International Journal of Information System Modeling and Design
Volume 13 • Issue 6

Energy-Aware VM Scheduler:
A Systematics Review

Ram Narayan Shukla, LNCT University, Bhopal, India*
Anoop Kumar Chaturvedi, LNCT University, Bhopal, India

ABSTRACT

Cost is the backbone of any business model to sustain and grow in the competitive business market. Considering this, higher energy consumption may affect the operational cost of any industry. Cloud computing is one of the most popular IT business models for service providers and their users. It is observed in many studies that the average cost of the services is highly dependent upon the run time and power consumption. More power consumption affects the environmental things as well. In this review, more than 100 research articles are considered to evaluate the energy consumption methods, issues, and challenges from 2009 to 2021. Based on the study, the energy-aware load balancing methods are classified into four major clusters. The most cited methods are compared based on method used, platform where it is deployed, different evaluation parameters like cost, run time, virtual machine utilization, service-level agreement violation rate, and quality of service (QoS). The results of this review are presented as open research issues and challenges for forthcoming research.

KEYWORDS

Auction, Cost, Energy Optimization, Execution Time, Load Balancing, SLA

INTRODUCTION

In today's digital era, where internet is the base for everything, it is resulting in the tremendous increase of the digital data. So much data is available that it is getting difficult to process or even handle properly. As a result, more and more datacenter, cloud servers and its infrastructure are being required.

According to a report, in the year of 2018, data center of whole world consumes approximately one percentage of whole world's electric power. That is roughly two hundred five Tera Watt hour of electricity in a year.

Since there is data center is responsible for almost everything, therefore it further requires much more components in it to function properly such as Cloud servers which are responsible for all the computational work (Panja, S., & Roy, B., 2021), Storage drives which keeps the data and allow user to access it, Cooling systems or Air conditioners which keeps the systems cool in a data center as in a data center due to over load system may get heated up, and networks (Ren, Y., et al., 2021).

Servers are the main component which serves the computation purpose; therefore, it consumes around forty-three percentage of energy. (i.e., 88.15 Tera Watt hour electricity). At second place comes

DOI: 10.4018/IJISMD.297631 *Corresponding Author

the Cooling and air conditioning system, it too consumes approximately forty-three percentage of data. Air conditioning is required in the data center so that heat generated by the servers of datacenter can be dissipated and system can become reliable for the consumer. It surely consumes considerable amount of energy but is much required, because when it comes to data center, it is a necessity (i.e., 88.15 Tera Watt hour electricity). Storage today is much required. Without storage any user will hardly be able to complete his work, therefore storage devices are generally responsible for around eleven percent of energy consumption (i.e., 22.55 Tera Watt hour electricity). Network is much important for cloud computing as it is responsible for connectivity. Without network, user won't be able to access and much of its purpose would be lost. Network consumes around three percent of energy (i.e., 6.15 Tera Watt hour electricity).

Whereas if we consider Steel industry steel industry use approximately eighteen to twenty five percent of whole energy usage in the industrial sector.

Therefore, in a data center most of the energy is consumed by Cloud servers and cooling & air conditioning system followed by the storage devices, which is further followed by network.

Steel industry uses approximately six percent of the total energy consumption (Jafari, V., & Rezvani, M. H., 2021). This is because of properties steel exhibit. Steel exhibit strong and anti-corrosion properties which makes it ideal for several purposes. Some of the examples of steel uses are in construction, automobile industry, manufacturing etc. China is the biggest consumer of steel in the world. Steel industry is the 2nd largest industry in the world which have a turnover of around nine hundred billion dollars, as well as steel is the main constituent of all the energy sources like solar, hydro and wind energy. Steel industry produces employment of over two million people.

Load Balancing is the strategy that permits you to make a proper balance of the measured of work that is being performed on different virtual or physical machine. Virtualization in cloud computing help to service provider to manage resources and enhances energy-efficiency (Kim & Seo, 2014). The energy costs for data centers are rising rapidly. This may lead to exceed the cost of purchasing server hardware in future (Kumar, 2007).

Although large Internet organizations (such as Google, Amazon, and Microsoft) have greatly decreases the energy consumptions of their multi-megawatt data-centers and servers, so far, they have focused on physical servers. In addition to large-scale data-centers, efficient operation is also extremely beneficial to small and medium-sized data-centers that constitute most of the energy consumption of data centers (Katsaros et al., 2013). High energy utilization also increases the service level agreement in data-centers (Zhou et al., 2018).

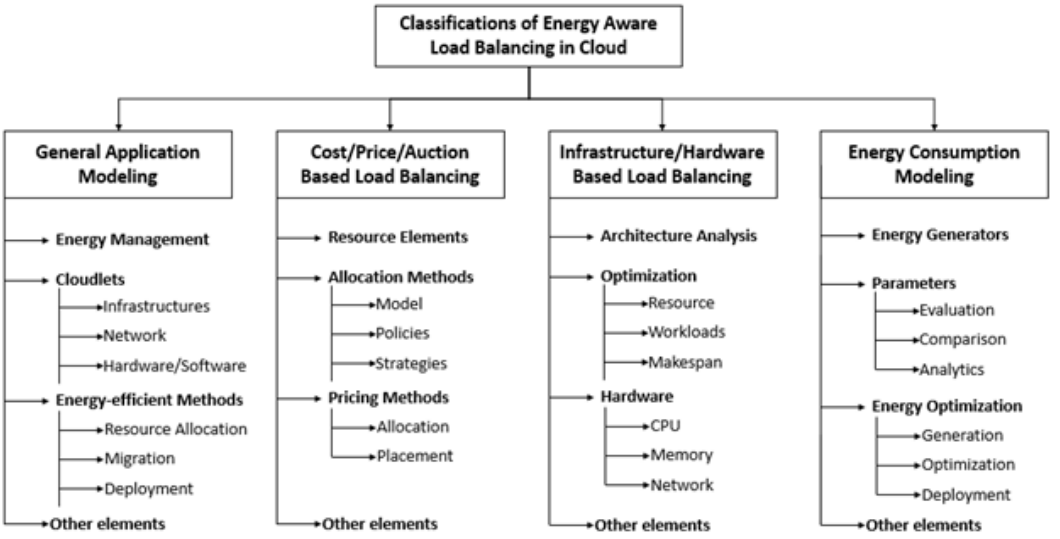
The objective of the systematics review is to focus to find the answer of the defined research questions and its impact towards energy efficient VM management techniques in cloud computing. Next section is all about the related work based on energy consumption and VM provisioning with the comparative analysis of selected articles published in high impact factor journals, for this Web of Science database is considered. The organization of the article is constructed to maintain the flow from introduction to related work, comparative analysis of different energy-aware load balancing methods, issues and challenges and conclusion. Figure 1 presents article selection scheme.

PREVIOUS WORK

The schematic literature survey have been conducted on different energy-aware load balancing technique in cloud. In this, more than 200 articles from reputed journals are considered, which are published from 2007 to 2021. Out of 300, about 50 articles have been considered to conduct different analysis based on parameters. In this review, the energy-aware load balancing techniques are classified into different categories as mentioned in Figure 2.

Figure 1. Article selection scheme

Figure 2. Classification of different load balancing methods based on energy utilization in cloud



General Application and Modeling

A Mapreduce method has been introduced by Li, Y. et al. for VM provisioning and migration named EECLOUD. The experimental effects are very satisfactory in terms of electricity financial savings over current methods (Li, Y. et al., 2012). A dynamic workload provision introduces by Xu, X et al., to resolve the resource allocations problem called EnReal. In particular, it provides a power consumption technique for applications used worldwide. Experimental evaluations show that EnReal is efficient and effective as compared to existing methods (Xu, X. et al., 2015). Author Djemame,

K. et al., worked on different IaaS layers (Djemame, K. et al., 2017). Authors Li, X. et al. limit the full power price in a cut-off date restricted energy-conscious workflow scheduling hassle with facts being geographically disbursed throughout facts centers. (Li, X. et al., 2020).

Cost/Price/Auction-Based Load Balancing

Talking about the cost investments, data center need investment of around ten million dollars to twenty-five million dollars annually. Whereas if we consider cost investment in data centers fifty seven percent of energy are consumed by the cloud servers which is approximately 2 million dollars on monthly basis. On yearly basis cloud servers cost 23 million dollars, on weekly basis it requires 470 hundred thousand dollars and on daily basis it costs 67 thousand dollars per day. Servers are the main technical component of any data center therefore; it's worth the cost to invest. Air conditioning and cooling system comes second at cost (Shende, M. P. 2012). investment as it is a crucial thing which is required for proper working of the data centers. It takes approximately eighteen percent of the total cost investment and costs around 6-7 million dollars on monthly basis, 1.5 hundred thousand dollars on weekly and on yearly basis it costs around 7.5 million dollars and on daily basis it costs around 21 thousand dollars. Power and power equipment require approximately thirteen percent of total cost investment which sums up to 5 hundred thousand dollars on monthly basis. On yearly basis power and power equipment requires approximately 6 million dollars, 1.2 hundred thousand dollars on weekly basis and on the daily basis it costs seventeen thousand dollars. Networking equipment take around eight percent of the total cost investment that is 3 hundred thousand dollars on the monthly basis. Whereas on yearly basis it would require 3.5 million dollars as cost investment, 70 thousand dollars on weekly basis and on the daily basis it would require ten thousand dollars cost investment. Other infrastructure takes around four percent of the total cost investment that is around 140 hundred thousand dollars on the monthly basis, 5 thousand dollars on daily basis, 40 thousand dollars on monthly basis and around 1.7 million dollars on yearly basis. This surely proves that data center requires a huge amount of capital involvement to be invested. Therefore, in data center cost investment is most required in cloud serves followed by cooling & air conditioning system which is further followed by power consumption networking equipment and other infrastructure respectively. As compared to other industries, energy industry is the one which requires the most investment of all industries. Energy industry consists of coal oil gas etc. These industries require heavy equipment for production, and needs to be replaced within a certain amount of time, due to safety hazards. After energy industry, transport industry requires the most cost investment. Transport industry includes air transportation which is heavily cost invested, followed by road transportation and then waterways.

Farahnakian, F. et al., advises a dynamic digital system consolidation set of rules to limit the wide variety of lively bodily servers on a facts middle so one can lessen power cost. Proposed consolidation technique makes use of the KNN regression set of rules to expect useful resource utilization in every host. The proposed consolidation technique can determine (i) while a bunch will become over-utilized, and (ii) while a bunch will become under-utilized. The results outcomes at the actual workload lines from greater than 1000 PlanetLab digital machines proves that Proposed approach minimizes power intake and continues required overall performance levels (Farahnakian, F. et al., 2013).

Gai, K. et al., proposed system is basically made for minimizing the energy cost for heterogeneous system and mobile-cloud. They have proven concept of energy saver system that can be deploy in embedded heterogeneous environment for the mobile-cloud (Gai, K. et al. 2018).

Author Aldossary, M. et al., introduces a unique Cloud device structure that enables a power conscious and green computing operation method estimate the electricity utilization (Aldossary, M. et al., 2019).

Infrastructure/Hardware-Based Load Balancing

Author presented a renewable-power-conscious cloud carrier and digital system migration to distribute power call for the use of dynamic and bendy cloud useful resource allocation methods, and assist

conquer the demanding situations of renewable power. A U.S.-extensive cloud community infrastructure display that, the use of easy migration techniques, as much as 30 percentage nonrenewable power may be changed with the aid of using renewable power, at the same time as ingesting most effective a small quantity of more assets and power to carry out call for relocation (Mandal, U., et al., 2013).

Author Rezai, H., and B Speily, O. R. present a framework and concepts for strength-green cloud computing environments. The set of rules employs a heuristic approach which enhance the strength performance in information center to ensure the Quality of Service (QoS). The performance of the proposed method is evaluated through the use of the maximum not unusual place cloud computing simulation toolkit, CloudSim (Rezai, H., and B Speily, O. R. 2017).

Authors Katsaros, G., et al., presented a carrier framework that lets in us to screen the strength intake of a Cloud infrastructure, calculate its strength performance, and examine the collected facts so that you can installed area and powerful digital VM management. (Katsaros, G., et al., 2013).

Author's Hongyou, L. et al., proposed two algorithms are primarily based totally at the truth that a couple of sources (which includes CPU, memory and community bandwidth) are shared via way of means of customers simultaneously in cloud statistics centres and heterogeneous tasks have distinct aid intake characteristics. Simulation consequences display that each algorithms correctly make use of the sources in cloud statistics centres, and the dynamic sources have true balanced utilization's, that display their promising power saving capability (Hongyou, L. et al., 2013).

Author Farahnakian, F. et al., presented a VM provisioning technique that takes into consideration each the modern and destiny usage of resources. The method is totally based on CPU and memory usage of PMs and VMs. They look at the effectiveness of digital and bodily aid usage prediction in VM provisioning overall performance the use of Google cluster and PlanetLab actual loads. The experimental outcomes show, the technique affords large development over different heuristic and meta-heuristic problems (Farahnakian, F. et al., 2016).

Author Zhou, Z., et al., cope with the hassle of lowering Cloud datacenter excessive strength intake with minimum SLA violation. Although there are numerous strength-conscious useful resource control answers for Cloud datacenters, current techniques awareness on minimizing strength intake of software walking with inside the VMs and hence won't clearly lessen strength intake with minimum SLA violation below a lot of workloads (Zhou, Z., et al., 2018).

Energy Consumption Modeling

Data Center in the whole world consumes approximately 1 percent to 1.5 percent of global energy produced, which is quite a big amount. Therefore, to power so much of energy, we cannot rely only on one source of energy, as there are different geographical and environmental condition for different places, therefore there are different source of energy which we can use according to need.

There are many forms of energy, with different source of them, but in today's digital word where transportability and manageability play a crucial role, there we need a refined form of energy. Therefore, electric energy is used the most now a days. In the modern world, efficient and cost-effective method of producing energy is one of the upcoming global concern. In Every industry even small scale to large scale industry require power (or so-called energy) to serve its purpose. Currently the world is facing power shortage mainly due to in-efficiency and its miss management. Electricity is the refined form of energy which in the modern world is used as there are many advantages of it such as Easy & convenient to store, Easy to handle, can be easily transported without much loss, Easy to control and Easy to utilize efficiently.

There are several different methods of energy / power generation which used in cloud/ datacenters are:

- **Thermal Power Generation method:** Thermal in this method is referred to heat. It uses heat to produce electricity. Dominantly only Coal is used to generate heat in this type of method. Coal is the part of non-renewable source of energy which is present in abundance. Coal has

high calorific value, therefore when burnt produces high amount of heat energy. Heat energy is hard to store and handle therefore we convert heat energy into electric energy. It involves low risk. There abundant amount of coal present on earth makes it one of the best choices for energy production. Coal is easy to transport and store. But despite there are many advantages of using coal, there are many disadvantages of it such as it induces a lot of pollution in the environment and especially in the nearby areas which lead to eight to sixteen percent increase of respiratory problems in the people nearby, and the fact that it's a non-renewable resource is a disadvantage for using it. NTPC (National Thermal Power Corporation) is a good example for it.

- **Nuclear Power Generation method:** Nuclear Power generation method is one of the clean energy production method. There are two types of nuclear power plant i.e., one which uses nuclear fission and nuclear fusion. Nuclear fusion technology currently is not used commercially and is under research and development phase. Nuclear fission is used for commercial electricity production for example NPCIL (Nuclear power Corporation of India Limited), generate electricity using nuclear active fuels such as Uranium 235 and Uranium 237. Advantages of using Nuclear Power plant for power generation are, produces huge amount of energy, it is zero emission clean source of energy, involves huge money investment but is worth the profit it generates. Whereas some of the disadvantages of nuclear power plant are, it involves high risk and skilled labor to operate, Uranium is not much handy and can be handled after taking all protective measure, small mistake can result in severe catastrophic.
- **Hydro Power Generation method:** Hydro power generation is the utilization of the flowing water energy for energy production. Dams are an example of hydro power generation which uses water energy to rotate the turbine and generate electricity. There are many advantages for this method as it is a renewable source of energy, doesn't cause any pollution to the nature, sustainable form of energy. Whereas some of its disadvantages are, it affects the marine life, geographical location is a big concern and it includes much initial investment and slower returns.
- **Geo Thermal Power Generation method:** In Geo thermal, Geo stands for earth and thermal stand for heat. It is one of the important pillars of sustainable development. Geo thermal power plant utilizes earth's heat energy to generate electricity. It is a very trustworthy energy production method. No constant fuel is required keep the process running. Since it has so many advantages and no environmental danger associated with it, it has a huge prospective of growth in future. Some of cons associated with Geo thermal are, it is geographically restricted, it is believed to prompt earthquakes and it is costly.
- **Battery Power Generation method:** It is a chemical energy production technique in which a chemical reaction is used to produce the energy. They are generally used as stored form of energy which can be utilized as backup. These are rechargeable and can be reused several times. These provides a reliable source of energy till quite a good amount of time. Battery power generation is most commonly used in automobile industry, electronic industry, to power daily gadgets etc. This field is much in research and development phase for better and longer period of battery. Performance.
- **Wind Power Generation method:** Wind power generation method is a renewable and sustainable form of energy. It converts wind's kinetic energy into electrical energy. It is eco-friendly, easier to setup, does not cost much, good alternative on a low scale perspective, it requires low maintenance, does not require much of land space to setup. But it is not producing power much in quantity and is somewhat dependent on the natural factor to produce electricity.
- **Bio Generation method:** In this method, the biomass is burned and its energy is used to generate electric power. It is advantageous as it helps in waste management, energy production, trustworthy source of energy. But some of its negative attributes are, it is costly to construct, it requires more land space and can have environmental effect too. It is a great method for, energy generation in small areas such as villages, where bio waste are much generated and other methods are not possible. At such places this method is much efficient to be utilized.

- **Solar Power Generation Method:** This method stands for utilizing sun's energy i.e., sunlight to convert it into electrical energy. It is a pure clean form of energy. It is sustainable as well as renewable source of energy. It is much efficient but some of its disadvantages are, it is costly, energy production somewhat depends on the weather conditions, if weather is sunny then good energy production and low in cloudy conditions. Space is a concern for it, as it requires large space to be installed. It is one the technology which scientist are working on to make it more economically and cost efficient, as it is one of the important pillars for sustainable development.

Table 1 presents the different tools for energy management in High performance computing. Some energy-based load balancing are discussed as follows.

Alzamil, I. et al., worked on the system architecture for resource provisioning which include energy parameters. In this system, author given a provision to service providers that they can take quick decision while they are working on energy optimization techniques at the time of Vm selection. They have developed a cloud based architecture which adapted the energy-aware strategy based on the current system architecture (Alzamil, I. et al., 2015).

Vafamehr, A. and Khodayar, M. offers the policies, resources, and models for power-conscious cloud computing and envisions a marketplace shape that pointed the effect of the fee of power deliver at the quality and fee of cloud services. Energy control practices for cloud carriers on the macro and ranges to enhance the fee and reliability to the cloud offerings are proposed (Vafamehr, A. & Khodayar, M. 2018). Zhang, X. et al. proposed an energy saver system that simplified an engine for cloud simulator to find not only the optimal VM, but can accommodate more VM for less physical machines which is energy efficient than related work in 2019. The system proposed by authors is also profit improvement system based on the energy parameters as they save energy up to 24% (Zhang, X. et al., 2019). Feng, H. et al., suggest a global-power-conscious virtual-machine placement VMP method to lessen, from more than one aspects, the entire power intake of statistics facilities (Feng, H. et al., 2021).

Figure 3 presents the energy necessities with the load, the power performance of a computing machine isn't a linear characteristic of the load; even if idle, a machine can also additionally use 50% of the energy similar to the whole load. Data accrued over a protracted time frame suggests that the everyday working vicinity for facts middle servers is with inside the variety 10%−50% of the total load.

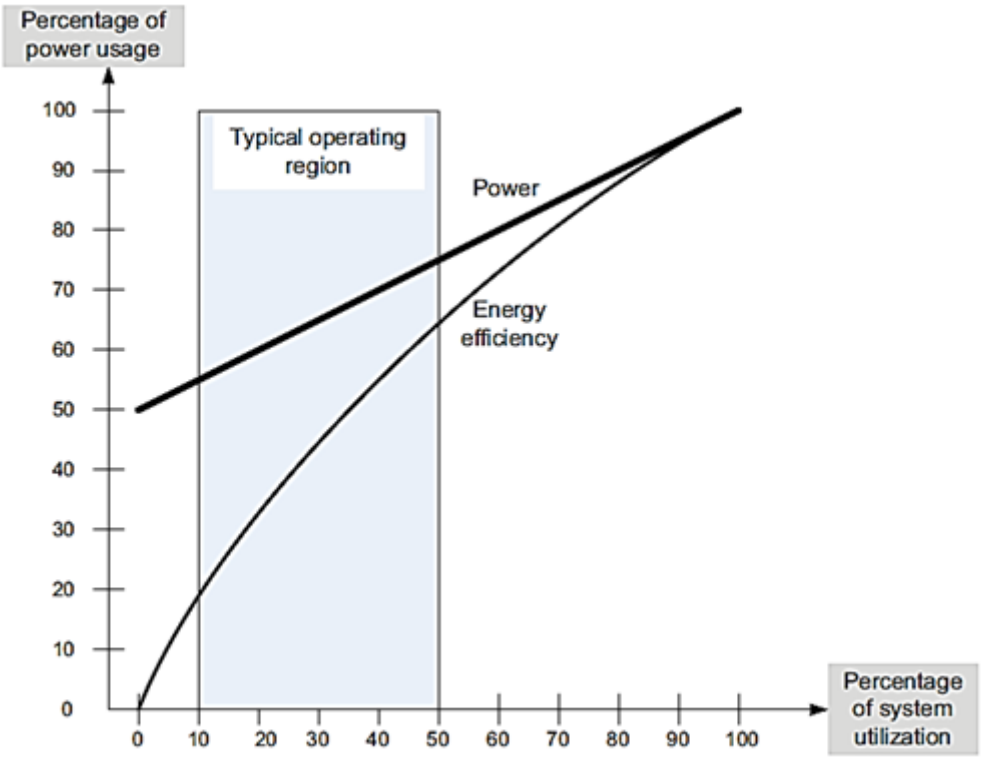
Wang, X., and Liu, Z., proved their placement method with the aid of using simulation effects. The simulation effects display that this set of rules cannot most effective boom aid usage rates, however additionally make the records middle greater power-efficient (Wang, X., and Liu, Z., 2012).

Cardosa, M. et al., describe a completely unique spatio tradeoff that consists of green spatial becoming of VMs on datacenters to gain excessive usage of gadget resources, in addition to balanced temporal becoming of servers with VM which have comparable run-times to make sure a server executables at an excessive usage at some point of its uptime. They endorse VM provisioning policy that explicitly include those tradeoffs. Further, they endorse strategies that actively scale MapReduce clusters to in addition enhance electricity

Table 1. Energy management in high performance computing (Czarnul, P., et al., 2019)

Vendor	Tool	Device type	Supported	Description
Intel	RAPL	CPU	Since Sandy Bridge generation	Used for performance vs maximum power measurements
AMD	APM	CPU	Since bulldozer	Developers guide describing the capabilities of the AMD TDP power cap
IBM	Energy scale	CPU	Since POWER6	Overview on POWER7 power management capabilities
NVIDIA	NVML/nvidia-smi	GPU	Most of tesla, Quadro, Titan, and GRID lines	Discussion on differences of using DVFS on CPU and GPU

Figure 3. Energy necessities with the load (Paya, A., & Marinescu, D., 2013)



intake even as making sure that jobs meet or enhance their predicted runtimes. The proposed algorithms gain electricity financial savings over present placement strategies, and an extra optimization approach in addition achieves financial savings even as concurrently enhancing activity performance (Cardosa, M. et al., 2012).

Author suggest a unique rolling-horizon provisioning structure for real-time provisioning in virtualized clouds. Then an undertaking-orientated electricity intake version is produced and examined. According to proposed provisioning structure, they broaden a unique electricity-conscious scheduling set of rules EARH for real-time, periodic, unbiased responsibilities. EARH employs optimization coverage and also can be prolonged to combine different electricity-conscious scheduling algorithms. Furthermore, they suggest techniques in phrases of useful resource scaling up and cutting down to make an awesome trade-off among undertaking's provision ability and electricity conservation. The simulation results artificial responsibilities in addition to responsibilities following the final model of the Google cloud data are performed to validate the prevalence of EARH with the aid of using evaluating it with a few baselines. The simulation effects display that EARH notably improves the provisioning first-rate of others and it's far appropriate for real-time undertaking provisioning in VM (Zhu, X. et al., 2014).

EWRR set of rules improved scheduler video display units and evaluates the going for walks VMs popularity for viable venture VM Migration. It observes VMs usage price to begin stay migration and over-applied Processing Element (PE). (Alnowiser, A. et al., 2014).

Author affords a singular stochastic framework for power performance and overall performance evaluation of DVS-enabled cloud. The proposed system makes use of digital system request arrival charge, failure charge, restore charge, and provider charge of datacenter as version inputs. This paper

offers analytic answers of 3 metrics. The proposed System may be used to assist the layout and optimization of power-conscious excessive overall performance cloud systems (Xia, Y. et al., 2014).

The load balancing and elastic algorithms additionally make the most a number of the maximum proper functions of server consolidation mechanisms mentioned within side the literature (Paya, A. and Marinescu, D 2015).

Author present a powerful capacity-sharing mechanism in cloud surroundings which could cause an international power sustainability coverage. A coalition sport idea changed into applied to version diverse interactions amongst providers. However, not like the present approaches, the proposed sport version seems for a hard and fast of low-power-price CPs in an association and presents a fare and appropriate sales for them. In addition, they keep in mind the call for versions of inner customers of a CP while sharing VM resources. In continue to this, an in depth evaluation of diverse charges and sales elements is presented (Hassan, M. et al., 2015).

Gai, K. el at., proposed EA-HRM2 that is hold up with the aid of using a first-rate set of rules Optimal Task Assignment (OTA) set of rules. The result reviews have proved the proposed method is powerful to save electricity while deploying heterogeneous embedded structures in cellular cloud structures (Gai, K. el at., 2016). He again proposed EPRF method to enhance his method in 2020 (Gai, K., et al., 2020).

Author supplied a hybrid energy-aware useful resource allocation technique to help requestors gather energy-efficient and happy manufacturing services. The problem description on energy-aware useful resource allocation in CMfg is first summarized. Then a close-by desire technique based totally mostly on fuzzy similarity degree is located forth to benefit appropriate candidate services. (Zheng, H. et al., 2017) where Fletscher, L. et al., presented fuzzy based solution (Fletscher, L. et al., 2018).

Author Liu, S. et al., proposed a DWT-MAC protocol, modifications consistent with the quantity of senders, that can make sure that it usually procedures the optimum value (Liu, S. et al., 2020).

On-call for electricity-green useful resource allocation version is designed primarily based totally in this version. Its capabilities attractiveness ratios which can be 11%–17% better than present answers and 9% decrease electricity consumption. (Liu, P. et al., 2018).

Hassan, H. et al., provide simulates the electricity intake as much as 18%, the share of closing date happy offerings as much as 14% and the common reaction time as much as 10% in evaluation with the second-first-class results (Hassan, H. et al., 2020).

Author's reviews on a power green interoperable cloud structure realized as a cloud toolbox that makes a specialty of decreasing the power intake of cloud packages holistically throughout all deployment models. The structure helps power performance at carrier production, deployment and operation. They speak their realistic revel in all through implementation required to facilitate production of power conscious cloud packages. They perform an overall performance assessment of the element on a cloud testbed. The outcomes display the overall performance of Virtual Machine production, in general constrained via way of means of to be had I/O, to be good enough for agile, power conscious software program development. They finish the VMIC is feasible, incurs minimum overall performance overhead relatively to the time taken via way of means of different factors of the cloud utility production life-cycle, and make tips on improving its overall performance (Armstrong, D et al., 2017).

Cloud provider, display that the supply of inexperienced electricity have a massive effect on most reliable electricity control rules and that the contribution of the community is a ways from being negligible (Addis, B., et al., 2014).

Authors Shafqat, S., et al., suggest a brand new mobile verbal exchange structure that integrates power conscious cloud with social conscious tool to tool verbal exchange (Shafqat, S., et al., 2019).

Authors introduces a reinforcement getting to know technique embedded in a clever agreement to similarly reduce the power value. Because the reinforcement getting to know technique is knowledgeable from the ancient knowledge, it is based on no request arrival and power supply.

Experimental effects on Google cluster strains and real-global energy charge display that our technique is capable of lessen the datacenters' value substantially as compared with different benchmark algorithms (Xu, C., et al., 2017).

Xu, H., et al., proposed an optimization model, a heuristic-primarily based totally set of rules known as greedy-primarily based totally load stability (GBLB) set of rules is developed. Since lowering the quantity of lively servers commonly will increase the quantity of VM migrations, we similarly reduce the quantity of VM migrations within side the proposed GBLB set of rules. Simulation consequences display that, as compared with different 3 famous algorithms, the proposed GBLB set of rules can lessen the quantity of lively servers and obtain the satisfactory load balancing stage on the fee of some extra migrations (Xu, H., et al., 2017).

Authors suggest a heuristic primarily based totally aid allocation of VM choice and a VM allocation method that goals to reduce the overall electricity intake and running charges even as assembly the client-stage SLA. Our test effects show good sized upgrades in cloud providers' earnings and electricity financial savings even as enhancing the SLA at a sure stage (Sabbir Hasan, M. and Huh, E. N. 2013).

It is then prolonged to a greater trendy multi-goal scheme which at the same time optimizes the sum electricity and throughput retaining a stability among them. In each of the schemes, quality-of-provider is assured in phrases of cease-to-cease signal-to-noise ratio. They count on the presence of estimation mistakes in channel country information. A set of rules to beautify equity amongst customers in those schemes is likewise supplied. Simulation outcomes are supplied to affirm the overall performance of proposed schemes in phrases of electricity efficiency, device throughput, outage probability, and equity to cease customers (Devarajan, R., et al., 2012)

Author advocate polynomial heuristics problems in three levels. Simulation experiments display that during all 3 instances a few heuristics can obtain consequences near optimal, i.e., result in desirable activity performance at the same time as retaining energy (Borgetto, D., et al., 2012).

COMPARATIVE ANALYSIS

Table 2 presents the comparative analysis of different energy-aware load balancing techniques in cloud computing.

OPEN RESEARCH CHALLENGES AND ISSUES

With the above research analysis and reviews, the following open areas are for the improvements of energy-aware load balancing systems:

- **Optimization:** Energy optimization in cloud, servers, and datacenters are the critical areas where VM need to be optimized. However, many research have been proposed by eminent researcher, but still there are some scope to be further enhancements (Resma, K. S. et al., 2021).
- **VM Usage:** Based on energy consumption, VM usage are very low. VM utilization can be increased by applying some optimization algorithms (Agarwal, A. et al., 2020).
- **SLA Violation Rate:** SLA violation rate is very high, due to this many cloud user distract to opt the services (Venkatadri, M., and Pasricha, A. (2019).
- **Fault-tolerant:** It is the backbone of VM utilization, however many researcher done some extraordinary works, still hybrid methods are not considered, and optimization for threshold is required (Gupta, P. et al., 2021).
- **Execution Time and Average Cost:** Cost is directly proportional to energy consumption and execution time. So run time and energy consumption of VM and datacenters are need to be minimize to lower the operating cost (Dewangan et al., 2018).

Table 2. Comparative analysis of different energy-aware load balancing techniques

Technique	Method	Platform	Challenges	Parameters				
				C	QoS	ET	VMu	SLA
EECLOUD (Li, Y. et al., 2012).	MapReduce, Dynamic VM placement and VM allocation strategies	Private cloud	Unable to find the remaining run-time for tasks which runs slow	✓	✓	✓	✓	
ESTT (Cardosa, M. et al., 2012).	MapReduce, VM provisioning policy	Cloudsim				✓		
EnReal (Xu, X. et al., 2015).	Dynamic VM deployment and scientific workflow	Cloudsim	Dynamic VM deployment needs in real cloud to validate the accuracy			✓		
EARMCD (Rezai, H., and B Speily, O. R. 2017).	heuristic approach	CloudSim	Resource optimization is missing	✓				
VPME (Zhou, Z., et al., 2018).	VM deployment based on CPU and Memory usage	Private cloud	Energy improvements amount is not measured	✓				✓
EA-HCM (Gai, K. et al. 2018).	Heterogeneous embedded systems	Cloudsim	Energy efficiency not measured	✓		✓		
OPEX Cloud (Aldossary, M. et al., 2019).	Cloud workload patterns	Cloud testbed	Cost of utilization in all scenarios are not measured, execution time varies	✓			✓	
SMART (Dewangan B.K. et al., 2019)	Antlion optimizer	Cloudsim	Fault tolerant optimization is missing	✓	✓	✓	✓	✓
SAG (Feng, H. et al., 2021).	virtual-machine placement and migration	Private cloud			✓	✓		
WARMS (Dewangan B.K. et al., 2021)	GreyWolf optimizer, Fuzzy logic	Cloudsim	Execution time can be increased	✓	✓		✓	✓

C- Cost, QoS- Quality of Service, ET- Execution Time, VMu- VM Utilization, SLA- Service level agreement

CONCLUSION

Web of Science (WoS) database is consider to find the energy-aware load balancing techniques from 2009 to 2021. Moreover, 370 research article filtered through WoS website database and about 100 research article considered for the review. The energy-aware load balancing methods are classified into four major categories, and based on that research are classified as well. The research for each classification are distributed from year 2009 to most recent work published in 2021. In this, we have used "Energy-aware VM Scheduling", "VM-Scheduling", as search key. The outcome of this review are major issues and challenges, which is discussed in the above section. The Exploration by Pike Research, a perfect innovation market knowledge firm, guarantees the appropriation of distributed computing could prompt a 38 percent decrease in energy utilization on the planet's server farms by 2020. It gauges the development of cloud computing will diminish energy utilization from the current

pace of 201.8 terawatt hours to a 2020 pace of 139.8 TWh, bringing about a 28 percent decrease in ozone harming substance outflows in the following five years. is review outcome is based on the comparative analysis done in table 2. The cloud researcher may consider the open research issues and challenges to enhance the existing energy-aware load balancing technique in cloud.

ACKNOWLEDGMENT

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Addis, B., Ardagna, D., Capone, A., & Carello, G. (2014). Energy-aware joint management of networks and cloud infrastructures. *Computer Networks*, *70*, 75–95. doi:10.1016/j.comnet.2014.04.011

Agarwal, A., Choudhury, T., Pasricha, A., & Chandra Satapathy, S. (2020). Extensive review of cloud resource management techniques in industry 4.0: Issue and challenges. *Software, Practice & Experience*.

Aldossary, M., Djemame, K., Alzamil, I., Kostopoulos, A., Dimakis, A., & Agiatzidou, E. (2019). Energy-aware cost prediction and pricing of virtual machines in cloud computing environments. *Future Generation Computer Systems*, *93*, 442–459. doi:10.1016/j.future.2018.10.027

Alnowiser, A., Aldhahri, E., Alahmadi, A., & Zhu, M. M. (2014, March). Enhanced weighted round robin (EWRR) with DVFS technology in cloud energy-aware. In *2014 International conference on computational science and computational intelligence* (Vol. 1, pp. 320-326). IEEE. doi:10.1109/CSCI.2014.62

Alzamil, I., Djemame, K., Armstrong, D., & Kavanagh, R. (2015). Energy-aware profiling for cloud computing environments. *Electronic Notes in Theoretical Computer Science*, *318*, 91–108. doi:10.1016/j.entcs.2015.10.021

Armstrong, D., Djemame, K., & Kavanagh, R. (2017). Towards energy aware cloud computing application construction. *Journal of Cloud Computing*, *6*(1), 1–13.

Borgetto, D., Casanova, H., Da Costa, G., & Pierson, J. M. (2012). Energy-aware service allocation. *Future Generation Computer Systems*, *28*(5), 769–779. doi:10.1016/j.future.2011.04.018

Cardosa, M., Singh, A., Pucha, H., & Chandra, A. (2012). Exploiting spatio-temporal tradeoffs for energy-aware mapreduce in the cloud. *IEEE Transactions on Computers*, *61*(12), 1737–1751. doi:10.1109/TC.2012.166

Czarnul, P., Proficz, J., & Krzywaniak, A. (2019). Energy-aware high-performance computing: Survey of state-of-the-art tools, techniques, and environments. *Scientific Programming*, *2019*, 2019. doi:10.1155/2019/8348791

Devarajan, R., Jha, S. C., Phuyal, U., & Bhargava, V. K. (2012). Energy-aware resource allocation for cooperative cellular network using multi-objective optimization approach. *IEEE Transactions on Wireless Communications*, *11*(5), 1797–1807. doi:10.1109/TWC.2012.030512.110895

Dewangan, B. K., Agarwal, A., Venkatadri, M., & Pasricha, A. (2019). Design of self-management aware autonomic resource scheduling scheme in cloud. *International Journal of Computer Information Systems and Industrial Management Applications*, *11*, 170–177.

Dewangan, B. K., Agarwal, A., Choudhury, T., & Pasricha, A. (2021). Workload aware autonomic resource management scheme using grey wolf optimization in cloud environment. *IET Communications*, *15*(14), 1869–1882. doi:10.1049/cmu2.12198

Dewangan, B. K., Agarwal, A., Venkatadri, M., & Pasricha, A. (2018, December). Autonomic cloud resource management. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 138-143). IEEE. doi:10.1109/PDGC.2018.8745977

Djemame, K., Bosch, R., Kavanagh, R., Alvarez, P., Ejarque, J., Guitart, J., & Blasi, L. (2017). Paas-iaas inter-layer adaptation in an energy-aware cloud environment. *IEEE Transactions on Sustainable Computing*, *2*(2), 127–139. doi:10.1109/TSUSC.2017.2719159

Farahnakian, F., Pahikkala, T., Liljeberg, P., & Plosila, J. (2013, December). Energy aware consolidation algorithm based on k-nearest neighbor regression for cloud data centers. In *2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing* (pp. 256-259). IEEE. doi:10.1109/UCC.2013.51

Farahnakian, F., Pahikkala, T., Liljeberg, P., Plosila, J., Hieu, N. T., & Tenhunen, H. (2016). Energy-aware VM consolidation in cloud data centers using utilization prediction model. *IEEE Transactions on Cloud Computing*, *7*(2), 524–536. doi:10.1109/TCC.2016.2617374

Feng, H., Deng, Y., & Li, J. (2021). A global-energy-aware virtual machine placement strategy for cloud data centers. *Journal of Systems Architecture*, *116*, 102048. doi:10.1016/j.sysarc.2021.102048

Fletscher, L. A., Suárez, L. A., Grace, D., Peroni, C. V., & Maestre, J. M. (2018). Energy-aware resource management in heterogeneous cellular networks with hybrid energy sources. *IEEE eTransactions on Network and Service Management*, *16*(1), 279–293. doi:10.1109/TNSM.2018.2866533

Gai, K., Qin, X., & Zhu, L. (2020). An energy-aware high performance task allocation strategy in heterogeneous fog computing environments. *IEEE Transactions on Computers*.

Gai, K., Qiu, M., & Zhao, H. (2018). Energy-aware task assignment for mobile cyber-enabled applications in heterogeneous cloud computing. *Journal of Parallel and Distributed Computing*, *111*, 126–135. doi:10.1016/j.jpdc.2017.08.001

Gai, K., Qiu, M., Zhao, H., Tao, L., & Zong, Z. (2016). Dynamic energy-aware cloudlet-based mobile cloud computing model for green computing. *Journal of Network and Computer Applications*, *59*, 46–54. doi:10.1016/j.jnca.2015.05.016

Gupta, P., Sahoo, P. K., & Veeravalli, B. (2021). Dynamic Fault Tolerant Scheduling with Response Time Minimization for Multiple Failures in Cloud. *Journal of Parallel and Distributed Computing*, *158*, 80–93. doi:10.1016/j.jpdc.2021.07.019

Hassan, H. O., Azizi, S., & Shojafar, M. (2020). Priority, network and energy-aware placement of IoT-based application services in fog-cloud environments. *IET Communications*, *14*(13), 2117–2129. doi:10.1049/iet-com.2020.0007

Hassan, M. M., Abdullah-Al-Wadud, M., Almogren, A., Song, B., & Alamri, A. (2015). Energy-aware resource and revenue management in federated cloud: A game-theoretic approach. *IEEE Systems Journal*, *11*(2), 951–961. doi:10.1109/JSYST.2015.2472973

Hongyou, L., Jiangyong, W., Jian, P., Junfeng, W., & Tang, L. (2013). Energy-aware scheduling scheme using workload-aware consolidation technique in cloud data centres. *China Communications*, *10*(12), 114–124. doi:10.1109/CC.2013.6723884

Jafari, V., & Rezvani, M. H. (2021). Joint optimization of energy consumption and time delay in IoT-fog-cloud computing environments using NSGA-II Metaheuristic algorithm. *Journal of Ambient Intelligence and Humanized Computing*, 1–24. doi:10.1007/s12652-021-03388-2

Katsaros, G., Subirats, J., Fitó, J. O., Guitart, J., Gilet, P., & Espling, D. (2013). A service framework for energy-aware monitoring and VM management in Clouds. *Future Generation Computer Systems*, *29*(8), 2077–2091. doi:10.1016/j.future.2012.12.006

Li, X., Yu, W., Ruiz, R., & Zhu, J. (2020). Energy-aware cloud workflow applications scheduling with geo-distributed data. *IEEE Transactions on Services Computing*, 1. doi:10.1109/TSC.2020.2965106

Li, Y., Wang, Y., Yin, B., & Guan, L. (2012, September). An energy efficient resource management method in virtualized cloud environment. In *2012 14th Asia-Pacific Network Operations and Management Symposium (APNOMS)* (pp. 1-8). IEEE.

Liu, P., Chaudhry, S. R., Huang, T., Wang, X., & Collier, M. (2018). Multi-factorial energy aware resource management in edge networks. *IEEE Transactions on Green Communications and Networking*, *3*(1), 45–56. doi:10.1109/TGCN.2018.2874397

Liu, S., Huang, G., Gui, J., Wang, T., & Li, X. (2020). Energy-aware MAC protocol for data differentiated services in sensor-cloud computing. *Journal of Cloud Computing*, *9*(1), 1–33.

Mandal, U., Habib, M. F., Zhang, S., Mukherjee, B., & Tornatore, M. (2013). Greening the cloud using renewable-energy-aware service migration. *IEEE Network*, *27*(6), 36–43. doi:10.1109/MNET.2013.6678925

Paya, A., & Marinescu, D. C. (2013). *Energy-aware application scaling on a cloud.* arXiv preprint arXiv:1307.3306.

Paya, A., & Marinescu, D. C. (2015). Energy-aware load balancing and application scaling for the cloud ecosystem. *IEEE Transactions on Cloud Computing*, *5*(1), 15–27. doi:10.1109/TCC.2015.2396059

Panja, S., & Roy, B. (2021). A secure end-to-end verifiable e-voting system using blockchain and cloud server. *Journal of Information Security and Applications*, *59*, 102815. doi:10.1016/j.jisa.2021.102815

Kumar, R. (2007). *US datacenters: The calm before the storm*. Gartner White Paper.

Ren, Y., Leng, Y., Qi, J., Sharma, P. K., Wang, J., Almakhadmeh, Z., & Tolba, A. (2021). Multiple cloud storage mechanism based on blockchain in smart homes. *Future Generation Computer Systems*, *115*, 304–313. doi:10.1016/j.future.2020.09.019

Rezai, H., & Speily, , O. R. (2017). Energy aware resource management of cloud data centers. *International Journal of Engineering*, *30*(11), 1730–1739.

Resma, K. S., Sharvani, G. S., & Somula, R. (2021). Optimization of cloud load balancing using fitness function and duopoly theory. *International Journal of Intelligent Computing and Cybernetics*.

Sabbir Hasan, M., & Huh, E. N. (2013). Heuristic based energy-aware resource allocation by dynamic consolidation of virtual machines in cloud data center. *Transactions on Internet and Information Systems (Seoul)*, *7*(8), 1825–1842. doi:10.3837/tiis.2013.08.005

Shafqat, S., Kishwer, S., & Qureshi, M. A. (2019, January). Energy-aware cloud architecture for intense social mobile (device to device) 5G communications in smart city. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 739-745). IEEE.

Shende, M. P. (2012). Survey on user behavior trust evaluation in cloud computing. *International Journal of Science*, *Engineering, and Technology Research*, *1*(5), 1–5.

Vafamehr, A., & Khodayar, M. E. (2018). Energy-aware cloud computing. *The Electricity Journal*, *31*(2), 40–49. doi:10.1016/j.tej.2018.01.009

Venkatadri, M., & Pasricha, A. (2019). Sla-based autonomic cloud resource management framework by antlion optimization algorithm. *International Journal of Innovative Technology and Exploring Engineering*, *8*(4), 119–123.

Wang, X., & Liu, Z. (2012, January). An energy-aware VMs placement algorithm in cloud computing environment. In *2012 Second International Conference on Intelligent System Design and Engineering Application* (pp. 627-630). IEEE. doi:10.1109/ISdea.2012.467

Xia, Y., Zhou, M., Luo, X., Pang, S., & Zhu, Q. (2014). A stochastic approach to analysis of energy-aware DVS-enabled cloud datacenters. *IEEE Transactions on Systems, Man, and Cybernetics. Systems*, *45*(1), 73–83.

Xu, C., Wang, K., & Guo, M. (2017). Intelligent resource management in blockchain-based cloud datacenters. *IEEE Cloud Computing*, *4*(6), 50–59. doi:10.1109/MCC.2018.1081060

Xu, H., & Yang, B. (2017). Energy-Aware Resource Management in Cloud Computing Considering Load Balance. *Journal of Information Science and Engineering*, *33*(1).

Xu, X., Dou, W., Zhang, X., & Chen, J. (2015). EnReal: An energy-aware resource allocation method for scientific workflow executions in cloud environment. *IEEE Transactions on Cloud Computing, 4*(2), 166-179.

Zhang, X., Wu, T., Chen, M., Wei, T., Zhou, J., Hu, S., & Buyya, R. (2019). Energy-aware virtual machine allocation for cloud with resource reservation. *Journal of Systems and Software*, *147*, 147–161. doi:10.1016/j.jss.2018.09.084

Zheng, H., Feng, Y., & Tan, J. (2017). A hybrid energy-aware resource allocation approach in cloud manufacturing environment. *IEEE Access: Practical Innovations, Open Solutions*, *5*, 12648–12656. doi:10.1109/ACCESS.2017.2715829

Zhou, Z., Abawajy, J., Chowdhury, M., Hu, Z., Li, K., Cheng, H., & Li, F. (2018). Minimizing SLA violation and power consumption in Cloud data centers using adaptive energy-aware algorithms. *Future Generation Computer Systems*, *86*, 836–850. doi:10.1016/j.future.2017.07.048

Zhu, X., Yang, L. T., Chen, H., Wang, J., Yin, S., & Liu, X. (2014). Real-time tasks oriented energy-aware scheduling in virtualized clouds. *IEEE Transactions on Cloud Computing*, *2*(2), 168–180. doi:10.1109/TCC.2014.2310452

Ram Narayan Shukla is pursuing a PhD from LNCT University, Bhopal in the Department of Computer Science and Engineering. Their research area is Cloud Computing.

Anoop Kumar Chaturvedi is working as a professor in the department of Computer Science and Engineering LNCT University Bhopal since last 10 years. Working in the field of cloud computing, wireless sensor networks and mobile ad hoc networks and published many research papers in the many reputed journals and conferences.