International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 11 (2018) pp. 9686-9691

An Energy Efficient Data Placement Technique for Cloud based Storage Area Network

Shabeen Taj G A

Assistant professor, Department of Computer Science and Engineering
Government Engineering College, Ramanagar, India.

Dr. G.Mahadevan

Professor of Computer Science and Engineering
AMCEC, 18th KM, Bannerghatta Road, Bangalore, India.

Abstract

Cloud based storage area network (CSAN) has recently emerged as a successful paradigm for provisioning IT infrastructure, resources, and for various scientific and data intensive application services. As, the wider adoption of CSAN has led to the formation of large scale data centers that consume enormous energy and have significant carbon footprints, energy efficiency is becoming increasingly critical and important for data centers and CSAN. An efficient way to minimize energy consumption and computing cost to CSAN is to incorporate energy aware data placement algorithm. This work present such design to minimize energy consumption using energy aware data placement technique that maximize the number of idle storage server to put in sleep mode. Performance of EEDP is evaluated using real-time scientific application workflow. The outcome shows EEDP technique minimize data access latency and cost over state-of-art technique.

Keywords: Cloud computing, Data placement, Energy efficency, ILP, SAN, SDN.

INTRODUCTION

Recently, cloud based storage and computing model has attained wide scope and adoption across IT infrastructure, services and resources adopting pay-as-you-go or pay-per-use basis. The increased adoption of clouds based storage virtualization technologies across various industries has led to the growth of large scale data centers across globes that offers cloud services. This progress resulted in excessive rise of energy or electricity consumption, increased data center ownership cost and carbon emission. As a result, energy efficacy is becoming a critical element for data centers and cloud computing environment. The study carried out in [1] shows that the amount of energy consumption is set to increase 76% by 2030 with data centers contributing to it. Therefore, it highlights the importance of minimizing energy in data centers and clouds computing environment. According to report presented in [2], the data center is estimated to consume energy as an average of twenty thousand households, and based on report presented in [3], the total estimated energy bill for data center in 2010 is 11:5 billion and is expected to double every five years. Along with electronic waste and huge amount of energy used to power these data centers, energy efficient data center technique have become one of the big challenges.

A critical issues of energy inefficiency in cloud storage environment is the idle power wasted when resources are underutilized. Along with low resources utilization, servers are switched on permanently even if they are not used and still consume around 70% of their peak energy. To overcome these issues, it is important to minimize or remove the energy wastage, to enhance efficiency and to change the way resource are utilized. This can be modeled by devolving an energy efficient resource/data allocation/placement design adopting storage area network which is the focus of this work. Along with these issues and challenges, the new design should scale considering various performance requirement and cloud service provider must also address user service level agreement (SLA) requirement of various data intensive and scientific application. Demanded services are cultured and completed since subscribers need to deploy their own services with the network architecture they choose and with having the control on both programs and infrastructure. As a result, data placement or resource allocation design on cloud based storage area network should be flexible enough to adapt and cater the growing cloud environment and user application requirements. Energy efficient data placement is composed in detecting and allocating resources to each incoming subscriber task in such way, that the subscriber needs are met with minimal resources is used and that optimize the energy of data center. Recently, number of energy efficient data placement approaches is presented.

In [4] designed an energy efficient data center by investigating different techniques and architectures. Typically, the cloud services were provided to the users with the help of datacenters with thousands of computing devices. Moreover, the research issues that were related to QoS improvement, and energy consumption minimization were addressed in this paper. From the investigation, it was observed that the optimizing the energy of datacenters was a highly complicated task due to the dynamic factors of workload placement, resource mapping, cooling schedule, traffic patterns, and inter-process communication. Furthermore, energy and resource optimization were an important tasks in SDNs, because the data centers consumed enormous amount of energy or power. In [5] suggested a new SAN prototype, namely, QuickSAN for reducing the computational overhead in Storage Area Networks (SANs). This prototype combined a network adapter into Solid State Disks (SSDs), which enables the direct communication with service storage services. The benefits of this paper were, it reduced the block transport overhead and increased the bandwidth for the small requests. In [6] recommended a SAN to handle the I/O traffic between the storage devices and servers. Here, the measures of cost and complexity of Fiber channel were considered.

The state of art Cloud based storage area network design consume enormous amount energy or power especially if storage resources are switched on permanently even if they are not utilized. An idle storage server consumes about 70% of its overall energy [7]. This result in wastage of energy as a result induce energy inefficiency. As a result incur higher cost. An efficient way to minimize energy consumption and computing cost to storage is to incorporate energy aware data placement algorithm. In [8] discussed the need for such dynamic placement strategy to enhance energy efficiency of cloud data centers by shutting and putting sleep idles.

To undertake the benefit of SAN and Cloud computing framework several hybrid [9], [10], and [11] and heterogeneous [12] approaches have been presented. The future SAN model should consider heterogeneity of storage in provisioning real-time services to users. This work present such design to minimize energy consumption using energy aware data placement technique that maximize the number of idle storage server to put in sleep mode. This work present an energy efficient resource allocation scheme by reducing the number of used storage server and maximizing the number of idle storage server to transition in to sleep mode. For meeting real-time application deadline a linear integer programming constraint is used to achieve optimal solution. Our model can be used to improve the energy efficiency of current Cloud based SAN infrastructure such as OpenStack [13] and OpenNebula [14].

The Contribution of research work is as follows:

- This work present energy efficient data placement technique (EEDP) on cloud based SAN network.
- Experiment are conducted on real-time work flow and performance is evaluated in terms of computation cost and latency.
- Our model brings a good trade-off between computation cost (energy minimization) and task completion (minimizing latency).
- The outcome shows significance performance over state-of-art architecture.

The rest of the paper is organized as follows. In section II the proposed energy efficient data placement technique for cloud based storage area network is presented. In penultimate section experimental study is carried out. The conclusion and future work is described in last section.

ENERGY EFFICIENT DATA PLACEMENT TECHNIQUE FOR CLOUD BASED STORAGE AREA NETWORK

Here we present an energy efficient data placement (EEDP) model for cloud based storage area network. The EEDP model is designed considering to bring a good tradeoff between user performance requirements and meeting energy constraints. The symbol and notation used in our model are described below.

a) Symbols and notation:

For easiness of understanding, the variable and constant used in our model are listed out here. y is the demand size in number of requested data items, x is the number of storage node in the data center, $e_u$depicts the energy consumption of D_u, w_{uv} is a bivalent variable depicting that D_u is allocated to a storage server v, g_v is a variable considered to depict whether the storage server v is utilized or not, $E_{v,\uparrow}$ depicts the maximum energy consumption of storage server v, $E_{v,present}$ depicts the present energy consumption of storage server $v(E_{v,present} = E_{v,idle} + \sum_q eq \ with \ D_q \ possess \ by \ storage \ server \ v)$, $E_{v,idle}$ depicts the energy consumption of storage server v when it is idle.

b) System model:

The system model is defined as follows, let y be the set of client demand in terms of the amount of data items D and types of data applications. Each $D_u$is characterized by energy consumption e_u and lifetime k_u. Each storage node v, from particular data center, has energy consumption bound $E_{v,\uparrow}$ which depends on cloud based storage area network providers.

Here we consider heterogeneous environment. The objective of our work is to achieve optimal data placement. In addition to y, the amount of demanded data items, we describe the number of storage servers, x present in the data center. The storage server is presumed to possess same energy consumption bound, $E_{v,\uparrow,}\{v = 1,2,3, \dots, x\}$. Dynamically, each storage server v serves a set of data items is characterized by its present energy consumption $E_{v,present}$.

The preliminary objective of our model is to minimize power consumption of the data centers of cloud based storage area network, we describe as critical solution factor g_v for each storage server v that is set to zero if storage server v is not chosen, one if it is selected. Along with, we describe the bivalent variable w_{uv} to depicts that D_u has been successfully placed in storage server v and w_{uv} set to one; otherwise w_{uv} is set to zero. The objective strategy to place all the demands or data items in a minimal number of storage nodes is defined as follows

$$\min R = \sum_{v=1}^{x} g_v \tag{1}$$

The Eq. (1) is subjected to set of linear constraint with respect to overall storage server capacity and also that data item are allocated to one storage server at a time or based on resource available.

Each storage server has an energy bound $E_{v,\uparrow}$ that cannot go beyond when serving data items and this arises based on resource capacity available

$$\sum_{u=1}^{y} E_{v,\uparrow} g_v - E_{v,present}, \forall v = 1,2,\dots,x \quad (2)$$

A cloud based storage area network (CSAN) service provider need to assure all demands with in predefined service level agreement and each demanded items will be allocated to one storage server at a time

$$\sum_{v=1}^{x} w_{uv} = 1, \forall u = 1,2,\dots,y \quad (3)$$

For storage server evaluating $E_{v,\uparrow} > E_{v,present}$ and $E_{v,present} \neq 0$, the total amount of used storage is lower bounded by $\left\lceil \frac{\sum_{v=1}^{x} E_{v,present}}{E_{v,\uparrow}} \right\rceil$. This adds the respective inequality to model as follows

$$\sum_{v=1}^{v} g_v \geq \left\lceil \frac{\sum_{v=1}^{x} E_{v,present}}{E_{v,\uparrow}} \right\rceil \quad (4)$$

where

$$g_v = \begin{cases} 0, if\ the\ storage\ server\ is\ not\ utilized; \\ 1, Otherwise. \end{cases} \quad (5)$$

$$w_{uv} = \begin{cases} 0, if\ the\ D_u\ is\ not\ placed\ in\ storage\ server; \\ 1, Otherwise. \end{cases} \quad (6)$$

Further, constraint in cloud based CSAN server CPU, Memory and storage is considered to our model. i.e.

$$\sum_{u=1}^{y} c_u w_{uv} \leq C_v g_v \quad (7)$$

where c_u is the demanded CPU by D_u. $C_v\ is\ the\ CPU\ capacity$ of CSAN server v.

$$\sum_{u=1}^{x} m_u w_{uv} \leq M_v g_v \quad (8)$$

where m_u is the demanded memory by D_u and $M_v\ is\ the\ memory\ capacity$ of CSAN server v.

$$\sum_{u=1}^{x} s_u w_{uv} \leq S_v g_v \quad (9)$$

where s_u is the demanded storage by D_u and S_v is the storage capacity of CSAN server v. In our work all these constraint are considered to be met and our objectives is to minimize energy and find optimal energy placement.

Algorithm 1: optimal energy efficient data placement strategy

Step 1: Start

Step 2: Initialize set of data items D, storage server v, storage server capacity S_v

Step 3: Initialize maximum energy consumption of storage server $E_{v,\uparrow}$, CPU capacity C_v, memory capacity M_v, demanded storage s_u, is the storage capacity S_v.

Step 4: User request set of data items D from storage server.

Step 5: Data items requested are sorted in descending order using Equation (1) and are stored in a list

Step 6: Higher energy consumption data items are placed on low energy server till all the data items are stored and placed in most engaging server.

Step 7: if all data items are successfully is placed in storage server.

Step 8: set w_{uv} to 1.

Step 9: else set w_{uv} to 0.

Step 10: if server is utilized.

Step 11: set g_v to 1.

Step 12: else set g_v to 0.

Step 13: Put unused server to sleep.

Step 14: End.

The optimal energy efficient data placement strategy is shown in Algorithm 1 which is composed of following steps, firstly, the data items requested are sorted based on decreasing order using Eq. (1). Secondly, the sorted data items are obtained and placed based on higher energy consumption data items in the storage server with least energy consumption factor till a data items down the sorted objects fits in this target storage server till all the data items are placed in most engaged storage server. This aid in freeing storage server to put them in sleep mode and use it for future demand. As a result, our model minimize computing cost and meeting latency requirement of real-time data intensive application which is experimentally proved in next section.

SIMULATION RESULT AND ANNALYSIS

This section presents performance evaluation of proposed Energy Efficient Data Placement (EEDP) over exiting methodology in terms of latency and computing cost. The experiment are conducted on windows 10 enterprises edition operating system, Intel I-5 quad core processor with 16GB RAM with 4 GB dedicated CUDA enabled GPU. This work consider real-time scientific and data intensive workflow application such as Inspiral and Montage. The workflow is obtained from [15]. The proposed and existing methodology is designed using JAVA 8 using eclipse neon IDE. The proposed EEDP technique performance is evaluated interm of workflow latency, and computing cost and is compared with existing model [16].

a) Energy efficient data placement Latency performance considering different real-time workflow:

Experiment are conducted to study the performance achieved by EEDP over existing approach [16] in term latency achieved for executing task. Here we considered two real-time work flow such as Inspiral_1000 and Montage_1000 workflow. The number of datacenter are varied from 25 to 100 and each datacenter is composed of 20 nodes. The user is fixed to 100 users. The experiment study shows that the proposed EEDP performs better than exiting approach in term of latency achieved. A latency minimization of 8.42%, 11.68%, 12.4%, and 12.78% is achieved by EEDP over existing approach when datacenter size is 25, 50, 75 and 100 respectively, considering Inspiral_1000 workflow as shown in Fig. 1. An average latency minimization of 11.36% is achieved by EEDP over exiting approach considering Inspiral workflow. Similarly, latency minimization of 14.52%, 17.70%, 19.96%, and 20.78% is achieved by EEDP over existing approach when datacenter size is 25, 50, 75 and 100 respectively, considering Montage_1000 workflow. An average latency minimization of 18.31% is achieved by EEDP over exiting approach considering Montage workflow as shown in Fig. 2. An overall latency minimization of 14.81% is achieved by EEDP over exiting approach considering different case studies.

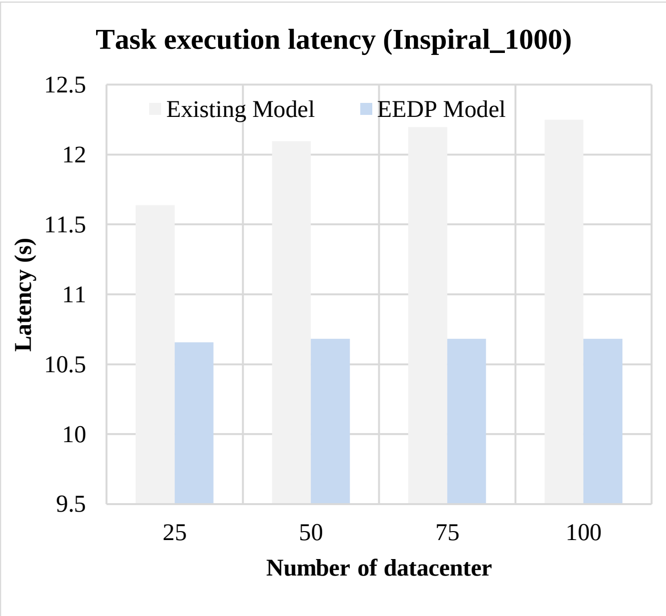

Figure 1. Latency performance considering Inspiral_1000 workflow

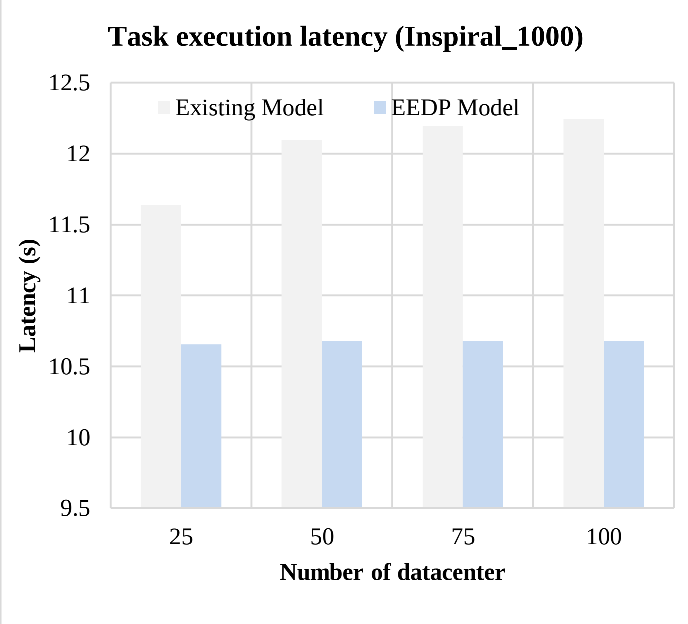

Figure 2. Latency performance considering Montage_100 workflow

b) Energy efficient data placement Computing cost performance considering different real-time workflow:

Experiment are conducted to study the performance achieved by EEDP over existing approach [16] in term computing cost for executing task. Here we considered two real-time work flow such as Inspiral_1000 and Montage_1000 workflow. The number of datacenter are varied from 25 to 100 and each datacenter is composed of 20 nodes. The user is fixed to 100 users. The experiment study shows that the proposed EEDP performs better than exiting approach in term of computation cost achieved. A computing cost reduction of 41.13%, 43.22%, 43.69%, and 43.93% is achieved by EEDP over existing approach when datacenter size is 25, 50, 75 and 100 respectively, considering Inspiral_1000 workflow as shown in Fig .3 An average computing cost reduction of 43.01% is achieved by BGDRP over exiting approach considering Inspiral workflow. Similarly, computing cost reduction of 45.05%, 47.09%, 36.58%, and 48.55% is achieved by EEDP over existing approach when datacenter size is 25, 50, 75 and 100 respectively, considering Montage_1000 workflow. An average computation cost reduction of 47.49% is achieved by EEDP over exiting approach considering Montage workflow as shown in Fig. 4. An overall cost reduction of 50.24% is achieved by EEDP over exiting approach considering different case studies.

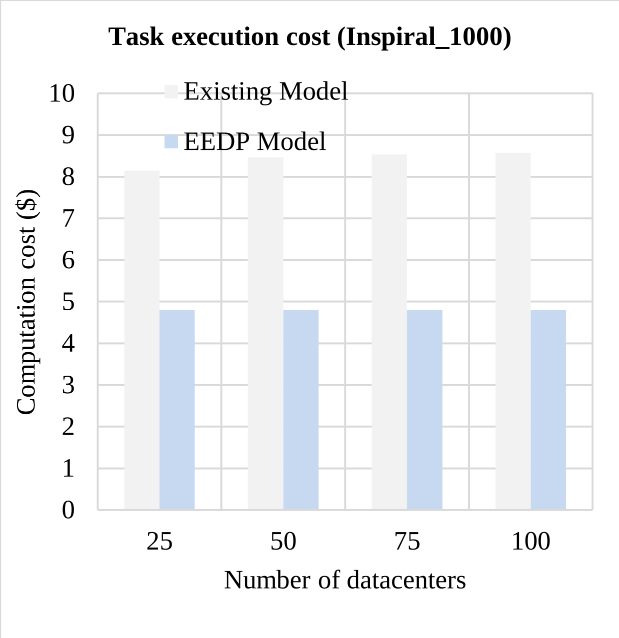

Figure 3. Task execution computing cost considering Inspiral_1000 dataset

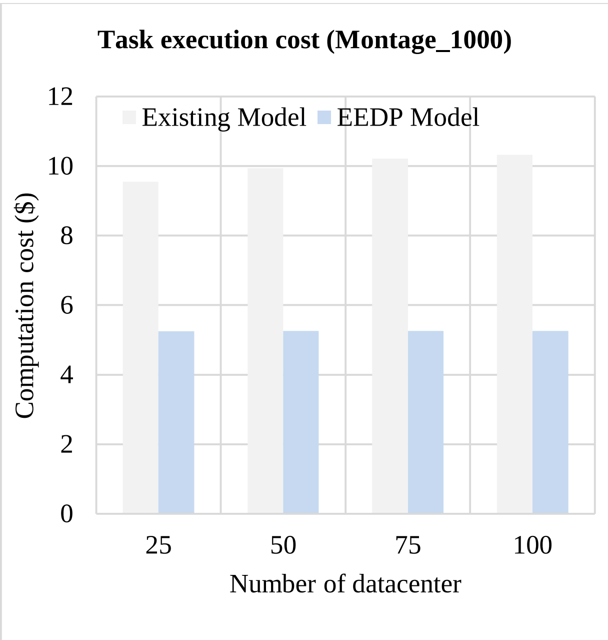

Figure 4. Task execution computing cost considering Montage_1000 dataset

CONCLUSION

Cloud based Storage area network (CSAN) consume enormous amount energy or power especially if CSAN resources are switched on permanently even if they are not utilized. Therefore minimizing energy will aid in minimizing computation cost and utilizing resource efficiently. This work present an energy efficient data placement (EEDP) model for CSAN. Experiment are conducted to evaluate performance of EEDP over existing approach using real-time workflow considering varied node/datacenter size with fixed user. The outcome shows an average performance improvement of 14.81%, and 50.24% is achieved by EEDP over existing model in terms latency, and cost respectively. The outcome shows EEDP technique minimize data access latency, and cost over state-of-art technique. This shows the proposed design minimize the number of required storage server to provision given workload. Thus aid in reducing overall computation cost of CSAN. The study shows the efficiency, scalability and robustness of our model. The significance of research work is no prior model has considered energy optimization for heterogeneous environment for cloud based storage network. The phenomenal growth of information technology has resulted in rapid development of storage technology. The future design we consider energy minimizing considering multi-core environment and different storage hardware such as SSD to further improve performance and minimize computing cost.

REFERENCE

[1] World energy outlook 2009 fact sheet. http://www.iea.org/weo/docs/weo2009/factsheetsWEO2009.pdf.

[2] Gartner report, financial times, 2007.

[3] James M. Kaplan, William Forrest, and Noah Kindle. Revolutionizing Data Center Energy Efficiency. Technical report, McKinsey & Company, July 2008.

[4] J. Shuja, et al., "Survey of techniques and architectures for designing energy-efficient data centers," IEEE Systems Journal, vol. 10, pp. 507-519, 2016.

[5] A. M. Caulfield and S. Swanson, "Quicksan: a storage area network for fast, distributed, solid state disks," in ACM SIGARCH Computer Architecture News, 2013, pp. 464-474.

[6] A. Verma, et al., "Performance analysis of Internet Protocol Storage Area Network (IP SAN) and its usage in Clustered Database," IJCSI International Journal of Computer Science Issues, vol. 10, pp. 97-104, 2013.

[7] Erica Naone. Conjuring clouds. Technology Review, 112(4):54{56, 2009.

[8] Intel's cloud computing 2015 vision. http://www.intel.com/Asset/PDF/whitepaper/cloud_vision.pdf.

[9] I. Foster, Y. Zhao, I. Raicu, S. Lu, Cloud Computing and Grid Computing 360-Degree Compared, in: Proceedings of the 1st Workshop on Grid Computing Environments, Austin, Texas, pp. 1, 2008.

[10] Rekha P M and Dakshayini M, "Dynamic network configuration and Virtual management protocol for open switch in cloud environment," 2015 IEEE International Advance Computing Conference (IACC), Banglore, 2015, pp. 143-148.

[11] N. Yoshino, H. Oguma, S. Kamedm and N. Suematsu, "Feasibility study of expansion of OpenFlow network using satellite communication to wide area," 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, Italy, 2017, pp. 647-651.

[12] J. J. Kuo, S. H. Shen, M. H. Yang, D. N. Yang, M. J. Tsai and W. T. Chen, "Service Overlay Forest Embedding for Software-Defined Cloud Networks," 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), Atlanta, GA, USA, 2017, pp. 720-730.

[13] Openstack. http://www.openstack.org/. last accessed March. 2018.

[14] Open nebula. http://www.opennebula.org/. last accessed April. 2018.

[15] Bharathi S, Chervenak A, Deelman E, Mehta G, Su MH, Vahi K. Characterization of scientific workflows. In: Workflows in Support of Large-Scale Science, 2008. WORKS 2008. Third Workshop on; p. 1±10, 2008.

[16] Lipeng Wan, Qing Cao, Feiyi Wang, Sarp Oral ”Optimizing checkpoint data placement with guaranteed burst buffer endurance in large-scale hierarchical storage systems,” Journal of Parallel and Distributed Computing, Volume 100, Pages 16-29, 2017.