ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

AN APPROACH TO HANDLE THE EVENTS OF CLOUD COMPUTING FOR BETTER ENVIRONMENTAL IMPACT

Ashish Kumar Trivedi [*1], Ajay Kumar Bharti[2] & Sohit Shukla[3]
[*1]Dept. of Computer Science & Engineering, Maharshi University of Information Technology
[2] Dept. of Computer Science & Engineering, Maharshi University of Information Technology
[3]Dept. of Information Technology Rajkiya Engineering College, Ambedkar Nagar

ABSTRACT

Cloud Computing is a computing paradigm where various tasks are assigned to a combination of connections, software and services that can be accessed by the user over a network. The growing demand of Cloud infrastructure has drastically increased the energy consumption of data centers, which has become a critical issue. High energy consumption not only translates to high operational cost, which reduces the profit margin of Cloud providers, but also leads to high carbon emissions which is not environment friendly so there is a need of energy efficient approach in cloud computing The research paper aims to reach a theoretical notion of sustainable development with proposing an incentive for reducing global warming through effective clustering techniques and methods. This paper aims to reduce cloud events by applying map reduce on large event clusters formed in cloud. The purpose of the paper is to develop a better methodology for handling the events of cloud computing and possibly clustering and reducing the similar types of events. This approach might lead to the reduction of carbon-dioxide gas (which is a greenhouse gas) by less usage of servers in cloud data centers. With the advent of IT services in cloud computing energy consumption it is necessary for the developing technology to progress towards sustainable development rather thrashing and harnessing energy from every possible means.

Keywords: *Cloud Computing, Clustering, Cloud Data centers, Clustering Algorithms, K-Means Clustering, Map-Reduce, Resource Identification and Clustering.*

I. INTRODUCTION

Cloud Computing is a most promising and emerging paradigm to overcome the challenges of SOA. Cloud Computing holds first place in Gartners's top 10 strategic technologies list [15]. It promises that resources like computing capacity and storage or services like databases or messaging systems can rapidly be acquired and released based on the current requirement s of deployed application. Basis of cloud computing is a combination of several well-known technologies like virtualization and concepts like service oriented architecture [16]. Cloud computing describes both a platform and type of application .A cloud computing platform dynamically provisions configures, reconfigures and de-provision servers as needed. Servers in the cloud can be physical machines or virtual machines. Advanced cloud typically includes other computing resources such as storage area network (SAN), network equipment's, firewalls and other security devices. [17]. Birman [1] has a notion for Cloud and network usage in it. There are also some environmental impacts [2] of cloud computing. Arif [3] has a valid description of Cloud Computing and its environmental impact. In this paper, we describe a method to reduce similar type of events in Cloud computing so that reduced events for cloud servers can be possible to minimize harmful impact on the environment. We have used the k-means clustering, k-medoid optimization and Map-Reduce concept. K-means clustering is the process of segmenting a group of data points into a minor number of clusters. Map-reduce is a framework for processing parallelizable problems across the large datasets. The aim of the research in the broader view is to reduce the global warming by minimizing the carbon emission. It is well-known that with the continuous processing on servers the heat is emitted which is allowed to reduce with the usage of air-conditioned server-rooms, this one way or another results into increase of carbon dioxide concentration in the environment. Therefore, to build a clean and green environment with cloud computing many redundant events can be reduced or removed by applying various clustering techniques which may lead to lesser and lesser energy usage and waste.

The organization of remaining paper is as follows: Section II is literature review describing various services and approaches used in cloud computing. Section III is event modeling for Identification and context description of event. Section IV is clustering in cloud whether section V describes K-means clustering on Event tuples. Section VI described K-Medoid optimization on Event context. Section VII is reduction of Events by Map-Reduce. At last, conclusion is provided in section VIII.

II.LITERATURE REVIEW

Cloud Computing has its preliminary foundation over providing the hardware, which helps the cloud to run referred as the data centers. Cloud data centers can be considered as a centralized repository for the storage and management of data. Within cloud data centers there are many cooling and electrical devices that do not directly provide services but are major contributors to the power consumption of a cloud data centers. The advantage of cloud is in terms of economic scale; when thousands of users share same facility the cost per user and server utilization is reasonable. More [4] provided the use of cloud computing for E-Governance services implementation. The services offered by cloud computing are Infrastructure-As-A-Service (IAAS), Platform-As-A-Service (PAAS), Software-As-A-Service (SAAS). Built on top of the data centers layers, IAAS layer provides the process of computing in infrastructure to the end users for example- storage capacity, CPU usage etc. PAAS allows the cloud user to deploy the application created by himself/herself on the cloud providers' infrastructures, in this way the user pays for platform software components which comprises its associated infrastructure cost such as operating systems, database, and middleware. SAAS allows the cloud user to access and use of providers' applications running in cloud infrastructures can be termed as a service on demand. Multi-tenancy is the core feature of SAAS. That means SAAS eliminates the need to install set up and run application on the individual computers. There are some more case studies over cloud computing [5], [6], [7] provided by many researchers. Gattulli [8] provided various routing algorithms for cloud computing.

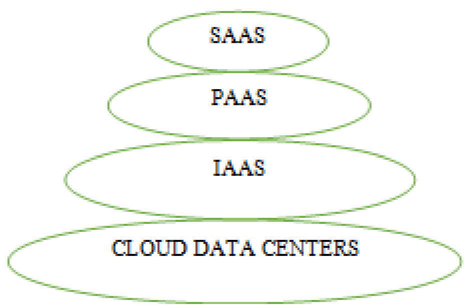

Fig. 1 Foundation Stones of Cloud Computing

III. EVENT MODELING

An event occurs at a particular instant in time and marks a change of state in the system. A stream processing query thus produces outputs when interesting events, such as meetings between groups of friends, occur. For example, it could automatically group lectures, department meetings, and social events into three distinct clusters of events. Similarly, the clustering process could categorize network anomalies into link failures, denial-of-service attacks, or false alarms. To perform such classifications, however, online stream clustering algorithms today rely only on event attributes: events in streams take the form of relational tuples whose attributes define a multidimensional space used during clustering. A meeting event, for example, could be characterized by the time, location, and number of attendees. Such event attributes, however, offer only limited information to help in the clustering process, thus circumscribing the quality of the resulting clusters.

In this paper, we present an approach that enhances the online clustering and thus event classification capabilities. In our approach, users specify complex context information for events.

ISSN 2348 – 8034
Impact Factor- 5.070

IV. CLUSTERING

The term 'Clustering' or 'Cluster Analysis' can be understood as the grouping of similar types of items in one set referred as Cluster. Many clustering algorithms [9], [10], [11], [12], [13] have been provided by many researchers. From the perspective of data mining it is one of the fundamental concepts to grasp. Various algorithms are used to find the significant characteristics in the items that can allow them to form a cluster. But the notion of 'cluster' cannot be defined efficiently which is the main reason for proposing various algorithms. In cloud computing the aim for forming the cluster of various events could not be fulfilled by using only one algorithm so we have performed a comparative study for the best suitable algorithm. The K-means clustering, resource identification and clustering and Map-Reduce technique has helped the comparative study to reach its desired goal.

V. K-MEANS CLUSTERING

K-means clustering divides n-observation into k- clusters in which each observation lies in a cluster with the nearest mean, serving as a blueprint of the cluster. It is often evident that k-mean clustering is well suited for large data sets. As cloud computing deals with massive databases k-means clustering is an efficient approach for handling the data in the form of clusters. K-means clustering allows a suitable structure for unstructured data in cloud computing. An extension of K-means is provided by Singh [13] using encryption.

VI. K-MEDOID

Partitioning Around Medoids or the K-medoids algorithm is a partitional clustering algorithm which is slightly modified from the K-means algorithm. They both attempt to minimize the squared-error but the K-medoids algorithm is more robust to noise than K-means algorithm. In K-means algorithm, they choose means as the centroids but in the K-medoids, data points are chosen to be the medoids. A medoid can be defined as that object of a cluster, whose average dissimilarity to all the objects in the cluster is minimal. The difference between k-means and k-medoids is analogous to the difference between mean and median: where mean indicates the average value of all data items collected, while median indicates the value around that which all data items are evenly distributed around it. The basic idea of this algorithm is to first compute the K representative objects which are called as medoids. After finding the set of medoids, each object of the data set is assigned to the nearest medoid. That is, object i is put into cluster $v_{i,}$ when medoid mv_i is nearer than any other medoid m_w.

The algorithm proceeds in two steps:

- BUILD-step: This step sequentially selects k "centrally located" objects, to be used as initial medoids.
- SWAP-step: If the objective function can be reduced by interchanging (swapping) a selected object with an unselected object, then the swap is carried out. This is continued till the objective function can no longer be decreased.

VII. MAP-REDUCE

Map Reduce model is a framework for processing similar events across the large datasets. Kim [14] has power efficient map-reduce description for integrated workload GPU In cloud computing it allows to run particular computation based on client query on the closest nodes (or computers) where data are stored in the form of clusters. Further these mapped tasks at each node are reduced. Hence this validates thesis approach by its name as Map-Reduce. There are various algorithms to facilitate this task such as the canopy method, CPCLUSTER Map-reduce algorithm and many more.

VIII. NOVEL APPROACH

The approach has the existence in the creation of the events in to the formal relational tuple placed in the process. Each and every basic events attribute value is to be converted in to the tuple and have a set relation with each others.

These relations between the attributes will become the base for formation of the clusters; apply the k-medoid clustering technique for the forrmation of the objects can be defined as that object of a cluster, whose average dissimilarity to all the objects in the cluster is minimal. The steps involve creating the approach is as follows:
Step 1: Take the process under consideration for the optimization.

Step 2: Find out the various events took placed within the process to complete, treat them as the attributes.
In our model,an event is a tuple in a stream that takes the form: (eid, time, a_1, ..., a_n), where eid is an attribute that uniquely identifies the event, time is the time when the event occurred, and a_1, ..., a_n are the value attributes of the event

Step 3: With the attributes created in the step 2 draw relational tuple table of the events.

Table 1: location

Eid	Time	Duration	Location
1	10:00am	90 minutes	Seminar hall
2	11:30am	30 minutes	Dinning Hall
3	12:00 pm	90 minutes	Auditorium

Table 2: Participants

Eid	Name	Type	College
1	Alice	Student	GRD
2	Bob	Student	BRD
3	Denish	Faculty	ABES

Step 4: Now apply the k-medoid clustering algorithm to find out the centrally located objects of each of the table (cluster) v_i
Step 5: Now we have the set of the objects from v_1 to v_n for each of the relational tuple table apply map –reduce algorithm on the set.
Step 6: With the step 5 we will create a cluster of the dissimilarity to all the objects in the cluster is minimal.

IX. CONCLUSION

With the application of K-Mediod optimization and Map-Reduce on Event stream reduces events in cloud leads to lesser usage of cloud servers. The aim of our research has reached its theoretical conclusion that with the lesser usage of servers or optimum usage of servers can be possible with the usage of Hadoop map-reduce concept it can allow. Hadoop map-reduce manages the tasks and events of Cloud Computing network effectively and efficiently. The optimum utilization of servers energy will allow lesser Carbon-dioxide emission that can save the environment from heating up or more precisely 'global-warming' could be prevented. With the increasing advent of cloud computing technology we have aimed to reach a sustainable development by contributing lesser harm from the cloud computing world.

REFERENCES

1. *K. Birman, "Networks and Cloud", CS5412 Spring (Cloud Computing: Birman), 2015.*
2. *R. K. Trivedi, R. Sharma, "Case Study on Environmental Impact of Cloud Computing", IOSR-JCE e-ISSN: 2278-0661, p-ISSN: 2278-8727 Volume 16, Issue 2, Ver. VI, PP 81-86, 2014.*
3. *M. Arif, T. Mahmood, "Cloud Computing and its Environmental Effects", International Journal of Grid Distribution Computing Vol.8, No.1, pp.279-286, 2015.*
4. *A. More, P. Kanungo, "Use of Cloud Computing for Implementation of e-Governance Services", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.3, pp.115-118, 2017.*
5. *Y. G. Patil, P. S. Deshmukh, "A Review: Mobile Cloud Computing: Its Challenges and Security", Vol.06, Issue.01, pp.11-13, 2018.*

6. *M. K. Saggi, A. S. Bhatia, "A Review on Mobile Cloud Computing: Issues, Challenges and Solutions", International Journal of Advanced Research in Computer and Communication Engineering, 2015.*
7. *R. V. Dharmadhikari, S. S. Turambekar, S. C. Dolli, P K Akulwar, "Cloud Computing: Data Storage Protocols and Security Techniques", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.2, pp.113-118, 2018.*
8. *M. Gattulli, M. Tornatore, R. Fiandra, A. Pattavina, "Low- Carbon Routing Algorithms for Cloud Computing Services in IP-over-WDM Networks", IEEE ICC Optical Network and Systems, 2012.*
9. *Malathy, G. R. Somasundaram, K. Duraiswamy, "Performance Improvement in Cloud Computing Using Resource Clustering", Journal of Computer Science 9 (6): 671-677, ISSN: 1549-3636, 2013.*
10. *D.K. Sharma, S.K. Dhurandher, A. Kumar, A. Kumar, A.K. Jha, "Cloud Computing based Routing Protocol for Infrastructure-based Opportunistic Networks", CAITFS Division of Information Technology, 2016.*
11. *S.N. Bushra, A.C. Sekar, "An Efficient Clustering Method for Incremental Cloud Data", IJARCSSE ISSN: 2277128X, 2014.*
12. *E. Sarkar, C.H Sekhar, "Organizing Data in Cloud using Clustering Approach", International Journal of Scientific & Engineering Research, Volume 5, Issue 5, 2014.*
13. *I. Singh, P. Dwivedi, T. Gupta, P. G. Shynu, "Enhanced K-means clustering with encryption on cloud", IOP Conf. Series: Materials Science and Engineering 263, 042057, 14th ICSET, 2017.*
14. *S.Y Kim, J. Bottleson, J. Jin, P. Bindu, S.C. Sakhare, J.S Spisak, "Power Efficient Map Reduce Workload Acceleration using Integrated GPU", IEEE First International Conference on Big Data Computing Service and Applications, 2015.*
15. *YongChul Kwon, Wing Yee Lee, Magdalena Balazinska "Clustering Events on Streams using Complex Context Information.", Data Mining Workshops, 2008. ICDMW '08. IEEE International Conference Dec 2008.*
16. *Stella Gatzuigrivas, Marc Schaaf, Michael Kaschesky, Guillaume Bouchard. Cloud-based Event-processing Architecture for Opinion Minning, 2011 IEEE World Congress on Services.*
17. *Greg Boss, Padma Malladi, Dennis Quan, Linda Legregni, Harold Hall, "cloud computing", 15 Feb2011.http://download.boulder.ibm.com/ibmdl/pub/software/dw/wes/hipods/Cloud_computing_wp_final8Oct.pdf.*