International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Energy Efficient Scheduling of Servers with Multi-Sleep Modes for Cloud Data Center

Mr.G. Prabhakar[1], P. Ruchitha[2], K. Sathvika[3], N.Sathwika[4]
[1]Assistant Professor, Department of CSE, Malla Reddy Engineering College for Women, Hyderabad, Telangana, India.
[2,3,4]UG-Students, Department of CSE, Malla Reddy Engineering College for Women, Hyderabad, Telangana, India.

ABSTRACT
In a cloud data center, servers are always over-provisioned in an active state to meet the peak demand of requests, wasting a large amount of energy as a result. One of the options to reduce the power consumption of data centers is to reduce the number of idle servers, or to switch idle servers into low-power sleep states. However, the servers cannot process the requests immediately when transiting to an active state. There are delays and extra power consumption during the transition. In this paper, we consider using stateof-the-art servers with multi-sleep modes. The sleep modes with smaller transition delays usually consume more power when sleeping. Given the arrival of incoming requests, our goal is to minimize the energy consumption of a cloud data center by the scheduling of servers with multi-sleep modes. We formulate this problem as an integer linear programming (ILP) problem during the whole period of time with millions of decision variables. To solve this problem, we divide it into sub-problems with smaller periods while ensuring the feasibility and transition continuity for each sub-problem through a Backtrack-and-Update technique. We also consider using DVFS to adjust the frequency of active servers, so that the requests can be processed with the least power. Our simulations are based on traces from real world. Experiments show that our method can significantly reduce the power consumption for a cloud data center.

Index Terms—Multi-modes, sleep state, cloud, data center, energy-efficient

INTRODUCTION
IN recent years, cloud data centers are expanding rapidly to meet the ever increasing demand of computing capacity. It is the powerful servers of the data centers that consume a huge amount of energy. According to a report, data centers consume about 1.3% of the worldwide electricity, which is expected to reach 8% in 2020 [1]. Meanwhile, much of the energy is wasted, because servers are busy only 10%~30% of the time on average, with most time in idle state. What's worse, a server can even consume 60% or more of its peak power when in idleness [2]. To handle the possible peak demand of user requests, servers are always overprovisioned, wasting a lot of energy as a result. Therefore, there is an urgent need to enhance energy efficiency for cloud data centers. The existing work has mainly focused on dynamic voltage frequency scaling (DVFS) and dynamic power management (DPM). The former is to adjust the voltage/frequency of CPU power according to the demand of computing capacity, while the latter reduces the total energy by putting servers into sleep states or turning off idle servers. However, a difficult issue is that the servers cannot process the incoming requests immediately when transiting to active state. There are delays and extra power consumption during the transitions, which have been ignored in the existing work. Besides, modern servers are usually designed

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

with several sleep states, and the sleep states with smaller transition delays consume more power when sleeping. In this paper, we study the issue of minimizing energy consumption of a data center by scheduling servers in multisleep modes and at different frequency levels to reduce

• The authors are with the Department of Computer Science and Technology, Harbin Institute of Technology, ShenZhen 518055, China; Key Laboratory of Internet Information Collaboration, Shenzhen 518055, China.

• E-mail: guchonglin@gmail.com, zhenlongli@hit.edu.cn, huanghejiao@hit.edu.cn, csjia@cityu.edu.hk. The corresponding author is Hejiao Huang. the total energy of active servers. That is, given the arrival of user requests, schedule the servers (to active state with different frequencies or to different sleep states), such that the total energy consumption of the data center can be minimized while satisfying the QoS requirement. The scheduling algorithm will determine:

1) how many of the active servers should be switched into which sleep state in each timeslot;

2) how many of the sleeping servers in sleep states should be woken up in each timeslot;

3) What frequency levels should the active servers be set to in each timeslot. The scheduling period of our problem consists of T small timeslots. We solve the problem in two steps. In the first step, we aim to minimize the total number of active servers to meet the QoS requirement by assuming that all servers run at the highest frequency. The problem is formulated as a constraint optimization problem with millions of decision variables due to the large number of timeslots. It is not feasible to solve the problem of such a large size using existing methods. We group multiple timeslots into a segment with equal length, and formulate the scheduling in each segment independently as an integer linear programming (ILP) subproblem. By using Cplex to solve each sub-problem, the optimal solution can be obtained for each segment.

However, the scheduling of the current segment doesn't consider the arrival of the requests in the next segment. It may lead to the situation that some servers are put into sleep at the end of this segment, but cannot be woken up immediately to cope with request burst at the beginning of the next segment. We propose a Backtrack-and-Update technique to solve this issue. In the second step, we make scaling of the frequency levels of the active servers, so that the requests can be processed with the least necessary power. In each timeslot, this problem can also be formulated into an independent ILP problem of a small size that the optimal solution can be obtained. Our simulations are based on traces from real world. Experiments show that our method can significantly reduce the total energy consumption for a cloud data center. The rest of this paper are organized as follows. In Section 2, the modeling and formulation of our problem will be given in detail. Section 3 gives the Backtrack-and-Update algorithm as solution. In Section 4, we set up the experiments and make evaluations. Section 5 reviews the related work. Finally, Section 6 concludes this paper.

EXISTING SYSTEM

In the existing system, DVFS mechanism scales the CPU chipset power through adjusting the voltage and frequency of CPU. That is, the processing capacity varies with different

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

power levels. Gandhi et al. in [17] combine DFS (Dynamic Frequency Scaling and DVFS to optimize the power allocation in server farm to minimize the response time within a fixed peak power budget. Gerards et al. in [18] try to minimize energy cost through global DVFS on multi-core processors platform while considering the precedence constraint in task scheduling.

PROPOSED SYSTEM

In the proposed system, the system studies the issue of minimizing energy consumption of a data center by scheduling servers in multi sleep modes and at different frequency levels to reduce the total energy of active servers. That is, given the arrival of user requests, schedule the servers (to active state with different frequencies or to different sleep states), such that the total energy consumption of the data center can be minimized while satisfying the QoS requirement.

Advantages

1. Automatic updation of Energy.
2. High Encryption.
3. Time-Saving.
4. No Redundancy.
5. Organized Storage Management

LITERATURE SURVEY:

Years	Title	Methodology	Research Proposal	Algorithm
2021	Advanced Energy Efficient Scheduling of Servers In Cloud	Created four Databases, one of them is responsible for user registration and the other three can be simulated as three servers. To manipulate the data in the Databases we have used JDBC and JSP.	63%	Advanced Backtracking
2020	An Energy-Efficient Task Scheduling Mechanism with Switching On/Sleep Mode of Servers in Virtualized Cloud Data Centers	the technique of switching idle servers off or to sleep mode to conserve energy.	58%	Backtracking
2017	A DVFS based energy-efficient tasks scheduling in a data center	we study the energy-efficient task scheduling, where the task execution time is unknown. We define a novel task model to describe the tasks and the energy consumption ratio to describe the effectiveness of different frequencies.	55%	Heterogeneous-Earliest-Finish-Time(HEFT) algorithm
2016	Peak efficiency aware scheduling for highly energy proportional servers	To evaluate PEAS, we use the Big House data center simulator. Big House is based on stochastic queuing simulation [30], a validated methodology for simulating the power-performance behaviour of data center workloads.	25%	Cluster-level scheduling techniques
2016	Energy conservation using dynamic voltage frequency scaling for computational cloud	The reason why organizations are moving to cloud rather than establishing their own infrastructure is to save money.	20%	Backtracking Algorithm ,K-Means Clustering Algorithm, Binary Search Algorithm, Merge Algorithm

ALGORITHMS USED

ADVANCED ENCRYPTION STANDARD(AES):

The more popular and widely adopted symmetric encryption algorithm likely to be encountered nowadays is the Advanced Encryption Standard (AES). It is found at least six time faster than triple DES.

A replacement for DES was needed as its key size was too small. With increasing computing power, it was considered vulnerable against exhaustive key search attack. Triple DES was designed to overcome this drawback but it was found slow.

The features of AES are as follows −

- Symmetric key symmetric block cipher

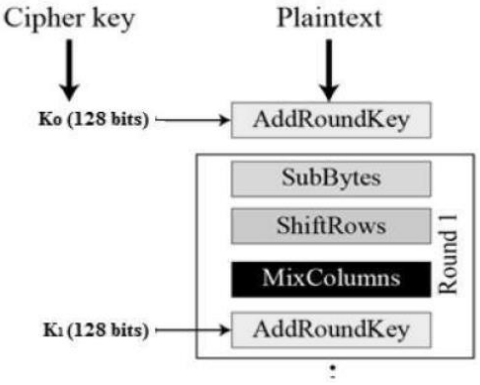
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

- 128-bit data, 128/192/256-bit keys
- Stronger and faster than Triple-DES
- Provide full specification and design details
- Software implementable in C and Java

ENCRYPTION PROCESS

Here, we restrict to description of a typical round of AES encryption. Each round comprise of four sub-processes.

The first round process is depicted below –



Byte Substitution (SubBytes)

The 16 input bytes are substituted by looking up a fixed table (S-box) given in design. The result is in a matrix of four rows and four columns.

Shiftrows

Each of the four rows of the matrix is shifted to the left. Any entries that 'fall off' are re-inserted on the right side of row. Shift is carried out as follows −

- First row is not shifted.
- Second row is shifted one (byte) position to the left.
- Third row is shifted two positions to the left.
- Fourth row is shifted three positions to the left.
- The result is a new matrix consisting of the same 16 bytes but shifted with respect to each other.

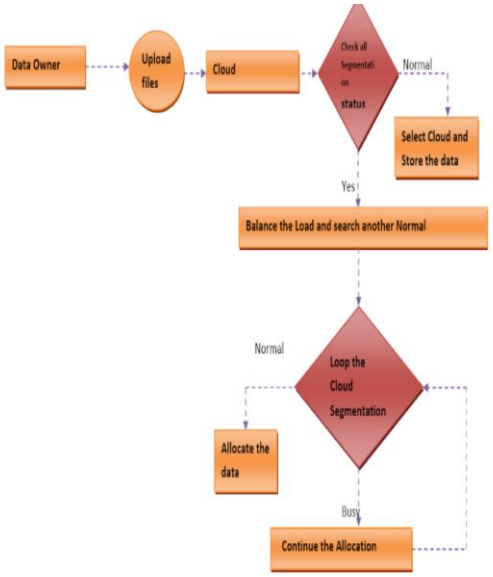
MixColumns

Each column of four bytes is now transformed using a special mathematical function. This function takes as input the four bytes of one column and outputs four completely new bytes, which replace the original column. The result is another new matrix consisting of 16 new bytes. It should be noted that this step is not performed in the last round.

Addroundkey

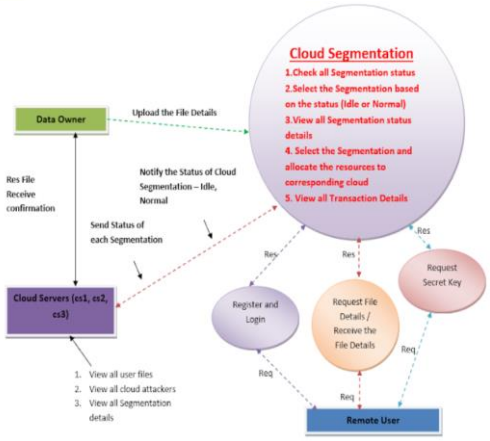
The 16 bytes of the matrix are now considered as 128 bits and are XORed to the 128 bits of the round key. If this is the last round then the output is the ciphertext. Otherwise, the resulting 128 bits are interpreted as 16 bytes and we begin another similar round.

FLOWCHART:



ARCHITECTURE:

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



.

RELATED WORK :

In general, there are three ways to green cloud data centers: dynamic voltage frequency scaling, dynamic power management, and the scheduling using renewable energy. A. Dynamic Voltage Frequency Scaling (DVFS) DVFS mechanism scales the CPU chipset power through adjusting the voltage and frequency of CPU. That is, the processing capacity varies with different power levels. Gandhi et al.

in [17] combine DFS (Dynamic Frequency Scaling) and DVFS to optimize the power allocation in server farms to minimize the response time within a fixed peak power budget. Gerards et al. in [18] try to minimize energy cost through global DVFS on multi-core processors platform while considering the precedence constraint in task scheduling. Elnozahy et al.

in [19] employ a DVS (Dynamic Voltage Scaling) and node On/Off method to reduce the aggregate power consumption of cluster during periods of reduced workload. They also use both DVS and requests batching mechanisms to reduce processor energy over a wide range of workload intensities [20]. Rossi et al.

in [21] build power models to estimate the energy consumption of user applications under different DVFS policies. Florence et al. I

n [22] first study the flow pattern of tasks of cloud, and then try to tune the incoming VM tasks with required frequency using DVFS. Lin et al. in [23] use DVFS to reduce the power consumption in task scheduling in mobile cloud computing environment, but with no consideration of On/Off servers. Chen et al.

in [24] combine the three approaches of request dispatching, service management and DVFS to improve energy efficiency for large scale computing platform. However, they assume the servers that provide various services are active all the time. Gu et al.

In [25] study the problem of dispatching requests to geo-distributed data centers using both On/Off switchings of servers and DVFS, but ignoring the delays during On/Off operations. Le Sueur et al. in [26] demonstrate that DVFS is effective in reducing energy consumption, even in situations where tasks with precedence constraints are scheduled. Aydin et al. in [27] schedule the frequency for independent tasks on multi-core and multiprocessor system with deadline, while Zhu et al. in [28] consider the precedence of the tasks on multi-core platform when reducing energy by adjusting dynamic voltage/speed. However, DVFS only adjusts CPU, leaving the remaining components running with no energy saved. B. Dynamic Power Management (DPM) In contrast, the DPM scheme can power down all the components of servers so that the total energy of servers can further be reduced. The commonly used policies of DPM are to power off or switch the idle servers into sleep states, and then wake them up when necessary. Xie et al. in [29] reduce the energy consumption of servers through consolidation of virtual machines, while considering the transition delays and

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

energy cost during transitions. They assume a transition only happen within scheduling timeslot, which is impractical in reality.

Lin et al. in [30] propose a right-sizing method to decide the minimum number of servers that should be online to serve the requests so as to reduce the total energy of servers. Gandhi et al. in [31] propose to wait a period of time before turning on servers through a distributed robust auto scaling (DRAS) policy for computing intensive server farms. SoftReactive tries to keep the right number of servers in idle by combination of timer and index-based routing [15].

SoftReactive is a special case for AutoScale that also sets timer to avoid the mistake of turning off a server just as a new arrival comes in. The latter uses a capacity inference algorithm to determine the right number of idle servers using the control knob nsys that can be tracked in Apache and load balancer to scale the number of servers periodically, lacking theoretical support. Niyato et al.

in [32] reduce the power consumption of server farm by formulating the problem as a Makov decision process using sleeping modes. Pinheiro et al.

in [33] take load balancing into account when reducing energy through dynamically turning on and off servers, but frequent turning on/off servers is harmful to the hardware and may reduce lifetime of servers. Meisner et al.

in [34] propose an energyconservation approach called PowerNap, which eliminates idle server power by quickly transitioning in and out of an ultra-low power state, but consider only one sleeping mode. Heath et al.

in [35] develop a model-based cooperative Web server to minimize energy consumption for heterogeneous clusters. Chen et al.

in [36] propose three strategies based on steady state queuing analysis, feedback control theory, and a hybrid mechanism to minimize operational costs while meeting the SLAs. Liu et al.

in [37] present an analytical framework for characterizing and optimizing the powerperformance tradeoff in SaaS cloud platform, and the energy saving is achieved by the scheduling of idle/busy states of virtual machines. Matthias et al. in [38] create a Markov model as an abstraction of queuing systems with sleep modes and add delayed activation and deactivation, so as to reduce the total server power. Raj et al. in [39] propose a simple method to reduce the total energy using only one type of sleep mode in combination with shutdown to save energy. However, using only one sleep mode is not enough to reduce total energy when making quick responses. C. Green Scheduling Using Renewable Energy There have been studies on green scheduling using renewable energy for cloud data center. Zhou et al. in [40] propose a carbon-aware scheduling framework that makes online decisions on geographical load balancing, capacity right-sizing, and server speed scaling using Lyapunov optimization techniques. In our previous work, we try to minimize the total carbon emissions under budget of energy cost when scheduling requests, servers with On/Off switching, and the usage of renewable energy among the data centers in geo-distributed locations [41]. To further green cloud data centers, we consider using ESDs (energy storage devices) to store different types of energy with fluctuating prices, so that both energy cost and total carbon emissions can be greatly reduced [6], [42]. To leverage the climate advantages of different locations, we propose a green plan that optimally deploys wind turbines, solar panels and ESDs for

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

future green cloud data centers [43]. In reality, the data centers may have their own wind or solar farms. The selfgenerated green energy can be used to power data centers directly or sold back to the power grid, so that the data centers can be greened with lower cost [44], [45]. However, the scheduling granularity in these work is usually very large, which is set to be from 10 minutes to 1 hour. Therefore, there is still large room to improve the energy efficiency through scheduling of servers with multi-sleep modes. As far as we know, there is not too much research about scheduling of servers with multi-sleep modes in energy saving. The possibility of using multiple sleep modes was first studied by Horvath in [46]. However, the importance of the transition power and delay of different sleep modes is not well considered. After that, Liu et al. in [47] propose a self-adaptive management of the sleep depths to reduce the total energy of servers. Unfortunately, the transition delay and power during sleep-down process have been ignored in their model. Different from existing work, we focus on minimizing the total power consumption of the cloud data center by switching servers between active state and multiple sleep states according to the varying incoming requests. The novelty of our problem is that we take into account of transition delays and power, and that consumed under different sleep modes in our scheduling, which greatly affect the decisions in energy saving for cloud data center.

RESULTS:

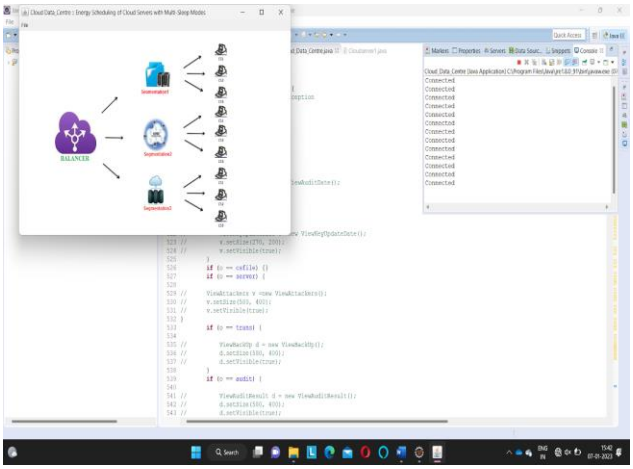
Round-Robin We endorse propelled device plan methodologies, Dynamic round Robin for modernized contraption mix. The target of these methods is to keep the measure of physical machines used to run each mechanized structure. This point is suspicious of the way in which that the measure of genuine machines used decidedly impacts to manage use. Dynamic round-Robin is proposed as an expansion to the round-Robin method[38]. the second standard of Dynamic round Robin is if a physical system is inside the retirement US for an enough huge bit of time, as opposed to looking ahead of time to the urging virtual machines to complete it with no other character's information, the genuine machine can be kept to transport loosening up of electronic machines to exceptional considerable machines, and shutdown after the relocation wraps up. This keeping up time impediment is recommended as retirement side [39]. A physical structure is in the retirement usa in any case furnished to complete each electronic machine after the leave angle couldn't can be constrained to transport its propelled machines and shutdown [40]. Our Dynamic round-Robin methodology uses the ones two musings with a specific true blue intend to join virtual machines passed on with the significant resource of the charming round Robin mechanical assembly [41]. The crucial proposition stream without which fuses progressively virtual machines to a leaving genuine machine so it will in general be closed down. the second one controlling rule hustles the affiliation system and draws in powerful round robin contraption to shutdown genuine machines, with the target that it may diminish the measure of significant machine used to run every single virtual gadget [42]. First Come First Server In First Come First Serve, the proposed strategy invigorates the scheduler saving the remarkable influenced time fundamentally based occupations into line. The client submitted occupations formed reliant on the rising curious for of the influenced time and it offers make again the preliminary financing with

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

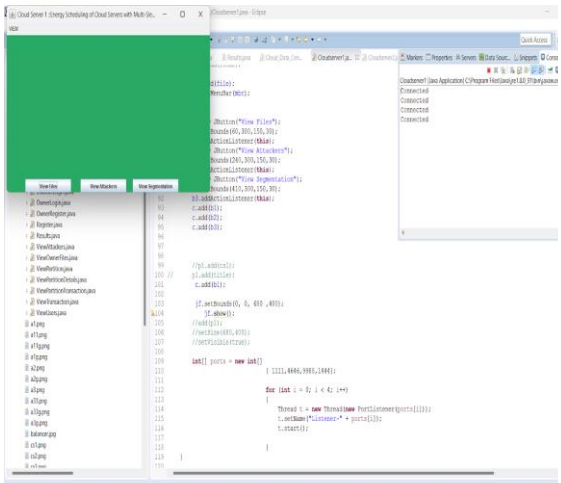
significance to all employments. It has any sort of effect to produce the consumer loyalty subject to the way wherein that the purchaser stipulations are moving in setting of the prevalent wishes [43]. It diminishes starvation using the dynamic dispersing of livelihoods to pick the remarkable significant occupations from a bit of the open and does now not cut down the execution of the machine. the street supervisor gives the property for the essential framework. the line supervisor is a pinch of the appropriated enrolling which offers with using the significant extent of focal concentrations inside the gathering. It screens the systems, which may be with everything considered by strategies for method for and through running the occupations by technique for adjusting the pile a portion of the meta scheduler and its change . It acclaimed the scheduler to plot the yield of the improvement that is collected returned with the guide of the street manager. The sorting out strategy and asset bundle structure is predicated upon resulting to propelling holding with noteworthy activities [45]. Little, medium and for the reason that a long time inside the past on a very basic level based are three unequivocal pursues kept as on creating requesting foreseen concerning shot time of the occupations [46]. The customer submitted employments are commitment to the organization sponsor it joins line boss that minding the occupations in rising dependent on affected time [47]. The occupations are performed depending upon careful assurance and commitment to cloud condition [48]. The awesome assignment lessens the time and accessibility of district in an a triumph path without repaying the high bore of the structure and supporter needs [49]. control careful endeavor based totally virtual device Consolidation set of rules on this demonstrated the general method to join propelled machines to genuine machines. in this we're having in the present 7 figurings. proper perfect here, the fact of the matter is to oblige this make length respect close-by the essentialness utilization of the cloud structure. In step-1 of set of rules 1, the sub-consider is insinuated using giving the coalition of errands and their due dates as information. The course of action of principles 2 type most of the errands in mountaineering request in their due dates. The push off immaterial point of confinement in step-2 of set of principles 2 will clear the errand with scarcest due date a spurring power from the sort of assignments and hold inside the line, AB to the progression 1 of set of standards 1. The need of count 2 is proportional as building up a Min-store of the assignments toward the beginning of their due dates. The Min-shop is manufactured depend upon the due date as key respect so the undertaking with unimportant due date to be cleared first [50]. The progression 2 of Algorithm-1 calls the subcalculation to portray all of the assignments into four groupings. The arranged undertaking line, AB is given as contribution of the calculation.

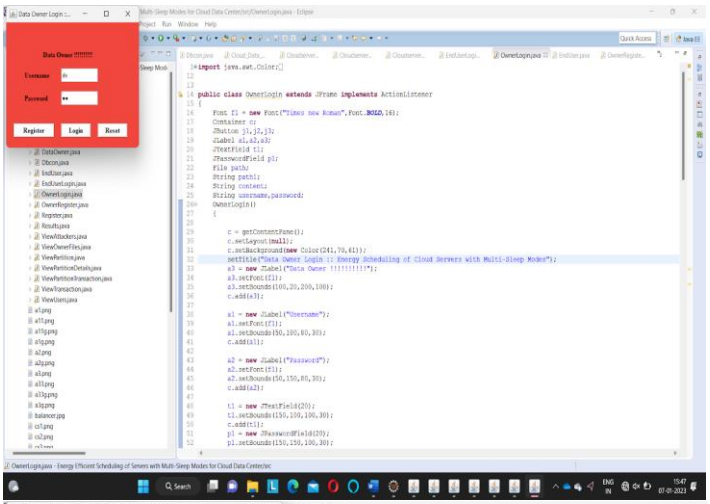
RESULTS:

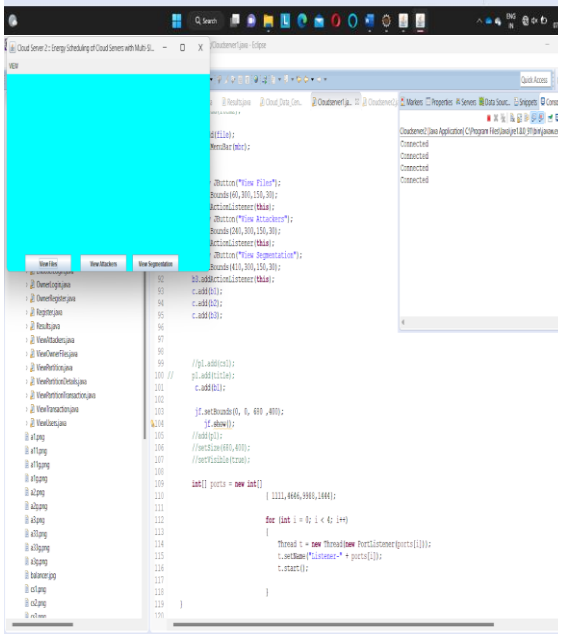
OUTPUT SCREENS:

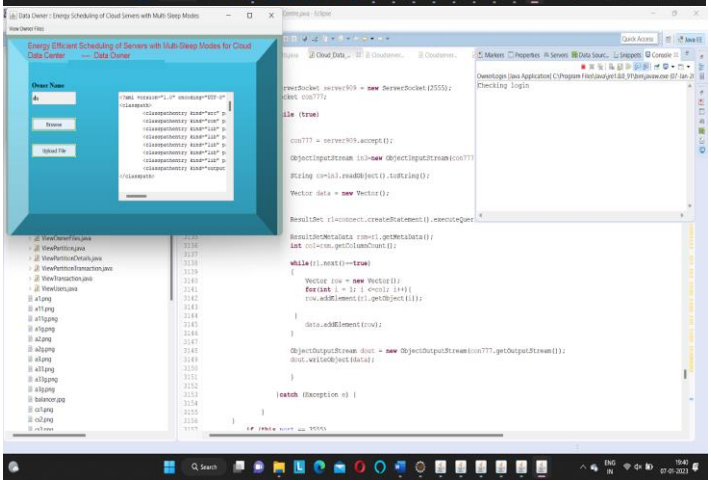


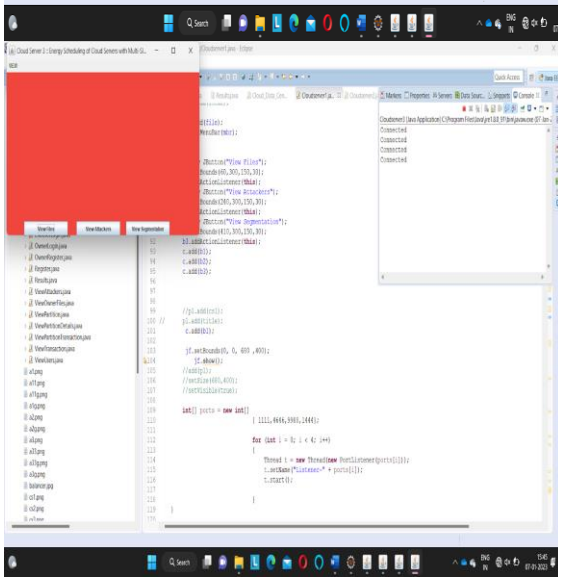
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



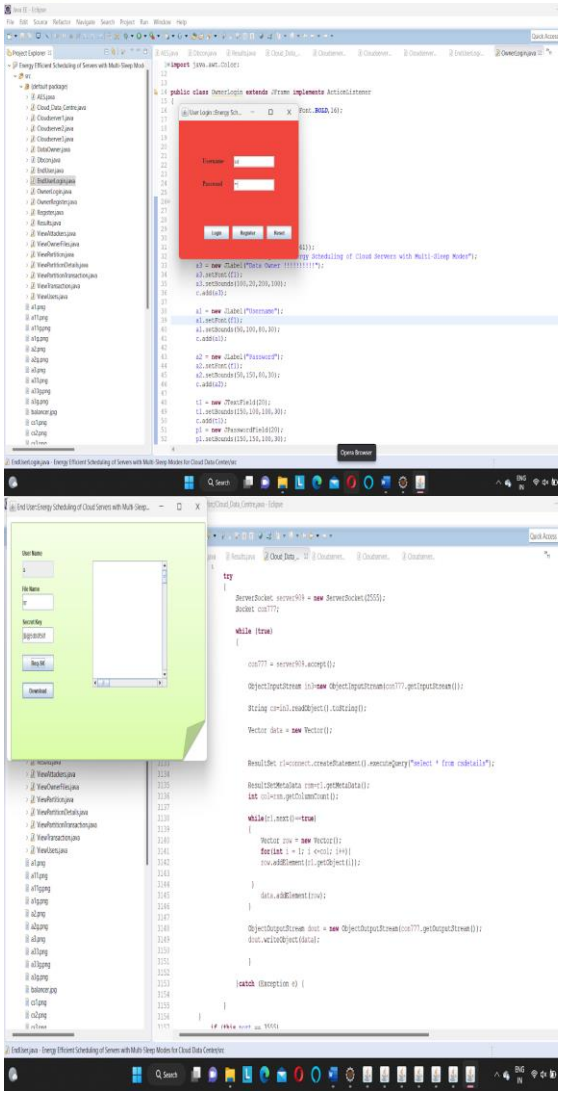








International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



CONCLUSION and FUTURE WORK:

In this project, we studied the problem of scheduling of servers with multi-sleep modes for cloud data centers. The servers can make transitions between one active state and different sleep states, which involves different sleep power and transition delays for the sleep modes. We proposed Backtrack-and-Update method to make schedule of the servers, deciding how many servers in each state should be switched to which states in each timeslot, so that the total power consumption can be minimized while satisfying the QoS requirement. The problem is too large to be solved by existing methods, so we divide the whole problem and then conquer them one by one while considering the ongoing transitions during the breakpoints. We also consider using DVFS to further reduce the energy caused by the over provisioned computing capacity. Experiments show that our scheduling using multi-sleep modes can significantly reduce the total energy with QoS of less than 10ms. Against the over-provisioned strategy of Always On, our method can reduce more than 58% of the total energy on average.

REFERENCES:

[1] P. X. Gao, A. R. Curtis, B. Wong, and S. Keshav, “It’s not easy being green,” ACM SIGCOMM Computer Communication Review, vol. 42, no. 4, pp. 211–222, Aug. 2012.

[2] A. Gandhi, M. Harchol-Balter, and M. A. Kozuch, “Are sleep states effective in data centers,” in Proc. IEEE IGCC, Jun. 2012, pp. 1–10.

[3] L. Rao, X. Liu, L. Xie, and W. Liu, “Minimizing electricity cost: Optimization of distributed internet data centers in a multielectricity-market environment,” in Proc. IEEE INFOCOM, 2010, pp. 1–9.

[4] Y. Zhang, Y. Wang, and X. Wang, “Greenware: Greening cloudscale data centers to maximize the use of renewable energy,” in USENIX International Conference on Distributed Systems Platforms and Open Distributed Processing, Dec. 2011, pp. 143–164.

[5] S. Wang, Z. Qian, J. Yuan, and I. You, “A DVFS based energyefficient tasks scheduling in a data center,” IEEE Access, vol. 5, pp. 13 090–13 102, 2017.

[6] C. Gu, H. Huang, and X. Jia, “Green scheduling for cloud data centers using ESDs to store renewable energy,” in Proc. IEEE ICC, Apr. 2016, pp. 1–6.

[7] IBM, “CPLEX Users Manual,” https://www.ibm.com/support/ knowledgecenter/SSSA5P 12.7.0/ilog.odms.studio.help/pdf/

usrcplex.pdf, 2017, [Online; accessed 25-December-2017].
[8] M. C. P. T. L. T. C. Hewlett-Packard Corporation, Intel Corporation, "Energy Management of ACPI," https://www.intel.com/content/dam/www/public/us/en/ documents/articles/acpi-config-power-interface-spec.pdf, 2017, [Online; accessed 25-December-2017].
[9] D. G. Feitelson, D. Tsafrir, and D. Krakov, "Experience with using the parallel workloads archive," Journal of Parallel and Distributed Computing, vol. 74, no. 10, pp. 2967–2982, Oct. 2014.
[10] "Logs of real parallel workloads from production systems," http: //www.cs.huji.ac.il/labs/parallel.
[11] D. D. Gutierrez, "The Intelligent Use of Big Data on an Industrial Scale," https://insidebigdata.com/2017/02/16/ the-exponential-growth-of-data/, 2017, [Online; accessed 25- December-2017].
[12] J. Li, Z. Li, K. Ren, and X. Liu, "Towards optimal electric demand management for internet data centers," IEEE Transactions on Smart Grid, vol. 3, no. 1, pp. 183–192, 2012.