iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42887

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

An Optimization Framework of Adaptive Computing-plus-Communication for Multimedia Processing in Cloud: A Review

Kiran Bhimrao Dabhade[1], C. M. Mankar[2]
[1, 2]*Department of Computer Science and Engineeringt, Sant Gadge Baba Amravati University*

Abstract: ***Clear trend within the evolution of network-based services is that the ever-increasing amount of multimedia system data concerned. This trend towards big-data multimedia system process finds its natural placement at the side of the adoption of the cloud computing paradigm, that looks the most effective solution to the strain of a extremely fluctuating work that characterizes this sort of services. However, as cloud data centers become a lot of and a lot of powerful, energy consumption becomes a significant challenge each for environmental concerns and for economic reasons. An effective approach to improve energy efficiency in cloud data centers is to rely on traffic engineering techniques to dynamically adapt the number of active servers to the current workload. Towards this aim, we propose a joint computing-plus-communication improvement framework exploiting virtualization technologies. Our proposal specifically addresses the everyday situation of data processing processing with computationally intensive tasks and exchange of a giant volume of data. The proposed framework not only ensures users the Quality of Service, but also achieves maximum energy saving and attains green cloud computing goals in a fully distributed fashion by utilizing the DVFS-based CPU frequencies***

Keywords*: Energy efficiency, Multimedia data processing, Cloud resource management, Load balancing, Dynamic voltage and frequency scaling (DVFS), Traffic engineering.*

I. INTRODUCTION

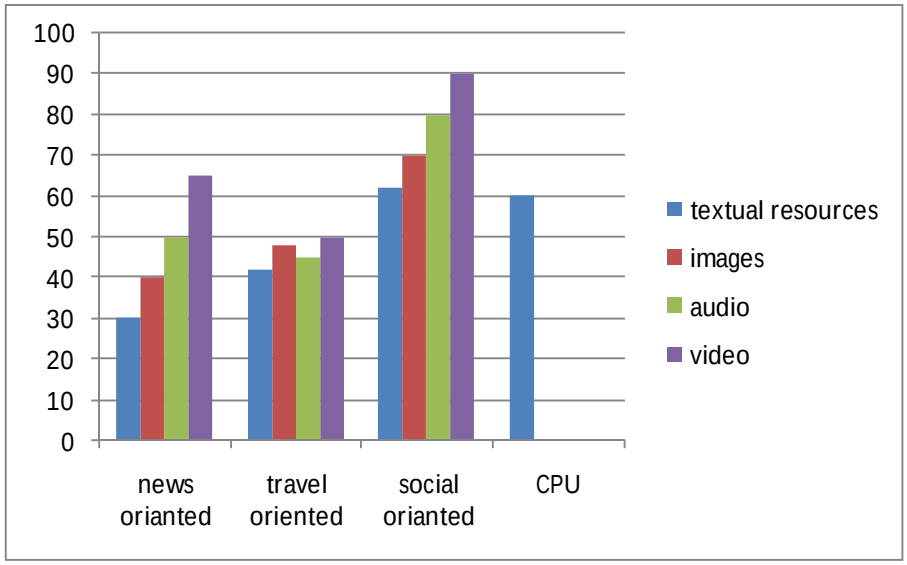
Internet based services are evolving towards an ever increasing amount of multimedia content, both in terms of number of resource[1]s and of their size (e.g., higher resolution videos) For example, the number of videos uploaded on YouTube recently reached a value in the order of 300 hours per minute, with videos that can be viewed both from small mobile devices or from full HD screens This trend determines an evolution towards workloads characterized by significantly increased computational and communication requirements and higher variability, with major fluctuations throughout the day. To cope with such demands, cloud computing seems a very promising approach, because it provides an elastic, pay-as you-go pricing model that can be used to address workload fluctuations, while the large data centers, that are typical of cloud infrastructures, can provide the computational power required to manage the huge amount of multimedia data of modern applications. This paper provide the solution to address this problem, we propose a new optimization framework, to reduce the energy consumption of computing, communication and infrastructure reconfiguration in a cloud data center. Our approach operates at the graininess of a data chunk, that will be either an image or a part of a video stream, reconfiguring the cloud infrastructure as required. We model SLA as a constraint on the computational and communication time to process a data chunk. During a nutshell, the most goal of this work is to introduce a completely unique framework that minimizes joint computing-plus-communication energy in cloud data centers. Our solution, which is especially designed to work with multimedia applications characterized by variable workload.

II. LITERATURE SURVEY

The widespread diffusion of the Mobile web opens several attention-grabbing design and management problems concerning the server infrastructures that may need to satisfy this and future client demand. Future Mobile Web-based services will have growing computational costs because even requests for the same Web resource will require the dynamic generation of different contents that will take into account specific devices, user profiles and contexts. [1] considers the evolution of the Mobile Web workload and of the technology trends of servers and client devices with the goal of evaluating possible bottlenecks and anticipating some management strategies for the server infrastructures that got to support the long run services for the Mobile web. They shows the comparison of predictive performance between Mobile Web-based service computational costs vs. CPU computational power in fig 1.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com



Graph 1: All workload scenarios

[2] was prepared as an account of work sponsored by the United States Government. While the document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of Californianor any of their employees, makes any guarantee, specific or implicit , or assumes any liability for the accuracy, completeness, or quality of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. reference herein to any specific business product, process, or service by its trade name, trademark, manufacturer, or otherwise, doesn't essentially represent or imply its endorsement, recommendation, or affirmative by the united states Government or any agency therefrom, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Recently dense trajectories were shown to be associate an video illustration for action recognition and achieved progressive results on a range of datasets [5] improves their performance by taking into account camera motion to correct them. To estimate camera motion, we match feature points between frames using SURF descriptors and dense optical flow, which are shown to be complementary. These matches are, then, accustomed robustly estimate a homography with RANSAC. Human motion is generally totally different from camera motion and generates inconsistent matches. to enhance the estimation, a human detector is used to remove these matches.. To improve the estimation, a human detector is employed to remove these matches. Given the estimated camera motion, they remove trajectories consistent with it. they also use this estimation to cancel out camera motion from the optical flow. This considerably improves motion-based descriptors, like HOF and MBH. Experimental results on four difficult action datasets (i.e., Hollywood2, HMDB51, Olympic Sports and UCF50) considerably outstrip this state of the art.

Object detection performance, as measured on the canonical PASCAL VOC dataset, has plateaued within the previous few years. The best-performing strategies are complicated ensemble systems that generally combine multiple low-level image features with high-level context. In [6]propose a simple and scalable detection algorithm that improves mean average precision (mAP) by more than 30% relative to the previous best result on VOC 2012—achieving a mAP of 53.3%. Their approach combines two key insights: (1) one can apply high-capacity convolutional neural networks (CNNs) to bottom-up region proposals in order to localize and segment objects and (2) once tagged training data is scarce, supervised pre-training for associate auxiliary task, followed by domain-specific fine-tuning, yields a big performance boost. Since they combine region proposals with CNNs, we call our method R-CNN: Regions with CNN features.

With the advent of future generation mobile communication technologies (5G), there is the potential to allow mobile users to have access to big data processing over different clouds and networks. The increasing numbers of mobile users come with extra expectations for customized services (e.g., social networking, smart home, health monitoring) at any time, from anyplace, and through any means that of connectivity.

©IJRASET: All Rights are Reserved

Because of the expected massive amount of complex data generated by such services and networks from heterogeneous multiple sources, an infrastructure is required to recognize a user's sentiments (e.g., emotion) and behavioral patterns to produce a top quality mobile user experience. To this end, [7] proposes an infrastructure that combines the potential of emotion aware big data and cloud technology towards 5G. With this proposed infrastructure, a bimodal system of massive data emotion recognition is proposed, wherever the modalities include speech and face video. Experimental results show that the proposed approach achieves 83.10 % emotion recognition accuracy using bimodal inputs. To show the suitability and validity of the proposed approach, Hadoop-based distributed processing is used to speed up the processing for heterogeneous mobile clients.

Today's data centers want economical traffic management to enhance resource utilization in their networks. In [8], Tian local area network et.al. study a joint tenant (e.g., server or virtual machine) placement and routing drawback to reduce traffic costs. These 2 complementary degrees of freedom—placement and routing—are mutually-dependent, however, are usually optimized individually in today's data centers. Leveraging and expanding the technique of Markov approximation, they propose an efficient online algorithm in a dynamic environment under changing traffic loads. The algorithm needs a really little number of virtual machine migrations and is simple to implement in practice. Performance evaluation that employs the real data center traffic traces under a spectrum of elephant and mice flows, demonstrates a consistent and significant improvement over the benchmark achieved by common heuristics. [9] was prepared as an account of work sponsored by the United States Government. whereas this document is believed to contain correct information, neither the u. s. Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, doesn't business represent or imply its endorsement, recommendation, or favoring by the united states Government or any agency therefrom, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Rahul Urgaonkar et.al. [10] investigate optimum resource allocation and power management in virtualized data centers with time-varying workloads and heterogeneous applications. previous work in this area uses prediction primarily based approaches for resource provisioning. In this work, we take an alternate approach that makes use of the queueing information available in the system to make online control decisions. Specifically, they use the recently developed technique of Lyapunov Optimization to design an online admission control, routing, and resource allocation algorithm for a virtualized data center. algorithm maximizes a joint utility of the typical application output and energy prices of the data center. their approach is adaptive to unpredictable changes in the workload and does not require estimation and prediction of its statistics.

With worldwide shipments of smartphones (487.7 million) exceeding PCs (414.6 million including tablets) in 2011 , and in the US alone, more users predicted to access the Internet from mobile mobile devices than from PCs by 2015, clearly there's a need to be able to use mobile devices and networks like we tend to use PCs and wireline networks nowadays. However, in spite of advances within the capabilities of mobile devices, a gap can still exist, and should even widen, with the necessities of wealthy multimedia applications. Mobile cloud computing can help bridge this gap, providing mobile applications the capabilities of cloud servers and storage together with the benefits of mobile devices and mobile connectivity, possibly enabling a brand new generation of really present multimedia applications on mobile devices: Cloud Mobile Media (CMM) applications. In [12], Sujit Dey et.al. consider early trends, and opportunities and advantages for brand new CMM applications and services. We analyze the challenges imposed by mobile cloud computing that need to be addressed to make CMM applications viable, including response time, user experience, cloud computing cost, mobile network bandwidth, and scalability to large number of CMM users, besides other important cloud computing issues like energy consumption, privacy, and security. They illustrate the challenges using Cloud Mobile Gaming (CMG), an approach that enables rich multiplayer Internet games on mobile devices, where compute intensive tasks like graphic rendering are executed on cloud servers in response to gaming commands on a mobile device, and the resulting video has to be streamed back to the mobile device in near real time, making it the most challenging of the CMM applications. Subsequently, they focus in this paper on developing adaptive mobile cloud computing techniques to address the CMG challenges. Specifically, they propose a rendering adaptation technique, which can dynamically vary the richness and complexity of graphic rendering depending on the network and cloud computing constraints, thereby impacting both the bit rate of the rendered video that needs to be streamed back from the cloud server to the mobile device, and the computation load on the CMG servers. Experiments conducted on a cellular network demonstrate that their proposed technique can significantly improve user experience, and ensure scalability of the CMG approach in terms of both network bandwidth and server computational need.

Yi-Wei Ma et.al. [13] proposes a power saving mechanism for embedded system and multimedia streaming service design based on digital signal processor. It at the same time achieves 2 main functions: a high-quality multimedia service in cloud computing, and an influence saving management mechanism for extending the period of time of hand-held devices used on multimedia streaming server applications.

Their paper enables dynamic voltage and frequency scaling power control mechanisms to enable frequency and voltage adjustment according to the system state of a device. The proposed mechanism increases the frame rate of multimedia services to 16.66%/s and mitigates power consumption by 1.5%.

Internet-scale distributed systems like content delivery networks (CDNs) operate many thousands of servers deployed in thousands of data center locations round the globe. Since the energy costs of operating such a large IT infrastructure are a significant fraction of the total operating costs, we argue for redesigning CDNs to incorporate energy optimizations as a first-order principle. [14] propose techniques to turn off CDN servers during periods of low load while seeking to balance three key design goals: maximize energy reduction, minimize the impact on client-perceived service availibility (SLAs), and limit the frequency of on-off server transitions to scale back wear-and-tear and its impact on hardware dependability. They propose an optimal offline algorithm and an online algorithm to extract energy savings both at the level of local load balancing within a data center and global load balancing across data centers.

They evaluate their algorithms using real production workload traces from a large commercial CDN. Their results show that it is possible to reduce the energy consumption of a CDN by 51% while ensuring a high level of availability that meets customer SLA requirements and incurring an average of one on-off transition per server per day. Further, they show that keeping even 10% of the servers as hot spares helps absorb load spikes due to global flash crowds and minimize any impact on availability SLAs. Finally, they show that redistributing load across highly proximal data centers can enhance service availability significantly, but has only a modest impact on energy savings.

There is growing incentive to reduce the power consumed by large-scale data centers that host online services such as banking, retail commerce, and gaming. Virtualization could be a promising approach to consolidating multiple online services onto a smaller range of computing resources. A virtualized server environment allows computing resources to be shared among multiple performance-isolated platforms called virtual machines. By dynamically provisioning virtual machines, consolidating the work, and turning servers on and off as required, data center operators will maintain the specified quality-of-service (QoS) whereas achieving higher server. utilization and energy efficiency. [15] implement and validate a dynamic resource provisioning framework for virtualized server environments wherein the provisioning problem is posed as one of sequential optimization under uncertainty and solved using a lookahead control scheme. The proposed approach accounts for the switch costs incurred whereas provisioning virtual machines and expressly encodes the corresponding risk within the optimization downside. Experiments victimisation the Trade6 enterprise application show that a server cluster managed by the controller conserves, on average, twenty sixth of the ability needed by a system control dynamic management whereas still maintaining QoS goals.

Many data-intensive services (e.g., planet analysis, sequence analysis, so on) have become progressively dependent on national cloud data centers (NCDCs) due to growing scientific collaboration among countries. In NCDCs, tens of thousands of virtual machines (VMs) are assigned to physical servers to provide data-intensive services with a quality-of-service (QoS) guarantee, and consume a massive amount of energy in the process. though several VM placement schemes are proposed to resolve this downside of energy consumption, most of those assume that each one the physical servers are homogenised. However, the physical server configurations of NCDCs usually dissent considerably, that ends up in varied energy consumption characteristics. [17] explore an alternative VM placement approach to minimize energy consumption during the provision of data-intensive services with a global QoS guarantee in NCDCs. they use an improved particle swarm optimization algorithm to develop an optimal VM placement approach involving a tradeoff between energy consumption and global QoS guarantee for data-intensive services. Experimental results show that their approach significantly outperforms other approaches to energy optimization and global QoS guarantee in NCDCs.

III.METHODOLOGY

The goal of the implemented test-bed is to demonstrate the effectiveness of framework in reducing computing-plus-communication energy compared to the other available techniques, and to support a sensitivity analysis with respect to its parameters. Fig. 2 shows that power reduction with DVFS.

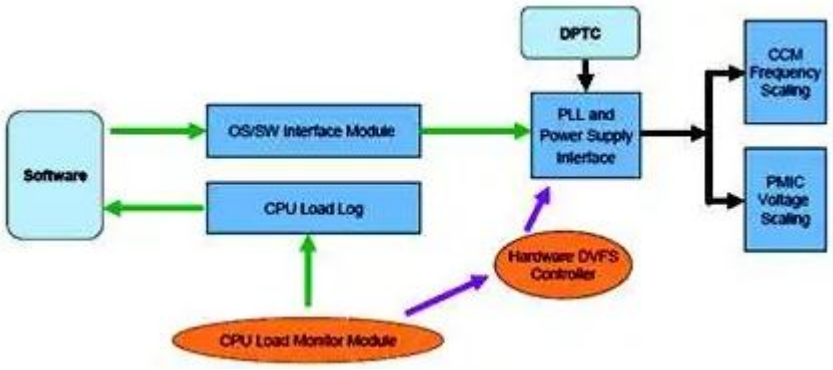


Fig. 2: Reduce power consumption with dynamic voltage and frequency scaling

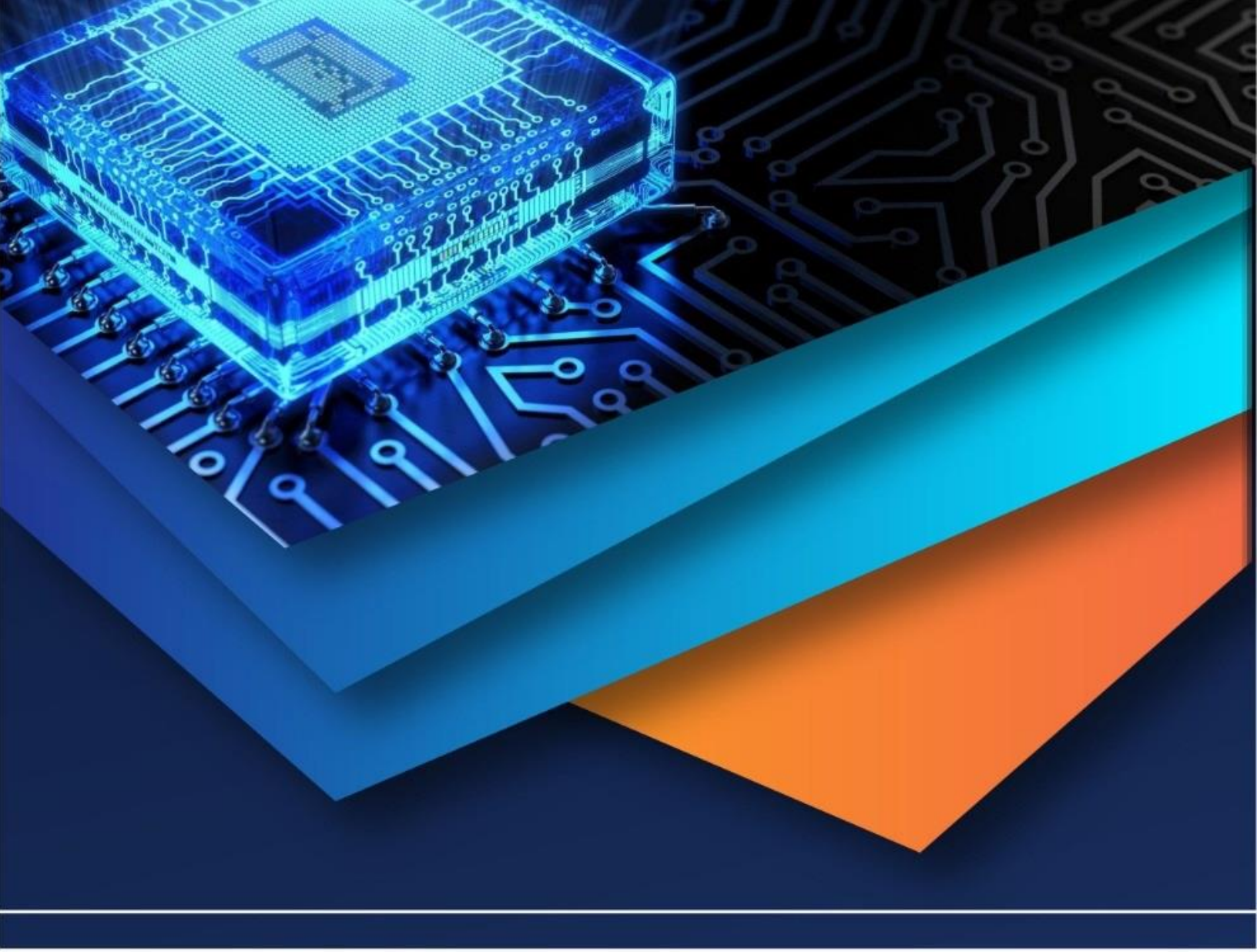
The TEST-DVFS testbed consists of the following modules: 1) Workload module: This module is developed to simulate various types of offered workloads that can be either parameter-based synthetic workloads or a real trace; 2) Component module: All the considered components of the system are implemented in this module; 3) Working module: The working module focuses on the energy model, scheduling types and network topology.

IV. CONCLUSIONS

Cloud data centers become more and more powerful, energy consumption becomes a major challenge both for environmental concerns and for economic reasons, to address this problem, this paper proposed model of social network website. Then we optimize server in such way that old data server should run on minimum cost. Also we find out the pattern of data consuming of users and according to that we generate the graphs. and after generating user data consumption data patterns we optimize server in such way that whichever server data have less traffic should stay on sleep mode.

REFERENCES

[1] C. Canali, M. Colajanni, and R. Lancellotti, “Performance evolution of mobile web-based services,” Internet Computing, IEEE, vol. 13, no. 2, pp. 60–68, 2009.
[2] EPA, “Report to congress on server and data center energy efficiency,” US Environmental Protection Agency, Tech. Rep., 2007.
[3] J. Whitney and P. Delforge, “Data center efficiency assessmenty – scaling up energy efficiency across the data center industry: Evaluating key drivers and barriers,” NRDC, Anthesis, Tech. Rep., 2014, – http://www.nrdc.org/energy/files/datacenter-efficiency-assessment-IP.pdf.
[4] M. Alizadeh, T. Edsall, S. Dharmapurikar, R. Vaidyanathan, K. Chu, A. Fingerhut, F. Matus, R. Pan, N. Yadav, G. Varghese et al., “Conga: Distributed congestion-aware load balancing for datacenters,” in SIGCOMM’14. ACM, 2014, pp. 503–514.
[5] H. Wang and C. Schmid, “Action recognition with improved trajectories,” in Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 3551–3558.
[6] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[7] M. S. Hossain, G. Muhammad, M. F. Alhamid, B. Song, and K. AlMutib, “Audio-visual emotion recognition using big data towards 5g,” Mobile Networks and Applications, pp. 1–11, 2016.
[8] J. W. Jiang, T. Lan, S. Ha, M. Chen, and M. Chiang, “Joint VM placement and routing for data center traffic engineering,” in INFOCOM’12. IEEE, 2012, pp. 2876–2880.
[9] R. Brown et al., “Report to congress on server and data center energy efficiency: Public law 109-431,” Lawrence Berkeley National Laboratory, 2008.
[10] R. Urgaonkar, U. C. Kozat, K. Igarashi, and M. J. Neely, “Dynamic resource allocation and power management in virtualized data centers,” in NOMS’10. IEEE, 2010, pp. 479–486.
[11] W. Zhu, C. Luo, J. Wang, and S. Li, “Multimedia Cloud Computing,” IEEE Signal Processing Magazine, vol. 28, no. 3, pp. 59–69, May 2011.
[12] S. Wang and S. Dey, “Adaptive Mobile Cloud Computing to Enable Rich Mobile Multimedia Applications,” IEEE Transactions on Multimedia, vol. 15, no. 4, pp. 870–883, 2013.
[13] Y.-W. Ma, J.-L. Chen, C.-H. Chou, and S.-K. Lu, “A Power Saving Mechanism for Multimedia Streaming Services in Cloud Computing,” IEEE Systems Journal, vol. 8, no. 1, pp. 219–224, March 2014.
[14] V. Mathew, R. K. Sitaraman, and P. Shenoy, “Energy-aware load balancing in content delivery networks,” in INFOCOM, 2012 Proceedings IEEE. IEEE, 2012, pp. 954–962.
[15] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, and G. Jiang, “Power and performance management of virtualized computing environments via lookahead control,” Cluster computing, vol. 12, no. 1, pp. 1–15, 2009.
[16] A. Beloglazov, R. Buyya, Y. C. Lee, A. Zomaya et al., “A taxonomy and survey of energy-efficient data centers and cloud computing systems,” Advances in computers, vol. 82, no. 2, pp. 47–111, 2011.
[17] S. Wang, A. Zhou, C. H. Hsu, X. Xiao, and F. Yang, “Provision of data-intensive services through energy- and qos-aware virtual machine placement in national cloud data centers,” IEEE Transactions on Emerging Topics in Computing, vol. 4, no. 2, pp. 290–300, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)