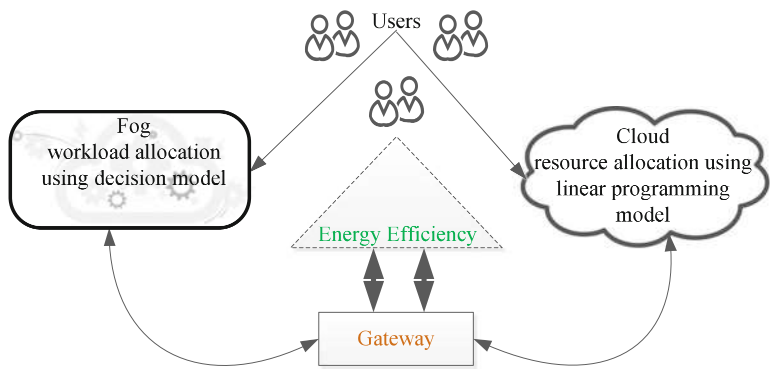


**Fig. 3** Operational framework of energy efficient Cloud–Fog Computing

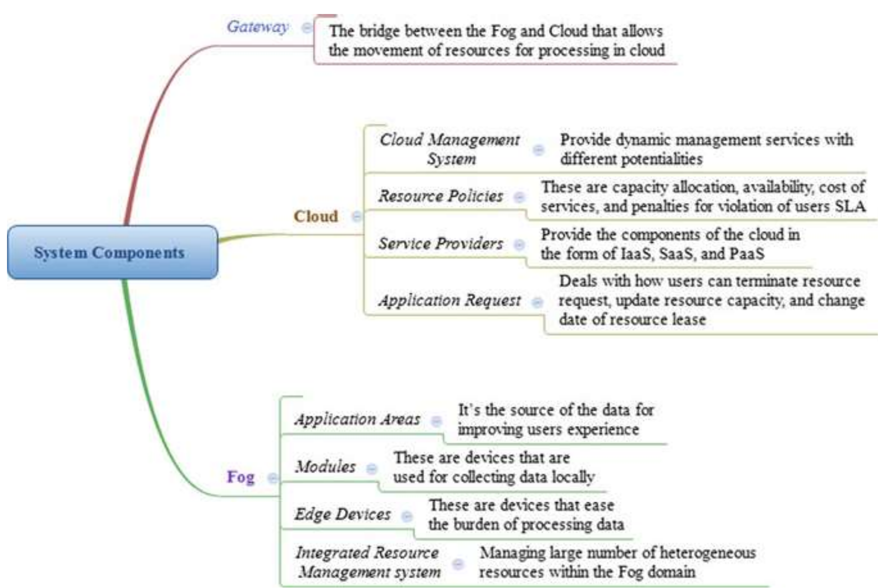


**Fig. 4** Components of the architecture

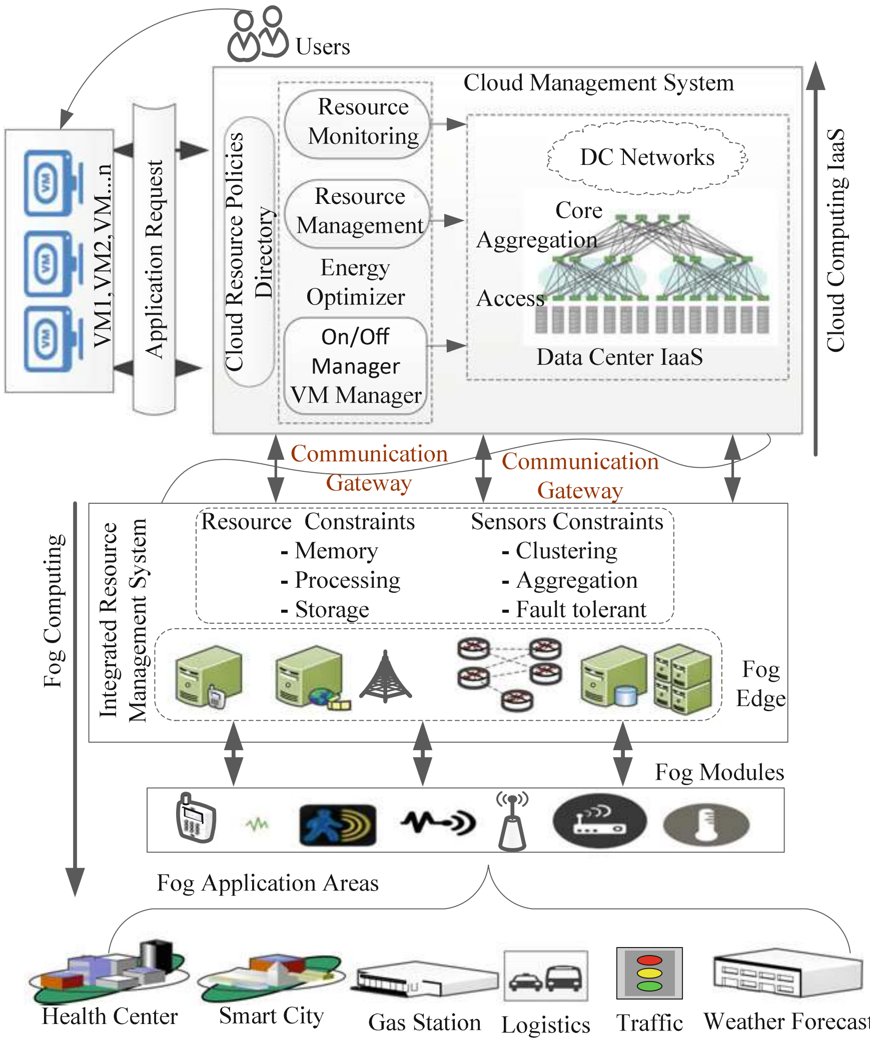


**Fig. 5** Integrated architecture of Cloud and Fog Computing

of the services provided by the Cloud IaaS. They sent application requests that are operated in distinguish VMs hosted on different servers while the service providers owned the applications that provide services to the users.

More so, users submit their request to the Cloud provider through the CMS, specifying the number of VMs needed for each application and their requirement in terms of memory, storage, CPU, and network. Other functional components of the CMS collect run-time statistics of the resources including availability, utilization, and status of the resources. This information will be published for decision making by the data center based on their resource policy requirement define in the service level agreement (SLA).

On the other part of the architecture, we push away the applications, data, and services from distributed nodes to the network edge. Therefore, this will allow the IRMS to perform its function and decide where to place the applications and data based on the application requirement.

### *4.3 Assumption of the Cloud–Fog Model*

The integration of Cloud–Fog Computing separates the workloads between the two components. This also leads to the efficient utilization of resources hence reduces the energy consumption. The detail assumptions of our architecture are as follows:

1. A request represents a user's computational request or task.
2. All requests come from Fog Computing environment considering three constraints such as CPU, storage, and memory.
3. The constraints are presented as $U(R_1)$, $U(R_2)$, and $U(R_3)$.
4. $U(R_i)$ is the cumulative resource requirement of the user request, Ø is the threshold of the fog resource capacity, where $\alpha$, $\beta$, and $\gamma$ represent resource requirement of a given task, where the task will be $T_1$, $T_2$, $T_3$, …, $T_n$.
5. Based on the assumption, the workload forwarded to the Cloud is in the form of tasks.
6. The resource allocation requests by users must provide the detail of resource usage for the task. Otherwise, if the user requests are more than the available resources, the request has to wait in the queue for the next allocation.
7. All hosts in Cloud–Fog consume energy in idle state and denoted as $P_{\min}$.
8. Hosts consume maximum energy at the highest stage of utilization which is denoted as $P_{\max}$.
9. Hosts put the request in the waiting queue if its CPU utilization is at pick level or it reaches its threshold.

### *4.4 Decision Model for Cloud–Fog Architecture*

In optimization, the resource constraints usually affect the solution. Therefore, how to select and decide where the resource will be run is an essential element for realizing efficient resource utilization and energy consumption in optimization, the resource constraints usually affect the solution. Therefore, how to select and decide where the resource will be run is an essential element for realizing efficient resource utilization and energy consumption. For the Fog Computing resources, the assumption is that the number of constraints which includes CPU, storage, and memory are presented as $U(R_1)$, $U(R_2)$, and $U(R_3)$. $U(R_i)$ is the cumulative resource requirement of the user request, Ø is the threshold of the fog resource capacity, and where $\alpha$, $\beta$, and $\gamma$ represent resource requirement of a given task, where the task will be $T_1, T_2, T_3, \ldots, T_n$.

Thus, the resource utilization function for single resource can be denoted as:

$$U(R_1) = \sum_{i=1}^{n} \alpha_i \tag{1}$$

$$\text{Single Decision Creteria} \begin{cases} U(R_1) < \theta; & \text{assign to Fog resources} \\ \text{Otherwise;} & \text{assign to Cloud resources} \end{cases}$$

The decision criteria for single-resource constraint take into consideration the size of the resource $\alpha_i$ which is being requested by the user. It checks the requested resource threshold Ø capacity whether it can be processed at the fog level else it will forward the request to the cloud.

For double-resource constraint's request, assuming we have a task that requires more than one resource at a time for execution. Thus, the resource utilization function can be denoted as:

$$U(R_1) = \sum_{i=1}^{n} \alpha_i \tag{2}$$

$$U(R_2) = \sum_{i=1}^{n} \beta_i \tag{3}$$

$$\text{Double Decision Creteria} \begin{cases} U(R_1) < \emptyset_1; \text{ and } U(R_2) < \emptyset_2; & \text{assign to Fog resources} \\ \text{Otherwise;} & \text{assign to Cloud resources} \end{cases}$$

The decision criteria for double-resource constraint also take into consideration the size of the resource $\alpha_i$ and $\beta_i$ which is being requested by the user. It checks the requested resource threshold Ø capacity whether it can be processed at the fog level else it will forward the request to the cloud. Whereas the multiple resource constraints request, assuming we have a task that requires more than two resource at a time for execution. Thus, the resource utilization function can be denoted as:

$$U(R_1) = \sum_{i=1}^{n} \alpha_i \tag{4}$$

$$U(R_2) = \sum_{i=1}^{n} \beta_i \tag{5}$$

$$U(R_3) = \sum_{i=1}^{n} \gamma_i \tag{6}$$

$$\text{Multiple Decision Creteria}\begin{cases} U(R_1) < \emptyset_1;\, U(R_2) < \emptyset_2;\ \text{and } U(R_3) < \emptyset_3; \\ \quad \text{assign to Fog resources} \\ \text{Otherwise;}\quad \text{assign to Cloud resources} \end{cases}$$

The decision criteria for multiple resource constraints take into consideration the size of the resource $\alpha_i$, $\beta_i$, and $\gamma_i$ which are being requested by the user. It checks the requested resource threshold Ø capacity whether it can be processed at the fog level else it will forward the request to the cloud. The Fog Computing resource capacity is given as

$$\begin{cases} 0 \le \theta \le f_i; & \text{allocate to fog } i \\ \text{otherwise;} & \text{allocate to cloud} \end{cases} \tag{7}$$

$N$ is the number of available fog devices $i$
$f_i$ resource capacity in terms of RAM, CPU, and storage.

### *4.5 Optimization Problem*

Resource allocation and energy efficiency of the overall Cloud datacenter can be denoted as linear programming (LP) problem to reduce the energy consumed E, as represented in Eq. (8).

$$\text{Minimize } E = \sum_{r=1}^{n}\sum_{i=1}^{r} E_i(\tau) \tag{8}$$

Subject to

$$E_i(\tau) = \left(P_{\max} - P_{\min}\right) * \frac{U_i(\tau)}{100} + P_{\min} \tag{9}$$

$$U_i(\tau) = \sum_{j=1}^{n} U_{ij} \le \text{Peakload at time } \tau, \quad \forall R_i \in R \text{ and } \forall t_j \in T \tag{10}$$

$$U_{ij} = 0, \quad \text{when the user request } j \text{ is not serve on node } R_i$$
$$U_{ij} = 1, \quad \text{when the user request } j \text{ is serve on node } R_i$$

Equations (8) and (9) indicate that minimization of energy consumption is proportionally subjected to the utilization of resources based on the user request for the time $\tau$. Our models relate to the overall Cloud Computing environment's energy consumption to total utilization represented by $U$ as can be formulated in Eq. (8). In the equation, $u$ is the percentage value of the processor utilization, $E$(u) is the energy consumed by CPU at the utilization $u\%$, and $P_{\max} - P_{\min}$ are the power consumption at maximum performance in watt and at idle, respectively.

### 4.6 Performance Evaluation

Fog does not replace the Cloud which have lots of advantages because of its central and scalable features. The Cloud datacenters offer unlimited virtual processing capabilities and an on-demand resource usage model. To explain our contribution, the storage, processing, and memory constraints have been considered in the decision criteria model. Therefore, we use the idea of resource modeling criteria as the key idea for the integration of Cloud and Fog to reduce energy consumption and maximize resource utilization of IaaS. To ensure the quality of our developed model, we cautiously perform some experiments based on assumption. From the assumption, all requests are coming from the Fog environment handle by IRMG. Therefore, the decision models will now decide where the request will be processed based on the threshold limit set by the model. In the following, we show some results obtained by using the decision model which includes single-, double- and multiple decision criteria.

From Fig. 6, it shows the uses of single-resource request from fog environment. When the requested resource by the user has not exceeded the fog capacity as define in the model, we can say that it has an upper and lower boundary of the requested resource. If resource happens to be at the lower boundary (less than $\theta$). Then, the request will be handled by the Fog schedulers. Likewise, when the request is greater than $\theta$ then the resource request will be forwarded to the cloud datacenter through the gateway. In the datacenters, the CMS will handle the request and allocate an appropriate resource for the user request. The request will be executed based on first-come first-serve. As shown in Fig. 6 anything above the $\theta$ will be processed in the cloud datacenter. This will improve resource utilization of the Fog and Cloud, as well as the energy efficiency of the datacenters.

Figure 7 shows the user request for double resource. In this case, the decision model will make a complex decision when the requested resources are not same. This is due to the nature of the application. Some applications require much memory than storage and vice versa. In this case, the decision criteria will not compare the two requests and in case of any of the request exceed Fog capacity $\theta$, then the request will be moved to the cloud datacenter. However, when the request has not been exceeded $\theta$, then it will be processed within the Fog environment. In addition, Figs. 8 and 9 show multiple decision criteria. In this case, the user's request is treated in the same way as single- and double-decision criteria. However, we have shown how the multiple resource requests will be processed in Fog. This can only be achieved when the user request has not exceeded the threshold for the three constraints. In case of any of the requests have been exceeded the threshold, the resource request will be moved to the Cloud datacenter.

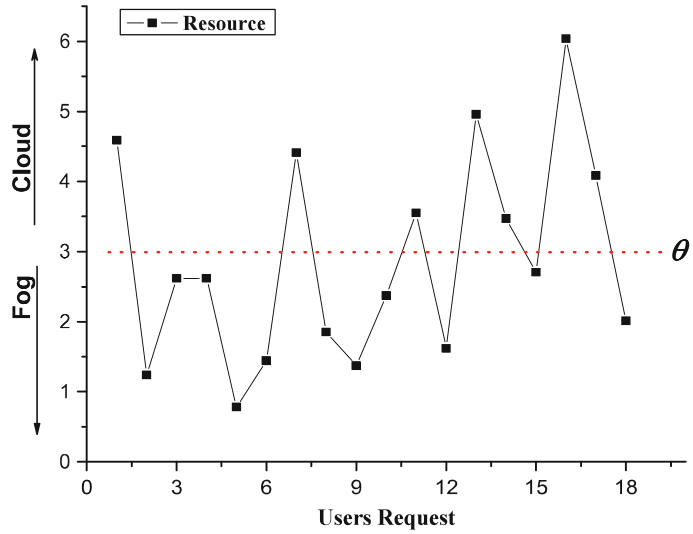


**Fig. 6** Single-decision criteria

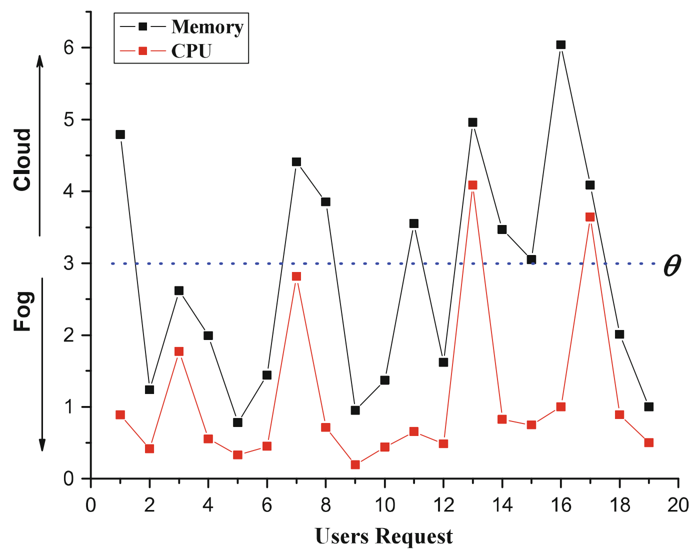


**Fig. 7** Double-decision criteria

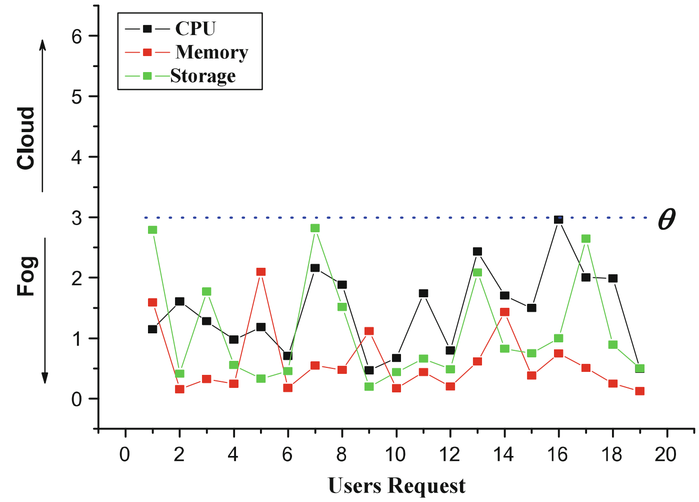


**Fig. 8** Multiple decision criteria

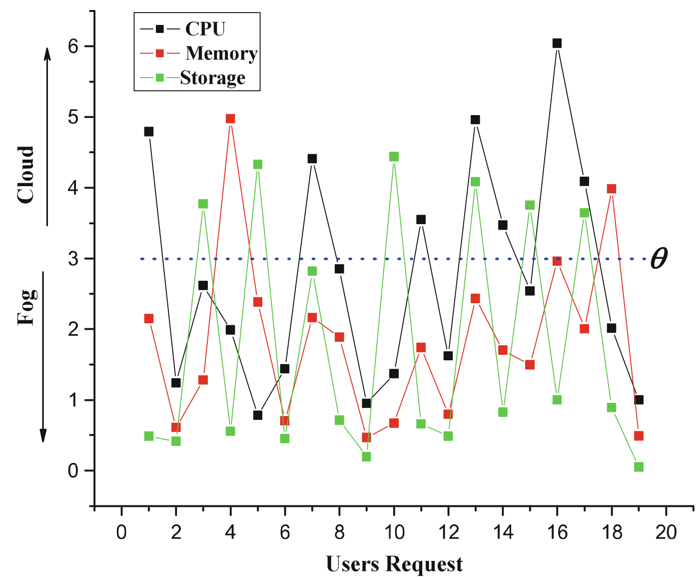


**Fig. 9** Multiple decision criteria

Based on this model and its analysis, we can say that the energy efficiency of the datacenter and resource utilization has been improved. This is due to the workload and application processing migration when the user's request exceeded the threshold of the Fog environment. Therefore, this triggers the interaction of the two-environment Fog–Cloud resource. But when the average workload is less than the threshold value then the user's request will be processed at the Fog environment. Therefore, only application that require much CPU, Memory and Storage will be sent to Cloud datacenter for processing. This allows the efficient use of the Cloud—Fog resource that eliminates the loss of data, massive communication bandwidth, and energy consumption.

## 5 Future Issues and Direction

Based on our discussion in the literature, there are important aspects of Fog and Cloud Computing paradigm due to their inherent finite resource capacity offered as a service to users. Therefore, there is a need to explore the issues that are left behind because of the compromise in solving the existing problems. This will bring advancement in finding solutions to the existing issues that require urgent attention of the researchers in this domain. Some of the issues and future research direction are discussed in the following section.

*Resource Provisioning* This deals with choosing and deployment of programs run time for applications which can be dynamic or static. It is found from the literature, very few research works have considered dynamic resource provisioning or using both static and dynamic in Cloud–Fog environment. Many important aspects such as latency, mobility, network traffic, and resource availability are still unexplored. To efficiently use both Cloud and Fog resources, application of different techniques on resource provisioning and service management will be a promising subject in the theme of Cloud–Fog research.

*Multitenancy Resource Management* Existing Fog Computing resources can be virtualized and at the same time allocated to numerous user requests. The current multi-tenants' research works do not take into consideration of Fog Computing resource scheduling of computational tasks and request based on the user application requirements (performance degradation, makespan and throughput). This can be carried out in future researches.

*Parameters* The parameters used by the implemented techniques to determine their performance are shown in the table. The choice of the parameters on Cloud–Fog user perspective or Cloud service provider are not the same. The provider is concerned more about efficient resource utilization while the users are focused on application performance. In that regards, the user's parameters include latency, response time, execution time, fairness, turnaround time and tardiness, and the provider parameters include energy consumption reduction, resource utilization, VM migration, workload, budget, and other dependency constraints (death line, reliability and priority constraint). Not all these parameters were used. There is a need to explore more on the parameters to see their performances in Cloud–Fog environment.

*The Sustainable of the Cloud–Fog Architecture* The Fog Computing sustainability reflects both economic and the effect of the environmental to certain degree. Though, sustaining the Fog Computing architecture has been struggling with issues such as the QoS, service availability, energy efficiency, and resource utilization. Additional research in this domain is highly required for the performance and accomplishment of Fog Computing. This is because the present research work presents a limited analysis of sustainability and reliability of the Fog Computing architecture.

*Interoperability* It is the ability to share information between systems and application in a meaningful way. This will enable users to understand why they are receiving and sending information over the network with different configurations. Although, the Cloud Computing is already standardized and the operational risk is minimal. Current future of the Cloud Computing is on the federation of different infrastructure across different sites. But when they are interacting with Fog nodes that are on premises platform which are entirely different in terms of processing and storage capabilities. This shows that the users are now operating in a highly diverse environment, different facilities located around the world. There is a need to support this diversity in that environment so that they can run their applications on Cloud environment without service interruption or resource luck-in. In general, interoperability will allow new efficiency that will not be interrupted by some constraints from the Fog environment. Therefore, there is a need to adapt standard or common interface that will create more innovation, support for different applications by different vendors and data-driven insight.

## 6 Conclusion

Energy efficiency is important for Cloud datacenters and Fog Computing. Fog Computing is becoming a desirable answer to application processing issues in the IoT paradigm. In this chapter, three parts of the developed model have been explained which include Cloud-only application module placement and a technique that pushes applications toward Fog devices. The technique allows the Fog to operate while disconnected from a larger network. It also supports the intelligent decision of processing, storing, memory, and actuating when enough resources are available. Nevertheless, we recommend that the usage of Cloud Computing paradigm has several advantages as a support infrastructure and large-scale reasoning, but part of the control and trust decisions should be shifted to the Fog and allow for innovative computing applications. The model has been tested using the origin platform. Some issues related to double and multi-resource constraints have not been addressed in this work. This limitation will be addressed in future work.

## References

1. Foster I et al (2008) Cloud computing and grid computing 360-degree compared. In: IEEE 2008 grid computing environments workshop, Austin, TX, pp 1–10
2. Beloglazov A, Abawajy J, Buyya R (2012) Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener Comput Syst 28(5):755–768
3. Usman M et al (2016) A conceptual framework for realizing energy efficient resource allocation in cloud data centre. Indian J Sci Technol 9(46):1–8
4. Madni SHH, Latiff MSA, Coulibaly Y (2016) An appraisal of meta-heuristic resource allocation techniques for IaaS cloud. Indian J Sci Technol 9:1–14
5. Madni SHH, Latiff MSA, Coulibaly Y (2016) Resource scheduling for infrastructure as a service (IaaS) in cloud computing: challenges and opportunities. J Netw Comp Appl 68:173–200
6. Yi S, Li C, Li Q (2015) A survey of fog computing: concepts, applications and issues. In: 15th proceedings of the 2015 workshop on mobile big data, Hangzhou, pp 37–42
7. Yannuzzi M et al (2014) Key ingredients in an IoT recipe: fog computing, cloud computing, and more fog computing. In: 19th IEEE international workshop on computer aided modeling and design of communication links and networks (CAMAD), Athens, pp 325–329
8. Bose R, Saha MK, Sarddar D (2015) Fog computing made easy with the help of Citrix and Billboard manager. Int J Comput Appl 121:19–23
9. Masip-Bruin X et al (2016) Fog-to-cloud computing (F2C): the key technology enabler for dependable e-health services deployment. In: 15th Mediterranean ad hoc networking workshop (Med-Hoc-Net), Vilanova i la Geltru, pp 1–5
10. Deng R et al (2015) Towards power consumption-delay tradeoff by workload allocation in cloud-fog computing. In: 5th IEEE international conference on communications (ICC), London, pp 3909–3914
11. Gital AY et al (2014) A framework for the design of cloud based collaborative virtual environment architecture. In: Proceedings of the international multiconference of engineers and computer scientists (IMECS), Hong Kong, pp 1–5
12. Usman MJ et al (2014) Modified low energy adaptive clustering hierarchy protocol for efficient energy consumption in wireless sensor networks. Int Rev Comput Softw (IRECOS) 9:1904–1915
13. Gital AY et al (2015) A zone-based CVE architectural model for improving scalability and consistency. Electrical Engineering and Information Technology (IEET), Hangzhou, China, pp 63–68
14. Deng R et al (2016) Optimal workload allocation in fog-cloud computing towards balanced delay and power consumption. IEEE Internet Things J 3:1171–1181
15. Al Faruque MA, Vatanparvar K (2016) Energy management-as-a-service over fog computing platform. IEEE Internet Things J 3:161–169
16. Tang B et al (2015) (2015) A hierarchical distributed fog computing architecture for big data analysis in smart cities. In: 2015 proceedings of the ASE big data & social informatics. ACM, Kaohsiung, Taiwan, pp 1–28
17. Stojmenovic I (2014) Fog computing: a cloud to the ground support for smart things and machine-to-machine networks. In: 26th Australasian telecommunication networks and applications conference (ATNAC), Southbank, VIC, pp 117–122
18. Costa R et al (2016) Smart cargo for multimodal freight transport: when "Cloud" becomes "Fog". IFAC-PapersOnLine 49:121–126
19. Vatanparvar K, Faruque A, Abdullah M (2016) Energy management as a service over fog computing platform. IEEE Internet Things J 3:163–169
20. Masip-Bruin X et al (2016) Foggy clouds and cloudy fogs: a real need for coordinated management of fog-to-cloud computing systems. IEEE Wirel Commun 23:120–128
21. Madsen H et al (2013) Reliability in the utility computing era: towards reliable fog computing. In: 20th international conference on systems, signals and image processing (IWSSIP), Bucharest, pp 43–46

22. Bonomi F et al (2014) Fog computing: a platform for internet of things and analytics. In: Big data and Internet of Things: a roadmap for smart environments. Springer International Publishing, pp 169–186
23. Ottenwälder B et al (2013) MigCEP: operator migration for mobility driven distributed complex event processing. In: 7th ACM international conference on distributed event-based systems, Arlington, Texas USA, pp 183–194
24. Lewis G et al (2014) Tactical cloudlets: moving cloud computing to the edge. In: IEEE military communications conference (MILCOM), Baltimore, MD, pp 1440–1446
25. Aazam M, Huh E-N (2014) Fog computing and smart gateway based communication for cloud of things. In: 2014 2nd IEEE international conference on future Internet of Things and cloud (FiCloud), Bercelona, pp 464–470
26. Aazam M, Huh E-N (2015) Dynamic resource provisioning through fog micro datacenter. In: IEEE international conference on pervasive computing and communication workshops (PerCom Workshops), St. Louis, MO, pp 105–110
27. Nishio T et al (2013) Service-oriented heterogeneous resource sharing for optimizing service latency in mobile cloud. In: ACM proceedings of the first international workshop on mobile cloud computing & networking, Bangalore, India, pp 19–26
28. Dsouza C, Ahn G-J, Taguinod M (2014) Policy-driven security management for fog computing: preliminary framework and a case study. In: IEEE 15th international conference on information reuse and integration (IRI), Redwood City, CA, pp 16–23
29. Dastjerdi AV et al (2016) Fog computing: principals, architectures, and applications. arXiv preprint arXiv:1601.02752
30. Liu F et al (2012) NIST cloud computing reference architecture: recommendations of the National Institute of Standards and Technology (special publication report 500–292). CreateSpace Independent Publishing Platform
31. Byers CC, Wetterwald P (2015) Fog computing distributing data and intelligence for resiliency and scale necessary for IoT: the Internet of Things (ubiquity symposium). Ubiquity, pp 1–12
32. Gupta H et al (2016) iFogSim: a toolkit for modeling and simulation of resource management techniques in Internet of Things, edge and fog computing environments. arXiv preprint arXiv:1606.02007
33. Bonomi F et al (2012) Fog computing and its role in the internet of things. In: ACM proceedings of the first edition of the MCC workshop on mobile cloud computing, Helsinki, Finland, pp 13–16
34. Perera C, Qin Y, Estrella JC, Reiff-Marganiec S, Vasilakos AV (2017) Fog computing for sustainable smart cities: a survey. ACM Comput Surv 50:32–43
35. Moreno-Vozmediano R, Montero RS, Huedo E, Llorente IM (2017) Cross-site virtual network in cloud and fog computing. IEEE Cloud Comput 4:46–53