Recommendation for Data Migration

Prof. Rohit Bag[1], Apurva Somvanshi[2], Prathamesh Sahasra Bhojane[3], Shifa Shaikh[4]
Department of Computer Engineering, Sandip Institute of Technology and Research Center, Nashik
rvbag@hotmail.com, apurva.2209.ms@gmail.com, ssahasrabhojane@gmail.com, sshifa322@gmail.com

Abstract- The purpose of cloud storage is to provide virtual storage so as to save memory and overhead of buying external storage devices. Apart from this the data can be accessed from anywhere via internet. We can avoid emailing files to individuals and instead send a web link to recipients through our email. All cloud providers provide certain amount of free storage. If the data exceeds up to the limit it will be charges as per the amount and type of data. This proposed system is going to work on how to reduce the overhead of the acquirer to check the rate of every cloud provider. Aim of "Recommendation for data migration" is to recommend the acquirer to store his or her data on the cloud that will have minimum charge. This saves the overhead work of cloud acquirer to search for optimal cloud storage according to his/her need. This system will dynamically recommend user the optimal cloud service according to the current market.

Key Words: Data rate, Data migration, Cloud provider, portal.

1. INTRODUCTION

We as the individuals store our data on cloud which may or may not exceed the limit. Hence, we are least concern about the costing. Costing is important for those sectors who produce a bulk of data daily. And this data is needed to be stored on daily bases. Use of secondary storage devices such as the CD, DVD or pen drives are not enough. These devices have various drawback such as they may get misplaced or damaged and storing large amount will require a greater number of storage devices which will in turn cost high.

Talking about multinational companies these are the one that produce large amount of data on regular bases. This data is then stored on to the drive. Now it is but obvious that the data will exceed the limit and the company will be charges a certain amount. Varies drives are available and they have different rates so the company has to manually calculate and figure it out that which storage provider will charge minimum.

This proposed system is going to eliminate this over head of manually calculating the costing plus getting the recommendation of least charge drive. It will act as a mediator between the cloud provider and the acquirer.

2. PROBLEMS ENCOUNTERED

- Acquirer has to communicate with every cloud provider individually.
- The acquirer has kept the update of change in the rates made by cloud service provider.
- Secondary storage devices will cost high.
- Chances of data loss will be more.
- Accessing of data will always need that storage device which contain required data.

3. EXISTING SYSTEM

Survey says that for common acquirers who merely store their data on drive do not worry about the cost. They stored data within the limit is free of cost. But in consideration we observed that the sectors who produce the data in bulk have to consider costing.

Currently all the acquirer checks the charge per amount of data and then store it accordingly. But they find is bit difficult to do it manually.
Study shows that:

- **Google Drive**
  - 15GB- Free
  - 100GB- $1.99 per month
  - 1TB- $9.99 per month

- **Dropbox**
  - 2GB- Free
  - Up to 16GB- Free, if you refer a bunch of friends. With a basic free account, you get an extra 500MB per referral and you can earn up to 16GB total through this method. It works the same for paid users, but you get 1GB instead of 500MB per referral.
  - 1TB- $9.99 per month
- **I Cloud**
  - 5GB- Free
  - 20GB- $0.99 per month
  - 200GB- $3.99 per month
  - 1TB- $9.99 per month

International Journal of Research in Advent Technology, Special Issue, ICATESM 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

- **One Drive**
- ➢ 15GB- Free
- ➢ 100GB- $1.99 per month
- ➢ 200GB- $3.99 per month
- ➢ 1TB- $6.99 per month

For the above study we may say that as rates vary so acquirer then has to figure out where the data is to be stored so as to get minimum charge.

4. PROPOSED SYSTEM

- Multinational companies have to provide details of their data that is to be stored on cloud to the portal. I.e. type of data to be stored and amount of data that is to be stored or is stored.
- Similarly, the cloud provider has to provide rates of data depending upon the data size. I.e. rate per MB/GB/TB etc.
- Both the data received from company and the cloud provider is sent to the proposed software by the portal. As the software is at developer side no entity except the portal will access the software.
- The job of the software is to figure out which cloud cost least for provided type and amount of data. I.e. if google cost 9.9 $ for 20 GB and drop box cost 2.9$ for same amount of data then software should inform the portal about the profit to store data on drop box.

Then the portal will provide a recommendation to the respective company to migrate their data to the cloud with least cost.

5. OBJECTIVE

- To save the memory required to install the software is saved as the software remains at developer side.
- To eliminate the overhead to check the rates of cloud providers as the portal will recommend you the cloud with least rate.
- Manual work will be eliminated for selecting optimal cloud service.
- Boost the profit buy saving the capital required to save the data.

6. METHODOLOGY

Below diagram shows the work flow of the proposed system. There are various companies namely A, B, C…N who communicates with the portal. Details regarding data i.e. Type and size are sent.
One the other side there are cloud providers who submit their charges to the portal.
With both the details the portal will communicate with the software and further it figures out which drive is suitable.

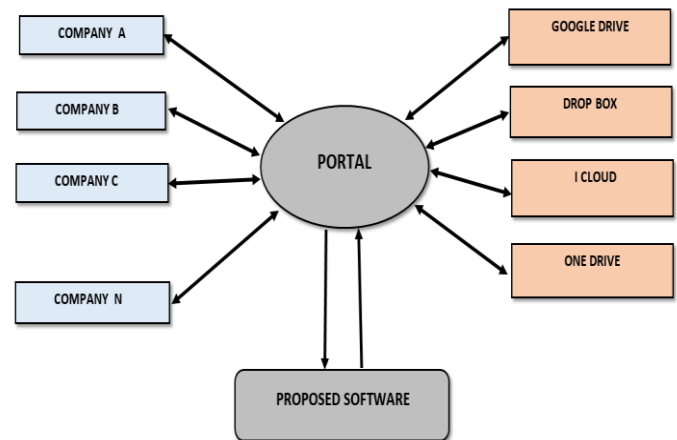Portal is informed about it and recommendation is provided.

FIG (1.1)

7. CONCLUSION

This proposed system will be highly beneficial for the cloud acquirer who use the cloud on daily bases. Use of this system will add up to their profit by saving the capital required to buy secondary storage devices. By elimination the manual work for selecting optimal cloud service and it will boost the profit.

REFERENCES

[1] Feng Youle, Zhu Liuzhang, "Analysis and Improvement of CEPH Dynamic Metadata Management[J]", Electronic Technology, September 2010

[2] T Bisson, J Wu, S A. Brandt, A Distributed Spin-down Algorithm for an Object-based Storage Device with Write Redirection [C] / /Proceedings of the 7th Work-shop on Distributed Data and Structures, pp. 459-468,2006.

[3] Shen Lianghao, Wu Qingbo, Yang Shazhou, "Research on Distributed Storage Energy Saving Technologies Based on Ceph[J]", Computer Engineering, August 2015.

[4] Age A Weil, Scott A Brandt, Ethan L Miller et al., "CRUSH: Controlled scalable decentralized placement of replicated data [C]", Proceedingsof the 2006 ACM /IEEE Conference on Supercom-uting(SC'06), November 2006.

[5] Forever. Ceph Way, Way Forever. Introduction of CRUSH data distribution algorithm based on Ceph. [EB/OL], October 2015.