An Event Diminishment Model to Optimize Cloud Environment

Ashish Kumar Trivedi, Ajay Kumar Bharti

Abstract: ***Cloud computing is to compute a task assigned to a set of connections, software and services that can be utilized by the user over a network. The trending need of Cloud infrastructure has drastically scale up the energy need of data centers, which has become a critical issue. In the row also lead to high carbon emission 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

In service oriented architecture Cloud Computing is a prominent and emerging paradigm. Kwon et al. [14] Cloud Computing holds first place in Gartners's top 10 strategic technologies list. According to Gatzuigrivas et al. [15], there are various resources say computing power, storage capacity and so on and the services that vary from database services, messaging services and many others can be availed and relinquished on the basis of immediate demand of deployed application. Cloud computing is basically a mixture of various established technologies say virtualization and approaches say service oriented architecture. Cloud computing broaden the area of platform used and the types of application offered. Cloud computing provides a means to configure, reconfigure and attach and detached servers as and when required.

Revised Manuscript Received on March 30, 2020.
*****Correspondence Author
Ashish Kumar Trivedi*, Research Scholar, Department of Computer Science, MUIT, Lucknow, (U.P.), India. E-mail.:ashishmtech@gmail.com
Dr Ajay Kumar Bharti, Professor, Department of Computer Science, MUIT, Lucknow, (U.P.), India. E. mail.: ajay_bharti@hotmail.com

In cloud environment servers may be physical or virtual computing power. Advanced form of Cloud computing involves various other resources like Firewall, Storage Area Network, network devices and other network security machines and equipments. 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 our approach, we proposed a model 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 associates a group of data points into a lesser number of clusters. Map-reduce are 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.

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 concenter 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 notable power utilizes of a cloud data centers. Economy in scale is one of the major benefits of cloud computing. When a large number of users access the common computing services, the user wise cost and utilization of servers are quite optimal. More [4] provided the use of cloud computing for E-Governance services implementation. Infrastructure-As-A-Service (IAAS), Platform-As-A-Service (PAAS), Software-As-A-Service (SAAS) are three types of services provided by the Cloud Computing. 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 him/her 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. SAAS removes the restriction to make environment for each 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.

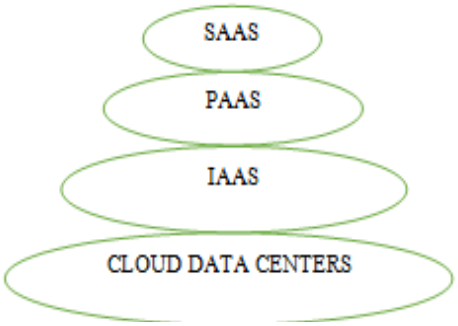

Fig-1: Foundation Stones of Cloud Computing

III. PROPOSED WORK

The change of the state of a system marked as an event to occur at a particular instant. In the way meeting a group of friend's events generates a stream processing query. On the other hand, It dynamically collects social activities, official meetings and different talks into different event clusters. Similarly clustering classifies networks related issues into DoS (Denial of Service) attack, communication failures etc.To assess such issues web based stream clustering algorithms heavily dependent on event characteristics. Stream oriented events looks like a relational database based tuple whose attributes may be considered as multidimensional that can be used at the time of clustering. We can consider a event related to meeting as an instance that can be characterized by various attributes like time of meeting held, location of meeting, department name and number of participants. These types of attributes inadequately supports in cluster formation process. This may degrade the quality of the formed clusters.

In our research paper, we proposed a framework by which using clustering and event classification reduction in events can be done consequently it heavily impact on green issues related to cloud servers.

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.

The overall work will be described in the following flowchart:

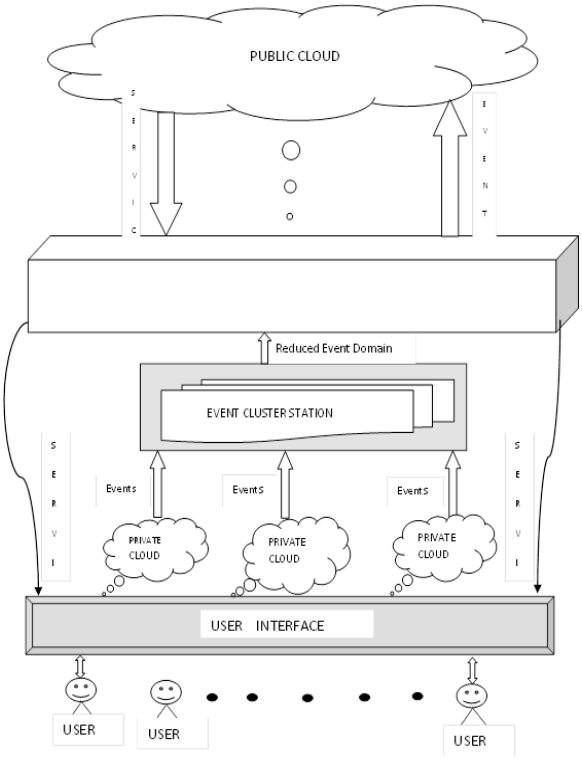

Fig-2: Event Reducing Model

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 other's. These relations between the attributes will become the base for formation of the clusters; K-medoid clustering algorithm is applied to construct object whose average dissimilarity is minimal with all other objects with in cluster. It 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 my proposed model, an event is treated as a relational tuple with the form (a1, a2, a3,a4,….., an), where a1 may be treated as event id that uniquely represents the event occurred, a2 may be time of event occurred and remaining a2,a3,….an are the attribute values of the event.

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

Table 1:

Event ID	Date and Time	Source
18487	12/08/19 12:54:59 pm	MSSQL$SQL SERVER
1309	12/08/19 12:54:27 pm	Asp.net
18487	12/08/19 12:51:46 pm	Logon

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 vn for each of the relational tuple table apply map–reduce algorithm on the set.

Step 6: We will create a cluster of objects, up to above mentioned step 5, with minimal dissimilarity to all other objects within the cluster.

The above approach can formed in the form of the algorithm provided below:

- The approach uses the events log in the form of formal relational tuple. Each and every basic events attribute value is to be converted in to the tuple.
- These relations between the attributes will become the base for formation of the clusters.
- Use the K-medoid clustering algorithm to create cluster in which all the objects have minimal average dissimilarity.

The steps involve creating clusters are as follows:

Step 1: Take the group of events from the event log for the optimization.

Step 2: Find out the various events took placed within the group of event to complete, treat them as the tuple.

An event is a tuple from the event log that takes the form: (Event ID, Date and Time, Source):

Event ID: an attribute that uniquely identifies the event.

Date and Time: the time when the event occurred.

Source: the origin of event occurrence.

Step 3: Process the event log to create relational tuple table of the events as defined in step 2.

Step 4: Following clustering algorithm is used to find out the centrally located event for cluster v_i

1.	From the data space D, select v representative events randomly and mark time of these as medoids.
2.	Remaining data items are non-medoids.
3.	Repeat till medoids stabilise/converge for all medoid items a_s for all non medoid items a_n calculate the cost of swapping $TCost_{sn}$ end end Select s_min and n_min such that $TCost_{s\_min,n\_min} = Min\ TCost_{sn}$ if $TCost_{s\_min,n\_min} < 0$, mark s_min as non medoid and n_min as medoid item. End.
4.	Generate k clusters v_1…….. v_k.

Once the clustering is done, all the key objects are assigned to a particular cluster based on the extent of similarity to the medoid data item of that cluster.

Step 5: Now events in clusters v_1 to v_n reduces multiple similar events into a single event by application of mapping and reducing function.

Map- takes Cluster v and Event ID in that cluster and return each Event ID and an associated count of occurrences of that Event ID.

```
Reduce –takes Event ID and its count of occurrences
in Cluster v and merge together all Event ID basis of
its count of occurrences.Map(Cluster vi , Event ID)
    for each Event ID  in Cluster vi
 number of counts=number of          occurrences of
                        Event ID;
return(Event ID, number of counts);
```

```
Reduce (Event ID, number of counts)
    for each value in number of counts
          R_Event ID=Event ID;
    return (R_Event ID);
```

Step 6: With the step 5 we will create a group of the dissimilar events which represent the reduced set of events.

IV. IMPLEMENTATION

In cloud computing, Event driven service oriented architecture is frequently used. In our approach we considered the event clustering to minimize the similar and duplicate type of events from different private clouds and then these events are used in event orchestration system which is accountable to identify services in the public cloud to provide these services at user level. Implementation of this technique impacts on the domain of huge and redundant events by reducing in number by large. Consequently it reduces the load of event orchestration engine. By this not only the performance of the system will improves also it improves the service availability to the user in an efficient and effective manner.

For the implementation of the model we have implemented the above algorithm on the Matlab with raw data as the log file from the cloud server.

Sample Log File:

Level Date and Time Source Event ID Task Category Information 8/12/2019 12:42:44 PM MSSQL$SQLEXPRESS 18487 (4) "The description for Event ID 18487 from source MSSQL$SQLEXPRESS [CLIENT: <local machine>] " Warning 8/12/2019 12:42:20 PM ASP.NET 4.0.30319.0 1309 (3) "The description for An unhandled exception has occurred. 8/12/2019 12:42:20 PM 8/12/2019 7:12:20 AM f077bfaeef104befaeb692a5a1496230 Application.evtx log file imported in the Matlab:

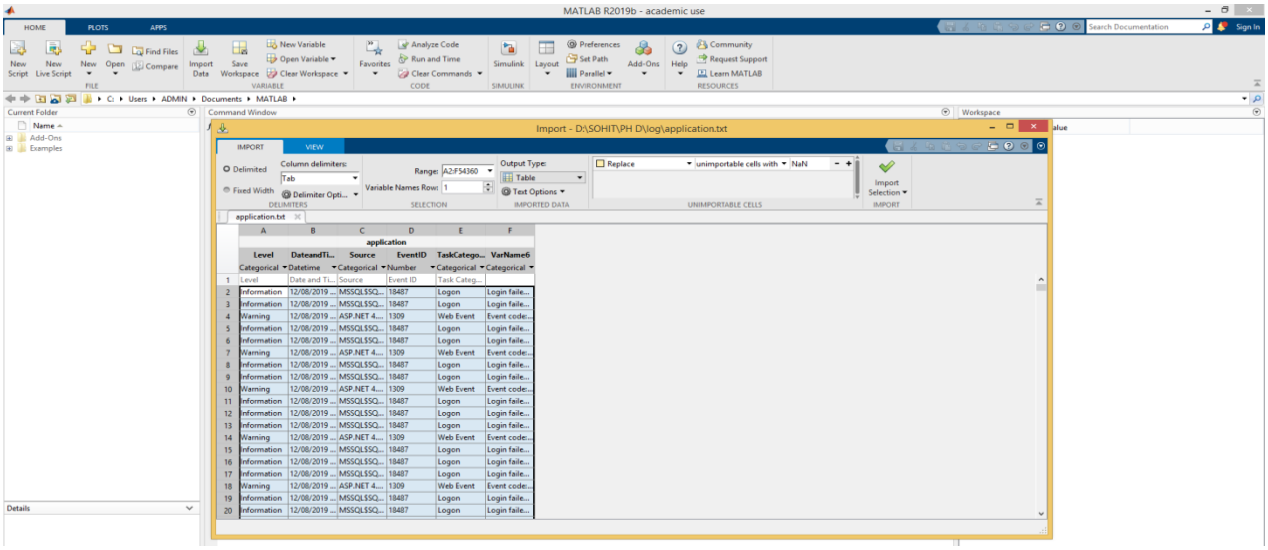

Fig-3: Raw data of application log

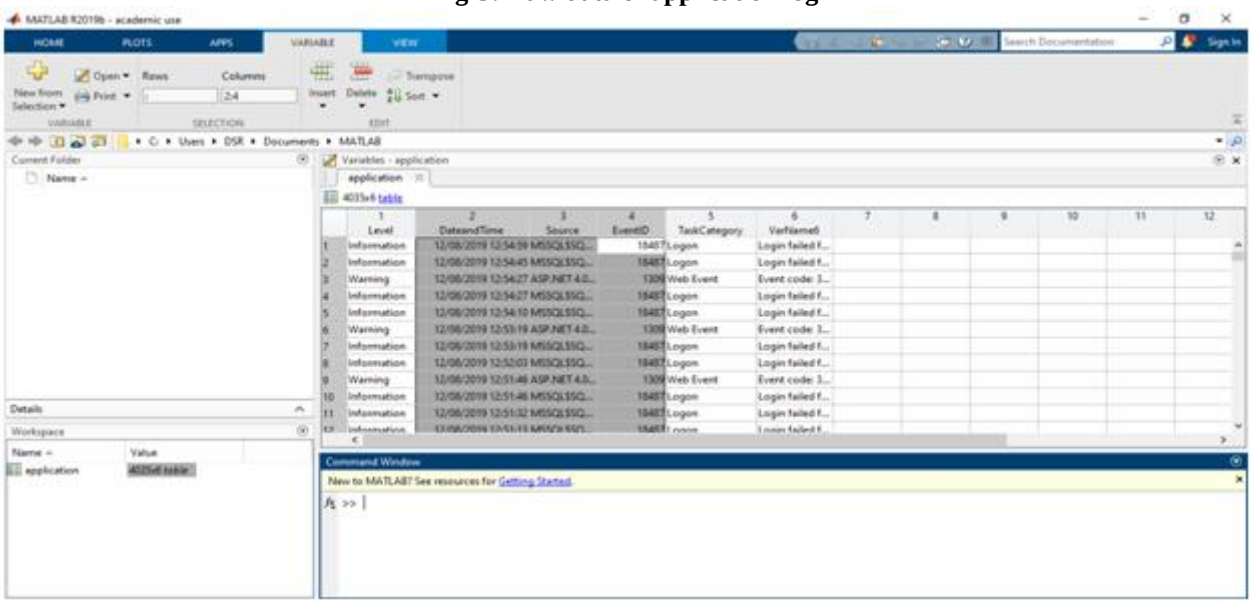

Fig-4: Pre-processing-1 application log

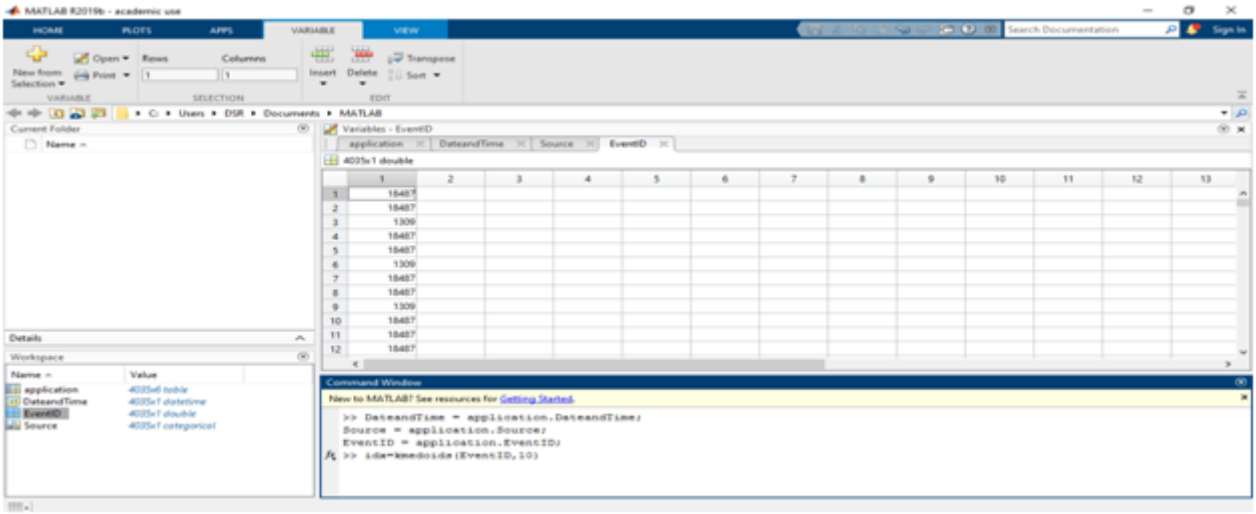

Fig-5: Pre-processing-2 application log

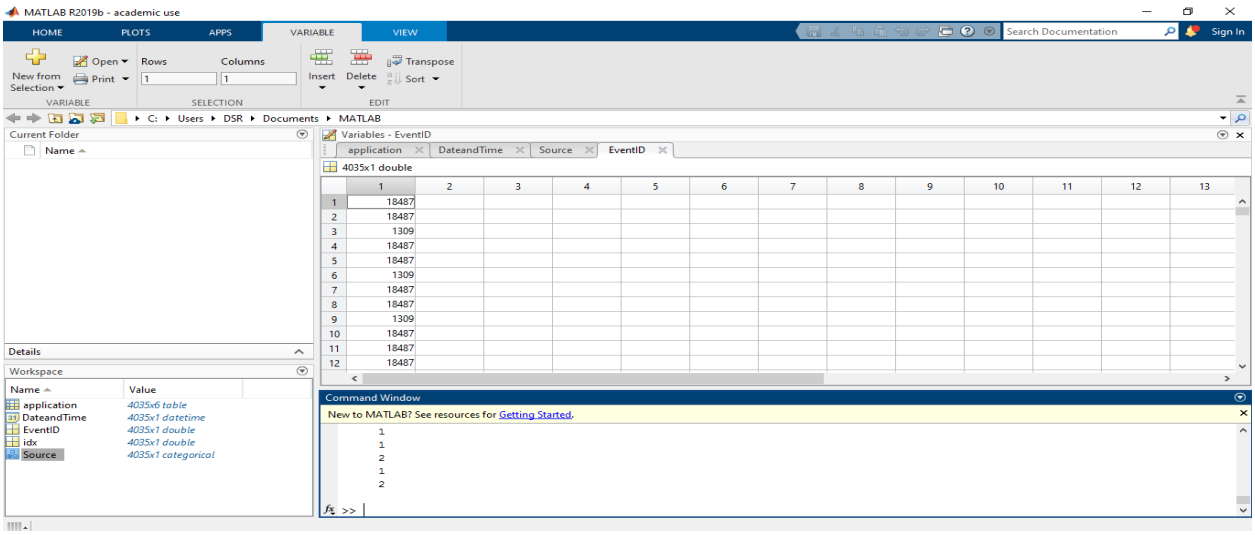

Fig-6: Implementation of K-medoids

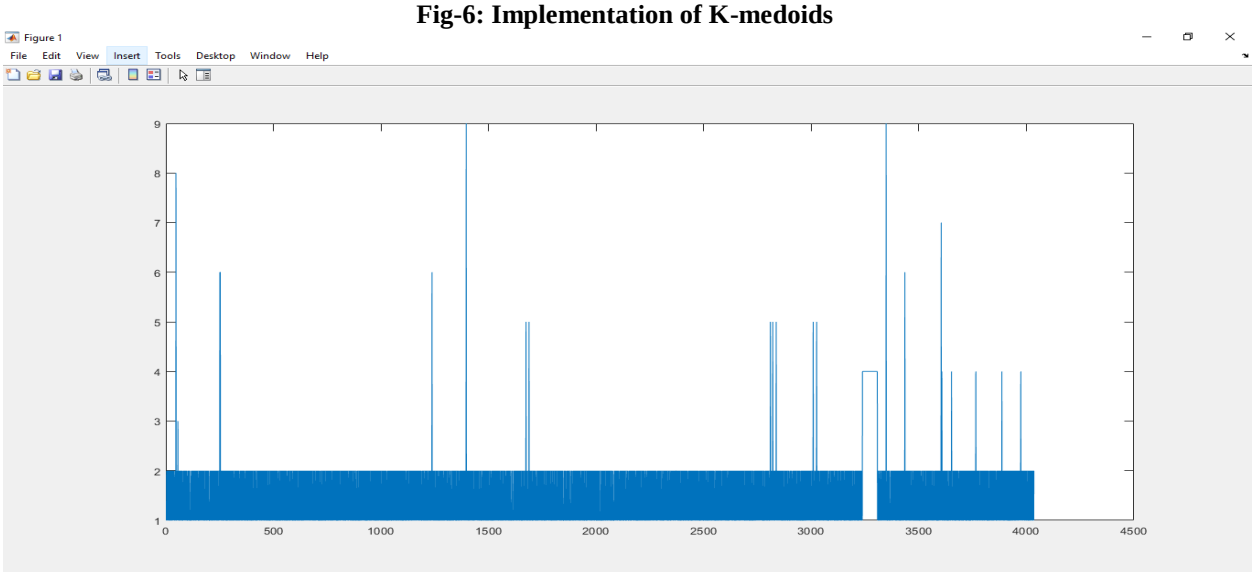

Fig-7: Events cluster graph

V. RESULT

For optimization of cloud servers; we have taken the events from the event log of cloud servers, preprocess the data of event logs to extract the events. Then apply the K-Medoid clustering algorithm on event log which provides cluster wise events. Cluster with least Medoid value is most suitable for implementation of Map-Reduce algorithm. With the application of the map-reduce algorithm we have reduced the events in the obtained cluster significantly by extraction of unique events.

VI. CONCLUSION

With the application of K-Mediod and Map-Reduce optimization 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 map-reduce concept it can allow. Map-reduce manage 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, (2015). "Networks and Cloud", CS5412 Spring (Cloud Computing: Birman).
2. R. K. Trivedi, R. Sharma, (2014). "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.
3. M. Arif, T. Mahmood, (2015). "Cloud Computing and its Environmental Effects", International Journal of Grid Distribution Computing Vol.8, No.1, pp.279-286.
4. A. More, P. Kanungo, (2017). "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.
5. Y. G. Patil, P. S. Deshmukh, (2018). "A Review: Mobile Cloud Computing: Its Challenges and Security", Vol.06, Issue.01, pp.11-13.
6. M. K. Saggi, A. S. Bhatia, (2015). "A Review on Mobile Cloud Computing: Issues, Challenges and Solutions", International Journal of Advanced Research in Computer and Communication Engineering.
7. R. V. Dharmadhikari, S. S. Turambekar, S. C. Dolli, P K Akulwar, (2018). "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.
8. M. Gattulli, M. Tornatore, R. Fiandra, A. Pattavina, (2012). "Low-Carbon Routing Algorithms for Cloud Computing Services in IP-over-WDM Networks", IEEE ICC Optical Network and Systems.

9. Malathy, G. R. Somasundaram, K. Duraiswamy, (2013). "Performance Improvement in Cloud Computing Using Resource Clustering", Journal of Computer Science 9 (6): 671-677, ISSN: 1549-3636.
10. D. K. Sharma, S. K. Dhurandher, A. Kumar, A. Kumar, A. K. Jha, (2016). "Cloud Computing based Routing Protocol for Infrastructure-based Opportunistic Networks", CAITFS Division of Information Technology.
11. S. N. Bushra, A. C. Sekar, (2014). "An Efficient Clustering Method for Incremental Cloud Data", IJARCSSE ISSN: 2277128X.
12. E. Sarkar, C. H Sekhar, (2014). "Organizing Data in Cloud using Clustering Approach", International Journal of Scientific & Engineering Research, Volume 5, Issue 5.
13. I. Singh, P. Dwivedi, T. Gupta, P. G. Shynu, (2017). "Enhanced K-means clustering with encryption on cloud", IOP Conf. Series: Materials Science and Engineering 263**,** 042057, 14th ICSET.
14. Yong Chul Kwon, Wing Yee Lee, Magdalena Balazinska (2008). "Clustering Events on Streams using Complex Context Information.", Data Mining Workshops ICDMW.
15. Stella Gatzuigrivas, Marc Schaaf, Michael Kaschesky, Guillaume Bouchard, (2011). Cloud-based Event-processing Architecture for Opinion Minning, IEEE World Congress on Services.
16. Rajender Kumar Trivedi, Rajani Sharma, (2014). "Case Study on Environmental Impact of Cloud Computing", OSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 2, Ver. VI, PP 81-86.
17. Nickolayev, O.Y., Roth, P.C., & Reed, D.A. (1997). Real-Time Statistical Clustering for Event Trace Reduction. *International Journal of High Performance Computing Applications, 11*, 144 - 159.
18. Conallen, James. (2003). Building Web Applications with UML.
19. Schwabe, D., Rossi, G., (1998). : An Object Oriented Approach to Web Applications Design. TAPOS 4(4).
20. Manolescu, I., Brambilla, M., Ceri, S., Comai, S., Fraternali, P., (May 2005). "Model-Driven Design and Deployment of Service-Enabled Web Applications", TOIT, Volume 5, number 2.
21. Trivedi A. K., Bharti A. K. & Shukla S., (2019). "An Approach To Handle The Events Of Cloud Computing For Better Environmental Impact" in Global Journal Of Engineering Science And Researches *[Trivedi,* 6(3): March 2019] ISSN 2348 – 8034 DOI-10.5281/zenodo.2584306.

AUTHOR PROFILE

Mr. Ajay Kumar Bharti is Professor and Dean in the faculty of computer science, MUIT, Lucknow, U.P., India. His research interest is in SOA, ICT and E-Governance, cloud computing. He has published number of research papers in International journal and conferences. He has also reviewed various reputed journals. He has more than 15 years of teaching and more than 6 years of research experience.

Mr. Ashish Kumar Trivedi currently pursuing Ph. D. from MUIT, Lucknow completed MCA (2001), M. Tech. (2006), M. Phil. (2010). He has published several research papers in reputed International/National journals and conferences, his main research work focuses on cloud computing and its techniques, He has more than 15 years of teaching experience.