A Novel Approach to Minimizing Service Latency in Edge Cloud Computing Systems

S. Ahmed[*1], Mrs. R. Fatima[2] & Dr. A. Khan[3]

[*1]Department of Computer Engineering, King Saud University, Riyadh, Saudi Arabia
[2]Assistant Professor, Department of Computer Engineering, King Saud University, Riyadh, Saudi Arabia
[3]Principal, King Saud University, Riyadh, Saudi Arabia

ABSTRACT

A lot of research has been done in the field of cloud computing in computing domain. For its effective performance, variety of algorithms has been proposed. The role of virtualization is significant and its performance is dependent on VM Migration and allocation. More of the energy is absorbed in cloud; therefore, the utilization of numerous algorithms is required for saving energy and efficiency enhancement in the proposed work. In the proposed work, green algorithm has been considered with meta heuristic algorithms, ABC (Artificial Bee colony .Every server has to perform different or same functions. A cloud computing infrastructure can be modelled as Primary Machines a set of physical Servers/host PM1, PM2, PM3… PMn. The resources of cloud infrastructure can be used by the virtualization technology, which allows one to create several VMs on a physical server or host and therefore, lessens the hardware amount and enhances the resource utilization. The computing resource/node in cloud is used through the virtual machine. To address this problem, data centre resources have to be managed in resource -effective manner for driving Green Cloud computing that has been proposed in this work using Virtual machine concept with ABC and Neural Network optimization algorithm. The simulations have been carried out in CLOUDSIM environment and the parameters like SLA violations,
Energy consumption and VM migrations along with their comparison with existing techniques will be performed.

Keywords: Virtual machine, VM migrations, Artificial Bee Colony (ABC), artificial neural network (ANN).

I. INTRODUCTION

The concept of cloud computing is based on the basic term of reusability of information technology capabilities. Cloud computing allows several organizations and customers to use variety of applications without installing and accessing their personal records on any portable computer with web access. With virtualization, cloud computing has addressed the vast client base of heterogeneous computing with same type of physical infrastructure. Resources of computer, storage resources and the different applications can dynamically provide charge as per use and later can be released if not needed. Such services are offered with an agreement i.e. SLA, which gives the desired Quality of service (QoS) to their users. In order to fully understand the capabilities of cloud computing, cloud providers need to ensure that they can adapt to their virtual machine (VM) transport to meet different buyer prerequisites while keeping customers away from the basic data centre. Cloud allows multiple services to be hosted on globally shared resource pools, where resources are allocated to on-demand services. It uses a virtualized environment to run the service, because there is no virtualized computing being inefficient and inflexible. But, it has some service performance degradation, and energy costs and a lot of power consumption. Most violations occur during real-time migration of virtual machines that affects SLA parameters such as availability, response time, throughput, network bandwidth, and so on.

Therefore, it is necessary to develop a new method for SLA-aware energy-efficient algorithms for resource allocation in the data centre . There are three services in cloud computing, termed as, SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service). VM method is one of the mostly used technique in cloud computing, that is used for implementing these three services along with VM migration techniques that is used for the maintenance of virtualized cloud computing data centers. Minimizing energy consumption of data center is the main research area for keeping the quality of service. For minimizing power developers has proposed green cloud computing that is based upon cloud computing and virtualization concept? This helps in decreasing power consumption. In this research work, we focused on VM migration based on green computing. In data centers it is always a need to migrate VM for some reasons. Whenever a physical machine is not able to fulfil the entire requirement of the virtual machine then there came the requirements of the migration of the VMs or the borrowing spaces from other hosts. In the same contrast, a lot of researchers have put their effort in order to minimize the SLA violation against different set of threshold values. Algorithms like Genetic; Particle Swarm Optimization have already being utilized in this scenario. The aim of this research

work is to optimize the selection process of the physical machine for VM migration and to crosscheck the migration using Artificial Neural Network. Artificial Neural Network is a classifier technique and used to classify the free and best suitable task for all VMs according to their optimized data that are optimized with ABC optimization technique. Artificial Neural Network classifies free and per allocated VMs and allocates the task to free VMs using their properties. In the proposed work, artificial neural network is used to make better performance of simulation. By using the Artificial Neural Network, the energy consumption rate can be reduced during the allocation process.
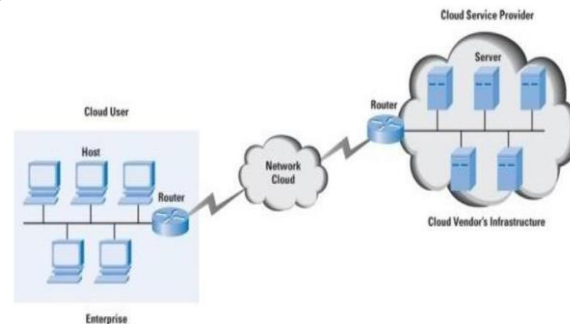

fig. 1 cloud architecture

II.   LITERATURE SURVEY

Bringing the cloud to the edge

Edge services become increasingly important as the Internet transforms into an Internet of Things (IoT). Edge services require bounded latency, bandwidth reduction between the edge and the core, service resiliency with graceful degradation, and access to resources visible only inside the NATed and secured edge networks. While the data center based cloud excels at providing general purpose computation/storage at scale, it is not suitable for edge services. We present a new model for cloud computing, which we call the Edge Cloud, that addresses edge computing specific issues by augmenting the traditional data center cloud model with service nodes placed at the network edges. We describe the architecture of the Edge Cloud and its implementation as an overlay hybrid cloud using the industry standard OpenStack cloud management framework. We demonstrate the advantages garnered by two new classes of applications enabled by the Edge Cloud - a highly accurate indoor localization that saves on latency, and a scalable and resilient video monitoring that saves on bandwidth.

The Case for VM-Based Cloudlets in Mobile Computing

Mobile computing continuously evolve through the sustained effort of many researchers. It seamlessly augments users' cognitive abilities via compute-intensive capabilities such as speech recognition, natural language processing, etc. By thus empowering mobile users, we could transform many areas of human activity. This article discusses the technical obstacles to these transformations and proposes a new architecture for overcoming them. In this architecture, a mobile user exploits virtual machine (VM) technology to rapidly instantiate customized service software on a nearby cloudlet and then uses that service over a wireless LAN; the mobile device typically functions as a thin client with respect to the service. A cloudlet is a trusted, resource-rich computer or cluster of computers that's well-connected to the Internet and available for use by nearby mobile devices. Our strategy of leveraging transiently customized proximate infrastructure as a mobile device moves with its user through the physical world is called cloudlet-based, resource-rich, mobile computing. Crisp interactive response, which is essential for seamless augmentation of human cognition, is easily achieved in this architecture because of the cloudlet's physical proximity and one-hop network latency. Using a cloudlet also simplifies the challenge of meeting the peak bandwidth demand of multiple users interactively generating and receiving media such as high-definition video and high-resolution images. Rapid customization of infrastructure for diverse applications emerges as a critical requirement, and our results from a proof-of-concept prototype suggest that VM technology can indeed help meet this requirement.

Migrate or not? Exploiting dynamic task migration in mobile cloud computing systems

Contemporary mobile devices generate heavy loads of computationally intensive tasks, which cannot be executed locally due to the limited processing and energy capabilities of each device. Cloud facilities enable mobile devices-clients to offload their tasks to remote cloud servers, giving birth to Mobile Cloud Computing (MCC). The challenge for the cloud is to minimize the task execution and data transfer time to the user, whose location changes due to mobility. However, providing quality of service guarantees is particularly challenging in the dynamic MCC environment, due to the time-varying bandwidth of the access links, the ever changing available processing capacity at each server and the timevarying data volume of each virtual machine. In this article, we advocate the need for novel cloud architectures and migration mechanisms that effectively bring the

computing power of the cloud closer to the mobile user. We consider a cloud computing architecture that consists of a back-end cloud and a local cloud, which is attached to wireless access infrastructure (e.g. LTE base stations). We outline different classes of task migration policies, spanning fully uncoordinated ones, in which each user or server autonomously makes its migration decisions, up to the cloud-wide migration strategy of a cloud provider. We conclude with a discussion of open research problems in the area.

Tactical Cloudlets: Moving Cloud Computing to the Edge
Soldiers and front-line personnel operating in tactical environments increasingly make use of handheld devices to help with tasks such as face recognition, language translation, decision-making, and mission planning. These resource constrained edge environments are characterized by dynamic context, limited computing resources, high levels of stress, and intermittent network connectivity. Cyber-foraging is the leverage of external resource-rich surrogates to augment the capabilities of resource-limited devices. In cloudlet-based cyber-foraging, resource-intensive computation and data is offloaded to cloudlets. Forward-deployed, discoverable, virtual-machine-based tactical cloudlets can be hosted on vehicles or other platforms to provide infrastructure to offload computation, provide forward data staging for a mission, perform data filtering to remove unnecessary data from streams intended for dismounted users, and serve as collection points for data heading for enterprise repositories. This paper describes tactical cloudlets and presents experimentation results for five different cloudlet provisioning mechanisms. The goal is to demonstrate that cyber-foraging in tactical environments is possible by moving cloud computing concepts and technologies closer to the edge so that tactical cloudlets, even if disconnected from the enterprise, can provide capabilities that can lead to enhanced situational awareness and decision making at the edge.
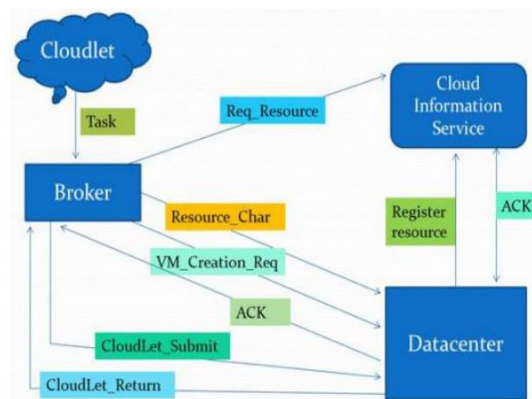

fig 2 cloud architecture

System model

Processing delay
We start by formulating Processing Delay based on the amount of VM servers hosted in each cloudlet. As a task processing environment, we assume that the task arrival rate at a single VM (i.e. the task arrival rate from a single user), which can be obtained from statistical observation,follows a Poisson process with rate λ. As mention before,each VM server only serves a single user. Therefore, given the number of hosted VMs in cloudlet ci, the average task arrival rate is given by the sum of the arrival rate at all individual VM servers [26], as shown below.

$$\Lambda_{c_i} = V_{c_i} \cdot \lambda.$$

Transmission delay
we will present a mathematical modelfor evaluating Transmission Delay. We utilize the Shannon-Hartley Theorem to approximate the channel capacity,, in both the uplink (user to cloudlet) and the downlink (cloudlet to user). We show the theorem here for ease of reference.

$$\Theta = B \cdot \log_2 \left(1 + \frac{S}{B \cdot N + I} \right)$$

III.   METHODOLOGY

The methodology is described in steps:

STEP 1: Start and design the simulation work frame.

STEP 2: Firstly, we Initialize VMs and define SLA parameters for VMs.

STEP 3: After that, Initialize Hosts and define features for the Host

STEP 4: For each VM in the allocation table at each host, initialize ABC algorithm

STEP 5: To apply Artificial bee colony algorithm for selection of possible physical machine.

STEP 6: Selection of appropriate machine from the output of ABC with help of artificial neural network. Evaluate the performance metrics.

STEP 7: Stop.

IV.   MODULES

- Edge Cloud Computing
- Service Delay minimization
- Resource management
- Virtualization

V.   MODULE DESCRIPTION

1.   Edge Cloud Computing

This service model makes it so the server environment is tailor made to user and an extension of the client environment, with its own local variables that change with the actions by the user and are different from user to user; this situation is favorable, for example, when you want to move the VM server to a different physical server, since this decision now depends only on whether or not it is better for a single user instead a group of users, which would obviously complicate the situation. Any task the user might have that is too demanding for its device gets sent through a wireless connection as a request to the corresponding VM server, which in turn processes the job and sends back its output, again wirelessly.

2.   Service Delay minimization

We propose an integration technique that utilizes both VM migration and Transmission Power Control to minimize the average Service Delay of the users. As mentioned before and mathematically explained in the previous section, the average Service Delay, i.e. the sum of the Processing Delay and the Transmission Delay, is a function of configurable parameters. Because of this, the main idea behind our proposal is finding the optimal values for these parameters in order to minimize Service Delay. In mathematical terms, this is reduced to solving the following optimization problem.

3.   Resource management

The decision of how often the procedure should be executed is an important one that should be considered case by case. There is definitely some overhead involved in the procedure, dominated by the time needed to migrate the VM servers in the last step, which may overpower the benefits brought by a lower Service Delay. For example dynamic scenarios benefit more from frequent execution of the proposal, while static scenarios are more suitable for low frequency execution. This relationship is not analyzed here because of how situation dependent it is, and also because our considered scenario is in a longer timescale (which would render the overhead delay irrelevant).

4.   Virtualization

The comparison of performances is done while varying the relevance of communication and computation, to show how parameters from both sides affect the Service Delay in the proposal and in the other methods. Moreover, the methods are also tested under three different applications, with different requirements. Through these scenarios, we will prove that the proposal is better suited to adapt to varying degrees of computation and communication burden, and that it is also more capable of dealing with different profiles of applications.

VI.   CONCLUSION

In this research work an artificial bee colony (ABC) optimization technique is proposed along with neural network for VM migrations in cloud computing.ABC algorithm is used for selecting the number of possible physical machines and the selection of appropriate machine from the output of ABC algorithm has been done by using Artificial neural network (ANN). The work is being simulated by using CLOUDSIM tool. As ABC algorithm has different alternatives, therefore, different variants are applied to VM migration problem. In this research work, Artificial Bee Colony (ABC) is applied and Performance of all these metrics has been compared with the existing work by means of energy consumption and number of migrations. From the research, it is being concluded that proposed work has performed better as compared to the existing work. The performance

parameters like SLA violation, Energy consumption, and number of migrations with respect to number of virtual machines have been measured. This research has shown that with the usage of VM, Number of migrations, energy consumption and SLA violation has been reduced.

VII.   FUTURE ENCHANCEMENT

In future to select the physical machines for VM migration different optimization algorithms like genetic algorithm along with particle swarm optimization (PSO) can be used in hybridization. The migrations can be checked by using ANN in conjunction with Fuzzy logic. This can be enhanced by adding methodsfor dealing with other dynamicity issues of ECC, such asthe susceptibility to failures (e.g. hardware problems, enemyattacks), the movement of the users, or the fluctuation in thedemands of the service model. Finally, heterogeneity couldbe added by accounting for more different applications inthe same scenario.

VIII.   REFERENCES

1. Zhang, Qi, Lu Cheng, and RaoufBoutaba. "Cloud computing: state-of-the-art and research challenges." Journal of internet services and applications vol.1, 2010, pp. 7-18.
2. Buyya, Rajkumar, Rajiv Ranjan, and Rodrigo N. Calheiros. "Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services." International Conference on Algorithms and Architectures for Parallel Processing. Springer, Berlin, Heidelberg, 2010.
3. Z. Xiao, W. Song and Q. Chen, "Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment," in IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 6, pp. 1107-1117, June 2013.
4. Qiang Li, QinfenHao, Limin Xiao, and Zhoujun Li, "Adaptive Management of Virtualized Resources in Cloud Computing Using Feedback Control," in First International Conference on Information Science and Engineering, April 2009, pp. 99-102.
5. Sarathe, Reena, Amit Mishra, and Shiv Kumar Sahu. "Max-Min Ant System based Approach for Intelligent VM Migration and Consolidation for Green Cloud Computing." International Journal of Computer Applications vol.136, 2016.