RESOURCE ALLOCATION IN MOBILE CLOUD COMPUTING: A REVIEW

C.Arun[1], K.Prabu[2]
[1]Research Scholar, PG & Research Department of Computer Science, Sudharsan College of Arts & Science, Pudukkottai, Tamilnadu, India.
[2]Associate Professor, PG & Research Department of Computer Science, Sudharsan College of Arts & Science, Pudukkottai, Tamilnadu, India.

Abstract- Mobile Cloud computing has become a new age technology. Mobile Cloud Computing (MCC) has been introduced as a possible solution to the inherited limitations of mobile computing. The combination of wireless communication, cloud computing and portable computing devices is called Mobile Cloud Computing. It can make it possible to access applications and associated data from anywhere. By using MCC, the processing and the storage of intensive mobile device jobs will take place in the cloud system and the results will be returned to the mobile device. Mobile cloud service providers are able to charge resources from cloud for storage and other computational purposes. But the mobile cloud computing have some issues resource poverty, like power consumption, and security. Resources in mobile cloud computing are radio and computing resources. These radio and computing resources is heart to the mobile cloud computing. Resource allocation is performed with the objective of minimizing the costs associated with it. Resources allocation means how much resources are allocated to the end user from available resources. Using the mobile devices for accessing the cloud it needs an efficient dynamic resource allocation for allocating the resources to the users.

Index Terms: Mobile computing, Cloud computing, Mobile device, Power consumption, Resource poverty, Security.

I. INTRODUCTION

The emergence of two different but important fields, mobile computing and cloud computing, has given birth to a new concept of mobile cloud computing (MCC) [1]. Mobile devices have the limitation of storage and processing power. In MCC, a mobile device is augmented by offloading its task and data into a resourceful cloud. A cloud is a rich collection of resources such as memory, storage, processing power, network, server, database, and applications. A cloud user employs these resources in a "pay as you use" or "elastic" manner. When the user sends a service request to the cloud, the cloud provider allocates the desired resource to the user. So, it is very important for the cloud provider to use a sound resource-allocation strategy to maintain the quality of service (QoS) of the cloud. The Mobile Cloud Computing Forum defines mobile cloud computing (MCC) as Mobile Cloud Computing at its simplest, refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. Mobile cloud applications move the computing power and data storage away from mobile phones and into the cloud, bringing applications and mobile computing to not just smartphone users but a much broader range of mobile subscribers [2]. MCC is basically the intersection of the fields of cloud computing and mobile networks. Mobile networks are basically networks that connect mobile users. The emergence of ultrafast mobile networks makes it necessary to bring the cloud domain to the mobile networking domain. This field is still in its primary stage of development. MCC basically enables the building and hosting of mobile applications over the cloud [3]. There are various resource sharing issues to be resolved such as live VM migration, and fault tolerance. There is a probability that migrating VMs become an overhead in MCC. Mobile devices usually have limited computing power and resources. In mobile applications that involve image processing, natural language processing, multimedia search, and so on, the lack of resources on the mobile device can be handled by renting services offered by the cloud in fig 1. Another scenario is where mobile devices can themselves become part of a cloud and offer their resources for rent to other mobile devices. Thus, collective resources can be made available, provided that they fall within the vicinity.

II. ADVANTAGES AND LIMITATIONS

There are many benefits in resource allocation while using mobile cloud computing irrespective of size of the organization and business markets. But there are some limitations as well, since it is an evolving technology. Let's have a comparative look at the advantages and limitations of resource allocation in cloud.

Advantages:

i. The next major benefit is that there is no limitation of place and medium. We can reach our applications and data anywhere in the world, on any system.
ii. The biggest benefit of resource allocation is that user neither has to install software nor hardware to access the applications, to develop the application and to host the application over the internet.

iii. The user does not need to expend on hardware and software systems.
iv. Cloud providers can share their resources over the internet during resource scarcity.

Limitations:

i. Peripheral devices like printers or scanners might not work with cloud. Many of them require software to be installed locally. Networked peripherals have lesser problems.
ii. Migration problem occurs, when the users wants to switch to some other provider for the better storage of their data. It's not easy to transfer huge data from one provider to the other.
iii. In public cloud, the clients' data can be susceptible to hacking or phishing attacks. Since the servers on cloud are interconnected, it is easy for malware to spread.
iv. Users rent resources from remote servers for their purpose; they don't have control over their resources.

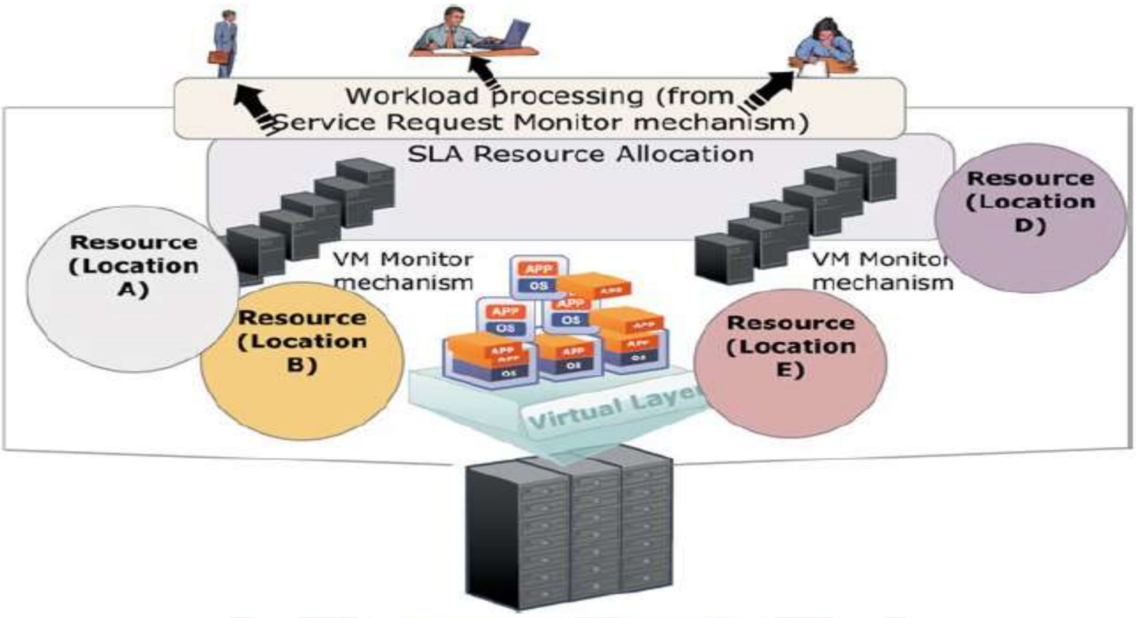

Figure 1. Resource Allocation in MCC

III. RELATED WORKS

L.Yang et. al., [4] proposed a framework to provide runtime support for the dynamic computation partitioning and execution of the application. Different from existing works, the framework not only allows the dynamic partitioning for a single user but also supports the sharing of computation instances among multiple users in the cloud to achieve efficient utilization of the underlying cloud resources. The framework has better scalability because it is designed on the elastic cloud fabrics. Based on the framework, they design a genetic algorithm for optimal computation partition.

To improves the QoS in mobile devices by choosing suitable cloud providers for a given user application M.H.Zarei et. al., [5] formulated the task allocation problem as a minimum nonlinear optimization problem and randomized algorithm was proposed in which the result is close to the optimal solution. This algorithm is faster and more scalable compared to optimal solution.

To solve the problem of resource allocation G.Wei et. al., [6] presented Game theory is used. A practical approximated solution proposed in two steps. First, each participant solves its optimal problem independently, without consideration of the multiplexing of resource assignments. A Binary Integer Programming method is proposed to solve the independent optimization. Second, an evolutionary mechanism is designed, which changes multiplexed strategies of the initial optimal solutions of different participants with minimizing their efficiency losses. The algorithms in the evolutionary mechanism take both optimization and fairness into account.

S.Vakilinia et. al., [7] proposed a model for wireless interfaces, mobile application profiles and cloud resources. First, an algorithm to allocate wireless interfaces and cloud resources has been introduced. The proposed model is based on the wireless network cloud (WNC) concept. Then, considering power consumption, application quality of service (QoS) profiles, and corresponding cost functions, a multi-objective optimization approach using an event-based finite state model and dynamic constraint programming method has been used to determine the appropriate transmission power, process power, cloud offloading and optimum QoS profiles. This algorithm save the mobile battery life and guarantees both QoS and cost simultaneously.

S.Kuribayashi et. al., [8] proposed to enhance the existing joint multiple resource allocation method, so as to provide the following two functions: (1)a function to take account of the total processing time of network delay and service processing time in allocating resources. (2) a function to prevent the degradation in service quality of other request types when requests that require a short network delay occur more than expected.

A novel resource allocation algorithm was proposed by Y.Liu et. al., [9] for secure mobile cloud computing systems. The mobile request for using cloud resource is classified according to its level of security requirement and the amount of required resource for remote computing. They formulated the resource allocation problem as a semi-Markov decision process under the average reward criterion through maximizing the long-term reward while meeting the system requirements of the blocking probability and the amount of resource requested with a security guarantee, the optimal resource allocation policy is calculated by using the linear programming.

W. Chen et. al., [10] introduced a heuristic algorithm to solve NP-hard problem. When no infrastructure is available, compared with the case when an application is executed solely on a single MD, this algorithm was reduced an application's response time significantly.

H.Raei et. al., [11] analyzed completely modeled three common resource allocation scheme (RAS), namely share-based scheme (SBS), reserve-based scheme (RBS), and hybrid-based scheme (HBS). The proposed models enable the cloudlet owner to properly decide which scheme is suitable for its conditions. The principal criteria for this decision are two important performance measures: request rejection probability and mean response delay. To model each scheme, an analytical performance model which consists of stochastic sub-models was proposed. The Markov Reward Model (MRM) is applied for obtaining the outputs of the sub-models.

M. Rahimi et. al., [12] proposed a novel framework to model mobile applications as a location-time workflows (LTW) of tasks; here user mobility patterns are translated to a mobile service usage patterns. M. Rahimi et. al., [12] proposed an efficient heuristic algorithm called MuSIC that was able to perform well, and scale well to a large number of users while ensuring high application QoS.

A.Jin et. al., [13] proposed a feasible and truthful incentive mechanism (TIM), to coordinate the resource auction between mobile devices as service users (buyers) and cloudlets as service providers (sellers). Further, TIM is extended to a more efficient design of auction (EDA). TIM guarantees strong truthfulness for both buyers and sellers, while EDA achieves fairly high system efficiency but only satisfies strong truthfulness for sellers. A.Jin et. al., [13] also show the difficulties for the buyers to manipulate the resource auction in EDA and the high expected utility with truthful bidding.

D.Ergu et. al., [14] proposed a model for task-oriented resource allocation in a cloud computing environment. Resource allocation task is ranked by the pairwise comparison matrix technique and the Analytic Hierarchy Process giving the available resources and user preferences. The computing resources can be allocated according to the rank of tasks.

D.Thai Hoang et. al., [15] formulated an optimization problem for dynamic resource sharing of mobile users in mobile cloud computing (MCC) hotspot with a cloudlet as a semi-Markov decision process (SMDP). SMDP is transformed into a linear programming (LP) model and it is solved to obtain an optimal solution. In the optimization model, the quality of service (QoS) for different classes of mobile user is taken into account under resource constraints.

Y.Ge et. al., [16] proposed a game-theoretic approach to optimize the overall energy in a mobile cloud computing system. Y.Ge et. al., [16] formulated the energy minimization problem as a congestion game, where each mobile device is a player and his strategy is to select one of the servers to offload the computation while minimizing the overall energy consumption.

Rakesh Kumar et. al., [17] proposed a model to considered the feedback sent by the mobile device and stored it in a directory maintained by the root server at cloud. The root server refers this directory while allocating the task to cloudlets. For the second problem they used the Gabriel architecture and crowd sensing framework collectively. The combination of these two quite efficiently processes the sensed information at the local level and passes the processed information to root server for decision making. Rakesh Kumar et. al., [17] also proposed metrics of yield factor for various parameters which will be calculated by the root server and based on those yield factor values root server can decide whether to scale up the cloud–cloudlet system or not.

H.Liang et. al., [18] proposed a service decision making system for inter-domain service transfer to balance the computation loads among multiple cloud domains. System focused on maximizing the rewards for both the cloud system and the users by minimizing the number of service rejections that degrade the user satisfaction level significantly. Also they formulated the service request decision making process as a semi-Markov decision process. The optimal service transfer decisions are obtained by jointly considering the system incomes and expenses.

A novel MCC adaptive resource allocation model H.Liang et. al., [19] was proposed to achieve the optimal resource allocation in terms of the maximal overall system reward by considering both cloud and mobile devices. To achieve this goal, they model the adaptive resource allocation as a semi-Markov decision process (SMDP) to capture the dynamic arrivals and departures of resource requests.

S.Barbaross et. al., [20] proposed a method to jointly optimize the transmit power, the number of bits per symbol and the CPU cycles assigned to each application in order to minimize the power consumption at the mobile side, under an average latency constraint dictated by the application requirements and consider the case of a set of mobile handsets served by a single cloud. The optimization leads to a one-to-one relationship between the transmit power and the percentage of CPU cycles assigned to each user. Also S.Barbaross et. al., [20] proposed a computation scheduling technique and verify the stability of the computations queue.

A.Beloglazov et. al., [21]Proposed energy-aware allocation heuristics provision data center resources to client applications in a way that improves energy efficiency of the data center, while delivering the negotiated Quality of Service (QoS).

IV. CONCLUSION

The combination of wireless communication, cloud computing and portable computing devices is called mobile cloud computing which permits users an online access. By using MCC, the processing and the storage of intensive mobile device jobs will take place in the cloud system and the results will be returned to the mobile device. Mobile cloud computing paradigm promises to be the ultimate outsourcing resources' and applications' solution starting from the individual users in small businesses scaling up to large enterprises and even governments, there are many resource consideration and constraints that should be addressed prior to the promulgation of services to the end-users. These resource considerations and constraints deal with the manipulation of resources requested by users and enable new mechanisms that face the QoS constraints and enable reliability in the offered services. In this paper, we discussed about previous of resource allocation models and different approaches. We demand to adapt the difficulties of the present approaches and overcome all the challenges of resource-allocation strategies to maintain proper QoS profitable deployment of mobile cloud services of the MCC system.

REFERENCES

[1]. J. S. Park and E. Y. Lee, Entropy-based grouping techniques for resource management in mobile cloud computing, Ubiquitous Information Technologies and Applications, 214, 773–780, 2013.
[2]. Y. Ge, Y. Zhang, Q. Qiu, and Y. H. Lu, A game theoretic resource allocation for overall energy minimization in mobile cloud computing system, in Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design, Rome, Italy, pp. 279–284, 2012.
[3]. D. Minarolli and B. Freisleben, Utility–based resource allocations for virtual machines in cloud computing, in IEEE Symposium on Computers and Communications, Kerkyra, Greece, pp. 410–417, 2011.
[4]. L.Yang, J.Cao, Y.Yuan, T.Li, A.Han, and A.Chan,"A Framework for Partitioning and Execution of Data Stream Applications in Mobile Cloud Computing", IEEE Fifth International Conference on Cloud Computing, DOI: 10.1109/CLOUD.2012.97
[5]. M.H.Zarei and M.A. Shirsavar," A QoS-aware Task Allocation Model for Mobile Cloud Computing", DOI: 10.1109/ICWR.2016.7498444
[6]. G.Wei , A.V. Vasilakos , Y.Zheng and N.Xiong," A game-theoretic method of fair resource allocation for cloud computing services", The Journal of Supercomputing November 2010, Volume 54, Issue 2, pp 252–269.
[7]. S.Vakilinia, D.Qiu and M.Ali," Optimal multi-dimensional dynamic resource allocation in mobile cloud computing", EURASIP Journal on Wireless Communications and Networking 2014.
[8]. S.Kuribayashi," Resource Allocation Method For Cloud Computing Environments With Different Service Quality To Users At Multiple Access Points", International Journal of Computer Networks & Communications (IJCNC) Vol.7, No.6, November 2015. http:// DOI: 10.5121/ijcnc.2015.7603 33.
[9]. Y.Liu and M.J. Lee," Security-Aware Resource Allocation for Mobile Cloud Computing Systems", International Conference on Computer Communication and Networks (ICCCN) DOI: 10.1109/ICCCN.2015.7288465
[10]. W. Chen, C.Lea and K.Li," Dynamic Resource Allocation in Ad-Hoc Mobile Cloud Computing", DOI: 10.1109/WCNC.2017.7925613
[11]. H.Raei and N.Yazdani," Analytical performance models for resource allocation schemes of cloudlet in mobile cloud computing", DOI 10.1007/s11227-016-1830-8
[12]. M. Rahimi, N.Venkatasubramanian, and A.V. Vasilakos," MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing",I EEE Sixth International Conference on Cloud Computing, DOI: 10.1109/CLOUD.2013.100
[13]. A.Jin, W.Song, P.Wang and P.Ju," Auction Mechanisms Toward Efficient Resource Sharing for Cloudlets in Mobile Cloud Computing", DOI 10.1109/TSC.2015.2430315
[14]. D.Ergu ,G.Kou,Y.Peng,Y.Shi and Y.Shi," The analytic hierarchy process: task scheduling and resource allocation in cloud computing environment", The Journal of Supercomputing June 2013, Volume 64, Issue 3, pp 835–848
[15]. D.Thai Hoang, D.Niyato, and P.Wang," Optimal Admission Control Policy for Mobile Cloud Computing Hotspot with Cloudlet", **DOI:** 10.1109/WCNC.2012.6214347

[16]. Y.Ge, Y.Zhang and Q.Qiu," A Game Theoretic Resource Allocation for Overall Energy Minimization in Mobile Cloud Computing System", DOI:dl.acm.org/citation.cfm?id=2333724
[17]. R. Kumar and S. Yadav," Scalable Key Parameter Yield of Resources Model for Performance Enhancement in Mobile Cloud Computing", DOI 10.1007/s11277-017-4035-4
[18]. H.Liang, , L.X. Cai, , D.Huang, XShen, and D.Peng," An SMDP-Based Service Model for Inter domain Resource Allocation in Mobile Cloud Networks",I EEE Transactions on Vehicular Technology Volume: 61, Issue: 5 , Jun 2012.
[19]. H.Liang, T.Xing, L.X. Cai, D.Huang, D.Peng, and Y.Liu," Adaptive Computing Resource Allocation for Mobile Cloud Computing", International Journal of Distributed Sensor Networks, Volume 2013.
[20]. S.Barbaross, S.Sardellitti, and P.D.Lorenzo," Joint allocation of computation and communication resources in multiuser mobile cloud computing", DOI: 10.1109/SPAWC.2013.6612005
[21]. A.Beloglazov, J.Abawajy and R.Buyya ,"Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing", Future Generation Computer Systems 28 (2012) 755–768

Authors Profile

Mr. C. Arun received his M.Sc degree from Bharathidasan University, Trichy, Tamilnadu and M.Phil from Bharathiyar University, Coimbatore, Tamilnadu, India. He is doing his Ph.D in Sudharsan College of Arts & Science, Pudukkottai, Tamilnadu, India. His Research interested is Mobile Computing and Cloud computing research area. He has published more than 12 technical papers at various National / International Conferences and Journals.

Dr. K. Prabu received his MCA and M.Phil from Annamalai University, Chidambaram, Tamilnadu, India. He received his Ph.D Degree in Computer Applications from Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India. He is now working as Associate Professor in PG & Research Department of Computer Science, Sudharsan College of Arts & Science, Pudukkottai, Tamilnadu, India. He is a Reviewer of 06 National/International Journals. His research interested is Adhoc Networks, Wireless & Mobile Computing, and Wireless Sensor Networks. He has published more than 50 technical papers at various National / International Conferences and Journals. He is a life member of ISTE, IACSIT, and IAENG.