International Journal of Electrical and Computer Engineering (IJECE)
Vol. 7, No. 4, August 2017, pp. 2176~2182
ISSN: 2088-8708, DOI: 10.11591/ijece.v7i4.pp2176-2182

Intelligent Hybrid Cloud Data Hosting Services with Effective Cost and High Availability

Madhu Bala Myneni, L V Narasimha Prasad, D Naveen Kumar
Departement of Computer Science and Engineering, Institute of Aeronautical Engineering, India

Article Info

Article history:

Received Feb 22, 2017
Revised May 27, 2017
Accepted Jun 11, 2017

Keyword:

Cost efficient
Data hosting
High availability
Intelligent hybrid cloud

ABSTRACT

In this Paper the major concentration is an efficient and user based data hosting service for hybrid cloud. It provides friendly transaction scheme with the features of cost effective and high availability to all users. This framework intelligently puts data into cloud with effective cost and high availability. This gives a plan of proof of information respectability in which the client has utilize to check the rightness of his information. In this study the major cloud storage vendors in India are considered and the parameters like storage space, cost of storage, outgoing bandwidth and type of transition mode. Based on available knowledge on all parameters of existing cloud service providers in India, the intelligent hybrid cloud data hosting framework assures to customers for low cost and high availability with mode of transition. It guarantees that the ability at the customer side is negligible and which will be helpful for customers.

Corresponding Author:

Myneni Madhu Bala,
Departement of Computer Science and Engineering,
Institute of Aeronautical Engineering,
Hyderabad, India 500073.
Email: baladandamudi@gmail.com

1. INTRODUCTION

Cloud data hosting services in the current web world existing steams, for example Windows Azure, Amazon S3, Google cloud storage has a place with incredible contrasts regarding working exhibitions and valuing arrangements. Choosing reasonable fitting excess system to store information with least cost and guaranteed accessibility assumes a noteworthy part is putting away of information. Many sellers build up their framework and continue updating them with recently rising innovations. The administrations can be gotten to by utilizing distinctive models like single specialist organization and numerous administrations suppliers. The issue with existing services is not secure means easy to hack and not ensured accessibility. Information hosting to cloud data storage disjoins the educating outline among many organizations and clients. Attributable clients related to its financial points of interest implies that the proprietor of the information moves its information moves to outsider cloud storage server which should apparently for a charge to store the information with it and give it back to the proprietor at whatever point required. As information era is far out driving information, it demonstrates expensive for little firms to regularly redesign their equipment at whatever point extra information is made. Additionally keeping up the stockpiles can be a troublesome errand. Putting away of client information in the favorable circumstances has many security concerns which should be widely researched for making it a solid answer for the issue of evading nearby stockpiling of information. Numerous issues like information validation and honesty outsourcing encoded information and related troublesome issues managing questioning over scrambled space were examined in this existing research. This research addresses the issue of executing a convention for getting a proof of information ownership. These proofs of data ownership schemes do not protect the data from exploitation by the collection. It just allows detection of altering or deletion of a remotely located file at an unreliable cloud

storage server. It just permits identification of altering or cancellation of a remotely found record at untrustworthy cloud data storage. In this paper efficient empirical based data hosting approach on hybrid cloud is proposed with the features of high availability and effective cost. This framework provides intelligent way of placing information into numerous mists with limited cost and ensured accessibility.

2. RESEARCH METHOD

Information Storage and incorporation has gotten a considerable measure at the information administration and application development stage. The issues incorporated with cloud applications and data storage and retrieval discussed in current research. Nowa days, the importance of rapid increase of distributed environments for large data sources and the respective benefits of cloud computing along with rapid increase of user's importance. The CAROM framework for combining the deployed replication and erasure coded schemes [1]. The content based multi hosting approaches for data hosting are implemented by using the CMO and the client adaptation algorithms to optimize both the performance and the cost. The experimental results shows the CMO algorithm effectively reduce the hosting cost [2]. The issue of multi cloud storage with an emphasis on accessibility and cost variables. As a result of more cost, the clients can't pick the services reasonably. There is a worry about moving vast measure of information into a solitary is like seller secure hazard. The management of information means the basic information divided into storing different mists and expecting with high information accessibility and security. The cloud computing highlights gives more advantages to the clients regarding minimal effort and accessibility of information is a primary variable. The failure of in time service availability is one more major issue in single cloud service [3].

A hybrid cloud environment is a mix of private, public and third party clouds with benefits of multiple deployments. Here this research addressed the issues related to single and multi-cloud platforms problems. Many new tools like Apache library provides a novel interface on various mists for helping organizations of multi- administrations. This technique helps in correspondence between various mists. Here the system in multi condition can't be thought about in light of the fact that it is demonstrated altogether different from the outcomes in two works. Here the creators manage the capacity of hub unsteadily. A comparable work on this framework, where replication in various server farms are examined and creator manage store the data in the essential server form. The research on cloud computing innovation has fundamental downside on seller secure. The benefit to designers won't permit to get benefit for nothing and does not permit to blend and match applications and administrations. Henceforth they presented a flow diagram with benefits for nothing of cost. The application and stacking the assets into this approach gives novel strategy to provisioning and robotizing administrations and furthermore applications run progressively on completely virtualized mists [4]. The security is a major concern for working with critical application and complex data. He proposed that the multiple clouds are compensations over single cloud. Trade of information from different mists and determination of mists in view of cost and administrations are discussed [5]. They checked on significant number of the ventures is needed in augmentation of embracing cloud computing innovation. The usage prompts to shakiness in range, for example, security and interoperability prompts to merchant secure. Consequently the institutionalization is presented which includes virtualization which assume critical part in cloud computing. The information facilitating plans said in our paper concentrate on various angles like seller secure, choosing appropriate information facilitating system, advancement of execution ensuring adaptable accessibility and security [6]. The author has given in previous publications the methodology for improving quality of data and comparison on computing techniques to add the intelligence with the help of machine learning algorithms [7], [8]. The CHARM is a one of the data hosting model in multiple cloud environments. This framework is applied to buid the model on various zones of clouds data throughout the world. By taking the reference from this framework, in this paper a hybrid cloud data hosting with high availability and less cost model is presenting on Indian data [9]. The effective framework based on component ranking for improving reliability of cloud applications. The mpact of component ranking is purely relay on two major factors as error propogation and internal failures. In their work, first they identified the significant components in application and then improving the reliability of these identified components leads to improvement of the reliability of cloud application. For improving significant components in the application more factors are influencing like timly response and availability of access [10]. In data hosting, one more influencing factor is secure stirage framework to provide secure access. In todays distributed storage environment, data sharing ensures whom to access and from whom to restrict access. They concentrated on overhead of execution when working with amazon S3. A successful execution of FADE model data storage framework ensures fine grained and stategy based control over distributed storage [11]. Social network data and cloud data contain important user personal information which needs high availability and security on this sensitive data. They addressed the security issue by using hierarchical attribute based encryption on cipher policy [12].

3. HYBRID CLOUD CHALLENGES

A hybrid cloud environment has few burning challenges that need to be considered: Network connectivity is one of the major challenges for connecting remote cloud service like public or private. Here not only the bandwidth, reliability and respective cost effectiveness are to be taken into account, but also the network topology also needs to be carefully designed. Manageability of different cloud services is another challenge in hybrid cloud environment. With multiple cloud services, their own management and environment setups. In hybrid cloud all these environments can be considered completely independent from each other. By having instances in different cloud services, there is no complete picture available showing the number of totally deployed instances. A transposition layer is a possible solution for this problem to provide a single interface for all clouds. The existing hybrid clouds mainly providing application and infrastructure services. The data integrators provided at the application layer. Thus a hybrid cloud is a combination of variety of clouds like private, public and a combination. All above discussed challenges are overcome with OpenStack standard API.

4. DATA HOSTING IN HYBRID CLOUD

4.1. Methodology

The methodology to data hosting scheme in hybrid cloud is shown in Figure 1.

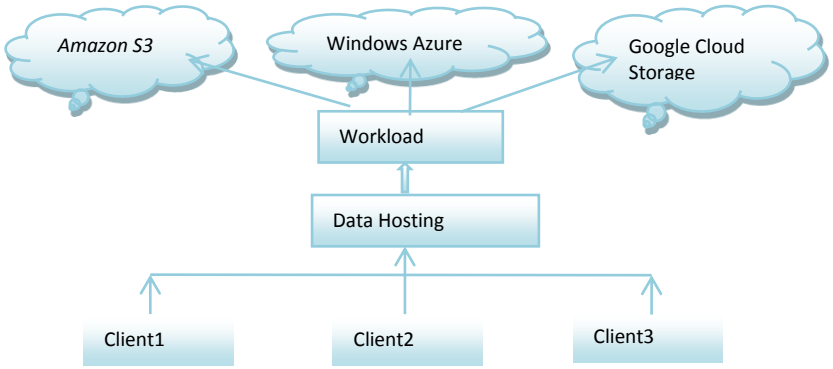

Figure 1. Hybrid cloud Data hosting Architecture

4.2. Storage Models of Majority Cloud Services

To know the pricing models of currently available cloud vendors in India, In this study three most current cloud storage services through the world: Windows Azure, Amazon S3, Google Cloud Storage are surveyed. The latest pricing models includes storage and bandwidth in India for 2017 are presented in Table 1.

Table 1. Prices of storage (in □ /GB/Month) and out going bandwidth(in □ /GB/Month)

	Windows Azure			AmazonS3	Google Cloud Storage
	Central India	South India	West India	Asia Pacific (Mumbai)	Standard Price
Storage (□ / GB / Month)	7.28	6.55	--	1.59	829.55
Band width (□ / GB / Month)	--	866.5	866.5	4.63	794.26

4.3. Hybrid Cloud

According to the majority of usage from above statistics multi clouds are available for data hosting. Data can be viewed by a user in some regions and availability of uploading data into multiple clouds is provided in intelligent way. So, the user can choose the suitable cloud according to his mode of transactions and available cloud with effective cost and high availability. Through this framework, the user can choose his required cloud with the comprehensive features of effective cost, easy use and high availability.

4.4. Transition Storage Modes

Transition storage modes are defined based on frequency of usage of files. Each cloud will specify how many ways they are providing different storage nodes. In general the storage modes are termed like Hot, Warm, Cool and Cold. The change in access frequency of the file falls under or rises above a threshold value, then need to change the mode of storage and save more money. These prices are determined time to time by the cloud vendors. From the existing clouds, the storage mode, input file's size, transaction type and active access frequency are collected. The available storage modes along with cloud storage providers in India are listed in Table 2.

Table 2. Storage models categorization

Storage class	Features	Providers
HOT	Available , durable, available and performance data storage for frequently accessed data	Amazon-S3, Microsoft Azure, Google Cloud Storage standard
WARM	Lower cost, less durability for frequently accessed non-critical reproducible data	Amazon-S3, Microsoft Azure, Google Cloud Storage
COOL	Storage class for data that is accessed less frequently	Amazon-S3, Standard I/A, Google Cloud Storage, Microsoft Azure
COLD	Secure, durable, low-cost storage service	Amazon Glacier

4.5. Data Hosting

In hybrid cloud environment, data hosting plays a major role to decide the storage mode and the relevant cloud service that the data to be stored in multi cloud environments. To choose the effective node here empirical algorithm is used as data hosting solution.

5. EXPERIMENTAL SOLUTION

The Empirical solution considers the four parameters as bandwidth; availability; storage space and operational prices are indicators of cloud performance. Each cloud value i is calculated based on these four factors. With this available information, the knowledge base is build. Among all available clouds, the most preferred n clouds are taken and then empirically interchange with the selected set for getting improved result. These four parameters are influence the effectiveness of cloud service. So different weights for each parameter is need to be provided due to the variation in size of a file effects operation price.

Algorithm 1: Empirical Model for Data hosting

```
Input: File size S, Access frequency Cr
Output:  min_cost Csm, X set of the identified cloud services
Csm ← infinity;
X ← {}
Ls ← sort(Pi) normalized αa_i + β(high to low )
for n = 2 to n  do
        Gs ← the first n clouds of Ls
        Gc ← Ls − Gs
    for m = 1 to n do
          Acur ← Availability(G_s)
          if Acur ≥ A then
          Ccur ← min(cost)
          if Ccur < Csm then
          Csm ← Ccur
          end
          else
          Gs ← sort(Gs,ai)   [ low to high ]
          Gc ← sort(Gc,Pi)   [ low to high ]
          for i = 1 to n do
                    f lag ← 0
                    for k = 1 to N - n do
          IF  G_c[k] > a G_s[i],  THEN
         swap (G_s[i], G_c[k])
          f lag ← 1
           break
```

```
                end
            end
          if flag = 0 then
          break
           end
          Acur ← availability(Gs )
          if Acur ≥ A then
           Ccur ← min(cost)
            if Ccur < Csm then
            Csm ← Ccur
            X ← Gs
            end
            break
            end
            end
        end
        end
        end
        return Csm, X
```

6. STORAGE MODES

The storage modes in cloud service are labeled as "cold", "cool", "warm", "hot" according to file transition frequency. For example, the storage mode will be changing from "hot" to "cold" according to the access frequency level reduces below the threshold value leads to the save money. The value is determined by the existing prices, which varies from time to time. The different storage mode cost provided by different vendors in India is shown in Table 3.

Table 3. Costs of different storage modes

	Amazon Web Service	Microsoft Azure	Google Cloud Storage
Hot	1.99	1.99	1.72
Warm	1.59	1.59	1.32
Cool	0.83	0.66	0.66
Cold	0.46	--	--

The storage mode is defined based on two dimensions like file_size and access_frequency. The each storage mode has respective pair of file_size and read_count, but for different pairs the storage modes are same. Hence explicit boundaries between different storage modes are existed. However, the change of the storage mode generates the cost and the bandwidth is more expensive than storage space for cloud storage services. For example, "the cost of one read access of a file can afford the file to be stored for around four months with no read access", thus, dealing storage mode transition is much careful and it affects the storage cost. The implementation of storage mode transition process is to get the data from the proxy clouds where the data is physically stored, and keeping the newly selected clouds with newly assigned mode of storage. The bandwidth and access operations like get,put and delete are parameters for implementation. The operation delete and ingoing bandwidth are cost free hence the transition cost T is collected on outgoing bandwidth, put and get operations. Hence outgoing bandwidth is more expensive than storage space. The transition cost is decided on mode of storage. So, the inequality in costs will rises:

$$\mathbf{M_f > M_P + T} \quad \textbf{(1)}$$

Where M_f is the previous storage mode cost and M_p is the new storage mode cost according to before and after transition. These two are calculated by considering read frequency provided. Consider, the file size as S, time period t and read frequency C_r ,set initial time t and then calculate M_f and M_p .

From Equation (1), the new storage mode will affects the transition cost can be observed. Here time period t of each transition is within one month. Then predict the cost of storage mode for each file using its predicted frequency of access with in the time interval t. If the storage mode is different from the earlier one and it uses Equation (1) again, means need of changing the storage mode of the file. The knowledge base of storage mode is pre calculated modes according to the no. of reads and writes, which can be calculated in advance, because it is only affected by the cost of storage mode with available cloud sources and their pricing policies according to timely changes. For finalizing the storage mode of each file, the determinant parameters

like access frequency and the size of the file having the same mode. Periodically the available cloud sources list will vary based on the availabilities, updated prices, strike out due to performance issue, and newly available cloud services appear. The updated cloud services data will be input to the Algorithm 2 and it returns the detailed transition process.

Algorithm 2: Data Storage mode transition

Input: Available Cloud services list T; ith file's storage mode M[i], current access frequency R[i], file size S[i]

Output: void

```
TSize ← Dimension of table T
TRead ← Access frequency of T
for each i  do
     for j in len(TSize)
          if S[i] ≥ TSize [j] then
          dS ← j
         else
           break
         end
  end
  for j in len(TRead)
      if R[i] ≥ TRead [j] then
          dR ← j
        else
          break
     end
 end
    if M[i] = T[dS][dR] then
       T ← Predicted_cost(M[i] )
    if M[i] > T[dS][dR] + T then
       transit from M[i] to T[dS][dR]
     end
  end
end
```

7. RESULTS AND DISCUSSIONS

7.1. Evaluation of Hybrid Cloud

Now a day, a lot of variety of information storage cloud suppliers is providing the services in keeping with specific regions like America, Asia, and sometimes many information centers are similar or completely different cloud suppliers. Therefore all the information centers will be accessed by the user with in a very bounded region, even the user can expertise completely diverse performance problems. The potential of some information centers is incredibly low whereas others are also unendurable. This framework chooses cloud services for hosting information from all the existing clouds that come across the performance claim, means they supply acceptable out turn and potential once they don't seem to be in outage. The performance of the service is not assured by storage mode. Since it's not a hidden method, it can reduce the importance of multiple transitions and batch wise implementation.

7.2. Datasets

The datasets are collected from multiple data centers like amazon, windows azure and google cloud. From each line of the data contains file access record with file access timestamp, file name, size of the file, and operation modes like get, put and delete. From storage mode table, the provision of cloud will assessed. With totally different file sizes varied from one computer memory unit to one GB and different browse counts varied from zero to a hundred with the step of zero.1 square measure accustomed produce the mental object of storage modes by mistreatment rule one. The knowledgebase includes the file browse.

7.3. Monetary Cost

The monetary cost of data hosting in clouds includes bandwidth, storage, operation and transition costs. So selection of clouds very randomly does not strictly increase the cost according to the availability of service.

8. CONCLUSION

This paper proposes, an intelligent approach for data hosting in hybrid cloud environment with effective cost and availability are the major features. This framework identifies two key components as bandwidth and key access modes to impact the data hosting in hybrid cloud. The empirical solution considers the major parameters asbandwidth, storage space required, operational prices and indicators of cloud performance. So different weights for each parameter is need to be provided due to the variation of a file effects operation cost. The storage mode is one more factor to influence the availability and operational cost of cloud service. The proposed algorithms ensures the intelligent way of data hosting according to user needs like storage mode and other available parameters for effective operational cost and high availability of users.

REFERENCES

[1] Ma Y, et al., "*An Ensemble of Replication and Erasure Codes for File Systems*," in INFOCOM. IEEE, 2013.
[2] Liu H, et al., "*Optimizing Cost and Performance for Content Multi homing*," SIGCOMM'12, 2012; 371-382.
[3] Mansouri Y, et al., "*R.: Brokering algorithms for optimizing the availability and cost of storage services*," In: Proceedings of the 2013 IEEE International Conference on Computing Technology and Science, Washington, USA, 2013;1; 581–589.
[4] Papazoglou M P, et al., "Blueprinting the Cloud," *IEEE Internet Computing*, 2011;74-79.
[5] Thandeeswaran R, et al., "Secured Multi-Virtual Infrastructure with Improved Performance", *Cybernetics and Information Technologies*, 12(2);11-22.
[6] Ortiz S Jr, "The Problem with Cloud Computing Standardization," *Computer*, 2011; 44(7); 13-16.
[7] Madhu Bala Myneni, et al., "*Correlated Cluster-Based Imputation for Treatment of Missing Values*," Proceedings of the First International Conference on Computational Intelligence and Informatics in: Advances in Intelligent Systems and Computing (AISC), 2017; 507; 171-178.
[8] Madhu Bala Myneni et al., "Comparative Analysis on Scene Image Classification using Selected Hybrid Features," *International Journal of Computer Applications*, 2013; 63; 0975-8887.
[9] Quanlu Zhang, et al., "CHARM: ACost-Efficient Multi-Cloud Data Hosting Scheme with High Availability", *IEEE Transactions on Cloud computing*, 2015; 3(3).
[10] Lixing Xue, et al., "*DCR: Double Component Ranking for Building Reliable Cloud Applications*," TELKOMNIKA, 2016;14(4);1565-1574.
[11] Naresh Vurukonda et al., "A Secured Cloud Data Storage with Access Privilages," *Indonesian Journal of Electrical Engineering and Informatics*, 2016; 4(3);219-224.
[12] Saikeerthana R et al., "Secure Data Storage and Data Retrieval in Cloud Storage using Cipher Policy Attribute based Encryption," *Indian Journal of Science and Technology*, 2015; 8(9); 318-325.