A CLOUD COMPUTING USING ROUGH SET THEORY FOR CLOUD SERVICE PARAMETERS THROUGH ONTOLOGY IN CLOUD SIMULATOR

Ashish Tiwari*, Amit Kumar Tiwari, Hukam Chand Saini, Amit Kumar Sharma, and Anoop Kumar Yadav

Central University of Rajasthan, Computer Science and Engineering,
NH-8,Bandarsindri, Kisangarh, Ajmer, Rajasthan
{ashishtiwari@curaj.ac.in,kumartiwariamit@gmail.com,
hukamchand@curaj.ac.in,amitsharma@curaj.ac.in,
anoop15690@gmail.com}

ABSTRACT :

Cloud Computing is an awesome technology which have go ahead and annex the computing world. The development of cloud computing embed its sprouting continually in the recent era. Cloud Computing have contrived clan very lavish comfortable to perform their chore. As its fundamental defnition says that As You Pay As You Go. In this delving work talk about the cloud simulator working in our algorithm which is prepared with the help of rough set theory. The algorithm is implemented in the cloud simulator in which cloudlets, datacenters, cloud brokers are created to perform the algorithms with the help of rough sets. The Ontology is the system consists of cloud services in which cloud service discovery system is maintained. At last there is the implementation of work done using the rough set in cloud simulator using net beans and sql as back end. The net beans is loaded with the cloud Sim packages in which some of the packages are prepared according to our algorithm and gives the expected output of the optimization using rough set as a new concept.

KEYWORDS :

Cloud Simulator, Rough Set Theory, Cloudlets, Datacenters , Cloud Ontology System and Parameters

1. INTRODUCTION

In the day today world Computing is moving in its seven heavens where it is making its place like a worm in a fruit. The most important thing to see here is its play a vital in businesses. By this rapid development in businesses most number of people wants to save their money, time,

space and performance etc. This increment done is due to the people requirements in the web applications, by these application they are enjoying the life and consuming the time and money. This may increment in the field of ebusiness by see the demand of people. In this world Cloud Computing is spreading very rapidly through internet. So from all over analysis the basic definition of the cloud computing can be given as the computing which tells about both the applications providing services with the help of internet and available scalable hardware and the software running on the systems may provide the services.[1] [2] Its basic definition tells as you pay as you go manner. There are many computing present in the clouds like public, private and hybrid. These computing play an important role in cloud computing concept. Here the cloud computing is dealing with the cloud service providers which play an vital role for the formation of the scheduling performed in the cloud computing.[4]
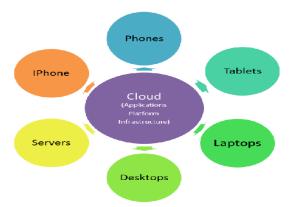

Fig. 1. Basic Structure of Cloud Computing.

There is a advanced study of the various cloud service providers and the parameters found after the survey may play an important role for the formulation. The Deep study of the standard organizations which are providing the various certification and protocols to perform their tasks in well ordered and organized manner.[3]

Cloud Ontology is the system that is maintained to determine the similarities between and among services.[9] Ontology when retrieving information about Cloud services. The relations among the different types of Cloud services and determining which service(s) would be the best or most appropriate service for meeting consumers service requirements.[7][8] On the calculated parameters there is the working of the cloud service providers and standard organizations are helping the services on that basis.
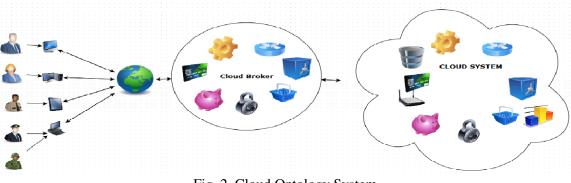

Fig. 2. Cloud Ontology System

2. PROBLEM STATEMENTS

The problem in cloud computing here deals with the cloud service providers selection of the formulated parameters. For some CSPs there may be certain services or activity that may be declining and not possible to be accessed at certain points during the day. This kind of unwanted selections defeats the purpose of the cloud. If we are investing a lot of projects in the cloud, then we expect the cloud service to be reliable. Even the top CSPs in the market have been facing the similar problems in the cloud environment. Here we have seen a lot of importance of CSPs for taking in account the features or characteristics such as Infrastructure, Platform, and Services. The method of selecting a Cloud Service Provider is evaluated on the basis of Which-Cloud Provider-Provides- What. The Basic Questions which comes in mind for the selection process are as follows.

- Why we will use this Provider?
- Does it satisfy our problems?
- How Costly It is May to Me?
- Is Data or information is safe in it?
- What Security features they are providing with it?
- What should be the limit of storage and Recovery of Data if Cloud or Office Destroyed as per act of God or Different accidents?
- Which type of access control is provided to the user as per safety purpose?
- How we can give the payments and at that level which security is applied?
- How will cloud Provider Checks that the user added in its Cloud is valid or not?

3. RELATED WORK

In this survey be having taken various parameters for the resource managements and users in benefits in terms of time and money.We have collected a set of parameter requirements and guidelines devised for cloud platform and service. The conformance of the cloud service platform and service considered for evaluation, with the requirements and guidelines is done by doing manual assessment. The algorithm internally calculates two scores based on the assessment results. One score is denotes the strength of parameter architecture of the cloud service, the second score denotes the fuzzy value associated with each service provider's. The requirements and guidelines has been collected in a technology and implementation through agnostic way.

Proper scheduling can have significant impact on the performance of the system. However, in general, the problem of mapping tasks on distributed services belongs to a class of problems known as NP-hard problems. For such problems, no known algorithms are able to generate the optimal solution within polynomial time. Even though the workflow scheduling problem can be solved by using exhaustive search, the complexity of the methods for solving it is very high. Upto now there are two major types of workflow scheduling, best-effort based and Quality of Services constraint based scheduling.

The idea of Adaptive Scheduling Algorithm is to divide the scheduling process into 2 phases: (1) a logical partition step and (2) a physical allocation step. The key motivation behind this multistage procedure is to reduce the complexity of scheduling problem. The Adaptive Scheduling Algorithm (ASA) [11] is proposed to find a suitable execution sequence for work

flow activities by additionally considering resource allocation constraints and dynamic topology changes. This algorithm utilises a multistage distribution algorithm to cope with network dynamics. [11]

The two major challenges exist in multimedia cloud are the resource cost and the service response time. [13] The author proposed queuing model to optimize the resource allocation for multimedia cloud in priority service scheme. There is also the formulation and solution of the resource cost minimization problem and the service response time minimization problem. The author has the priority service scheme in multimedia cloud datacenter as three concatenated queuing systems, which are schedule queue, computation queue and transmission queue. [13]

4. TECHNIQUES USED

Now the techniques of job scheduling and allocation to the cloud service providers gains a lot of attraction of the cloud researchers. The algorithms for task ordering on the basis of their arrival and task mapping are already proposed. We are proposing a similar kind of approach but to a level above that of job scheduling. It is a responsibility of the cloud computing service provider to provide the adequate service level satisfaction, so we are proposing to devise an algorithm which made possible the cloud middleware to determine capability of CSP by using Rough Set analysis on the basis of level of satisfaction of service. Rough set model can handle with objects and its characteristics.
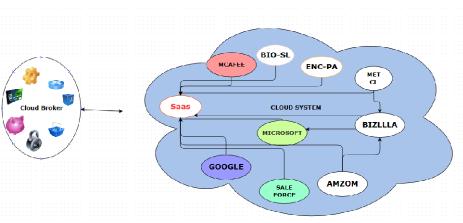

Fig. 3. Cloud Ontology System

4.1 Cloud Computing Parameters

On the above mentioned CSPs we have noticed out some of the important Services and the parameters according to that services? The Services and parameters are explained with its importance are as follows. The first one is Data Operation which deals with the different data operation such as searching of data either on the basis of content or location. Moreover some security and access mechanism issues are also discussed. Lastly, it also touches the data persistance methods. The second is Legal Issues in which we have found that it deals with the legal issues that could be generated when data is moved to the cloud, issues regarding NDA and other agreements between CSP and the customer and other issues that are resolved under laws and litigations of a country. The Third one is Risk Management in which we may deal about

Organizations shall develop and maintain a cloud oriented risk management framework to manage risk as defined in the master agreement or industry best practices and standards. The Fourth one is the Compliance And Audit in which Customer and provider both must understand the importance and should follow the implications on existing compliances, audits and other best practices. The fifth one is the Inter-portability and Portability is explained to interoperability features provided shall support security of the data exchanges and messages at protocol level, policy level and identity level. There shall be a mechanism to support integration of services, each of which has a different security mechanism. There are many more parameters found out but from all these are the best surved in our survey.

4.2 Cloud Service Providers

As per the Survey we have studied more than 30 CSPs and now going to represent the study on that CSP in detailed manner which are on the top of the levels according to the survey of Users and Functionalities given by gaining mind share. The Cloud Service providers are as follows according to their rating are as follows. Amazon Web Services (AWS), Second is Google Apps (AppEngine), Third VMWare vSphere (Virtualization), Fourth Rackspace Mosso, Fifth Cisco and so on.

4.3 Standard Organization

There are many standard organizations for cloud management and certifications provided by them for the safety purpose. Some of the Organizations are explained in detailed manner as follows.

The Cloud Security Alliance is made to provide the excellent security in Cloud Computing. The main aim is to make proper researches, Understanding, awareness, Education and many more functionalities. According to Cloud Security Alliance mission statement is "To promote the use of best practices for providing security assurance within Cloud Computing, and provide education on the uses of Cloud Computing to help secure all other forms of computing."

The Distributed Management Task Force (DMTF) it is an organization built in 1992 whose aim was on IaaS (Infrastructure as a Service) which is giving high performance, scalable, flexible infrastructure.

The National Institute of Standards and Technology (NIST) NIST is a non-regulatory federal agency whose goal is to promote innovation and United States competitiveness by advancing standards, measurement science, and technology. They are focused on helping federal agencies understand cloud computing. There are many more organizations in which we have gone through the survey and got help.

5. ROUGH SET MODEL

Rough set theory is a new mathematical approach to imperfect knowledge. The problem of imperfect knowledge has been tackled for a long time by philosophers, logicians and mathematicians. Recently it became also a crucial issue for computer scientists, particularly in the area of artificial intelligence.[14] The rough set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and

pattern recognition.[15] The formulation of upper and lower approximation are under Crisp Sets. The approximation of set may be Fuzzy Sets. The equivalence class of R determined by element x will be denoted by R(x). The indiscernibility relation in certain sense describes our lack of knowledge about the universe. Equivalence classes of the indiscernibility relation, called granules generated by R, represent elementary portion of knowledge we are able to perceive due to R.

Formal definitions of approximations and the boundary region are as follows:

1. R-lower approximation of X

$$R_*(x)S = \bigcup_{x \in U} \{R(x) : R(x) \subseteq X\} \qquad (1)$$

2. R-upper approximation of X

$$R^*(x) = \bigcup_{x \in U} \{R(x) : R(x) \cap X \neq \phi\} \qquad (2)$$

3. R-boundary region of X

$$RN_R(x) = \bigcup_{x \in U} \{R^*(X) - R_*(X)\} \qquad (3)$$

Rough set analysis will be on the basis of level of satisfaction of service Parameters. We know that all service providers are not all-rounder's i.e. they cannot be counted non-stop utilization for all kind of services. By applying Rough set model we are able to categorize each CSP on the basis of its speciality in service providing.

For each attribute $\rho\gamma i$ the value of

$$\begin{cases} \rho\gamma i \geq 1 & orTrue \\ 0 & ,Otherwise \end{cases} \qquad (4)$$

The value 1 and 0 above represents that either a CSP is specialized in a particular service or not. If the resultant value is 1 that means the CSP can be considered for the request otherwise not. We also categorise the attributes on the basis of relevance to tenant and CSP's. The relevance generates a threshold value for each attribute out of a scale of 10.

6. METHODOLOGY

As per the Rough Set representation, we have represented the CSP and their attributes in a tabular form called Information System. The rows of the table contain the list of cloud service providers and the columns consist of the attributes of the respective cloud service provider. When we talk about attributes that means the parameters we have identified in our study (Data Operations, Risk Management etc.)

$$\mu = \sum_{i=1}^{n} \delta i \qquad (5)$$

Algorithm

Input: Set of Jobs J and Set of Service Providers S with their associated resource and demand vectors Ri and Cv respectively, set of parameters. The Cloud Service discovery System (CSDS) is maintained. The values through Ontology (O) process we have calculated the values and put the actual values at mentioned in Matrix 1 and Matrix2.

Output: Optimal Allocation set of jobs to service Providers.

```
Method: S' = S
For all q belongs to set of jobs j.
{
        For (Cloud Ontology)
        {
                O = Select (Web-Pages formed by CSDS)
        }
        Select (Web-Pages formed by CSDS)
        {
                Find (ServiceUtility (Su) for Web-Pages)
        }
        S'=Su
        While ( S' ≠ φ )
        {
                r'= Select (q, r, S')
                S' = S'- r'
        }
        Select (q, r, S')
        {
                For each r, find maxcost(r) from available resources S'
                Rating is provided to the Resource(r) on the basis of Su of Ontology.
        }
}
```

Fig. 4. Algorithm Proposed by Us.

In rough set model we represent the CSP and their attributes in a tabular form. The rows of the table contain the list of cloud service providers and the columns consist of the attributes of the respective cloud service provider. This table 1 is converted into table 2. As per the rough set analysis processes table 2 is divided into lower and upper approximations. The lower one we cannot consider for assigning to the jobs because these are not suitable according to user standards. A very simple information system is shown in the following table 1, which contains rating of CSPs attributes ratings in a Cloud environment. These rating are just for example, as we can find the original ratings for original CSPs by performing some surveys. Here we considered seven service providers with four attributes: Application Security, Legal Issues, Virtualization and Compliance Audit.

```
The four Parameters are Virtualization, Security,Data Operation,Legal Issues
0        6        9        9
8        3        5        0
5        0        2        3
1        8        3        9
6        7        4        1
6        0        9        8
9        5        2        4
9        2        8        1
```

Fig. 5. Information System Table 1

```
Information About CSP & Parameters
The four Parameters are Virtualization, Security,Data Operation,Legal Issues
1        8        3        9
6        7        4        1
9        2        8        1
```

Fig. 6. Information System Table 2

```
Information About CSP & Parameters
0.11           1              0.38           1
0.67           0.88           0.5            0.11
1              0.25           1              0.11
This is The Final Fuzzy Value 0 = 2.49
This is The Final Fuzzy Value 1 = 2.15
This is The Final Fuzzy Value 2 = 2.36
```

Fig. 7. Fuzzy Cost Value.

7. CONCLUSION AND FUTURE WORK

In this research work we have proposed the scheduling for cloud service providers. After identifying the service requirements, the user can submit their job to the cloud and there the middleware can implement this mathematical model to rate CSPs and on the basis of their capabilities. In this system we are trying to develop the Ontology system which may give the actual values of the CSPs on the basis of the parameters taken for more efficient scheduling. The scope of this research can be extend further, in which apart from taking the cumulative cost function, we can also generate the cost on the basis of individual capability of CSP for an individual attribute. For example if a user needs a cloud that CSP whose cost function is maximum for security attribute rather than maximum value from the fuzzy costs.

REFERENCES

[1] J. Octavio, Gutierrez-Garcia, Kwang Mong Sim: A family of heuristics for agent based elastic Cloud bag of tasks concurrent scheduling. Future Generation Computer Systems, SciVerse Science Direct, doi:10.1016/j.future.2012.01.005.
[2] Daji Ergu, Gang Kou ,Yi Peng ,Yong Shi, Yu Shi on Pipeline: The analytic hierarchy process: task scheduling and resource allocation in cloud computing environment. In: Springer Science+Business Media, LLC 2011.

[3]   Jan Komorowski: Rough Sets: A Tutorial In .Department of Computer and Information Science Norwegian University of Science and Technology (NTNU) 7034 Trondheim, Norway.

[4]   Rodrigo N. Calheiros: CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services. In: Computing and Distributed Systems (GRIDS) Laboratory Department of Computer Science and Software Engineering The University of Melbourne, Australia.

[5]   Ashish Tiwari,Dr.A.Nagaraju and Mehul Mahrishi : An Optimized Scheduling Algorithm for Cloud Broker using Adaptive Cost Model: 3rd IEEE (International Advanced Computing Conference-2013).

[6]   Smitha Sundareswaran : A Brokerage-Based Approach for Cloud Service Selection In : 2012 IEEE Fifth International Conference on Cloud Computing.

[7]   Taekgyeong Han and Kwang Mong Sim : An Ontology-enhanced Cloud Service Discovery System : Proceedings of the International MultiConference of Enginers and Computer Scientists 2010 Vol I, IMECS 2010,March 17-19,2010,Hong kong .

[8]   Haytham Tawfeek al Feel : OCSS: Ontology Cloud Storage System : 2011 First International Symposium on Network Cloud Computing and Applications.

[9]   Use of Ontology in Cloud Computing, http://cloudontology.wikispaces.asu.edu/Use+of+Ontology+in+Cloud+Computing.

[10] Fangpeng Dong and Akl, S.G. : An Adaptive Double-layer Work ow Scheduling Approach for Grid Computing In : High Performance Computing Systems and Applications, 2007. HPCS 2007. 21st International Symposium on.

[11] A. Avanes and J. Freytag : Adaptive Workflow Scheduling under Resource Allocation Constraints and Network Dynamics In : Proc. VLDB Endow, 1(2), 1631-1637,2008.

[12] Radulescu, A. and Van Gemund, A. J C : Low-cost task scheduling for distributed memory machines In : Parallel and Distributed Systems, IEEE Transactions 2002.

[13] Nan, Xiaoming and Yifeng He and Ling Guan : Optimal resource allocation for multimedia cloud in priority service scheme In : Circuits and Systems (ISCAS), 2012 IEEE International Symposium.

[14] Rough Set Theory, http://www.en.wikipedia.org/wiki/Rough-set

[15] Rough sets, http://rst2009.disco.unimib.it/.

[16] Jan Komorowski ,Rough Sets: A Tutorial , Department of Computer and Information Science Norwegian University of Science and Technology (NTNU) 7034 Trondheim, Norway.

[17] Amit Kumar Sharma ,MADAM ID FOR INTRUSION DETECTION USING DATA MINING ,IJRIM , Volume 2, Issue 2 (February 2012).

[18] Mehul Mahrishi and Dr.A.Nagaraju ,Optimizing Cloud Service Provider Scheduling by Rough Set model,International Conference on Cloud Computing Techno and Management (ICCEAM-12).

[19] Mehul Mahrishi and Dr.A.Nagaraju ,Rating Based Formulation for Scheduling of Cloud Service Providers , National Confrance on Emerging treands in IT.

[20] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences, 11(5):341356, 1982.

[21] Mehul Mahrishi ,Globally Recorded binary encoded Domain Compression algorithm in Column Oriented Databases,Global Journals of Computer Science and technology Vol. 11 issue 23, Page 27-30.

[22] Amit Nathani, Sanjay Chaudharya, Gaurav Somani ,Policy based resource allocation in IaaS Cloud, Generation Computer Systems 28 (2012) 94103 doi:10.1016/j.future.2011.05.016.

[23] Daji Ergu, Gang Kou ,Yi Peng ,Yong Shi, Yu Shi ,The analytic hierarchy process: task scheduling and resource allocation in cloud computing environment, Springer Science+Business Media, LLC 2011.