FINITE HORIZON MARKOV DECISION PROCESS BASED FUZZY OPTIMIZATION FOR RESOURCE ALLOCATION IN SDN ENABLED VIRTUAL NETWORKS IN IAAS CLOUD ENVIRONMENT

G. SENTHIL KUMAR[1], Dr.M.P. CHITRA[2]

[1]Research Scholar, Sathyabama Institute of Science and Technology, Chennai, India[1]

[2]Professor and Head, Department of ECE, Panimalar Institute of Technology, Chennai, India[2]

[1]gsenthilkumarphd@gmail.com, [2]chi_mp2003@yahoo.co.in

ABSTRACT

Network technologies are dealing with a massive urge to break through the fundamental endorsements of networks. Software-Defined Networking (SDN) has been leading cloud Data Centres (DC's) which states with different policy adaptation for ensuring resource management, concerning about Network Virtualization (NV) performance that is capable of finding the related hardware components to map a Virtual Machine (VM) or a virtual link which represents Virtual Network Embedding (VNE) problem. To overcome the VNE problems our work as proposed a Finite Horizon Markov Decision Process Based Fuzzy Optimization. Fuzzy inference provides an linguistic variables and set of rules to obtain best policy from the available cloud resource and predicts the execution cost for every network function virtualization. This stage also deals with uncertainities and imprecision.Based on the priority and schedulable ability the Finite Horizon Markov Decision Process dynamically allocates the resource for NFV components. Thus, our work obtains a substantial amount of energy utilization by optimizing the use of local host services and will therefore provide greater policy control for physical DCs.

Keywords: *Software- Defined Networking (SDN), Network Virtualization (NV), Virtual Machine (VM), Virtual Network Embedding (VNE), Finite Horizon Markov Decision Process, Fuzzy Optimization, Resource Allocation.*

1. INTRODUCTION

Cloud computing increases the popular computing paradigm. Some cloud-based providers are accessible where a dynamic and virtualized distributed data, power and network services are provided to customers over the Internet at a discount while using the method [1]. The combination of machine learning and cloud computing reinforce their association and improve overall efficiency of resource provide in the IaaS service. Machine Learning could soon become an integral part of could environment (learning the behaviours of human interaction with resource-learning from the past, present and future) that could improve its functional aspects. The intelligent could collecting more resource of data from cloud user and cloud provider bring about transformational change in the existing technological framework. Using private clouds give customers the option of providing information centres tailored to the internal requirements of a particular business group and public clouds accessible to the general public on the Internet. Network concert and resource obtainability can be the close-fitting bottleneck for any cloud. It is seen as an opportunity for network service providers who are building their own clouds using distributed cloud architecture. Services are offered below numerous deployment models such as Network as a Service, Infrastructure as a Service, Software as a Service and Platform as a Service [2].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The CC Data Centre Networks (DCN) resource allocation and scheduling system handles all cloud supplier DCN services and manages customer requests, state the RA, meet QoS terms on the network, and reduce performance differences while reducing the service provider costs and monitoring the amount of power consumed [3]. Virtualization provides a promising approach of consolidating multiple online services in fewer computing resources within a data center. This technology allows a single server to be shared among many performance-isolated platforms called virtual machines (VMs). Virtualization also allows the on-demand or utility computing a just–in–time resource provisioning model in which computing resources such as CPU, memory, and disk space are made available to applications only as needed and not allocated statically based on the peak workload demand. [4].

A key feature of the future Internet would be the allocation of substrate network resources to the virtual networks requested. The key area of network virtualization is therefore Virtual Network Embedding (VNE). Each virtual node is mapped to a physical node, and each virtual connection is built into a substrata route to incorporate / assign / map a virtual network in the substratum / physical network. Based on the efficiency of the substrate and the specifications of the VNs, a set of Virtual Networks (VNs) can be installed above the physical network (or substrates).

The NP-hard (9, 10) virtual network incubator problem requires fulfilment from different constraints [5]. Therefore, VNs can be accessed from any standard network topology in any order at different times. There are also limited resources in the substratum network. Therefore, a VN with resource limitations must be incorporated or mapped to a Substrate Network (SN) with Finite resources [6]. In order to support these types of decisions, we present a Finite Horizon Markov Decision Making Process (FHMDP) based on a Fuzzy optimizer model. To adding the machine learning algorithm to a stage where there will be no human intervention necessary. The model we proposed to validate learn from data is seen as the bright future of science and being the next level of evolution in automation. The power of cloud computing and machine learning incorporate into termed “the intelligent cloud”.

The remainder of this paper provides technical details on the problem formula and the MDP model, including some cost modelling specifics, and summarizes some calculation tests. Rest of the paper is organized as follows. Section 2 defines the problem related works while Section 3 explains the proposed methodology. Section 4 presents simulation results. Section 5 concludes the paper.

2. RELATED WORKS

Adil Razzaq et al (2012) [7] Proposed a new vertex mapping method which reduces the entire exhaustion of substrates nodes while ensuring a good total use of resources. An outcome shows that BLA supports densely linked media, whereas GNM provides improved results for sparsely connected substrates. On the other hand, for both sparsely and densely connected substrates, HBNRM gives almost similar or better VN maps compared to BLA and GNM.

Ines HOUIDI et al (2008) [8] analysed the distributed VN mapping algorithm therefore planned, implemented and evaluated. The substrates nodes manage independent and intelligent agents that exchange messages and collaborate in order to implement the proposed algorithm. In order to transmit and exchange messages between agent dependent substratum nodes a VN Mapping Protocol was used. Results from the distributed mapping performance assessment indicate that viable solutions can be found to the challenge of developing VNs from a common physical network.

Siva Theja Maguluri et l (2012) considered a load balancing stochastic model and cluster planning. The development of non-pre-emptive VM configuration policies based on a frame is a main contribution. Simulations show that long frame life is good from the perspective of output as well as good delay [9]. Resource allocation, scheduling and load balancing are the three important things which improve the quality of service in cloud computing. Applications from the same user may need to interact and change data or information, thus, we may abstract the applications/services request from same user as a virtual network (VN). To improve performance and resource efficiency, it is critical that the VN request be optimally provisioned given the current resource state of the datacenters [10].

To do this and choose an optimum resource, optimum schedule and through this balancing the load can be obtained using ABC-SA method. The main contribution is to implement a hybrid optimization algorithm by integrating the functionality of Simulated Annealing (SA) into Artificial Bee Colony (ABC) algorithm to do the efficient scheduling according to the task size, priority of the request and closest distance between client nodes to a server in the cloud environment [11].

The computation offloading problem in a mixed fog/cloud system by jointly optimizing the offloading decisions and the allocation of computation resource, transmit power and radio bandwidth, while guaranteeing user fairness and maximum tolerable delay. This optimization problem is formulated to minimize the maximal weighted cost of delay and Energy Consumption (EC) among all UEs, which is a mixed-integer non-linear programming problem. Due to the NP-hardness of the problem, we propose a low-complexity suboptimal algorithm to solve it, where the offloading decisions are obtained via semidefinite relaxation and randomization and the resource allocation is obtained using fractional programming theory and Lagrangian dual decomposition [12].

The cloud computing environment provides infinite number of computing resources such as CPU, memory and storage to the users in such a way that they can dynamically increase or decrease their resources and its use according to their demands. In resource allocation model having two basic objectives as cloud provider wants to maximize their revenue by achieving high resource utilization while cloud users want to minimize their expenses while meeting their requirements. However, it is essential to allocate resources in an optimized way between two parties. In some situations, single cloud may not satisfy all the requirements of the users. To achieve this objective, two or more cloud providers cooperatively work together to satisfy the user's requirements. These cooperative cloud providers should share and optimize the computational resources in a reasonable technique to make sure that no users get much resource than any other users and also improve the resource utilization [13][14].

These different scheduling algorithms are used based on the number of tasks and number of virtual machines in multi cloud environment architecture. If the number of tasks is equal to number of virtual machines ELB (Equal Load Balanced) scheduling algorithm is used [15].

3. PROBLEM METHODOLOGY AND SYSTEM MODEL

3.1. Resource Allocation in SDN Virtual Network in IAAS Cloud Computing

Cloud computing provides on-demand resources and removes the boundaries of resources' physical locations. By providing virtualized computing resources in an elastic manner over the internet, IaaS providers allow organizations to save upfront infrastructure costs and focus on features that discriminate their businesses. The growing number of providers makes manual selection of the most suitable configuration of IaaS resources, or IaaS services, difficult and time consuming while requiring a high level of expertise [16]. If it provides good results than it is awarded otherwise it is further iterated to obtain the accurate result.

Performance of the cloud infrastructure is highly depending upon the task scheduling and load balancing. Therefore, number of load balancing algorithms and technique are proposed by researchers throughout the world whose aim is to distribute the workload fairly among all the virtual machine while attaining the objective [17].

The resource allocation plays major role in cloud computing, but fluctuating demand and dynamic nature of resources are affecting the performances are efficiency and scalability. In this paper, we propose a Hybrid Optimization Algorithm for Resource Allocation (HOA-RA) in IaaS cloud environment with Software Defined Network (SDN) enabled virtual networks. The first contribution is to introduce a Social Cognitive Optimization (SCO) algorithm for SDN structure, which imparts flexibility to remove the multiple layers in control plane. Then, to combine the network virtualization principles with Cuckoo Search Optimization (CSO) algorithm to share physical framework to connect with unmistakable master relationship to streamline booking. The flexible access requires gainful Natural Language Processing algorithm (NLP) to improve the system assets; the SDN control plane utilized for proficient association of virtual structures [18].

3.2. System Model

The proposed model provides a solution for resource allocation in cloud computing under VNE problems. The System Model has 3-layer

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

architecture namely Cloud User usually referred as Clients, Resource Allocation Controller a Data Centre Servers. The Resource allocation Controller is responsible for Virtual Machine (VM) reservation, VM Monitoring, Service request controller and management, Service Pricing, Energy consumption monitoring. Main scope of the work is to provide a prediction-based scheduling so as to utilize the finite resource present in substrate network efficiently.

Finite Horizon Markov Decision Process Based Fuzzy optimization is implemented on service request controller and management module. Initially we obtain a schedule plan of the resource with the help of fuzzy optimization based on uncertainties may be from user or network issues etc than based on that time series analysis is done by the FHMDP and with the help of schedule plan it provides a decision to achieve the allocation of resources to the Cloud Users / Clients when they are requesting. Figure 1shows this proposed system model.

Two basic prerequisites for running cloud efficiently and cost effectively – scalable and low-cost resources (computing and storage) **Cognitive Knowledge**: The millions of processes that happen every day, all provide a source of information for the machine to learn from. The whole process will provide applications in the cloud with sensory capabilities. The applications will be able to perform cognitive functions and make decisions.

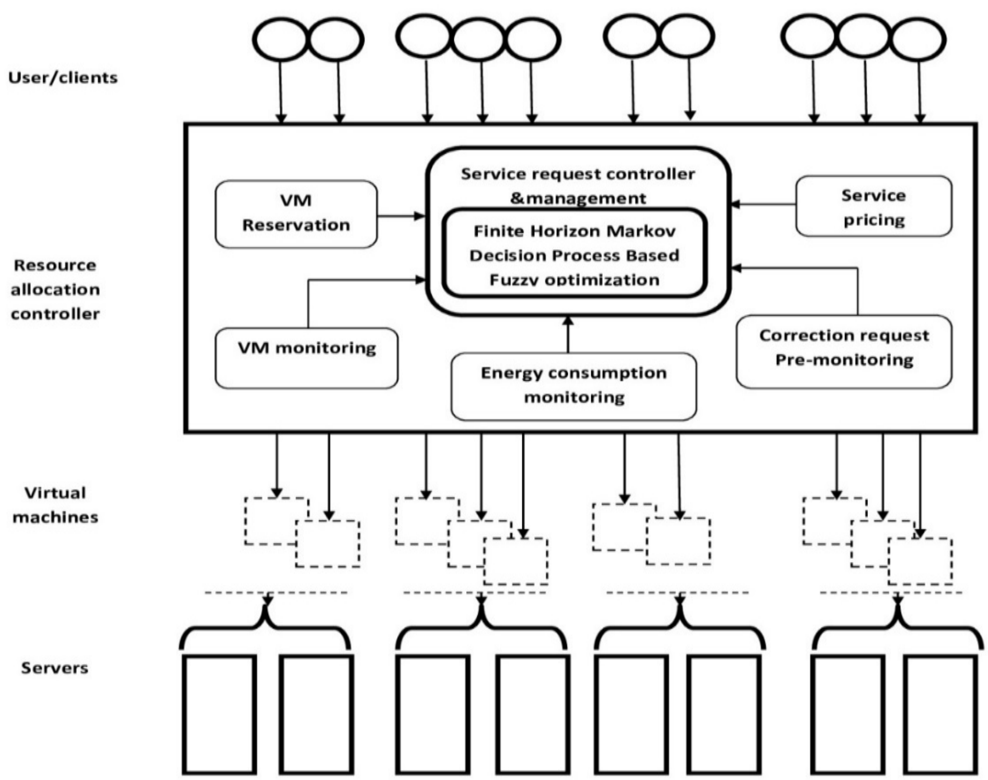

Figure 1: Finite Horizon Markov Decision Process Based Fuzzy Optimization For Resource Allocation In SDN Virtual Network

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Virtualization platform with fuzzy inference provide an linguistic variables and set of rules to obtain best policy from the available cloud resource and predicts the execution cost for every network function virtualization. Two steps criteria used:

a) Fuzzy real time resource allocation optimization controller

b) Finite Horizon Markov Decision Process for Scheduled Plan

3.3. Fuzzy real time resource allocation optimization controller

Fuzzy real time optimization provides a cost minimization and maintaining the schedule planform the user and predicting the plan under various circumstances. The detailed step is provided below:

Step 1: At a discrete time period r we are going to predict the future output of the resource based on the users priority at a prediction horizon h_p. Considering the prediction based outputs such as $\hat{x}(r+i|r)$, for i=1….h_p is determined by the equation stated above in equation ,now depending upon the previous requirement or demand of the RE may be at time instant rand the corresponding predictive output may be given as V(r+i|r) for i=1…h_c, where h_c indicates the control horizon verifying $h_c \leq h_p$

Step 2: Through increasing the output index, the optimal control sequence is calculated. Often this is a quadratic function that involves future control entries and tracking errors by:

$$e(r+i|r) = x_d(r+i|r) - \hat{x}(r+i|r) \quad (1)$$

Where $x_d(r+i|r)$ indicates the reference sequence to be followed that has to be known as priori. The cost function I is given in eqn below:

$$i = \sum_{i=1}^{h_p} \|e(r+i|r)\|_Q^2 + \sum_{j=1}^{h_C} \|v(r+i|r)\|_R^2 \quad (2)$$

Where $\|x\|_M^2 = x^T Mx. Q \; and \; R$ defines positive weighing matrices, $Q \geq 0$ and $R \geq 0$. The J minimization is often carried out subject to certain control input constraints and status variables. The limitations can be defined in the following form [11]:

$$\mathbb{U} = \{v \epsilon R^{n_v \times 1} | v_{min} \leq v \leq v_{max}\} \quad (3)$$

$$\mathbb{X} = \{x \epsilon R^{n_x \times 1} | x_{min} \leq x \leq x_{max}\} \quad (4)$$

V_{max}, v_{min}, x_{max}, x_{min} are known to be constant vector. $\mathbb{U}$ and $\mathbb{X}$ is said to be constraint subsets fuzzy optimization aim towards obtaining an output response closest to the referred response such as [v (r | r), v(r+1|r) ….v(r+h_c-1) |r)].

Step 3: Only in the next sample period of the system the first sample of the calculated optimal control sequence [k, k+1] is applied.

Step 4: At every sampling point all the previous steps are repeated.

Optimization of the uncertainties of the model stated in step 3 finds out the solution by inhabiting fuzzy optimization.

Depending on the problem form that is modelled, the fuzzy sets with the assumed membership functions identify only certain variables. The membership function includes lower and high limits along with a purely monotonous function that decreases continuously.

3.4. Finite Horizon Markov Decision Process for Scheduled Plan

A plan can be described as the collection of chances of a particular State action, i.e., π = {Q (b|T)} $_{b, T}$ which is stated for each and every0020state with corresponding action pairs. For maximizing the state value, the policy π is assumed to be most favourable, i.e., π^* = arg max U_π (T), ∀T. Therefore, in order to find the optimum strategy, the optimum status function is necessary U^* (T) = max V_π(T), so best action is considered at every state [12]. Defining the optimum function of the action value P∗ (T, b) = max P_π (T, b), from (2) and (3), so optimal function of state-value function is represented as

$$U^*(T) = \max_b P^*(T,b) = \max_b \mathbb{E}[r_t + \gamma U^*(T')|T,b] \quad (5)$$

The idea of optimal state-value U^* (T) stimulates the search for optimal policies considerably. Considering that the goal of optimizing future benefits for the successor state is already presented to the best value, U^* (T,0) is excluded from the requirement in (4). The best local acts in each state therefore ensure optimal strategy,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$b^* = arg \max_b \mathbb{E}[r_t|T,b] + \gamma U^*(T')|T,b] \quad (6)$$

In our question, first the consumer arrives, then we decide to work or wait (refer the cloud), which means that at decision-making time the reward u and the reward −η are established. The optimal operation at State S_n is therefore given by (5)

$$b_n^* = \begin{cases} serve \; if \; z + \gamma U^*(T_{n+1}) > -\eta + \gamma U^* * (T_n) \\ wait \; otherwise \end{cases} \quad (7)$$

Where $h_n = -\eta + \gamma [U^*(T_n) - U^*(T_n+1)]$ Denotes the state T_n maximum threshold.

The value iteration technology can be used for the optimum position values needed by the optimal policy as shown in (7).

The entire Finite Horizon Markov Decision Process for Scheduled Plan, work as follows: First cloud resources are identified and predicting the parameters such as Execution Cost for every Network Functional Virtualization (NFV), and then Energy Cost is calculated. Second Cluster the NFV components into Asynchronous partitions. After cluster the resource components obtain predictive plan for making reservation with local minima point of each partition. If any uncertainties, scheduling will be changed by Fuzzy interference and Decision based on the priority. This entire process is shown in Figure 2.

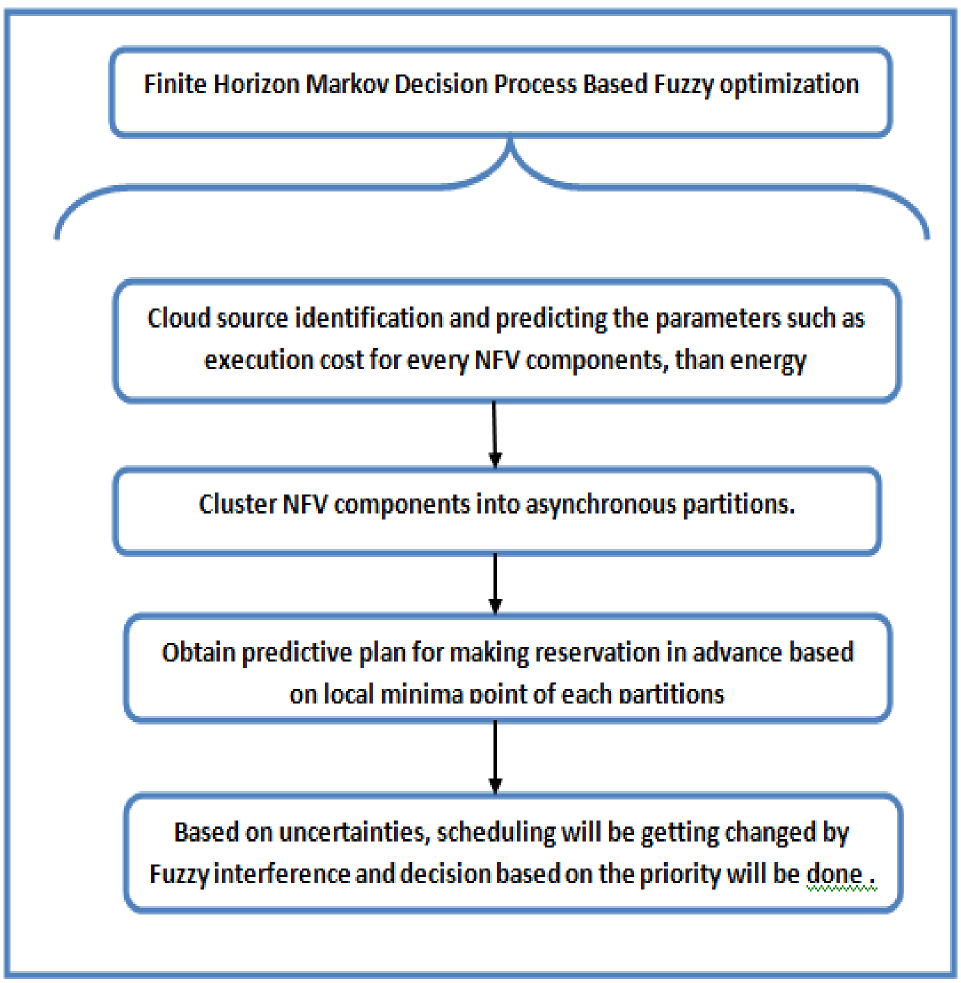

Figure 2: Finite Horizon Markov Decision Process Based Fuzzy Optimization

The approach used by the Monte Carlo algorithm for learning optimal strategy is in algorithm 3.4.1. Since we concentrate in this paper on supporting the FHMDP approach against the traditional fixed-threshold approach, the choice of FHMDP method has not been elaborated. The Monte Carlo method is chosen in this work because of its quick and fast convergence under the law of many. Many Other approaches based on RL were used to learn the best policy, but the results of the contrast with the fixed-threshold approach were not changing at any cause. Notice that as the dimensionalities of the state space can be monitored in the problem under consideration, the tabular RL methods work well; therefore, functional methods such as gradient-descent and profound RL methods are not needed.

3.4.1. Algorithm for Finite Horizon Markov Decision Process (FHMDP)

1. Initially opt for: $\gamma \in [0, 1]$, $\eta \in R$;
2. Input: Returns(S): is an array for all iterations to save the returns of states;
3. Initiate: U (T) ← 0, ∀T;
4. for loop = 0, 1, 2 ... do
5. Initialize: T ← T_0;
6. Generate an event: until the loops stops take action;
7. W (T) is the amount of reduced incentives for all states in the events from T to terminal status.;
8. Add W (T) to proceeds (T);
9. U (T) ← standard (proceeds (T));
10. if U (T) converge for all T then
11. break
12. U^* (T) ← U (T), ∀T;
13. end if
14. end for
15. Use the estimated U^* (T) to find optimal actions using

4. RESULTS AND DISCUSSIONS

Our work has undergone multiple review and comparative study to determine its quality. This section demonstrates the efficiency of our method. Automated resource provisioning techniques enable the implementation of elastic services, by adapting the available resources to the service demand. This is essential for reducing power consumption and guaranteeing QoS and SLA fulfilment, especially for those services with strict QoS requirements in terms of latency or response time, such as web servers with high traffic load, data stream processing, or real-time big data analytics. Elasticity is often implemented in cloud platforms and virtualized data-centres by means of auto-scaling mechanisms. These make automated resource provisioning decisions based on the value of specific infrastructure and/or service performance metrics.

The QoS levels agreed between the service provider and user are defined by means of Service Level Agreements (SLAs), in such a way that service level failures can result in cost penalties for the service provider and a potential loss of clients. On the other hand, service elasticity enables power consumption to be reduced, by avoiding resource overprovisioning. Over-provisioning is a typical and simple solution adopted by many service providers to satisfy peak demand periods and guarantee QoS during the service lifetime. However, this results in a waste of resources that remain idle most of the time, with the consequent superfluous power consumption and CO_2 emissions. The use of service elasticity mechanisms enables a reduction in the number of resources needed to implement the service and, along with other efficient techniques for server consolidation, virtual machine allocation, and virtual machine migration, it can lead to important energy savings for the datacentre or cloud provider [19].

We equate FHMDP-Fuzzy optimization methods with three algorithms, which provide heuristics of web applications and online services for the complex resource use of VMs with workloads. The main purpose of these algorithms is to prioritize data both on top and below, so that the complete use of the CPU of a nodule connecting these limits is retained. When the upper boundary is reached, VMs for load balance are rescheduled and when the host is used below the lower line, VMs for consolidation are rescheduled. The algorithms dynamically adapt the consumption threshold on the basis of a MAD, Interquartile Range and Local Regression (LR) move towards to estimate the use of the CPU. In addition, we also consider the THR process, which tracks the use of CPUs and migrates VM if the current use is greater than 80 percent of the total CPU capacity available on the host. The proposed FHMDP base Fuzzy Optimization solution focused on the allocation of resources to SDN Networking.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.1. Average SLA violation percentage

This calculation reflects the percentage that was not allocated to a request for average CPU output, resulting in a declining performance [5]. Equation (8) is determined as part of the discrepancy in the total lifetime of the MIPS, where n is the VMs number, between the required MIPS and the real MIPs.

$$SLA = \frac{\sum_{i=1}^{n} \int U_r(t) - U_r dt}{\sum_{i=1}^{n} \int U_r(t) dt} \quad (8)$$

The table shows FHMDP-based Fuzzy Optimizer, Threshold, Mean Absolute Deviation, IQR and linear regression methods for random workload breaches of SLA.

Our work is more effective than other methods to reduce the percentage of the violation rate of SLA. The results obtained can be explained by the fact that the FHMDP-based Fuzzy Optimizer by the overloaded prediction escapes the SLA breach. Therefore, by considering the current resource requirement, the SLA breach can be reduced.

The Table show the average Service Level Agreement (SLA) present in random working users. The FHMDP-Fuzzy Optimization is discontinuity the overloaded prediction comparing with THR, MAD, IQR an LR methods. And the Figure 3 show the comparison of the methods.

Table 1: Average SLA Present In The Random Working Users

FHMDP-Fuzzy Optimization	THR (%)	MAD (%)	IQR (%)	LR (%)
8.45	11.89	9.89	9.99	13.98

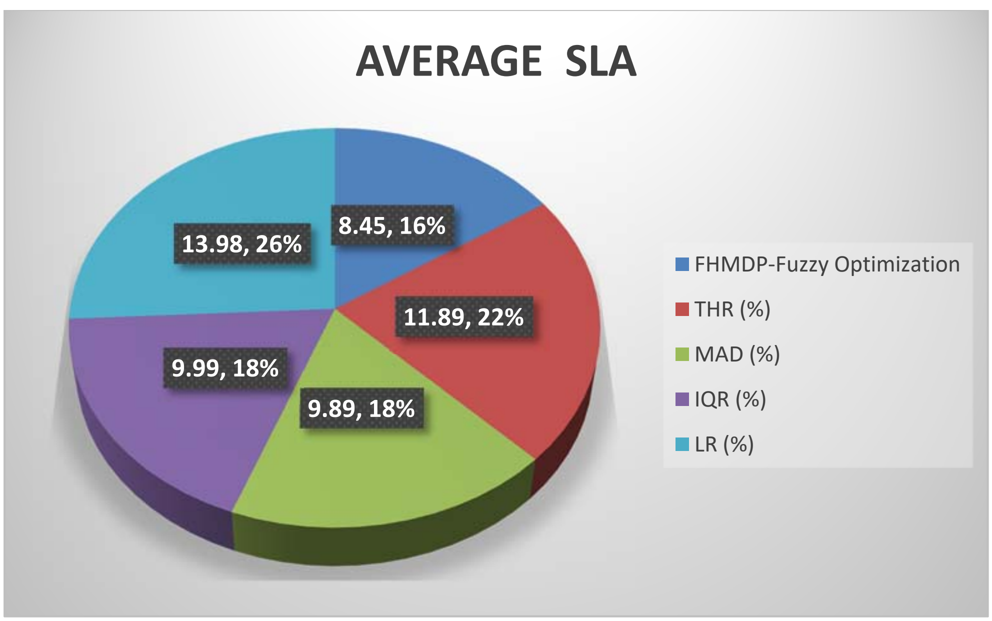

Figure 3: Average SLA Present In Random Working User

4.2 Average Energy Consumption and Comparison

In addition to that our work has experimented based on the map created connecting the virtual and the physical machine. The experiment has chosen the following 3 dissimilar policy: Non-Power Aware Policy (NPAP) generates the most CPU resources for the virtual machine in the physical machine; Central processing unit Disk Policy (CDP) sees CPU as the judge standard for the physical engine to select;

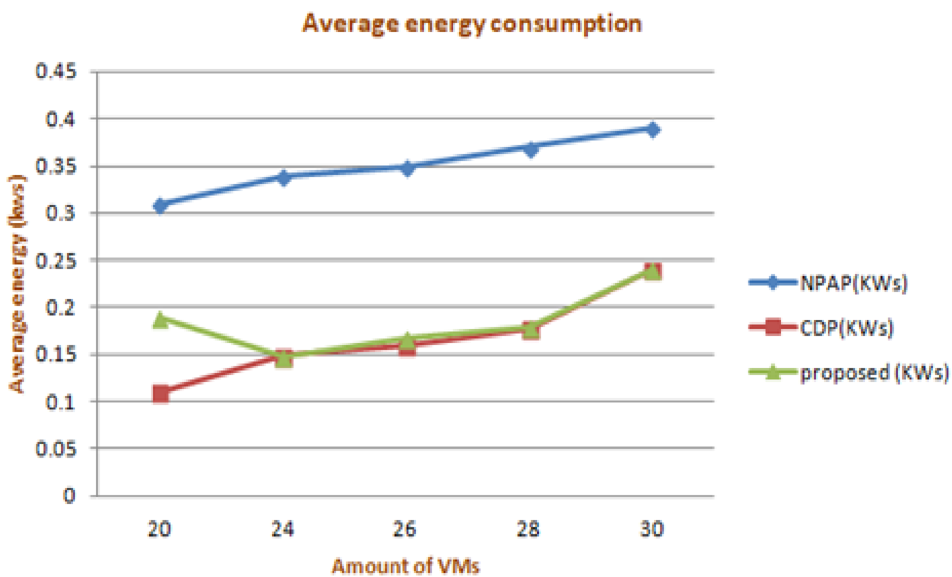

Figure 4: Comparison Of Average Energy Consumption

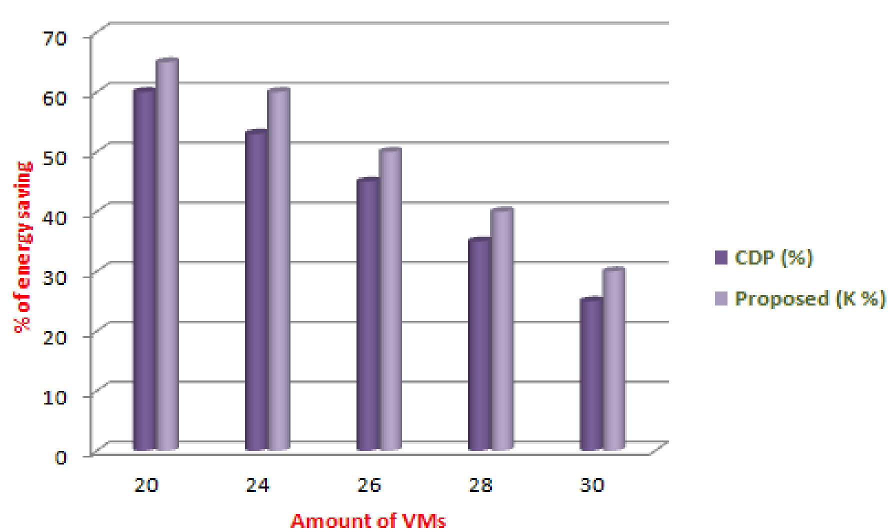

Figure 5: Comparison Between Reductions Of Energy

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

How to reduce the massive amount of energy consumption in cloud computing data centres. To address this issue, many power-aware Virtual Machine (VM) allocation and consolidation approaches are proposed to reduce energy consumption efficiently. However, most of those existing efficient cloud solutions save energy cost at a price of the significant performance degradation. A novel VM allocation algorithm called" PPRGear", which leverages the Performance-to-Power ratios for various host types. By achieving the optimal balance between host utilization and energy consumption, PPRGear is able to guarantee that host computers run at the most power-efficient levels (i.e., the levels with highest Performance-to-Power ratios) so that the energy consumption can be tremendously reduced with little sacrifice of performance. Our extensive experiments with real world traces show that compared with three baseline energy-efficient VM allocation and selection algorithms, PPRGear is able to reduce the energy consumption up to 69.31% for various host computer types with fewer migration and shutdown times and little performance degradation for cloud computing data centres [20]. The average power usage of the data centre is shown in figure 5. This change is primarily due to increasing energy consumption due to the increasing load of the digital system.

5. CONCLUSION

This paper addresses the efficient Resource Allocation and Energy consumption in SDN enabled Virtual Network in IaaS cloud environment by proposing a physical machine scheduling strategy based on Finite Horizon Markov Decision process based Fuzzy Optimization. Our proposed work provides a solution to solve the incorporation of the virtual network problems such as NP-hard where several constraints have to be satisfied. Constraints that may be in terms of Capacity, Location, Bandwidth, Link propagation delay etc which are taken into consideration for solving VNE problems. This uses Finite Horizon Markov Decision Process to study the strategy of host power mode detection without prior environmental and workload awareness. Our system can also adjust the amount of dynamic hosts to the current requirements for resources. Experimental results show that under all possible condition proposed algorithm reduce the CPU energy consumption using 3 policy Non-Power Aware Policy (NPAP) generates the most CPU resources for the virtual machine in the physical machine; Central processing unit Disk Policy (CDP) increase the average resource utilization. Thus our model is capable of minimizing the energy cost, consumption as well as SLA violation rate efficiency

REFERENCES

[1] M. Abu Sharkh, M. Jammal, A. Shamis, "Resource Allocation in a Network-Based Cloud Computing Environment: Design Challenges", IEEE Communications Magazine, 0163-6804/1, November 2013.

[2] Wei, Athanasios V, Yao Zheng, Naixue Xiong and Athanasios V. Vasilakos. "A game-theoretical equal resource distribution for cloud computing services.", The supercomputing journal 54, No. 2: 252-269, 2010.

[3] Mamata Rath "Resource provision and QoS support with added security for client-side applications in cloud computing". Intenatinal Journal of Information Technology, volume 11, pages357–364, 2019.

[4] Leontiou, Nikolaos, Symeon Papavassiliou, and Dimitrios Dechouniotis, Spyros Denazis, "Hierarchical load management control system and the allocation of resources to cloud computing services.", Computers & Electronics Technology 67 : 235-251, 2018.

[5] Mohamed Abu Sharkh, Abdelkader Ouda and Abdallah Shemi, "The Cloud Computing Data Center System for the Department of Electrical & Computer Engineering", Western University London, 2013.

[6] Shalabh Bhatnagar Emmanuel Fern'andez-Gaucherand Michael C. Fu Ying He Steven I. Marcus, "A Markov decision process model for capacity expansion and allocation", 38th IEEE Conference on Decision and Control, 1999.

[7] Adil Razzaqu, Markus Hidell and Peter Sjodin, "Virtual Network Embedding: A Hybrid Vertex Mapping Solution for Dynamic Resource Allocation", Journal of Electrical and Computer Engineering Volume 2012, Article ID 358647, 17 pages, 2012.

[8] Ines HOUIDI, Wajdi LOUATI, Djamal ZEGHLACHE, "A Distributed Virtual Network Mapping Algorithm", Communications, 2008. ICC '08. IEEE International Conference on, 2008

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[9] Siva Theja Maguluri and R. Srikant "Heavy Traffic Optimal Resource Allocation Algorithms for Cloud Computing Clusters", amc, arxiv, 206.1264v1,2012

[10] G. Sun, V. Anand, H. Yu, D. Liao and L. Li, "Optimal provisioning for elastic service oriented virtual network request in cloud computing," 2012 IEEE Global Communications Conference (GLOBECOM), Anaheim, CA, pp. 2517-2522, 2012.

[11] Muthulakshmi, B., and K. Somasundaram. "A hybrid ABC-SA based optimized scheduling and resource allocation for cloud environment." Cluster Computing 22, no. 5: 10769-10777, 2019

[12] Du, Jianbo, Liqiang Zhao, Jie Feng, and Xiaoli Chu. "Computation offloading and resource allocation in mixed fog/cloud computing systems with min-max fairness guarantee." IEEE Transactions on Communications 66, no. 4 : 1594-1608,2018

[13] Das, Himansu, Ajay Kumar Jena, J. Chandrakant Badajena, Chittaranjan Pradhan, and R. K. Barik. "Resource allocation in cooperative cloud environments." In Progress in Computing, Analytics and Networking, pp. 825-841. Springer, Singapore, 2018.

[14] Papagianni, Chrysa, Aris Leivadeas, Symeon Papavassiliou, Vasilis Maglaris, Cristina Cervello-Pastor, and Alvaro Monje. "On the optimal allocation of virtual resources in cloud computing networks." IEEE Transactions on Computers 62, no. 6 (2013): 1060-107,2013.

[15] Thirumalaiselvan, C., and V. Venkatachalam. "A strategic performance of virtual task scheduling in multi cloud environment." Cluster Computing 22, no. 4 : 9589-9597,2019

[16] Kumar, Mohit & Sharma, "Dynamic load balancing algorithm for balancing the workload among virtual machine in cloud computing", Procedia Computer Science. 115. 322-329,2017.

[17] Soltani, S., Martin, P. & Elgazzar, K.,"A hybrid approach to automatic IaaS service selection", J Cloud Comp **7,** 12 (2018).

[18] G. Senthilkumar., Dr.M.P. Chitra., "Hybrid Optimization Algorithm for Resource Allocation in SDN Enabled Virtual Networks in IaaS Cloud Environment", Volume 11 | 10-Special Issue Pages: 787-802 (2019)

[19] Moreno-Vozmediano, R., Montero, R.S., Huedo, E. et al. Efficient resource provisioning for elastic Cloud services based on machine learning techniques. J Cloud Comp **8,** 5 (2019)

[20] Ruan, X., & Chen, H., "Performance-to-Power Ratio Aware Virtual Machine (VM) Allocation in Energy-Efficient Clouds", IEEE International Conference on Cluster Computing. doi:10.1109/cluster.2015.46,2015.